Contemporary Issues in Income Distribution Research

Author:
Bradbury, Bruce

Publication details:
Working Paper No. 115
Reports and Proceedings
0733409105 (ISBN)

Publication Date:
1994

DOI:
https://doi.org/10.26190/unsworks/914

License:
https://creativecommons.org/licenses/by-nc-nd/3.0/au/
Link to license to see what you are allowed to do with this resource.

Downloaded from http://hdl.handle.net/1959.4/45225 in https://unsworks.unsw.edu.au on 2024-04-24

SPRC Reports and Proceedings
SOCIAL POLICY RESEARCH CENTRE No. 115 June 1994

Contemporary Issues in Income Distribution Research

edited by
Bruce Bradbury

THE UNIVERSITY OF NEW SOUTH WALES

For further enquiries about the work of the Social Policy Research Centre
or about purchasing our publications, please contact the Publications Officer, SPRC,
University of New South Wales, NSW, 2052, Australia.
Telephone (02) 385 3857. Fax: (02) 385 1049. Email: sprc@unsw.edu.au

ISSN 1036 2835
ISBN 0 7334 0910 5

As with all issues in the Reports and Proceedings Series, the views expressed in this publication do not represent any official position on the part of the Social Policy Research Centre. The Reports and Proceedings Series is produced to make available the research findings of individual authors, and to promote the development of ideas and discussions about major areas of concern in the field of social policy.

Foreword

In December 1993, the Social Policy Research Centre (SPRC), in conjunction with the Centre for Applied Economic Research at the University of New South Wales, hosted a one-day seminar on *Contemporary Issues in Income Distribution Research.* The six papers presented at the seminar covered a wide range of substantive and methodological issues in income distribution research, with a considerable empirical and international flavour. This volume contains revised versions of five of these papers, together with the comments from the Discussants.

In his paper, Michael Wolfson examines the nature of income distribution 'polarisation' - the so-called 'disappearing middle' thesis - and presents new evidence on trends in income inequality and polarisation in Canada. Bruce Bradbury then considers the extent of the change in poverty in Australia, but focuses on a different methodological issue - how to measure poverty changes without detailed knowledge of the relative needs of people in different family types. The paper by Atkinson, Gardiner, Lechene and Sutherland (presented by Holly Sutherland at the seminar) examines a number of methodological issues in the measurement and comparison of low incomes in yet another two countries, the United Kingdom and France.

The volume concludes with two papers which move beyond snapshot measures of inequality to a longer-term picture of incomes and well-being. Gábor Kõrösi, Russell Rimmer and Sheila Rimmer draw on data from the Australian Longitudinal Study to consider how gender and union membership affects the labour market outcomes of young Australians over time. Bruce Headey and Peter Krause consider the (often very loose) relationships in Australia and Germany between income, health and well-being and consider the implications of these for stratification theory.

Participants at the seminar were also fortunate to have Professor Tim Smeeding, Professor of Economics at Syracuse University and Director of the Luxembourg Income Study, provide a preview of the forthcoming OECD report, *Income Distribution in OECD Countries: The Evidence from the Luxembourg Income Study* (co-authored with Professors Atkinson and Rainwater). Unfortunately, we are unable to include his paper in the present volume as the OECD report is not expected to be released until later this year.

Together, these papers represent a significant contribution to the already substantial body of research on income distribution. The SPRC has played its part in the development of this area of research - both nationally and internationally - and I am delighted with our involvement in this seminar and the publication of its proceedings.

Peter Saunders
Director

Contents

Foreword i

Michael C. Wolfson
Divergent Inequalities: Theory, Empirical Results and Prescriptions 1

N. Anders Klevmarken
Comments 23

Bruce Bradbury
Poverty Measurement with Bounded Equivalence Scales: Australia in the 1980s 27

Michael C. Wolfson
Comments 53

A. B. Atkinson, Karen Gardiner, Valérie Lechene and Holly Sutherland
Comparing Low Incomes in France and the United Kingdom: Evidence from Household Expenditure Surveys 61

Deborah Mitchell
Comments 103

Gābor Kõrösi, Russell J. Rimmer and Sheila M. Rimmer
Contributions from Gender and Unions to Earnings Differences Among Young Australians: The Analysis of a Panel 107

Elizabeth Savage
Comments 129

Bruce Headey and Peter Krause
Inequalities of Income, Health and Happiness: The Stratification Paradigm and Alternatives 133

Peter Travers
Comments 173

Divergent Inequalities: Theory, Empirical Results and Prescriptions

Michael C. Wolfson[1]
Statistics Canada and Canadian Institute for Advanced Research

1 Motivation

This paper is principally about methods of income distribution analysis, particularly the foundations for claims about the extent or trend in inequality. The role of income inequality, and social inequalities more generally, has been central to public policy discussions for millennia - at least since Plato warned that too wide a gulf between rich and poor would lead to civil war. There has been a resurgence of interest in this topic in the last decade, as inequality appears on the rise in the advanced economies.

Moreover, there appears to be a shift in thinking as to the underlying causes. The optimistic view in the immediate post World War II period was that the rising tide of technical and economic progress was lifting everyone. However, we have endured more than a decade of economic stagnation in Canada and the US, and a new phenomenon of the 'disappearing middle class' has been identified. Previously the source of optimism, technical change - in the forms of de-industrialisation, computerisation, and global technology diffusion - is increasingly seen as the culprit for rising social inequality.

Underlying these changing views, there has been a major increase in analysis of income distribution trends. Unfortunately, the expansion of analysis has been accompanied by a somewhat undisciplined expansion in statistical methods. The result, occasionally, is prose conclusions which are not supported by the statistics cited. The sources of such divergences between evidence cited and conclusions claimed are the topic of this paper.

---

1 I am greatly indebted to Tony Atkinson for suggesting a collaboration with James Foster to probe more deeply the question of measuring polarisation, to James Foster for our joint work in developing the measurement concepts, to Brian Murphy, Geoff Rowe and Milorad Kovacevic for support on the empirical and statistical work, and to my CIAR Economic Growth Program colleagues for prompting a critical review of some of the recent literature on wage inequality. I remain solely responsible for any errors or omissions, and for the views expressed. A substantially shortened version of this paper is Wolfson (1994).

One major divergence concerns the fundamental meaning attached to the notion of inequality in the distribution of income. Further divergences between conclusions and evidence concern the particular statistical measures, populations and income definitions used to capture the intended concept. A final concern is the margin of error in the measures commonly used to support claims of trends in income inequality. We shall consider these points in turn.

2 First Divergence: Fundamental Concepts

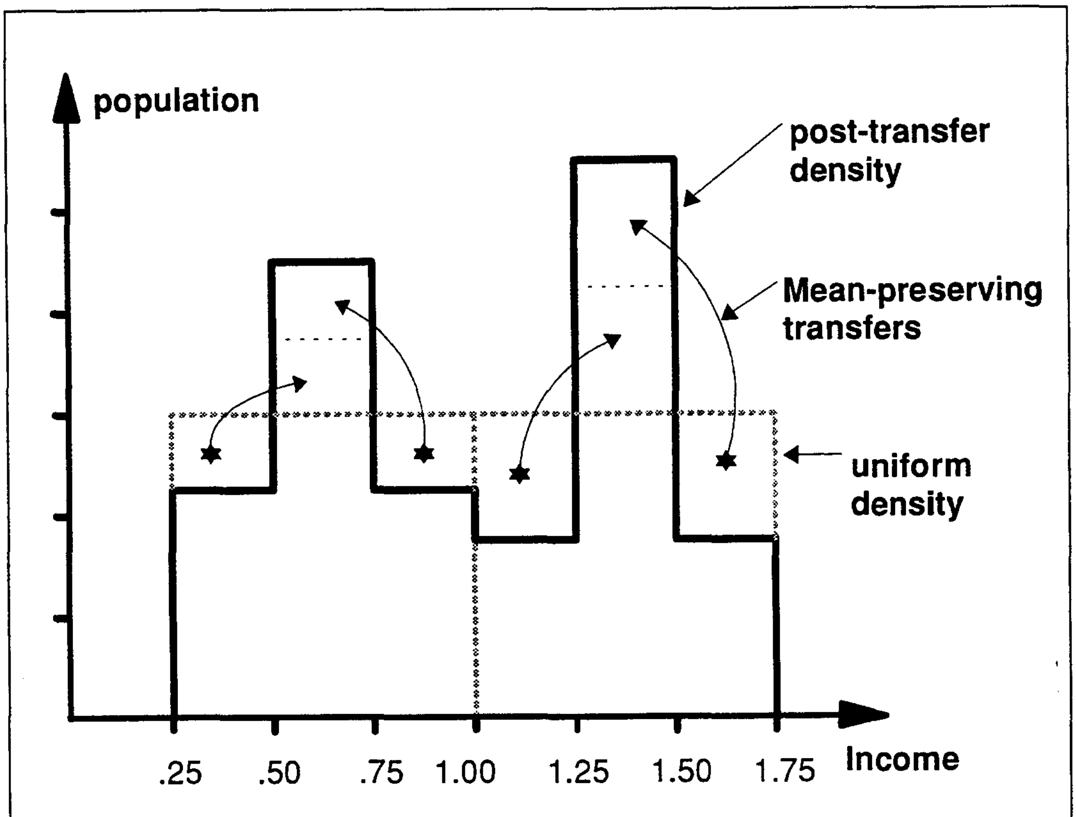
A significant innovation in discussions of income inequality is the addition, since the early 1980s, of the 'disappearing middle class' (e.g. Kuttner, 1983; Thurow, 1984). This concept is typically equated with the concept of increased income inequality. However, such an equation entails a fundamental conceptual error. With apologies to Levy and Murnane (1992), a few quotes from their recent survey of trends in US earnings inequality indicate the problem:

> ... a polarization of the earnings distribution means a decline in middle class jobs .... Despite the variety of scalar (inequality) measures, none seems well suited to the proposition of a vanishing middle class. That proposition refers to a polarization in which observations move from the middle of the distribution to both tails. Standard inequality measures cannot distinguish this polarization from other kinds of inequality .... If the middle of the male earnings distribution was hollowed out, that fact would be registered by scalar inequality measures (Levy and Murnane, 1992: 1338, 1339 and1351).

Levy and Murnane thus recognise, correctly, that polarisation - a shorthand for the phenomenon of the disappearing middle class - is not like the usual notion of inequality. However, they continue to use conventional scalar measures of inequality to assess the extent and trend in polarisation.

Is this a problem? I shall argue that it is, first on theoretical and then on empirical grounds. Theoretically, Figure 1 from Wolfson (1989) (an earlier version was in Love and Wolfson, 1976) should dispel any doubts. This graph shows two hypothetical income distribution density functions. The first is a uniform or rectangular density over the interval 0.25 to 1.75, shown by a dashed line. The second density, shown by a solid line, is clearly bi-modal, and has a somewhat depleted middle. We would argue that according to any sensible definition of polarisation or disappearing middle, this latter density is the more polarised. Is it also more unequal?

Figure 1: Polarisation and Inequality



The answer is unequivocally no. The second bi-modal density has been constructed such that according to **any** inequality measure that is consistent with the Lorenz criterion, the 'gold standard' for the concept of inequality, it is more equal. In other words, the bi-modal density has a Lorenz curve that is closer to the 45 degree line than the Lorenz curve for the uniform density. The formal proof follows simply from the fact that the bi-modal distribution can be 'derived' from the uniform distribution (in several ways, one of which is) by two sets of progressive mean-preserving redistributive transfers in the sense of Atkinson (1970), as indicated by the arrows in Figure 1.

One set of equalising income transfers is from some of the individuals in the 0.75 to 1.00 part of the income range (let's call them Ps) to an identical number of individuals in the lowest part, 0.25 to 0.50 (the Qs). The Ps give the Qs portions of their incomes equal to half the difference between their incomes, 0.25 on average, so both the Ps and the Qs move to the 0.50 to 0.75 income range in the bi-modal distribution. Similarly, a subset of individuals in the highest part of the income distribution with incomes between 1.50 and 1.75 (the Ms, say), give an average of 0.25 of their income to an equally sized set of individuals in the upper-middle part of the distribution (the Ns), with incomes from 1.00 to 1.25. As a result of this set of progressive transfers, the Ms and Ns both end up in the same 1.25 to 1.50 income range of the bi-modal distribution.

Thus, by construction, the bi-modal distribution is at the same time more polarised and more equal than the uniform distribution from which it was derived. Polarisation and inequality are thus demonstrably different concepts, as first pointed out in Love and Wolfson (1976), and reiterated in OECD (1993).

This result leaves open the question of what statistics should be used to measure polarisation. In the literature on the disappearing middle, in addition to inequality measures, some authors have used quintile income shares, while others have used the fraction of the population in various income ranges defined in terms of the mean or median income - such as the proportion of the population with incomes within 25 per cent of the median. In fact, Figure 1 has been constructed in a particularly nasty way for these kinds of statistics.

Since the distribution is symmetric, the mean is equal to the median which is 1.0. It can be shown that the **income** share of the middle third of the bi-modal distribution is lower than the income share of the middle third of the uniform distribution, while the income share of the middle two-thirds rises in the transition to the bi-modal distribution. Thus, the income shares of various middle quantile groups are **not** necessarily consistent with any sensible formalisation of the concept of polarisation or disappearing middle. In turn, this means that the large number of papers purporting to analyse the disappearance of the middle class which use inequality indicators such as quintile shares (e.g. Levy, 1987; Beach, 1988) are simply unable to detect the phenomenon they claim to be studying.

Moreover, the share of the **population** with 'middle level incomes' is similarly perverse in this example, going up or down depending on how 'middle' is defined. This is easily seen by inspecting Figure 1. The population with incomes within 25 per cent of the mean = median clearly falls, but the population with incomes within 50 per cent of the mean = median rises. Thus, statistics that count the share of the population with 'near middle' incomes are also **not** necessarily consistent with a sensible definition of polarisation. For example, Thurow (1984) considered the proportion of the population with incomes between 75 and 125 per cent of the median in his analysis of the disappearing middle class, while Blackburn and Bloom (1985) in a similar analysis focused on the proportion with incomes between 60 and 225 per cent of the median.

This is an unsatisfactory situation. There is an expanding literature seeking to analyse the phenomena of inequality and the disappearing middle class, accompanied by an incoherent variety of statistical indicators. These are only rarely accompanied by a justifiable sense of unease about what precisely they are measuring. Moreover, we have just seen that perhaps the most basic axiom underlying the formal axiomatic theory of inequality measurement - the Pigou-Dalton condition of transfers which is formally equivalent to the Lorenz curve criterion - is inconsistent with the concept of polarisation, or the roughly equivalent notions of spreadoutness from the middle, or bi-modality that lie at the heart of the disappearing middle class phenomenon.

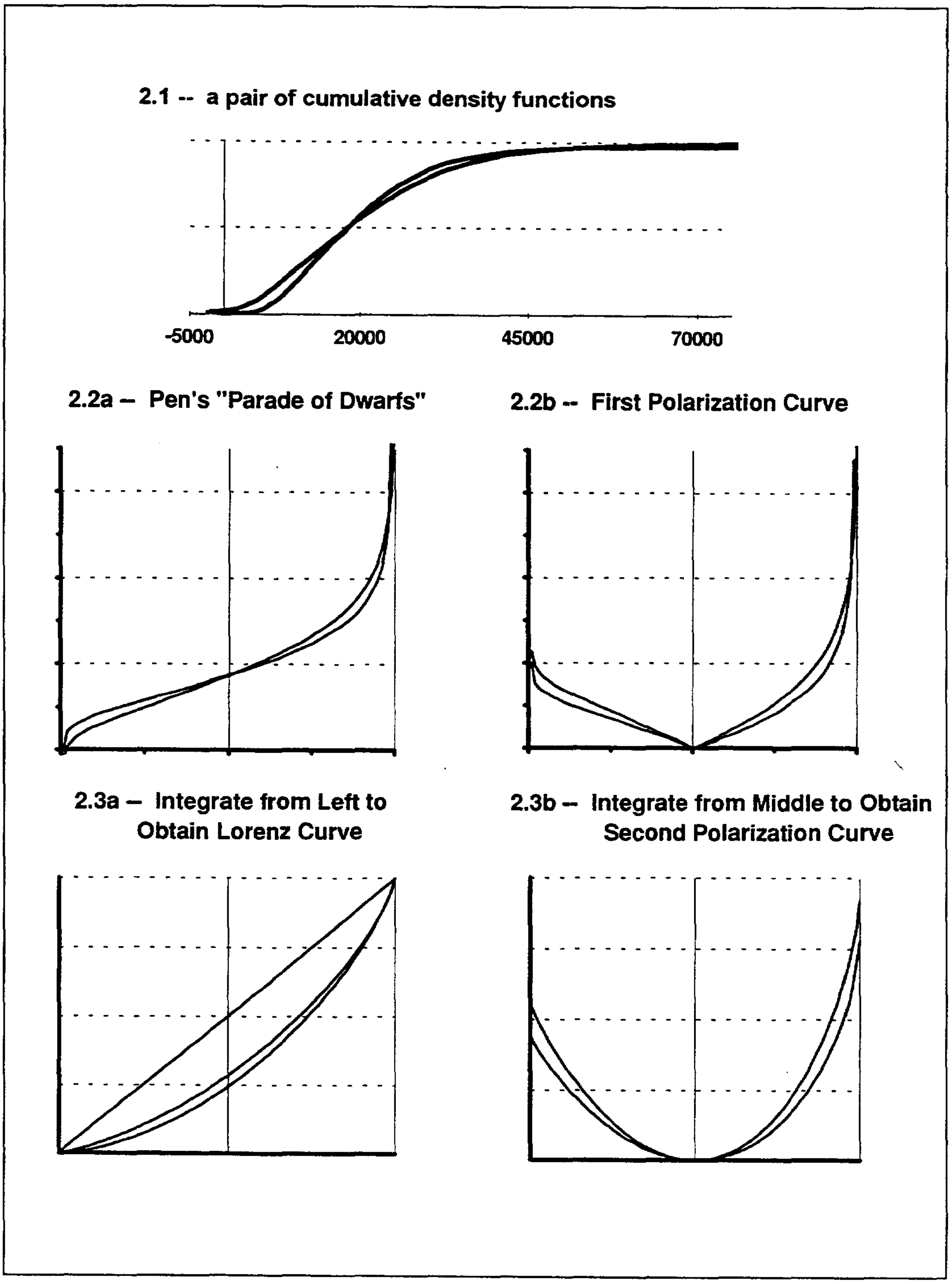
In this context, Foster and Wolfson (1992) sought to provide a formalisation of these latter concepts to give them the same rigour as we have for inequality measures. There is space here only to sketch the lines of development. It turns out that there is quite a nice duality or complementarity between polarisation and inequality. This is shown via a sequence of graphs in Figure 2. The strand of development in both cases starts with a cumulative density function (cdf) for the distribution of income (Figure 2.1). For inequality measures, we can think graphically of one intermediate step to arrive at the Lorenz curve. This step involves 'exchanging' the axes of the cumulative density function so population percentiles are ranged along the horizontal axis and incomes along the vertical. The result is Jan Pen's (1973) 'parade of dwarfs (and a few giants)', where Figure 2.2a shows the parade after dividing each individual income by the mean income. This normalised 'parade' curve is then integrated moving right from the origin to obtain the usual Lorenz curve, Figure 2.3a.

Formalising our concept of polarisation can follow a similar path of graphical transformations of the initial cumulative density function (cdf), but with a few key differences. After exchanging the axes of the cdf as above, individuals' incomes along the vertical axis are divided by the median (rather than the mean) income. The resulting median-normalised 'parade' is next cut at the mid-point of the horizontal axis, the 50th population percentile. The horizontal axis is then shifted up to touch the curve at this point, which is (by definition) the median income. The portion to the left (i.e. the first half of the 'parade' curve for the 50 per cent of the population with incomes below the median) is then rotated around the horizontal axis. The result is a curve looking a bit like a lopsided gull, Figure 2.2b. It shows, for any population percentile along the horizontal axis, how far its income is from the median, thus giving an indication of how 'spread out' from the middle (50th percentile) the distribution of income is. A less spread out distribution (i.e. one with a larger middle class) will have a curve that is lower.

For reasons that are intuitively similar to the notion of second order stochastic dominance, we can integrate this curve out from the mid-point along the horizontal axis (where, by construction, the height of the curve is zero) to get what Foster and Wolfson (1992) have called the (second or final) polarisation curve, Figure 2.3b. We have shown that one distribution has an unequivocally smaller middle class (is more spread away from the median) if, and only if, its polarisation curve is everywhere higher. This polarisation curve thus plays the same role for the concept of polarisation as the Lorenz curve plays for inequality.

It can then be shown that the area under this polarisation curve is a scalar index of polarisation, just as the Gini coefficient, as (twice) the area between the 45 degree line and the Lorenz curve, is a scalar index of inequality. However, as with the

Figure 2: Inequality and Polarisation, Parallel Strands of Graphical Development



Gini and Lorenz curves, it is still possible to have crossing polarisation curves. Thus, polarisation curves induce only a partial ordering over income distribution densities with respect to the sizes of their middle classes, while the area under the polarisation curve induces a complete ordering.

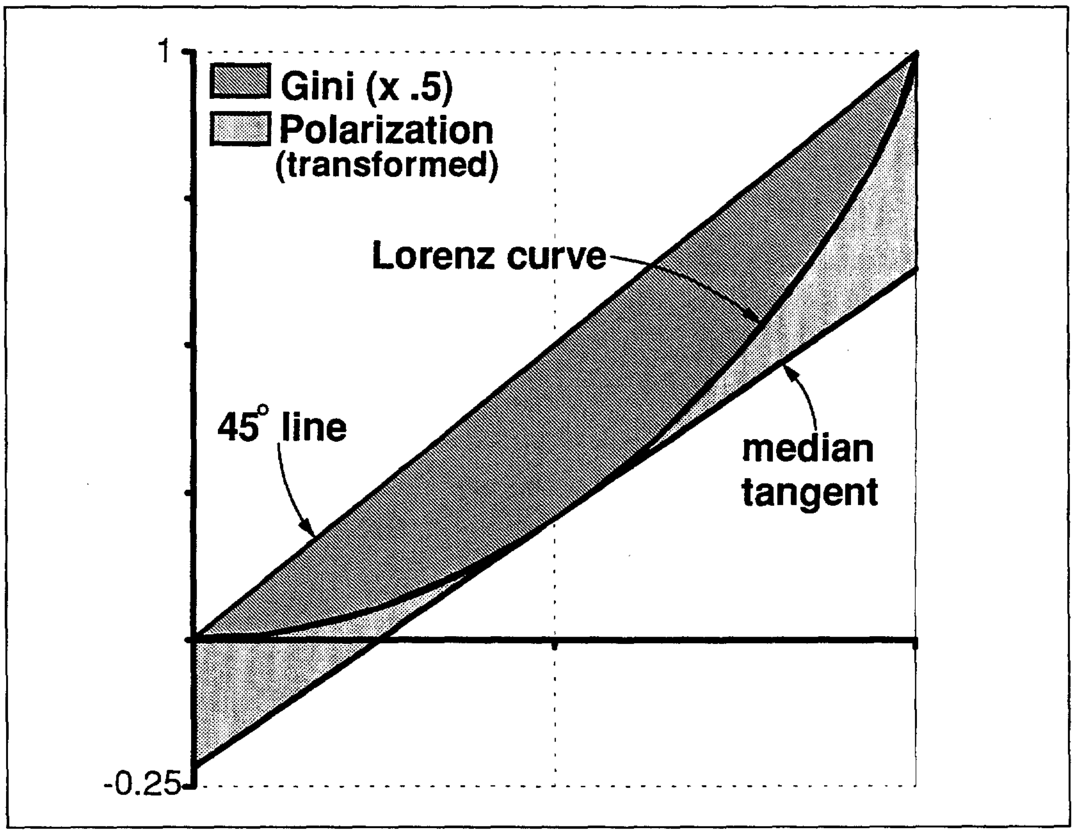
Moreover, these two strands of development, with their common starting point in the cumulative income distribution density function, can now be brought together again in a very nice extension of the Lorenz curve. Figure 3 shows a typical Lorenz curve. The key addition is the tangent line to the Lorenz curve at the 50th population percentile, with the vertical axis extended down to meet this tangent. It turns out that the polarisation curve just described is closely related to the Lorenz curve. If we first renormalise the vertical axis of the polarisation curve (Figure 2.3b) by multiplying by the ratio of the mean to the median, and then tilt the horizontal axis until it has the same slope as the tangent to the Lorenz curve at the 50th population percentile, the result is that this transformed polarisation curve is identical to the Lorenz curve!

In turn, the area under the polarisation curve in Figure 2.3b, let's call it P* corresponds to the lightly (dot) shaded area in Figure 3, and is a scalar indicator of the extent of polarisation or the size of the middle class. The lightly shaded area in Figure 3 between the tangent line and the Lorenz curve is T - Gini / 2, so that the P* of Figure 2.3b is (T - Gini / 2) / mtan; where mtan = 'median tangent' = m / μ = the slope of the tangent to the Lorenz curve at the 50th population percentile; m = median; μ = mean; and T = the area of the trapezoid defined by the 45 degree line and the median tangent = the vertical distance between the Lorenz curve and the 45 degree line at the 50th percentile = 0.5 - L(.5) = the difference between 50 per cent and the income share of the bottom half of the population (which latter, L(.5), we refer to as the 'median share'). P* has a minimum of zero for a perfectly equal distribution of income, and a value of 0.25 for a perfectly bimodal distribution with half the population at zero income and the other half at 2 μ .[2] In order to have an index with a similar range to the Gini (i.e. in the [0,1] interval if there are no negative incomes), we arbitrarily define our scalar polarisation index P = 4 P* = 2 (2 T - Gini) / mtan.

This diagram makes it immediately clear where the conflicts between inequality and polarisation arise, and why the concepts have so often been confused. If there is an 'equalising transfer' of income (in the sense of the Pigou-Dalton condition of transfers) from an individual above the median to an individual with income below the median (and the transfer is not so large that it causes either to cross the

2 Note that 0.25 is not necessarily the maximum. P* could exceed 0.25 if half the population had a negative average income. For the same reason, the Gini can exceed 1.0. For any given median share and mtan both positive, P* is minimised and approaches zero for a trimodal distribution where one individual has a very large negative income, another has a very large positive income, and everyone else has the same income in between.

Figure 3: A New Measure of Polarisation Based on the Lorenz Curve



median), then both inequality and polarisation decline - the Lorenz curve moves closer to the 45 degree line as does the tangent line at the 50th population percentile. By virtue of this class of examples, there are clearly many situations where inequality and polarisation rankings will agree.

The two concepts will disagree, however, when there are equalising transfers entirely on one side of the median - exactly as in Figure 1 earlier. In these cases, the median tangent curve is unaffected by the transfer, but the portion of the Lorenz curve on the affected side of the median moves closer to the 45 degree line. Such a shift in the Lorenz curve necessarily reduces the Gini coefficient, and correspondingly increases the polarisation measure P.

Such a divergence between inequality and polarisation could, of course, be merely a theoretical curiosum. An important question is whether in practice we may see divergent trends in the two kinds of attributes of income distributions. The answer is yes, and illustrations are provided later.

The demonstration that inequality as formalised is not always in accord with the concept of polarisation reopens the question of the axiomatic foundation of inequality measures. Specifically, it raises questions about the Pigou-Dalton condition of transfers. As noted by Amiel and Cowell (1989; fn. 14), Pigou was doubtful about its validity. Moreover, in their survey of almost one thousand undergraduate economics students (most before they had studied this topic), a majority rejected this axiom as part of their concept of inequality. At the very least, this suggests that in order to capture the concerns of the general public, summary measures based on concepts like polarisation should be given equal space along with Lorenz-consistent inequality measures when describing trends in income distribution. Indeed, polarisation as formalised here may be closer to the general public's vernacular concept of inequality than formal measures of inequality based on Pigou-Dalton-Lorenz-Gini concepts.

3 Second Divergence: Construct Validity

We have just shown a fundamental conceptual divergence between key ideas related to inequality - specifically the Lorenz curve / Pigou-Dalton concept - and to polarisation. But this is not the only 'divergent inequality' problem in extant statistical analyses. In social survey methodology, there is a well accepted criterion of 'construct validity'. Essentially, this means that the specific statistical measure being used, or set of survey questions, properly captures the intended underlying concept. For example, a set of questions only on wage rates would not be a valid way of determining the underlying construct or concept of labour force participation, while questions on whether or not the individual was working for pay or was actively searching for such work would be much closer. Many purported statistical measures of income inequality in current use fail this criterion of construct validity.

Let us consider only the one most widely agreed concept, that inequality is related to the partial ordering induced by the closeness of the Lorenz curve to the 45 degree line (notwithstanding the criticisms just made from the viewpoint of polarisation). The most serious failure of construct validity is when there is inconsistency - a more equal distribution according to the Lorenz ranking being measured as more **un**equal. Unfortunately, a number of summary measures are in wide use that fail on this count. The most common is the variance of logs (e.g. Karoly, 1992; Davis, 1992; Katz and Murphy, 1992). As pointed out both in Love and Wolfson (1976) and in Cowell (1977), for **equal**ising transfers above about 2.7 times the mean income (often at about the 95th percentile), the variance of logs will indicate an **in**crease in inequality.[3]

3 One factor that could account for the continued popularity of the variance of logarithms is its convenient decomposition. However, the Theil-Entropy, Theil-Bernouilli and squared Coefficient of Variation (defined below) also have 'nice' decompositions and do not suffer from this same construct **in**validity.

Another commonly used set of inequality indicators is ratios or differences in quantile income cut-points or their logs. For example, Davis (1992) focuses on the log of the ratio of a higher percentile wage to a lower percentile wage (e.g. 90th to 10th, 90th to median, and median to 10th), while the OECD (1993) and Atkinson et al. (1993) analyses are based on the unlogged ratios. These measures also fail the basic criterion of construct validity. A simple numerical example is a society with three individuals having incomes (1, 5, 9). If the middle individual gives one unit of income to the bottom individual, the resulting distribution (2, 4, 9) is clearly more equal according to the Lorenz curve criterion. However, the ratio of the 90th percentile income to the median increases, indicating an opposite direction of change.

This simple numerical example is admittedly extreme. It is much more likely that many of the situations where these income percentile ratios change are associated with crossing Lorenz curves. A solution to this form of construct invalidity is simply not to refer to these statistics as inequality indicators or measures of inequality. They do in fact have a straightforward interpretation: the ratio of any given income percentile to the median is simply the height of Jan Pen's (1973) median-normalised 'parade curve' at that population percentile (Figure 2.2a).

These examples of failures of 'construct validity' for purported inequality measures represent a second kind of divergent inequality: between statistical measure and underlying concept, or in other words between **indicator** (that which we are actually measuring) and **indicatum** (that which we intend to measure). There is really no reason at all to continue to use the variance of logs for inequality analysis; and ratios of income quantile cut-points should definitely not be called inequality measures, for the simple reason that they can give misleading results with respect to the underlying ordering implied by Lorenz curves. Again, it is reasonable to ask whether this is merely a theoretical curiosum; and again some empirical results presented below suggest it is not.

Given that many inequality measures in wide use are inappropriate and should be banished from rigorous analysis, what statistics should be used? It is generally impractical to produce myriad graphs of Lorenz curves on tracing paper (or a computer screen) and compare them visually. Our best advice - until drafting this paper - was a carefully chosen small set of inequality and polarisation measures.

For Lorenz inequality (i.e. indicators consistent with the partial ordering induced by Lorenz curves), a reasonable choice, to use Cowell's (1977) terminology, is one each of a bottom-sensitive, middle-sensitive, and top-sensitive inequality measure, each strictly consistent with Lorenz curve orderings.[4] If all three measures agree in a comparison of two income distributions, we can then be moderately sure that their Lorenz curves do not cross (at least to any substantial

4 A slightly different formulation of sensitivity to transfers at various points in the income spectrum, with similar conclusions for the choice of inequality measures, was developed in Love and Wolfson (1976).

extent). However, if they do disagree, we know the Lorenz curves do cross, and hence that no unambiguous ranking is possible.

Our preferred set of bottom-, middle- and top-sensitive inequality measures is the exponential measure (Exp), Gini, and (squared) coefficient of variation (CV) respectively. The only unfamiliar measure may be the exponential which was introduced in Wolfson (1986).[5] It was introduced precisely because it has the advantage over other bottom-sensitive measures like the Theil-Entropy, Theil-Bernouilli (also referred to as the Mean Logarithmic Deviation), and members of the Atkinson (1970) family that it does not explode with zero or near-zero incomes.

For the concept of polarisation, the obvious candidate is the P measure defined above.[6] A much simpler and more convenient set of measures is the proportion of the population with incomes, say, between 75 per cent and 150 per cent of the median, as well as a number of other similar ranges (e.g. 60 to 225 per cent). However, as noted in connection with the discussion of Figure 1, such individual measures are **not** necessarily consistent with the formal concept of polarisation that has been developed, namely in terms of the polarisation curve in Figure 2.3b.

Moreover, crossing polarisation curves are possible, just as are crossing Lorenz curves. Thus the complete ordering induced by our P statistic is not necessarily unambiguous. Also, despite the very close relationship between the Lorenz curve and the polarisation curve shown in Figure 3, crossing Lorenz curves do **not** necessarily imply similarly crossing polarisation curves (or vice versa). Intuitively from Figure 3, recall the polarisation curve depends not only on the 'curvature' of the Lorenz curve, but also on the slope and height of the tangent to the Lorenz curve at the median. More formally, if we define the polarisation curve at population percentile p as P(p) and the Lorenz curve as L(p), then $P(p) = (\mu / m) [L(p) - L(.5)] + (.5 - p)$. The clear implication is that the 'gold standard' for polarisation rankings should be inspection of the polarisation curves P(p).

One final point concerns ratios like the 90th percentile income to the median or 10th percentile income, for example as highlighted in Atkinson et al. (1993). As already discussed, these are **not** necessarily consistent - and can indeed be inconsistent - with the Lorenz curve-based partial ordering. It can also be shown that they are not necessarily consistent with the polarisation curve-based partial ordering either. These statistics therefore have no redeeming features for the

---

5 $Exp = \Sigma p_i \exp(- y_i / \mu)$ where p_i is the proportion of the population in the i-th income group, y_i is the average income in that group, and μ is the overall mean income.

6 The P measure is, like the Gini coefficient, in some sense middle-sensitive. Related families of polarisation measures that are analogues of bottom- and top-sensitive inequality measures should be quite easily definable. Neither P nor G is additively separable, but an additively separable class of summary polarisation measures analogous to Atkinson's (1970) class of inequality measures, for example, could be constructed.

measurement of formal concepts of inequality or polarisation. They simply describe a few ordinates on a median-normalised version of Pen's 'Parade of Dwarfs' curve (Figure 2.2a).

4 Further Divergences: Statistical Problems

A further set of divergent inequalities may arise for statistical reasons. One concern is sampling variability. The vast majority of analyses of trends in income inequality completely omit any consideration of the underlying sampling errors. A notable exception is Karoly (1992), but even here no account is taken of the fact that the underlying US March Current Population Survey sample has a complex multi-stage cluster design. Love and Wolfson (1976), using a method of half sample replication, estimated the variance of the Gini coefficient, and found that the complex sample design resulted in variances for family income about one and two-thirds times what one would have expected if the data had come from a simple random sample, i.e. a design effect of about 1.7.

More recently, Kovacevic (1994) has used an estimating equation approach at the level of sample clusters to estimate variances of inequality and polarisation measures, taking full account of the complex sample design. Essentially, the implied 95 per cent confidence intervals for summary distributional statistics like the Gini coefficient and polarisation P for the wage distribution data from the Canadian Survey of Consumer Finance used in this analysis (sample sizes on the order of 50,000 individuals) suggest as a general guideline that **at most** the first two digits of any of the widely used inequality measures have any statistical reliability; for a top-sensitive measure like the squared coefficient of variation, only the first digit has any statistical reliability.[7]

There are many other statistical questions that can influence results.[8] Given current interest in the distribution of wages, one important question is the population chosen for analysis. The general consensus in studies like those of Davis (1992) and Karoly (1992) is that wage inequality has been increasing over the past decade or more. However, the populations included in the analyses range from all individuals with positive labour income (both employment and self-

7 These are somewhat wider confidence intervals than those estimated by Karoly (1992) for the US Current Population Survey, particularly for the squared CV, though recall she did not take account of the complex sample design.

8 For example, no account is taken here of Rowe's (1994) recent finding that respondents frequently report their incomes in rounded amounts (e.g. to the nearest $100 or $1,000). Rowe estimates for similar earnings data to that used here that this rounding behaviour results in response errors in addition to and of the same magnitude as the sampling errors estimated by Kovacevic (1994).

employment) to only full-time, full-year male workers.[9] Moreover in Davis (1992), because of the limited international data available, the comparisons among countries may be contaminated by the quite different populations covered in each country's data. Thus, to carry on the theme of this paper, inequalities may diverge for the simple reason that like populations are not being compared with like.

5 Empirical Results

Four potential sources of divergence in inequality results have been noted:

- conceptual differences between Lorenz inequality and polarisation;
- construct validity problems for purported inequality measures which are not Lorenz consistent;
- differing populations of interest; and
- sampling variability.

In this section, empirical results from a time series of the Canadian Surveys of Consumer Finance are used to illustrate and assess these topics.

The divergent inequalities are illustrated with two sets of time series of distributional statistics: one for the distribution of labour income for full-time male workers, and one for the distribution of labour income for all individuals with annual labour income of at least 5 per cent of the average wage. These two populations are denoted 'FT Males' and 'All ELFPs' (ELFP = effective labour force participant) respectively. Labour income includes wages and salaries, military pay and allowances, and self-employment income (which may be negative). FT Males were age 18 to 64, worked at least 48 weeks, and indicated that they mostly worked full-time. All ELFPs include females, were also age 18 to 64, but had no other restrictions on their weekly hours or annual weeks of work.

Data for these two populations and for selected years are shown in Table 1. Two sets of statistics are given. The first is a set of inequality measures. The first three are the top-sensitive (squared) coefficient of variation (CV), the middle-sensitive Gini coefficient, and the bottom-sensitive Exponential measure. These are augmented by two further bottom-sensitive measures, the Theil-Entropy and the

9 There is a further problem of interaction between the use of bottom-sensitive inequality measures and the choice of population. If all strictly positive earners are included as in Karoly (1992), compared with a somewhat higher *de minimus* threshold like 5 per cent of the average wage, measures like the Theil-Entropy and Theil-Bernouilli could show spurious changes due to fluctuations in the sub-populations with only a few dollars of earnings. This problem has been encountered using Canadian data.

Table 1: Selected Inequality and Polarisation Indicators, Canada, 1967 - 1991

	Year				
	1967	1973	1981	1986	1991
All ELFPs					
Inequality					
Squared CV	0.582	0.609	0.544	0.620	0.667
Gini	0.363	0.379	0.378	0.396	0.403
Exponential	0.446	0.452	0.450	0.458	0.461
Theil-Entropy	0.246	0.261	0.252	0.278	0.289
Theil-Bernouilli	0.310	0.315	0.315	0.345	0.352
Variance of Logs	0.871	0.782	0.790	0.857	0.863
Polarisation					
population share (%) in income range					
75-150% median	42	37	36	32	32
60-225% median	66	64	62	59	59
range of income/median covering middle					
40-60% population	0.338	0.400	0.395	0.434	0.437
30-70% population	0.698	0.792	0.815	0.909	0.887
20-80% population	1.125	1.258	1.313	1.404	1.411
median share	0.243	0.229	0.226	0.213	0.210
median / mean	0.897	0.869	0.882	0.860	0.852
Polarisation (P)	0.338	0.376	0.385	0.415	0.417
FT Males					
Inequality					
Squared CV	0.335	0.275	0.235	0.289	0.332
Gini	0.264	0.246	0.242	0.263	0.272
Exponential	0.413	0.407	0.405	0.412	0.414
Theil-Entropy	0.139	0.117	0.109	0.131	0.140
Theil-Bernouilli	0.168	0.123	0.122	0.148	0.148
Variance of Logs	0.499	0.277	0.287	0.357	0.329
Polarisation					
population share (%) in income range					
75-150% median	59	60	58	56	54
60-225% median	81	84	84	79	79
range of income/median covering middle					
40-60% population	0.193	0.196	0.204	0.221	0.239
30-70% population	0.421	0.420	0.430	0.473	0.498
20-80% population	0.712	0.705	0.711	0.762	0.807
median share	0.318	0.327	0.327	0.313	0.307
median / mean	0.916	0.919	0.935	0.944	0.911
Polarisation (P)	0.221	0.217	0.222	0.236	0.250

Theil-Bernouilli measures.[10] Finally, the last statistic is the variance of logarithms - included only to show that it is **not** an inequality measure.

The second set of statistics is related to the concept of polarisation. The first five statistics count the proportion of the population with 'middle class' labour incomes, though from two different perspectives. The first pair, denoted 'population share (%) in income range (of) 75-150% or 60-225% median', give the numbers of individuals with incomes between 75 and 150 per cent of the median, and those with incomes between 60 and 225 per cent of the median, respectively. These statistics measure the size of the middle class defined in terms of a range of median-normalised incomes. The next three statistics effectively exchange the axes by defining the middle class in 'people space' rather than 'income space'. These statistics are based on symmetric percentile ranges of the population: within 10, 20, and 30 per cent of the 50th percentile, denoted '40-60% population', '30-70% population', and '20-80% population' respectively. For each of these 'people space' ranges, the corresponding range of incomes they span, divided by the median, is the statistic given. Thus, for example, if the figure for '40-60% population' is .338 (as shown for All ELFPs in 1967), this is the 60th percentile income minus the 40th percentile income divided by the median. Even though Figure 1 above shows that any one of these statistics may be misleading by itself, agreement amongst a set is more likely to indicate an unambiguous change in polarisation as we have formalised the concept.

The last three polarisation-related statistics all derive from the polarisation/Lorenz curve shown in Figure 3. The first of these is the 'median share' mentioned earlier: the share of income accruing to the bottom half of the population. This in turn is exactly the height of the Lorenz curve halfway along the horizontal axis, i.e. at the 50th percentile. Also, 0.5 - median share is the distance between the 45 degree line and the Lorenz curve at the 50th percentile, hence the area T of the trapezoid enclosing the Lorenz curve in Figure 3. The second statistic is the ratio of the median to the mean income, m / μ. In addition to the graphical interpretation of this being the slope of the tangent to the Lorenz curve at the 50th percentile, this ratio is also an indicator of the skewness of the distribution. Finally, the last statistic is the polarisation measure P defined above. Higher P means more polarisation, and a smaller middle class.

Let us now turn to an examination of Table 1 to explore the varieties of divergent inequality. For the time being, we ignore sampling variability and assume the distributions have been observed with infinite precision.

The first kind of divergence is between Lorenz inequality and polarisation. Generally, measures indicating the two concepts move in the same direction. But

---

10 Theil-Entropy = $\Sigma (y_i / \mu) \ln (y_i / \mu)$; and Theil-Bernouilli = $- \Sigma p_i \ln (y_i / \mu)$ where p_i is the proportion of the population in the i-th income group, y_i is the average income in that group, and μ is the overall mean income.

from 1973 to 1981 for both labour force definitions, all of the Lorenz consistent inequality measures decline or are constant; at the same time, almost all of the polarisation measures increase.

The second kind of divergence is the construct validity of the variance of logarithms. There are four instances of apparent divergences. For All ELFPs in the top half of Table 1, the variance of logarithms moves in the opposite direction to all the Lorenz-consistent inequality measures over the periods 1967 to 1973, and 1973 to 1981. A similar divergence is shown for the FT Males population from 1973 to 1981 and 1986 to 1991. We describe these as 'apparent' divergences because the inequality measures are only indicative; the 'gold standard' with respect to inequality is the relationship of the distributions' respective Lorenz curves.

Table 2 gives the underlying Lorenz curves for the 'All ELFP' series of populations. The columns of signs between the Lorenz curve ordinates for pairs of years indicate whether (+) or not (-) the Lorenz curve to the left is above the one to the right at that population percentile.

In both cases of divergence between inequality and the variance of logarithms for the All ELFPs populations, the Lorenz curves do cross. However, this crossing is marginal when comparing the 1967 and 1973 Lorenz curves, occurring somewhere between the 95th and the 97th percentile. To the right of this point (thinking of the Lorenz curves), the 1973 Lorenz curve is .001 higher than the 1967 Lorenz curve, even though all the inequality measures show 1973 being more unequal than 1967. However, for the bottom 95 per cent of the population, the 1973 Lorenz curve is everywhere below the 1967 curve, sometimes by as much as .016. This is therefore the clearest example of construct invalidity for the variance of logarithms.

A third kind of inequality divergence concerns the choice of population. As already noted, researchers like Davis (1992) have focused on full-time male populations. In principle, however, such a focus on only a subset of the working population may neglect the impacts of contemporaneous trends, such as increasing female labour force participation, an increase in part-time work, and changes in self-employment. In general, inequality among FT Males is lower and more stable than among All ELFPs. The clearest divergence in trends associated with choice of population is in the 1967 to 1973 period. Inequality moves significantly in opposite directions for the two populations. Moreover, polarisation is stable or declining over this period for FT Males, while it clearly increases for All ELFPs.

The final kind of divergent inequality is where authors interpret the data as showing trends where no trends exist because the changes are not statistically significant. Table 1 shows three digits for most of the statistics, while it was noted earlier that taking account of sampling error would leave at most two digits statistically significant, and often only one for top-sensitive measures like the

Table 2: Lorenz Curve Ordinates (%) at Selected Population Percentiles, All ELFPs, Canada, 1967-1991

Population Percentile	1967		1973		Year 1981		1986		1991
5	0.4	o	0.4	o	0.4	o	0.4	o	0.4
10	1.3	o	1.3	+	1.2	o	1.2	+	1.1
15	2.6	+	2.5	+	2.4	+	2.2	o	2.2
20	4.4	+	4.1	+	4.0	+	3.7	o	3.7
25	6.7	+	6.2	+	6.0	+	5.5	o	5.5
30	9.4	+	8.7	+	8.4	+	7.8	+	7.7
35	12.5	+	11.7	+	11.3	+	10.5	+	10.3
40	16.1	+	15.0	+	14.7	+	13.6	+	13.4
45	20.0	+	18.7	+	18.4	+	17.2	+	17.0
50	24.3	+	22.9	+	22.6	+	21.3	+	21.0
55	29.0	+	27.4	+	27.2	+	25.8	+	25.5
60	34.0	+	32.4	+	32.3	+	30.8	+	30.4
65	39.5	+	37.9	o	37.9	+	36.3	+	35.9
70	45.3	+	43.8	-	43.9	+	42.3	+	41.8
75	51.6	+	50.3	-	50.5	+	49.0	+	48.3
80	58.4	+	57.3	-	57.7	+	56.3	+	55.6
85	65.8	+	65.0	-	65.7	+	64.5	+	63.6
90	74.0	+	73.6	-	74.6	+	73.6	+	72.7
95	83.8	+	83.7	-	84.8	+	84.0	+	83.4
97	88.4	-	88.5	-	89.7	+	88.9	+	88.3
99	94.4	-	94.5	-	95.3	+	94.8	+	94.5

squared CV. More specifically, the following table from Kovacevic (1994) gives consistent estimates of the standard errors for the 1991 data in Table 1 taking account of the underlying complexities of the sample design, as well as the resulting design effects (measured as the ratios of the estimated variances to the variances that would have been estimated had the survey been based on a simple random sample):

Measure	Value	Standard Error	Design Effect
CV^2	.667	.0487	1.87
Gini	.403	.0027	3.54
Exp	.461	.0012	3.53
P	.417	.0054	2.18

These standard errors suggest that for All ELFPs, the only clearly statistically significant pairwise changes in Table 1 are increases over the 1967 to 1973 and 1981 to 1986 periods for the Gini, Exp and P measures.

These results based on summary measures of inequality actually raise a more difficult set of issues. The significant increases in inequality over the 1967 to 1973 and 1981 to 1986 periods are supported by inspection of the underlying Lorenz curves in Table 2, which do not cross substantially. However, the conclusion of no significant trend over the 1973 to 1981 period may be inappropriate because the underlying Lorenz curves do cross. The virtual constancy of the Gini and Exp mask offsetting shifts in the underlying distributions over the period. Moreover, these changes in the Lorenz curves are likely statistically significant. Analysis by Kovacevic (1994) indicates that for the 1991 data, the 95 per cent confidence intervals for the Lorenz curve ordinates in Table 2 are less than 0.2 per cent for the bottom third, 0.3 per cent at the median, and about 0.5 per cent in the top third of the population.[11] Thus, all the digits shown in Table 2 for the bottom third of the population, and the first two digits for the remainder of the population, are probably significant.

Given the limitations of summary measures of inequality in the situation of crossing Lorenz curves, and their sampling variability, what can we reliably conclude about trends in the Canadian distribution of labour income? Scanning across the rows of Table 2, the Lorenz curves are steadily falling for the first two-thirds of the population, indicating an increase in inequality. However, between the 75th and 95th percentiles, 1981 is a blip in this trend by showing lower inequality. Above the 95th percentile, this 'blip' diffuses out to the adjacent years, so that the 1986 Lorenz curve ordinate for the top five per cent, for example, is higher than the corresponding ordinates for 1991, 1973, and 1967 as well.

The story with regard to polarisation is a bit different, and clearer. Polarisation increased continually from 1967 to 1986. This is indicated in Table 1 by the falling proportions of workers with middle range incomes (both 75-150 per cent and 60-225 per cent of the median); by the generally widening range of incomes (as a proportion of the median) required to enclose the middle 20 per cent, 40 per cent and 60 per cent of the population (i.e. 40-60, 30-70, and 20-80th population percentiles); and by the polarisation measure P.

These trends are shown more rigorously in Table 3 which gives the polarisation curve ordinates for the same set of population percentiles as for the Lorenz curves in Table 2. With minor exceptions (beyond the third digit shown), the polarisation curves are steadily rising to 1986 and based on the standard errors estimated by

11 Similarly estimated 95 per cent confidence intervals for the polarisation curve ordinates in Table 3 range from 0.4 per cent for the first population quintile, declining to 0.08 per cent at the 45th percentile, 0.0 at the median, then rising to 0.08 per cent at the 55th percentile, 0.5 per cent at the 80th percentile, and 1.0 to 1.2 per cent in the top decile.

Table 3: Polarisation Curve Ordinates (%) at Selected Population Percentiles, All ELFPs, Canada, 1967-1991

Population Percentile	1967		1973		Year 1981		1986		1991
5	18.4	-	19.1	-	19.8	-	20.7	-	20.8
10	14.4	-	15.1	-	15.7	-	16.6	+	16.6
15	10.8	-	11.5	-	12.1	-	12.8	-	12.9
20	7.8	-	8.4	-	8.9	-	9.5	-	9.7
25	5.4	-	5.8	-	6.2	-	6.6	-	6.8
30	3.4	-	3.7	-	3.9	-	4.3	-	4.4
35	1.9	-	2.1	-	2.2	-	2.4	+	2.4
40	0.9	-	0.9	-	1.0	+	1.0	-	1.1
45	0.2	+	0.2	-	0.2	+	0.2	-	0.3
50	0.0	o	0.0	o	0.0	o	0.0	o	0.0
55	0.2	+	0.2	-	0.2	-	0.2	-	0.3
60	0.8	-	0.9	-	1.9	-	1.0	+	1.0
65	1.9	-	2.3	-	2.3	-	2.4	-	2.5
70	3.4	-	4.0	-	4.1	-	4.4	-	4.4
75	5.4	-	6.5	-	6.6	-	7.2	+	7.1
80	8.0	-	9.6	-	9.8	-	10.7	+	10.6
85	11.2	-	13.4	-	13.8	-	15.2	+	15.0
90	15.4	-	18.3	-	18.9	-	20.8	+	20.7
95	21.3	-	25.0	-	25.5	-	27.9	-	28.3
97	24.4	-	28.5	-	29.0	-	31.6	-	32.0
99	29.1	-	33.4	-	33.4	-	36.5	-	37.3

Kovacevic (1994), these changes are statistically significant. But according to the indicators in Table 1, polarisation appears to have remained generally stable over the period from 1986 to 1991. Table 3 generally supports this impression of stability over the most recent 1986 to 1991 period, but does show repeatedly crossing polarisation curves.

Tables 2 and 3 together show that both Lorenz curves and polarisation curves can and do cross in practice, so many rankings are ambiguous. These tables also show that the two kinds of curves do **not** cross in the same ways or at the same places, notwithstanding their close relationship.

Finally, this experience of analysing observed trends in summary measures of inequality and polarisation has been somewhat frustrating. Often, changes in summary measures are not statistically significant. But sometimes apparent stability among a threesome of inequality measures deliberately chosen to be sensitive to changes in inequality throughout the income spectrum can mask substantial changes in the underlying distribution of labour income, though these are changes that involve crossing Lorenz curves.

The main argument for the use of summary measures in the first place has been their convenience compared to the tedium of actually comparing the full Lorenz curves, even though this is the 'gold standard' for judging changes in inequality. This tedium argument loses force, however, with current computing power, and the widespread availability of microdata.

For example, we can draw on the notion of data visualisation. Table 2 can be thought of as the basis for a contour graph where time is displayed along the horizontal axis and population percentiles are arrayed along the vertical axis. Then, within the graph, contour lines would show the population quantiles corresponding to receipt of shares of, say 10, 20, 30, ... per cent of total income at each point in time. For example, inspection of the percentile data underlying Table 2 indicates that the proportions of the population needed to account for the first 10 per cent of income were 31, 32, 32, 34, and 34 per cent respectively for each of the years considered. In other words, these are the points along the horizontal axes of the sequence of Lorenz curves where each Lorenz curve has a height of 0.1. Visual inspection of such a contour graph would immediately reveal whether or not there were crossing Lorenz curves over many periods, not just in pairwise comparisons. The contours could also be drawn with lines whose thickness corresponded to their 95 per cent confidence intervals.

This new kind of Lorenz contour graph, as well as the corresponding polarisation contour graph, could at least become a standard complement to the usual summary measures of inequality and polarisation.

6 Concluding Comments

Widely used summary statistical indicators of inequality or the 'disappearing middle class' are potentially misleading. First, the fundamental concepts of inequality and polarisation are distinct and do not always rank distributions the same way. Second, some measures like the variance of logarithms do not measure what most people think: they should be banished from **inequality** analysis. Beyond these fundamental problems of clarity of concept and construct validity, claims made about trends in inequality may be inappropriate because they fail to account for sampling variability, or they should be more clearly circumscribed when only a sub-population like full-time male workers is being considered.

For all of these cases of potential divergence between evidence cited and conclusions claimed, examples have been given to show the salience of the problems. The implications can be summarised in a handful of 'do's and don'ts' for income distribution analyses:

- add P and/or related polarisation measures to the suite of statistical indicators used for analysis;
- banish the variance of logs and income ratios like the 90th to the 10th percentile from all discussions of inequality;
- consider only two digits of any inequality statistic (and only one digit for top-sensitive measures like the coefficient of variation) to be meaningful, unless you can show via careful estimation of confidence intervals taking full account of the complexities of the underlying sample design that more precision is statistically significant;
- try to use comprehensive and consistent populations for comparison, or at least present these results as background when focusing on sub-populations;
- continue to examine the underlying Lorenz and polarisation curves, for example as in Tables 2 and 3, as the 'gold standard' for unambiguous rankings; and
- more generally, try to take advantage of modern computing power to produce more comprehensive suites of statistical indicators and new kinds of tabular or diagrammatic methods for visualisation of trends.

References

Amiel, Y. and F. Cowell (1989), *Measurement of Inequality: Experimental Test by Questionnaire*, Discussion Paper No. TIDI/140, October, Suntory-Toyota International Centre for Economics and Related Disciplines, London School of Economics, London.

Atkinson, A.B. (1970), 'On the measurement of inequality', *Journal of Economic Theory*, 244-63.

Atkinson, A. B., L. Rainwater and T. M. Smeeding (1993), 'Income Distribution in OECD Countries: The Evidence from the Luxembourg Income Study (LIS)', mimeo, presented at the conference 'Contemporary Issues in Income Distribution Research', University of New South Wales, Sydney, Australia, December 3, 1993.

Beach, C.M. (1988), *The 'Vanishing' Middle Class?: Evidence and Explanations*, Queen's Papers in Industrial Relations No. 88/5, Industrial Relations Centre, Queen's University at Kingston.

Blackburn, M. L. and D.E. Bloom (1985), 'What is happening to the middle class?', *American Demographics*, January, 7(1), 18-22.

Cowell, F.A. (1977), *Measuring Inequality*, Philip Allan Publishers, Oxford.

Davis, S.J. (1992), *Cross-Country Patterns of Change in Relative Wages*, NBER Working Paper No. 4085, National Bureau of Economic Research, Cambridge, Massachussetts.

Foster, J. and M.C. Wolfson (1992), 'Polarization and the Decline of the Middle Class: Canada and the U.S.', mimeo, Vanderbilt University and Statistics Canada.

Karoly, L. (1992), 'Changes in the distribution of individual earnings in the United States: 1967-1986', *Review of Economics and Statistics*, 74(1), February, 107-15.

Katz, L. and K. Murphy, (1992), 'Changes in relative wages, 1963-1987: supply and demand factors', *Quarterly Journal of Economics*, 107(1), February, 35-78.

Kovacevic, M.S. (1994), 'Estimation of Standard Errors for Measures of Income Inequality and Polarization - Estimating Equations Approach', mimeo, Methodology Branch, Statistics Canada, Ottawa, January.

Kuttner, B. (1983), 'The declining middle', *The Atlantic Monthly*, July, 252, 60-71.

Levy, F. (1987), 'Changes in the distribution of American family incomes, 1947 to 1984', *Science*, 236, May, 923-7.

Levy, F. and R.J. Murnane (1992), 'US earnings levels and earnings inequality: a review of recent trends and proposed explanations', *Journal of Economic Literature*, September, 30(3), 1338-81.

Love, R. and M.C. Wolfson (1976), *Income Inequality: Statistical Methodology and Canadian Illustrations*, Catalogue 13-559 Occasional, Statistics Canada, Ottawa, March.

OECD (1993), *Employment Outlook*, Paris.

Pen, J. (1973), 'A parade of dwarfs (and a few giants)', in A.B.Atkinson, ed., *Wealth, Income and Inequality*, Penguin, Middlesex, 73-82.

Rowe, G. (1994), 'Income Statistics from Survey Data: Effects of Respondent Rounding', mimeo, Social and Economic Studies Division, Statistics Canada, Ottawa, January.

Thurow, L. (1984), 'The disappearance of the middle class', *New York Times*, February 5, sect 3, p. 2.

Wolfson, M.C. (1986), 'Stasis amid change - income inequality in Canada 1965-1983', *Review of Income and Wealth*, 32(4), December, 337-69.

Wolfson, M.C. (1989), 'Inequality and polarization: is there a disappearing middle class in Canada?', in *Proceedings of the Statistics Canada Symposium on Analysis of Data in Time*, October 1989, Statistics Canada, Ottawa.

Wolfson, M.C. and J. Foster (1993), 'Inequality and Polarization - Concepts and Recent Trends', paper presented to the Winter Meetings of the American Statistical Association, Ft. Lauderdale, January.

Wolfson (1994), 'When inequalities diverge', *American Economic Review*, forthcoming, May.

Comments on paper by Michael Wolfson

N. Anders Klevmarken
Department of Economics
Uppsala University

Michael Wolfson raises four issues which are important when evaluating changes in income inequality:

- polarisation;
- consistency with Lorenz curve ranking ('construct validity');
- the population concept (populations which are compared might be defined differently or they might change over time); and
- sampling variability.

I will offer you a few comments on each issue.

1 Polarisation

Wolfson has jointly with a co-author suggested a measure of polarisation (P) which is derived very nicely and it has an interesting relation to the Lorenz curve and the Gini coefficient. I have not seen the original report which presumably contains a more rigorous mathematical derivation so my intuition about the properties of the polarisation measure is primarily based on Figure 3. My understanding is that P=0 if everybody has the same income, **and** if everybody except one individual has zero income. The latter case is somewhat odd. Furthermore, suppose there are three groups of income earners. If the middle group decreases, P will increase, but what happens to P if the middle group disappears? (In this case there is no well defined tangent in Figure 3).

It would be nice to have a measure which is defined on the interval (0, 1). Although I have no formal proof my intuition is that P-maximum now is 1/4.[1]

There are a few aspects of polarisation which might be of interest but are not discussed in the paper: we are used to thinking that a relatively high

1 Wolfson's revised paper in this volume now includes more discussion of the bounds of the P index (ed.).

inequality is acceptable if income **mobility** is high. The same could be said about polarisation. If high polarisation is associated with high mobility it is a lesser evil. This raises the issue of the causes of polarisation. One might distinguish between a few different causes, for instance, technical change combined with a schooling system which does not allow for retraining might lead to permanent polarisation (segregation), while temporary leaves of absence from the labour force and decreases from full-time to part-time (maternity leave, sick leave, child care, unemployment) might lead to temporary polarisation. There is now in many Western countries a trend from a general social policy towards a means tested policy. Will this lead to a permanent polarisation?

2 Consistency with Lorenz Curve Ranking

Wolfson suggests that the variance of logs index should be banned from inequality analysis because it does not always satisfy the principle of transfers or Pigou-Dalton efficiency (i.e. it does not always increase if income goes from poor to rich). However, the variance of logs satisfies this principle for many well-known distributions **and** if a transfer takes place below or across the geometric mean. Hart (1980) suggested: 'The chances of a perverse result are small enough to be ignored in any practical work on income distribution, as shown by Creedy (1977). Thus the principle of transfers argument in favour of entropy or welfare economics measures cannot be given much weight.' Hart prefers the variance of logs because of its decomposition properties, it is a simple moment estimator and it is a convenient measure when modelling earnings functions.

It is true that Wolfson is able to demonstrate that conflicting changes do occur, but are they significant? There are no standard errors to measure the sampling variability of the variance of logs in Table 1.

3 The Population Concept

It is a well known problem that differences in concepts and population definitions make cross-country comparisons difficult (c.f. the LIS-project!). But also comparisons in the time domain might meet with difficulties. For instance, in 1990 the Gini index for equivalised disposable income in Sweden was 0.231 (official statistics based on tax return data). After the tax reform in 1991 it was 0.261. However, 0.015 units (0.264-0.231) were the result of a broadened tax base (previously exempt benefits became taxable).

4 Sampling Variability

Wolfson rightly notes that sampling variability might be important. For this very reason it would have been good to have estimates of standard errors in Table 1.

Standard errors of most inequality measures and the most common sampling designs can now be computed (see for instance, Nygård and Sandström, 1981; Sandström, 1983; and Frank Cowell's computer program INEQ).

A more basic issue is what kind of inference we are interested in, an inference to a finite population or to a 'superpopulation'? A 'normal' increase in the GINI does not necessarily imply that the income process works in a less egalitarian way.

References

Creedy, J. (1977), 'The principle of transfers and the variance of logarithms', *Oxford Bulletin of Economics and Statistics*, 39, 152-8.

Hart, P. (1980), 'The statics and dynamics of income distributions: a survey', in N. A. Klevmarken and J. A Lybeck, eds, *The Statics and Dynamics of Income*, Tieto Limited Clevedon, Avon.

Nygård, F. and A. Sandström (1981), *Measuring Income Inequality*, Acta Universitatis Stockholmiensis, Stockholm Studies in Statistics 1, Almquist and Wiksell International Stockholm.

Sandström, A. (1983), *Estimating Income Inequality. Large Sample Inference in Finite Populations*, Research Report 1983: 5, Department of Statistics, University of Stockholm.

Poverty Measurement with Bounded Equivalence Scales: Australia in the 1980s

Bruce Bradbury
Social Policy Research Centre
University of New South Wales[1]

1 Introduction

Whilst many concepts of poverty exist, a definition that encompasses most is that **people are in poverty when they have access to a level of economic resources which is particularly low relative to their needs.** It is reasonable to suppose that, all other things equal, a lower level of poverty in society is a universally held social goal.

Before poverty levels can be compared empirically at different points in time (or between different societies), there are a number of more specific issues that must be resolved. These include those listed below.

- **The measure of economic resources.** Whilst income is commonly used, there is a strong case for including other resources such as wealth (possibly including human capital). When income alone is used, the time period of measurement becomes particularly important. A very short time period may overstate poverty as many people may have fluctuating incomes with consumption financed by dissaving. Major imperfections in capital markets, on the other hand, mean that lifetime incomes are also not appropriate. An entirely different approach is to use multi-dimensional measures from the outset, with people defined to be in poverty if they lack some combination of socially important goods or capabilities.

- **The unit to which poverty applies.** There are two issues here. First is the question of how economic resources are shared in the household sector. Empirical analysis typically assumes that resources are shared so that all

1 The author would like to thank Nanak Kakwani for his comments on an earlier draft of this paper.

members of the same (co-resident) family have the same living standard. This ignores intra-family inequalities, as well as between-household sharing. The definition of family is also not unambiguous. The second is a question of counting. Are **families** in poverty, or the **individuals** who live in families. Many researchers count families, but from the perspective of social rights it probably makes most sense to count individuals. The latter is necessary if assumptions about intra-family sharing are to be directly addressed.

- **How to define 'particularly low' incomes**. This is usually done by setting a (more or less arbitrary) poverty line. People with incomes above this level are not counted for the measurement of poverty. Once a poverty line is set, consideration needs to be given to how it is adjusted over time or between countries. An **absolute** poverty line is adjusted in accordance with some price index so that it represents the same consumption opportunities. A **relative** poverty line, on the other hand, is adjusted in line with some income norm (such as average income) so that it can provide a measure of the extent to which people can enjoy the normal consumption patterns of their society. Given a poverty line and a method of adjustment, different methods of counting poverty may be employed. These include the simple **headcount** (the number of people below the line as a proportion of the population), the **poverty gap** which is a measure of how far below the line people are, and more sophisticated measures which take account of the undesirability of inequality among the poor.

- **How to measure relative needs**. It is commonly assumed that some people require a greater level of (family) income than others to attain the same living standard. This is typically addressed by using an equivalence scale, which describes relative income needs, and by undertaking poverty calculations in terms of **equivalent income** - income divided by the equivalence scale. Equivalent income is thus a measure of income relative to needs.

This paper provides estimates of the change in poverty in Australia between 1981-82 and 1989-90. Poverty is measured in terms of family annual after-tax incomes as defined by the ABS, with the individual as the counting unit. It is assumed that living standards are the same for all members of the same family. Rather than define a particular poverty threshold or counting method, the paper draws on the results of Atkinson (1987) to provide estimates which permit (in some circumstances) conclusions about the direction of poverty change for a range of poverty lines and a wide class of counting methods. Attention is, however, restricted to absolute poverty lines.

The main contribution of the paper is to generalise the approach of Atkinson to the situation where relative needs (equivalence scales) are not precisely known. This is important because, despite a great deal of effort, economic research has not been successful in calculating equivalence scales which command wide-spread acceptance. Indeed, just as appropriate poverty thresholds might be expected to vary between observers, it might be reasonable to expect that evaluations of relative needs (or even 'deserts') might vary - irrespective of the conclusions of consumption research.[2] At the same time, however, there **are** widely accepted notions of relative need that can be used to provide bounds on equivalence scales. This paper shows how these bounds can be used in turn to provide bounds on the extent to which poverty has increased or decreased.

In the next section, the generalised approach to poverty measurement is described. Section 3 then shows how a decomposition of equivalent income can be used to provide upper and lower bounds for poverty increases (assuming some bounded range of equivalence scales). This is then used in Section 4 to examine changes in Australian poverty between 1981-82 and 1989-90. The main conclusion is that the measurement of poverty changes is reasonably sensitive to the choice of equivalence scale. It is shown that plausible scales can be chosen which would lead to either an increase or a decrease in poverty over the 1980s in Australia.

2 Generalised Poverty Measurement When the Equivalence Scale is Known

Begin by assuming that we know the relative needs of people in family type i compared to the people in some reference family type. Denote this equivalence scale by m_i, with the scale for the reference family type denoted m_1 and set equal to unity.[3] Equivalent family income, $z_i = y_i/m_i$, can then be used as a measure of resources relative to needs for each of the individuals in the population.

If we fix some level, t, of equivalent income as a poverty threshold, a general poverty index for an individual in a family with equivalent income z can be defined as $p(z,t)$. This index takes some positive value when z is less than t and zero when equivalent income is equal to, or higher than, the poverty line. In other words, we do not care abut the distribution of the incomes of people above the poverty line (though for relative poverty measurement these incomes may influence the choice of t). The poverty index for the population as a whole can then be defined as the average of personal indices $p(z,t)$. That is,

2 Part, by no means all, of the divergence in equivalence scale estimates stems from the scope of the welfare comparison considered appropriate. Should estimates of the 'costs of children' only focus on their demands on household consumption, or should this be offset by the benefits of parenthood (and how should the latter be measured). These issues are discussed further in Bradbury (1992).

3 For simplicity, it is assumed throughout that the equivalence scale is the same at all income levels. The first (but not the second) order dominance conditions described below can be easily generalised to the case where the equivalence scale varies with income.

$$P = \int_0^t p(z,t) f(z) dz \tag{1}$$

where $f(z)$ is the distribution function of equivalent income. Atkinson (1987) shows that most poverty indices in common use can be represented in this form.[4] The headcount measure, for example, is obtained by setting $p(z,t) = 1$ when $z < t$ and 0 otherwise. The poverty gap (averaged over the whole population) is obtained by setting $p(z,t) = t\text{-}z$ when $z < t$. Other measures such as the Foster et al. (1984) measures, which give more weight to people well below the poverty line, can also be expressed in this form.

Atkinson groups poverty indices into two classes. Measures in class *I* satisfy the following condition.

$$\partial p(z,t)/\partial z \leq 0 \tag{2}$$

That is, an increase in the equivalent income of a person's family will reduce the person's poverty (or make no difference). Measures in class *II* must belong to class *I* and also satisfy

$$\partial^2 p(z,t)/\partial z^2 \geq 0 \tag{3}$$

This implies that a given increase in equivalent income will lead to a greater decrease in poverty the lower is the equivalent income of the person who receives it. In other words, measures in this second class also reflect some degree of concern about the inequality of living standards among the poor. The headcount measure clearly does not belong to class *II* since the poverty index will be decreased most by increasing the incomes of people just below the poverty line rather than by increasing the incomes of people who would still be below the poverty line after the increase. The poverty gap is the limiting case of measures which belong to *II* as the gap is decreased equally by an (infinitesimal) increase in equivalent income for any people below the poverty line (i.e. $\partial^2 (t-z)/\partial z^2 = 0$).

Note that all these derivatives are defined in terms of equivalent income. If property (3) is satisfied, this does **not** necessarily imply that a given increase in **actual** income will decrease poverty most if it is directed at people with the lowest level of equivalent income. This is because a dollar increase in income increases equivalent income by only $1/m_i$ dollars.

4 The results in Atkinson's paper are developed in terms of actual incomes and homogeneous populations (and are expressed as the negative of the expressions used here). Given the assumptions used here, however, his methods can equally be used to derive results in terms of equivalent incomes.

If two people have identical incomes, but one has greater needs, implying a smaller z, the largest reduction in poverty **may** be achieved by giving some additional income to the person with the higher living standard because they can use the income more 'efficiently'. When needs vary, but family sizes are identical, this 'paradox of targeting' will occur unless a sufficiently convex (i.e. inequality averse) poverty function is used (see Atkinson and Bourguignon, 1987, Keen 1992). In the more common case where needs vary because of family size differences, this paradox will usually not occur. This is because a dollar increase in the income of a family will increase the welfare of **each** family member by $1/m_i$ dollars. If needs increase at a slower rate than family size, minimisation of the overall poverty index will then require a direction of expenditure towards the larger family.

Whatever the interpretation in terms of actual incomes, however, finding general conditions for changes in the class *I* and *II* poverty measures is straightforward and can also incorporate the situation where there is ambiguity about the level of the poverty threshold. Consider two income distributions with cumulative distribution functions $F(z)$ and $G(z)$ and assume that the poverty threshold lies in some range $[t^-, t^+]$. Then a necessary and sufficient condition for there to be a reduction (or no change) in poverty for all poverty measures satisfying *I* when we move from distribution F to G is that

$$D(z) \leq 0 \ \forall z \in [0, t^+] \tag{4}$$

where $D(z) = G(z) - F(z)$.[5] This first degree dominance condition thus requires that the headcount poverty rate be lower (or equal) in the second distribution for any poverty line up to maximum possible poverty line (t^+). If this is the case, then poverty will have decreased (or not changed) for any poverty measure which can be written in the form of equation (1), and which belongs to class *I*. Note that this is a stronger requirement than would be required for the headcount measure itself. If the headcount were the only poverty measure under consideration, it would only be necessary for D(z) to be negative for z between t^- and t^+.

If it is desired to restrict the class of poverty measures to those in class *II* then a second degree dominance condition can be derived. Distribution G will have less (or equal) poverty than distribution F for all measures in class *II* if and only if

$$S(z) \leq 0 \forall z \in [0, t^+] \tag{5}$$

5 This result follows from integrating equation (1) by parts for each distribution. It also requires continuity of $p(z,t)$. This is violated for the headcount measure, but this measure is discussed explicitly in the text.

where $S(z) = [C_G(z) - C_F(z)]$ and $C_F(z) = \int_0^z F(y)dy = \int_0^z (z-y)f(y)dy$ (with $C_G(z)$ defined similarly). The function $C(z)$ can thus be represented as either the area under the distribution curve up to equivalent income z, or as the poverty gap averaged over the whole population.[6] In other words, **all poverty measures in class II will decrease (or remain the same) if the poverty gap is lower (or equal) for all poverty lines between zero and t^+.** Note that if the first degree condition is satisfied, then so will this second degree condition, but the reverse is not true. Again, if attention is restricted to the poverty gap measure only, then the condition in equation (5) need only apply for poverty lines in the range (t^-, t^+).[7]

The headcount and poverty gap measures thus hold a special place among the possible ways of counting poverty as their dominance over a range of incomes can be used to imply dominance for a much wider class of measures. However the above approach does assume that we do know the relative needs of people in different circumstances. But precise knowledge about relative needs does not exist. We turn now to a consideration of how information on the approximate bounds of equivalence scales can be used to provide upper and lower bounds on the functions $D(z)$ and $S(z)$.

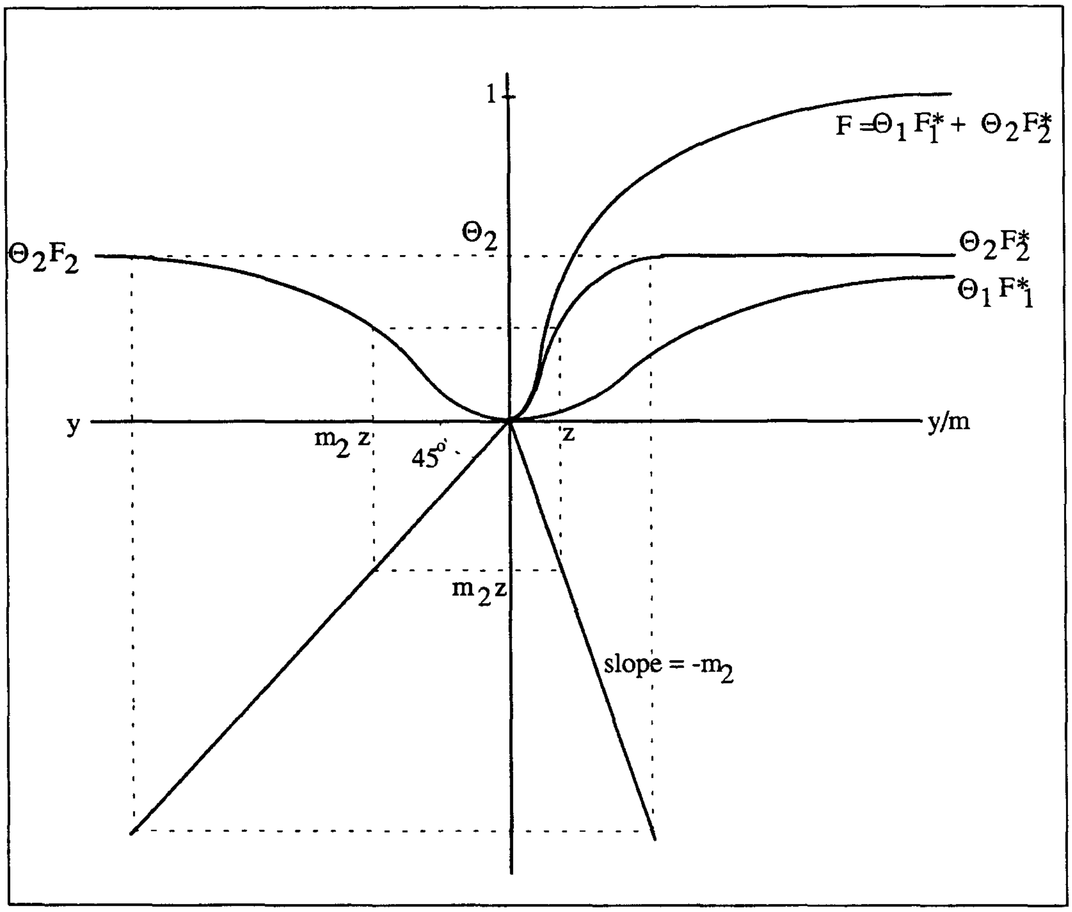
3 Using Bounded Equivalence Scales

To do this, it is necessary to decompose the indices described above. The method used here is illustrated in Figure 1 which describes the decomposition of $F(z)$ when there are two family types, with m_1=1. The overall distribution of equivalent income, $F(z)$, is equal to the weighted sum of $F^*_1(z)$ and $F^*_2(z)$, the equivalent income distributions for each family type. The weights, Θ, are the proportion of the sample in each family type. For family type 2, however, the proportion of people with equivalent income below z will be equal to the proportion of people with **actual** income below m_2z. Hence the overall distribution of equivalent income can be written as a function of the actual income distributions in each family type.

6 Integration by parts can be used to show the equivalence of these two formulations.

7 If t^+ is set equal to the highest equivalent income in either distribution, then this condition becomes identical to that for general second order social welfare dominance, and is equivalent to the requirement that generalised Lorenz curves not cross.

Figure 1: Decomposition of the Equivalent Income Distribution



More generally, denoting the proportion in family type category i as Θ_i (with $\sum_i \Theta_i = 1$), we can write $F(z) = \sum_i \Theta_i F_i(m_i z)$ where $F_i(m_i z)$ is the proportion of people in family type i with actual income below $m_i z$. Using this, the first order dominance condition can be written as

$$D(z) = \sum_i D_i(m_i z) \leq 0 \ \forall \ z \in [0, t^+] \tag{6}$$

where $D_i(m_i z) = \Theta_{iG} G_i(m_i z) - \Theta_{iF} F_i(m_i z)$. This is a particularly useful decomposition, because it means that the contribution of each family type to the overall headcount increase (D_i) can be calculated for each value of actual income, prior to a decision about the values of m to be employed. Note that these contributions depend not only on the increase in poverty in each family type at each income level, but also on the change in the population distributions. This is shown more clearly by writing D_i at income level y as

$$D_i(y) = \overline{\Theta_i}\Delta F_i(y) + \overline{F_i(y)}\Delta\Theta_i \tag{7}$$

where $\overline{\Theta_i} = (\Theta_{iG} + \Theta_{iF})/2$, the average proportion of the population in family type i, $\Delta F_i(y) = G_i(y) - F_i(y)$, the within-group increase in poverty, $\overline{F_i(y)} = (G_i(y) + F_i(y))/2$, the average poverty rate and $\Delta\Theta_i = \Theta_{iG} - \Theta_{iF}$, the increase in the proportion of the population in family type i. The first term in this equation thus shows the increase in poverty due to the within-group increase in poverty, whilst the second term shows the increase in poverty due to the increase in the relative size of the family type. In other words, it reflects the fact that if family types with relatively high poverty rates become more common, then poverty will increase.

A similar decomposition can be written for the second order condition. In Figure 1 it can be seen that the area under the F curve up to z will equal the sum of the areas under the $\Theta_1 F^*_1$ and $\Theta_2 F^*_2$ curves, and that the area under the F_2 curve up to $m_2 z$ will be m_2 times the area under the F^*_2 curve.[8] Hence the second degree social welfare dominance condition becomes

$$S(z) = \sum_i S_i(m_i z)/m_i \leq 0 \forall z \in [0, t^+] \tag{8}$$

where $S_i(m_i z) = \Theta_{iG}C_{iG}(m_i z) - \Theta_{iF}C_{iF}(m_i z)$ and $C_{iF}(m_i z)$ is the poverty gap (in actual rather than equivalent income) at poverty line $m_i z$ in family type i in distribution F. These poverty gaps are divided by m_i so as to express the poverty gap in terms of equivalent dollars.

Now consider the case where the m_i parameters are not precisely known, but lie between lower and upper bounds m_i^- and m_i^+. Family type number 1 serves as the reference group, with $m_1 = m_1^- = m_1^+ = 1$. In order for distribution G to imply a lower (or equal) poverty level than distribution F for all possible equivalence scales the following conditions must then be satisfied.

8 If m_2 is not constant with income, the curve in the lower-right quadrant of Figure 1 will be curved rather than straight. This will not fundamentally alter the testing for first degree dominance, but will significantly complicate the testing for second degree dominance, as it will no longer be possible to multiply the area under the distribution curve by a constant scaling factor.

First degree: $\sum_i D_i(m_i z) \leq 0 \;\; \forall\, z \in [0,t^+],\; \forall\, m_i \in [m_i^-, m_i^+]$

Second degree: $\sum_i S_i(m_i z)/m_i \leq 0 \;\; \forall\, z \in [0,t^+],\; \forall\, m_i \in [m_i^-, m_i^+]$ (9)

Since both z and m_i are real numbers, it is not possible to test these restrictions over all possible values. However, provided the income distributions are 'smooth enough' this test can be approximated by testing at some finite number of points. This is the conventional procedure for homogeneous populations, where a finite number of income points z are examined.

More specifically, consider an ordered set of income points spread across the range of incomes $Y^* = \{0 < y_2 < y_3 \ldots < y_{n^*}\}$. The choice of points is arbitrary, but for optimal estimation are chosen here as equally spaced percentile points in the distribution of y/m^* where m^* is an 'average' equivalence scale taken from near the middle of the sets of possible equivalence scales (in the empirical analysis here this is calculated over the combined F and G distributions).

For each family type assume that between each pair of points (y_j, y_{j+1}) the function $D_i(y)$ can be approximated by a quadratic curve. This curve is uniquely determined by the magnitudes of y_j, y_{j+1}, $D_i(y_j)$, $D_i(y_{j+1})$ and $S_i(y_{j+1}) - S_i(y_j)$.[9] Since the $D_i(y)$ segments are quadratic, their maximum and minimum values will occur at either their end points or at the single point where $\partial D/\partial y = 0$. Any such interior maximum or minimum points can then added to the set Y^* to form sets Y_i for each family type.

Using this assumption of quadratic segments, the first degree dominance test can then be implemented as: for each $z \in Y^*$ calculate the maximum and minimum of $\sum_i D_i(y_i)$ over all $y_i \in Y_i$ such that $m_i^- z \leq y_i \leq m_i^+ z$. The second degree tests are calculated in the same way. Because the extreme values $m_i^- z$ and $m_i^+ z$ might not lie in Y_i, the fitted quadratic curve is used to calculate the value of the function at these points.

The computational burden of testing these hypotheses depends very much upon the restrictions imposed on m. Three different types of constraints can defined.

9 Since D is the difference between the two distribution functions, this result follows from the fact that a quadratic segment for each distribution curve will be determined by its end points and the area underneath it.

Independent Constraints

In this structure, the equivalence scales for each group are independent of each other (though for first degree dominance they may depend upon z). A simple example for the family types 1: couple with no children, 2: couple with one child and 3: couple with two children, might be:

$$m_1 = 1; \quad m_2 \in [1.1, 1.5]; \quad m_3 \in [1.2, 2.0] \tag{10}$$

In this case the first order minimisation problem for each value of z can written as,

$$\textstyle\sum_i \min_{m_i^- z \leq y_i \leq m_i^+ z} D_i(y_i)$$

(with analogous representations for the maximisation and second order problems). Because the minimisation is undertaken separately for each family type, the number of comparisons to be made is not particularly large.

However independent restrictions can produce implausible results. It could be the case, for example, that at some value of z, minimum values for $D_i(y_i)$ are found at m_2=1.3 and m_3=1.2. Usually some sort of constraint on the ranking of equivalence scales is also required.

Ranked Constraints

For the above family types, a simple ranked constraint might be specified as

$$m_1 = 1; \quad m_2 \geq 1; \quad m_3 \geq m_2 \tag{11}$$

This form of equivalence scale constraint is very similar to the approach developed by Atkinson and Bourguignon (1987).[10] In their case, however, the ranking is on the basis of the marginal utility of income, rather than on the equivalence scale. The targeting paradox arises because these are not necessarily identical.

One implication of their method is that for overall welfare dominance to occur, (first or second degree) dominance **must** occur for the most needy family type. This is because the constraints on the individual welfare functions do not prohibit a zero marginal utility of income for all groups other than the most needy. A similar requirement applies here, and one aspect of this can be seen by considering

10 For papers which further develop their approach see Atkinson (1992), Bourguignon (1989) and Jenkins and Lambert (1992).

Figure 1. As m_2 increases, the lower part of the overall F distribution will increasingly be dominated by that for family type 2, and if the population size is finite, will eventually be completely determined by it. Hence a necessary requirement for poverty dominance at low values of z is that poverty dominance exists for the neediest family.

Even if we ignore the very bottom of the distribution, the fact that infinite needs for the largest family type are not ruled out means that the possibility that they are all in poverty cannot be discounted. If this were the case then this family type could have a contribution to the increase in poverty equal to its growth in population proportion.

The unrestricted upper bound on the equivalence scale for one family type thus imposes quite a significant limitation on the methodology. A more complicated structure for the equivalence scale, taking account of both bounds and ranking, is required to avoid this limitation.

Nested Constraints

For these same three family types, a simple nested structure might be given by

$$m_1 = 1; \quad m_2 \in [1.1, 1.5]; \quad m_3 \in [m_2, 2.0] \tag{12}$$

or

$$m_1 = 1; \quad m_2 \in [1.1, 1.5]; \quad m_3 \in [m_2, m_2 + (m_2 - 1)] \tag{13}$$

The latter formulation assumes that the additional cost of the third child is non-negative, but also not more than the additional cost of the second child.

These additional constraints make the computation considerably more intensive, because it is now impossible to split the minimisation into separate operations for each family type. With constraints of the form (12) or (13), a nested iterative process which computes the value of the $D(z)$ and $S(z)$ statistics for every possible combination of m_2 and m_3 is required. If n is the average number of income points over which each equivalence scale constraint ranges, then this calculation will need to be evaluated approximately $n * n^{I-1}$ times (where I is the number of family types). The only consolation is that the constraints may make n relatively small. However with current computational capabilities such a search is quite feasible. The results obtained from two empirical examples are described in the next section.

4 Poverty Changes: 1981-82 to 1989-90

4.1 An Illustrative Worked Example

This section works through an evaluation of the first order conditions for a simple three family type model. Only couples with no, one or two children are considered and it is assumed that the equivalence scales take the form shown in equation (13). Some basic statistics for these families in 1981-82 and 1989-90, based on the decomposition described in equation (7), are shown in Table 1. Over the 1980s, the proportion of people in larger families fell, whilst smaller families increased. This reflects the combination of falling fertility and the ageing of the population.

First Order Dominance

Figure 2a shows the (headcount) poverty rates at different income levels for these three family types, averaged over the two years. Poverty is much higher among couples with no children because this includes many retired people. As a point of reference for this figure, the basic Australian pension for couples without children was around $11,600 per annum in 1989-90.

Figure 2b shows the changes in the income distribution between the two years for each of the family types. For most family types, at most poverty line thresholds, poverty decreased over the period, though there are exceptions. However to calculate the overall change in poverty, this needs to be combined with the changes in population distribution shown in Table 1. This is done in Figure 2c.

The interpretation of Figure 2c is based on equation (7) which described the contribution of each family type to poverty in terms of the increase in within-group poverty and the increase in population proportion. Consider, for example, couples with no children at the income level of $13,850 (about 1.2 times the base pension level). Since they are the reference group for the equivalence scale, this is also their equivalent income level. Figure 2b shows that at this income level, their poverty rate declined by 2.6 percentage points. However, this decline is more than offset by the fact that at this income level couples without children have a relatively high poverty rate (Figure 2a), and this is a family type that has grown in prevalence over the period. Using equation (7), the contribution of couples without children to the headcount measure of poverty at an equivalent income of $13,800 is thus $0.405 \times -0.026 + 0.211 \times 0.052 = 0.0004$, or a marginal increase. This is the result shown in Figure 2c.

Table 1 Summary Information for Three Family Types

Family Type	**2+0**	**2+1**	**2+2**	**Total**
1981-82 proportion (Θ_{iF})	0.379	0.193	0.427	1.000
1989-90 proportion (Θ_{iG})	0.432	0.194	0.375	1.000
Mean proportion ($\overline{\Theta_i}$)	0.406	0.193	0.401	1.000
Increase in proportion ($\Delta\Theta_i$)	0.052	0.001	–0.052	0.000
1981-82 mean income[(a)] ($000)	27.1	32.0	32.6	29.6
1989-90 mean income ($000)	28.5	33.3	34.4	30.8
Ratio	1.051	1.040	1.055	1.040

Note: a) All incomes are in $1989-90. Proportions are calculated on a per person basis. Total mean income is per family.

Figure 2a: Average Within-Group Poverty Rates, 1981-82 and 1989-90 ($\overline{F_i(y)}$)

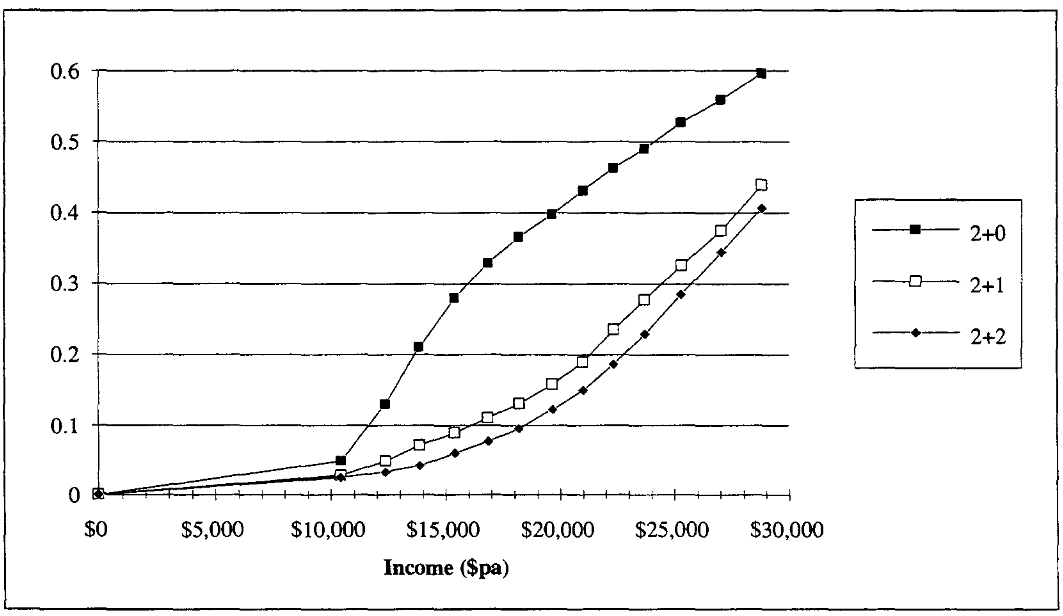


Figure 2b: Within-Group Increase in Poverty ($\Delta F_i(y)$)

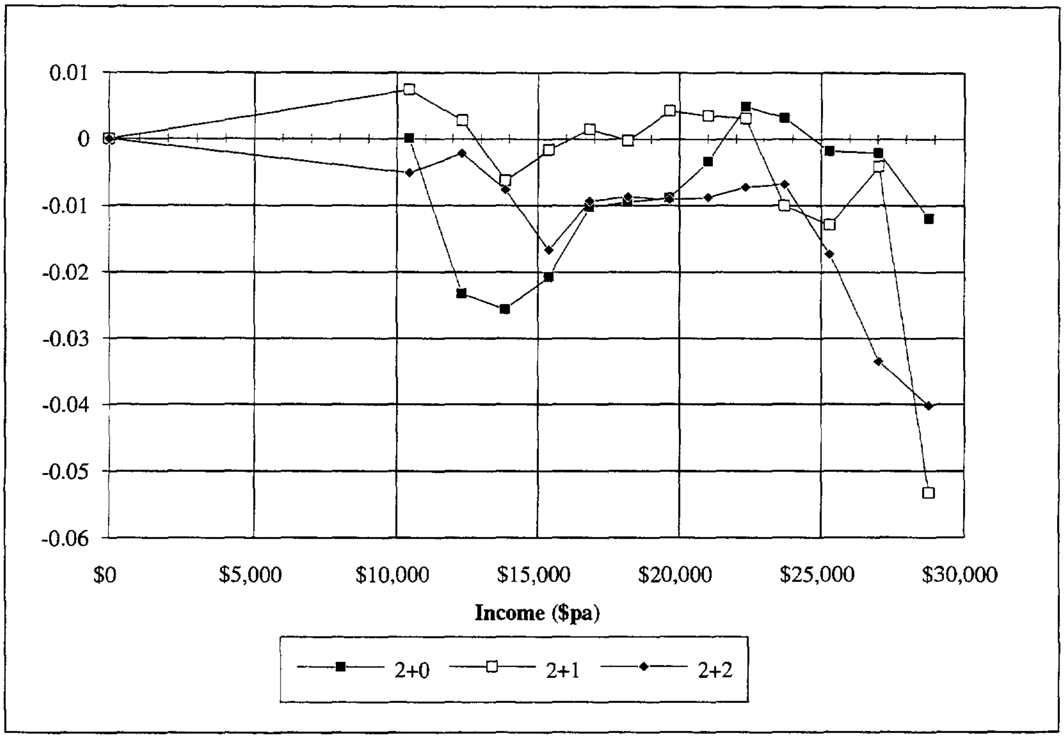


Figure 2c, however, can also be used to find the range of poverty changes which the allowable set of equivalence scales will permit. Consider again the equivalent income poverty threshold of $13,850. If this is the poverty line for the reference family, then equation (13) requires that the poverty line for a couple with one child lie between $15,250 and $20,775 (1.1 and 1.5 times $13,850). Figure 2c shows that the largest contribution to the poverty increase in this range is at the upper end, where m_2=1.5. However we also need to take account of couples with two children. Their contribution to poverty changes is strongly negative (mainly because their population proportion fell) and varies significantly with income level. If the equivalence scale were permitted to vary between 1.1 and 2.0, the greatest contribution to poverty increases would be found at the **lower** end of this range, with m_3=1.1. However we also have the constraint that $m_2 \leq m_3$, so this combination of choices is not feasible. Taking this constraint into account, the largest increase in poverty with a poverty line set at $13,850 equivalent dollars turns out to be at $m_2 = m_3$=1.216 (at the sixth point on the chart).

This calculation can be replicated for every other equivalent income point, and a minimum bound calculated in a similar fashion. The results of such a calculation are shown in Figure 3a, for equivalent income poverty lines up to $19,600 p.a. The

Figure 2c: Group Contributions to Overall Poverty Changes ($D_i(y)$)

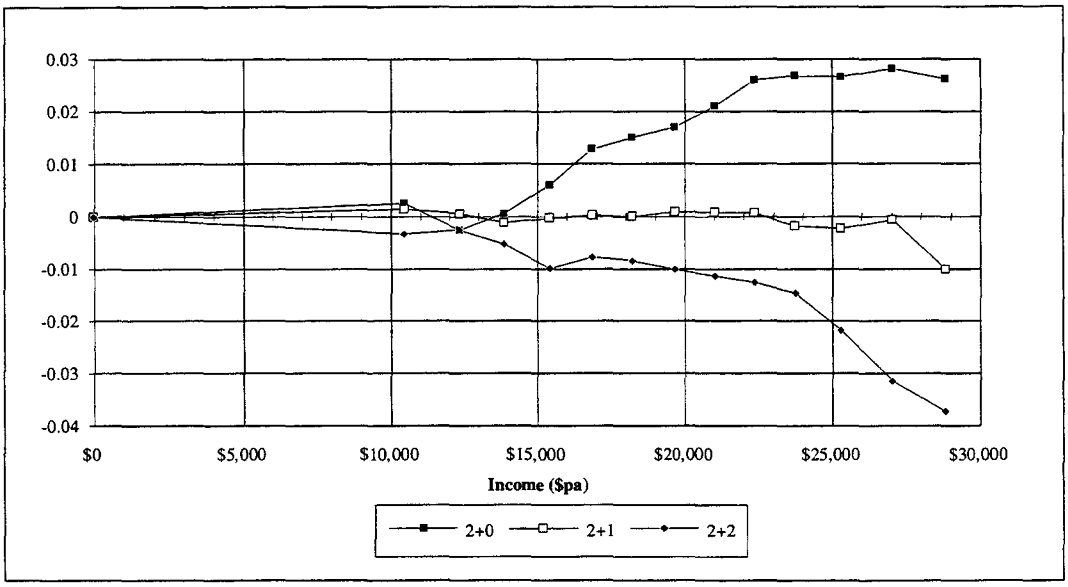
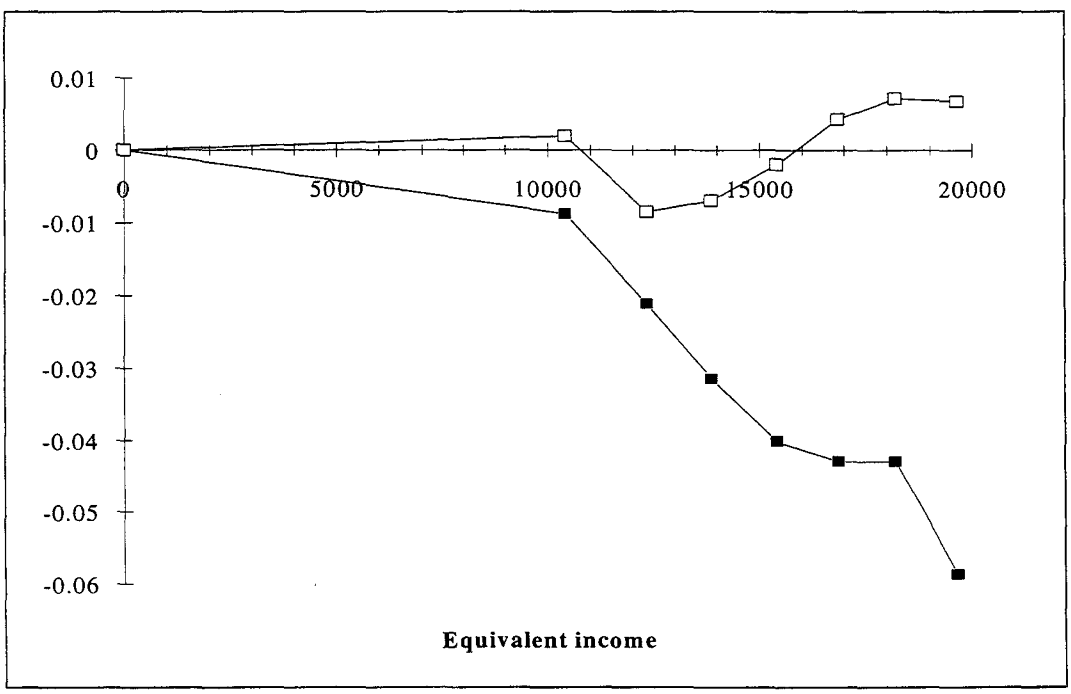


Figure 3a: Three Family Types: Maximum and Minimum Bounds for the Increase in the Headcount Poverty Measure.



upper line in the figure shows the highest increase in poverty possible for some set of equivalence scales satisfying the restrictions in expression (13), whilst the lower line shows the smallest increase in poverty that could be found. If both these lines were always below the axis, we could conclude that: for all poverty measures in Atkinson's class I and for all poverty lines below \$19,600 pa[11] and for all equivalence scales satisfying (13), absolute poverty had decreased for these three family types. Unfortunately this is not the case. Both at very low and at higher poverty lines equivalence scales can be found which would imply an increase in poverty over the period. For example, at a poverty line of \$18,200 an equivalence scale of m_2=m_3=1.10 would imply an increase of 0.4 percentage points in the headcount poverty measure. (A poverty decrease of 4.3 percentage points results from an equivalence scale of m_2=1.5, m_3=1.955).

There are two caveats to this conclusion. First, these estimates take no account of the fact that this data is drawn from a sample rather than the whole population. The small increases in poverty might not be statistically significant. A second possible concern is that the search for equivalence scale bounds is undertaken at only a small number of points (20 over the whole income distribution) and so may miss some important patterns in the income distribution.

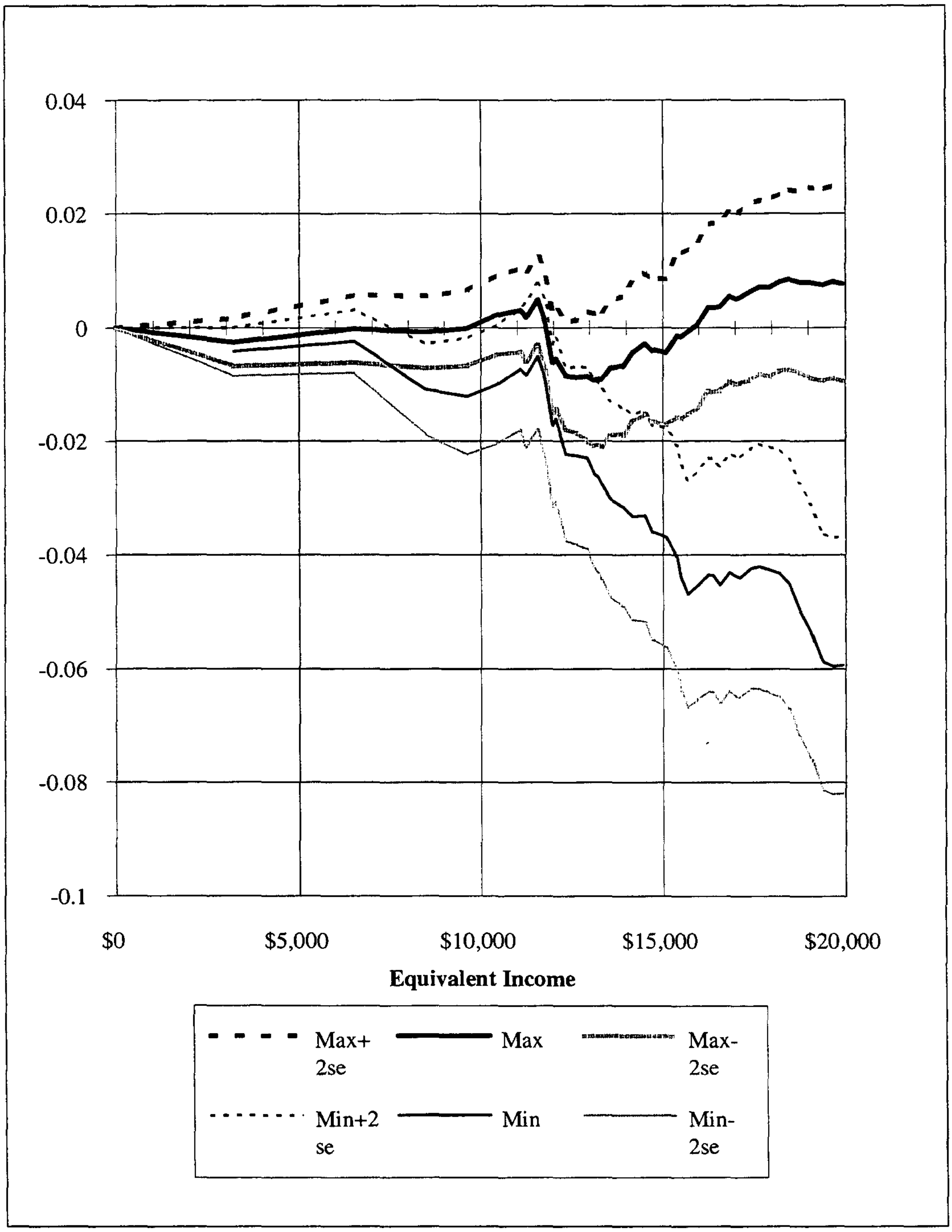
Whilst the exact calculation of standard errors from weighted data is not straightforward, it is relatively easy to obtain approximate estimates of standard errors. From equation (6) an approximation to the variance of the difference in poverty rates for a given set of m_i can be written as

$$\begin{aligned} V(D(z)) &= \textstyle\sum_i \Theta_{iF}^2 U_{iF} + F_i^2(m_i z)V_{iF} + \Theta_{iG}^2 U_{iG} + G_i^2(m_i z)V_{iG} \\ \text{where } U_{iF} &= V(F_i(m_i z)) = F_i(m_i z)(1-F_i(m_i z))/n_{iF} \\ \text{and } V_{iF} &= V(\Theta_{iF}) = \Theta_{iF}(1-\Theta_{iF})/n_{+F} \end{aligned} \tag{14}$$

and where n_{ij} is the unweighted number of families of family type i in distribution j and n_{+j} is the total number of families in distribution j. This formula is an approximation because: 1) it ignores any possible correlation between the proportion of the sample in a given family type and the income distribution of that family type, 2) it does not take account of the correlation between the Θ's (since they must add to one), and 3) it does not take account of the effect of the non-

[11] This is about 1.7 times the basic pension rate.

Figure 3b: Three Family Types: Maximum and Minimum Bounds for the Increase in the Headcount Poverty Measure (100 point distribution with confidence bands).



random sample design. In addition, it should be remembered that this is an estimate of variance conditional upon the choice of m. In other words, the fact that different samples from the same population may lead to different choices for the equivalence scales which provide the bounds on the poverty estimates is not incorporated into these estimates of variation.

Figure 3b replicates the results of Figure 3a, but calculated for a greater number of income points (100 across the whole distribution) and also including curves showing the maximum and minimum poverty increases +/- two standard errors. Whilst the additional detail of this figure does change the patterns slightly from that shown in Figure 3a, the results shown in the earlier figure are still well within the standard error bounds.

The width of the confidence intervals in Figure 3b, does however allow us to draw stronger conclusions about trends in poverty. Examining the figure, it can be seen that there is **no** poverty line/equivalence scale combination at which we can say with approximate 95 per cent confidence that headcount poverty has increased. That is, the lower bounds of the confidence intervals always lie below zero. On the other hand, there are some poverty lines and equivalence scale combinations where poverty for these three family types has unambiguously fallen.

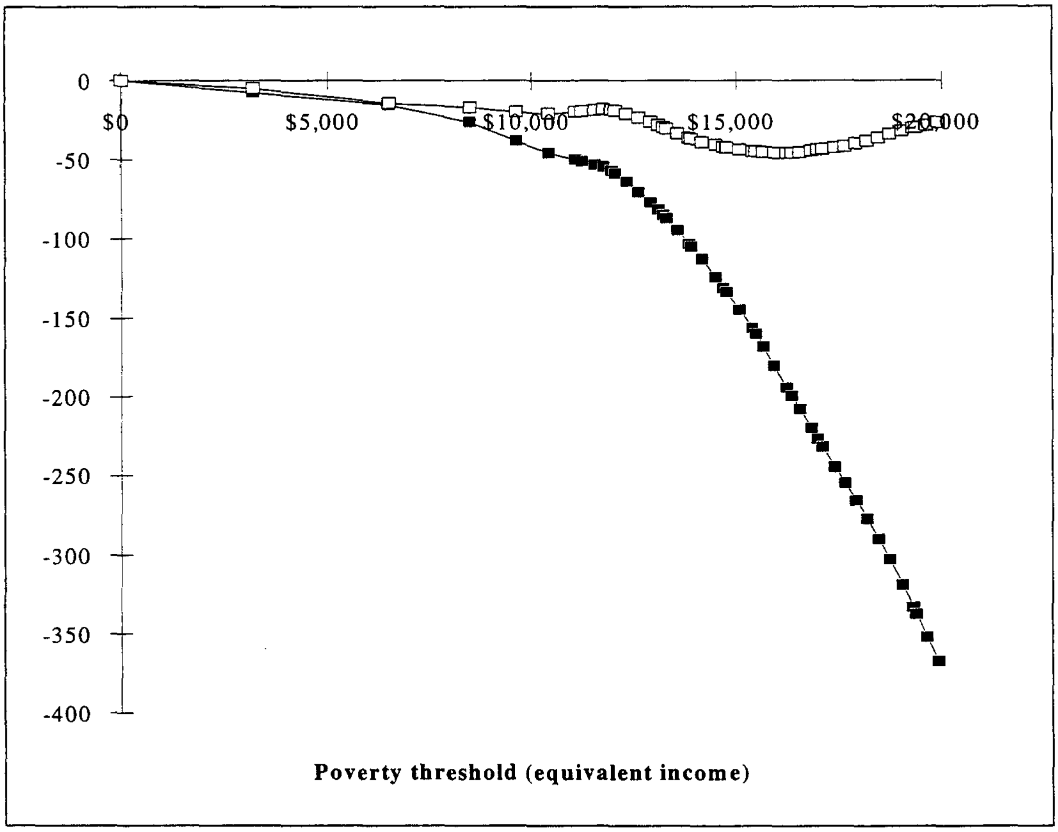
Second Order Dominance

A similar analysis can be carried out to test for a decrease in all poverty measures which satisfy property *II*. This requires that the difference in poverty gaps be always negative. Figure 4 shows the results from this analysis in a form comparable to Figure 3b (standard errors have not yet been calculated). Because property *II* narrows the class of poverty measures which can be applied, stronger conclusions can be derived. Thus for any equivalence scale and any poverty threshold (up to $20,000) the poverty gap has decreased. This differs from the first order condition because the fall in the headcount measure of poverty at very low income levels is more than sufficient to offset the higher headcount poverty rate at around $11,000 pa.

4.2 Changes for all Families

What happens when this approach is generalised to all family types? Table 2 illustrates a simple structure of equivalence scales which encompasses the whole population. The family definition here follows that used by the Luxembourg Income Study and includes people related by 'blood, marriage (including de-facto) or adoption' and is differentiated by the number of adults and children (only one couple or sole parent nucleus may be represented in each family). Children are defined as not married people under the age of 18 living with other family members.

Figure 4: Three Family Types: Maximum and Minimum Bounds for the Increase in the Poverty Gap



Two-adult households are typically married couples, but may also include siblings, or a sole parent together with their adult child. For families with more than two adults (usually couples plus adult children), only the total number of people is used to define the equivalence scale.

The income measure is family after-tax income,[12] the unit of counting is the individual, and it is assumed that all people in the same family have the same living standard (and hence equivalence scale).[13]

12 Income tax is imputed on the basis of observed characteristics (in 1981-82 by the SPRC, and in 1989-90 by the ABS). Family income is defined as the sum of the income of all persons aged 15 or over in the family. The income of full-time school students is not counted in either year. Families where the head or spouse 'changed status' during the year (the standard ABS exclusion for annual income measurement) are excluded in both years.

13 This approach can, in principle, be generalised to the case where each individual in the family has a different living standard.

Table 2: A Simple Equivalence Scale Structure

Adults	Children	People	i	m_i^-	m_i^+
2	0		1	1.00	1.00
2	1		2	1.05	1.40
2	2		3	m_2+0.04	$m_2+(m_2-1)$
2	3		4	m_3+0.03	$m_3+(m_3-m_2)$
2	4+		5	m_4+0.02	$m_4+(c_1-3)(m_4-m_3)$
1	0		6	0.60	0.90
1	1		7	m_6+0.05	1.10
1	2		8	m_7+0.04	$m_7+(m_7-m_6)$
1	3+		9	m_8+0.03	$m_8+(c_2-2)(m_8-m_7)$
3+	0+	3	10	1.10	1.80
		4	11	m_{10}+0.10	$m_{10}+(m_{10}-m_9)$
		5	12	m_{11}+0.10	$m_{11}+(m_{11}-m_{10})$
		6+	13	m_{12}+0.10	$m_{12}+(c_3-5)(m_{12}-m_{11})$

Notes: c_1 is the maximum over the two distributions of the average number of people in families with two adults and four or more children. c_2 and c_3 are similarly defined for the other family types.

The equivalence scale structure of Table 2 is 'partially nested', in that the equivalence scales for families with a given number of adults is nested, but there are no cross-restrictions across these groups. Additional restrictions could be imposed, such as that the additional cost of each child to a sole parent was at least as great as the cost to a married couple, but in the light of the small population size of sole parent families, this additional restriction would be unlikely to greatly alter the results (and would substantially increase the computational burden).

Some basic information for these 13 family types is shown in Table 3.

Note that the population proportions shown are proportions of people, whereas the mean incomes are generally expressed per family unit. Significant changes in family structure over the 1980s are clearly evident in this table. Ageing and delay of childrearing has meant that an extra 3.3 per cent of the population are now living in couple-only families. Larger families have fallen in predominance whilst single

Table 3: Summary Statistics

Family Type	Population proportion (Θ) (proportion of individuals)			Mean income (per family)		
	1981-82	1989-90	Increase	1981-82	1989-90	Ratio
2+0	17.6%	20.9%	3.3%	$27,097	$28,485	1.051
2+1	9.0%	9.4%	0.4%	$32,035	$33,317	1.040
2+2	19.8%	18.1%	-1.6%	$32,627	$34,416	1.055
2+3	12.0%	10.5%	-1.4%	$33,578	$34,214	1.019
2+4+	5.4%	4.5%	-0.9%	$31,936	$34,773	1.089
1+0	10.8%	11.4%	0.5%	$15,005	$15,014	1.001
1+1	1.5%	1.8%	0.3%	$13,946	$14,191	1.018
1+2	1.4%	1.8%	0.4%	$15,111	$15,190	1.005
1+3+	1.0%	1.1%	0.1%	$16,275	$15,247	0.937
3	5.8%	6.2%	0.4%	$43,328	$44,175	1.020
4	6.5%	6.7%	0.2%	$50,953	$51,367	1.008
5	5.0%	5.1%	0.1%	$52,528	$50,403	0.960
6+	4.3%	2.4%	-1.9%	$54,828	$49,242	0.898
Total	100.0%	100.0%	0.0%	$27,801	$28,025	1.008

All Families	1981-82	1989-90	Per cent Increase
Number of families (000)	5,440.0	6,414.2	17.9
Number of people (000)	14,523.3	16,357.0	12.6
Per family income ($000 1989-90)	27,800.9	28,024.7	0.8
Per capita income ($000 1989-90)	10,413.4	10,989.5	5.5
Sample size (families)	15,298	15,105	

adult and sole parent families have grown.[14] Mean incomes, on the other hand, increased for 10 out of 13 family types.

Nonetheless, the increase in real disposable incomes recorded by the income surveys is quite small when compared to National Accounts aggregates. Real household disposable income per capita grew by 8.9 per cent between 1981-82 and 1989-90 whilst the surveys suggest only a 5.5 per cent growth. Two reasons stand out as potential explanations for this difference (though more detailed research on this question is clearly warranted). First, there was significant growth in non-cash forms of income over the 1980s (occupational superannuation in particular), and these are not included in the surveys (but are in the National Accounts). Second,

14 It should be noted that sole parent families which 'formed' during the financial year are excluded from the analysis in both years.

the income share of capital grew significantly, and capital incomes are known to be under-recorded in income surveys. This suggests some caution in the interpretation of results.

Perhaps of more interest in the present context is the large difference between per capita and per family income growth. The shrinking family size means that per family incomes have increased much less than per capita incomes. The change in average living standards, one might presume, would lie somewhere between these two indicators.

Turning to poverty changes, Figure 5 shows the headcount poverty bounds for poverty lines up to (approximately) $20,000, and the bounds for poverty gaps are shown in Figure 6.[15] The conclusions from these two figures are clear. It is **not** possible to say that poverty either increased or decreased for any reasonable equivalence scale. For any given poverty line, an equivalence scale can be chosen to imply either an increase or a decrease in the headcount measure (and similarly for the poverty gap). Given the small increase in mean incomes and the significant variations in family sizes this result should not be surprising.

For poverty lines around the pension level ($10–$12,000) the gap between the maximum and minimum poverty increase is about three percentage points. This is about six times the standard errors associated with these estimates.

Looking further at the first order results, Table 4 shows (for selected equivalent income levels) the equivalence scales which produce the variation in poverty rates shown in Figure 5. For this particular set of income data, the larger increase in poverty generally results from an equivalence scale which assumes lower needs for couples with many dependents.

Table 5 takes one of these poverty lines and shows the within family poverty rates in each year using the two extreme equivalence scales. These results are calculated from the original data (rather than from the 50 point summary distribution) and so the minimum poverty rate increase differs slightly from that shown above. Note also that this is quite a high poverty line, and so the proportion of people below it is high in both years.[16]

---

15 These are calculated from a 50 point summary of the income distribution of each family type. The calculations took 168 CPU seconds when written in SAS and run on an IBM 3090 computer. SAS routines for carrying out these calculations are available from the author.

16 Also note that the estimates here include some groups, such as the self-employed, for whom incomes may be significantly understated. If families where the head or spouse (if present) were working in their own business at the time of the survey are excluded from the analysis then trends in poverty are much bleaker – because their incomes rose faster than the average during the 1980s. The headcount measure is not significantly negative, and the poverty gap shows an unambiguous increase in poverty.

Figure 5: Maximum and Minimum Bounds for the Increase in the Headcount Poverty Measure, 1981-82 to 1989-90.

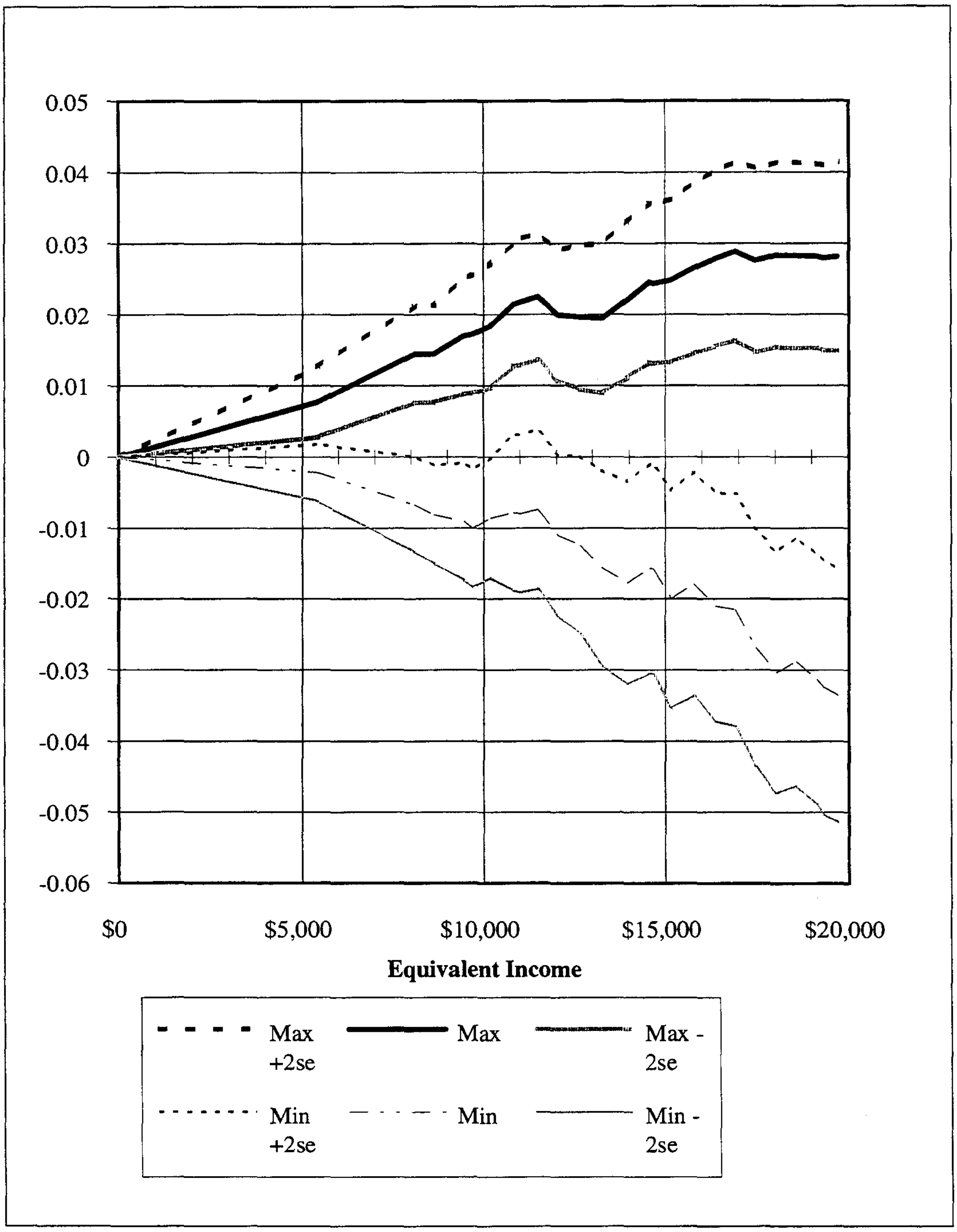
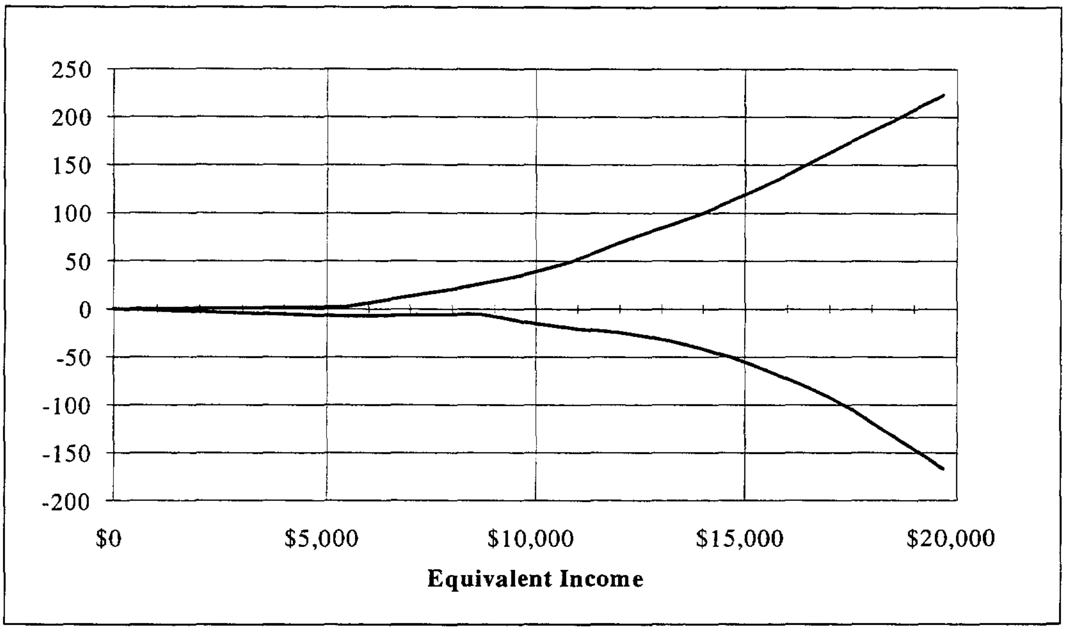


Figure 6: Maximum and Minimum Bounds for the Increase in the Poverty Gap, 1981-82 to 1989-90.



5 Conclusion

There are several conclusions that can be drawn from this analysis. First, it is possible to make useful general statements about (absolute) poverty changes without being confined to a fixed equivalence scale. The search methodology used here is computationally feasible for quite detailed equivalence scale specifications. Estimates of poverty changes are indeed quite sensitive to the choice of equivalence scale. Whilst discussion here has focused on poverty (and on first order results in particular) the method can easily be generalised to consider social welfare indices defined over the whole population by choosing a poverty line above the maximum income level.

It may also be possible to use the general concept of decomposition by family type and searching over possible equivalence scales in a number of other applications. These might include the measurement of relative poverty and inequality and the placing of bounds on additively separable social welfare and inequality indices. Algorithms for relative poverty measures, however, are likely to be more complicated than used here because the poverty line for the reference family type also varies with the equivalence scale.[17] One relatively straight-forward application

17 Relative poverty measures based on equivalence scale independent measures of community living standards (such as the Henderson methodology as used in Australia) can clearly be analysed using the method described here.

Table 4: Equivalence Sales Yielding Maximum Increase and Decrease in Poverty (for three poverty lines)

	Selected Poverty Lines					
	$10,184		$15,165		$19,692	
	Smallest increase	Greatest increase	Smallest increase	Greatest increase	Smallest increase	Greatest increase
Increase in head count rate	-0.9%	1.8%	-2.0%	2.5%	-3.4%	2.8%
Equivalence scales						
2+0	1.00	1.00	1.00	1.00	1.00	1.00
2+1	1.37	1.13	1.40	1.05	1.40	1.05
2+2	1.74	1.26	1.80	1.09	1.80	1.11
2+3	2.06	1.39	2.20	1.12	2.16	1.14
2+4+	2.51	1.56	2.75	1.14	2.42	1.18
1+0	0.60	0.90	0.60	0.76	0.80	0.60
1+1	0.65	1.10	0.67	1.10	0.92	1.03
1+2	0.69	1.18	0.74	1.44	0.96	1.30
1+3+	0.72	1.21	0.85	1.64	1.00	1.33
3	1.24	1.80	1.80	1.68	1.80	1.52
4	1.40	2.57	2.60	2.15	2.60	1.65
5	1.56	3.35	3.40	2.56	3.40	1.75
6+	1.82	3.45	4.73	2.66	4.73	1.90

of the present approach would be to permit equivalence scales to vary **within** the family to reflect unequal sharing of family incomes.

The substantive conclusions regarding poverty changes are also of interest. Though the limitations of the survey data need to be recognised, the general picture for poverty trends over the 1980s are not particularly encouraging. Equivalence scales can be chosen which would lead to an estimate of either an increase or a decrease in poverty (of similar magnitude). It must be recognised however, that part of the increase in poverty is related to demographic changes, in particular the ageing of the population and the growth of sole parent families. There is no space for a more detailed description of this disaggregation here.[18] Another potential area for further research is in the use of simulation methods to adjust income surveys for their data deficiencies. The comparison with National Accounts data suggests that such an adjustment might lead to a slightly brighter picture.

18 In the present context, equation (7) provides a natural framework for considering within and between-group changes separately.

Table 5: Headcount Poverty Rates at $15,165pa. (Direct Calculation Without Interpolation)

	Low Poverty Increase				High Poverty Increase			
	m_i	Headcount 1981-82	Headcount 1989-90	Difference	m_i	Headcount 1981-82	Headcount 1989-90	Difference
		%	%				%	%
2+0	1.00	28.1	25.6	-2.5	1.00	28.1	25.6	-2.5
2+1	1.40	19.9	19.5	-0.3	1.05	9.7	9.6	-0.1
2+2	1.80	37.1	33.4	-3.8	1.09	8.1	6.8	-1.3
2+3	2.20	57.3	59.2	1.9	1.12	8.8	8.7	0.0
2+4+	2.75	79.5	76.6	-2.9	1.14	10.4	14.2	3.8
1+0	0.60	38.3	38.6	0.3	0.76	46.0	49.2	3.2
1+1	0.67	45.5	44.7	-0.7	1.10	67.4	67.7	0.4
1+2	0.74	40.8	35.4	-5.4	1.44	80.6	81.9	1.3
1+3+	0.85	39.3	41.2	1.9	1.64	84.7	91.1	6.4
3p	1.80	17.2	21.4	4.2	1.68	15.0	19.5	4.5
4p	2.60	28.6	31.1	2.5	2.15	14.7	18.5	3.8
5+	3.40	51.1	60.2	9.1	2.56	27.9	34.3	6.4
6+p	4.73	82.4	82.6	0.1	2.66	23.6	44.7	21.1
Total		39.9	37.7	-2.3		21.2	23.7	2.5

References

Atkinson, A. B. (1987), 'On the measurement of poverty' *Econometrica,* 55(4),749-764.

Atkinson, A. B. (1992), 'Measuring poverty and differences in family composition', *Economica,* 59,1-16.

Atkinson, A. B. and F. Bourguignon (1987), 'Income distribution and differences in needs', in G. R. Feiwel, ed., *Arrow and the Foundations of the Theory of Economic Policy,* Macmillan, London.

Bourguignon, F. (1989), 'Family size and social utility, income distribution dominance criteria', *Journal of Econometrics,*42, 67-80.

Bradbury, Bruce (1992), *The Welfare Interpretation of Family Size Equivalence Scales,* SPRC Discussion Paper No. 37, Social Policy Research Centre, University of New South Wales, Sydney.

Foster, J. E., J. Greer and E. Thorbecke (1984), 'A class of decomposable poverty measures', *Econometrica,* 52, 761-76.

Jenkins, Stephen J. and Peter J. Lambert (1992), *Ranking Income Distributions When Needs Differ,* Paper presented at the Twenty-Second General Conference of the International Association for Research on Income and Wealth, Flims, Switzerland, 31 August - 5 September.

Keen, Michael (1992), 'Needs and targeting', *Economic Journal,* 102, 67-79.

Comments on paper by Bruce Bradbury

Michael C. Wolfson
Statistics Canada and Canadian Institute for Advanced Research

This is basically a good paper. I particularly like it for the following reasons: first, the analysis recognises that comparison of incomes across families of differing sizes and compositions is problematic. The standard method of positing an 'equivalent adult unit' (EAU) scale and re-expressing total family income as 'equivalent' family income has no empirical foundation. The choice of EAU scale is arbitrary, and a wide range is encountered in practice (Buhman et al., 1988).

I would go further than Bradbury's comment that 'economic research has not been successful in calculating equivalence scales which command wide-spread acceptance' and argue that we shall never have an empirically grounded EAU scale, for example for use in poverty or income distribution analysis. The reason, most generally, is the impossibility of making inter-personal comparisons of utility. The way the question is posed in those econometric analyses that do seek to answer the question of what EAUs actually are is of the form 'what income y_i would leave this family type i just as well off as a family of type j with income y_j?' This is fundamentally an unanswerable question.

Of course, one can adopt a neo-classical approach and pretend to answer it as follows. First suppose that family expenditure decisions are based on maximisation of a utility function. Then with sufficiently rich data, estimate a system of family expenditure functions; 'integrate' the resulting expenditure system to recover the underlying utility function; and use that indirectly estimated utility function to derive the EAUs. Unfortunately, this suffers from basic flaws, as we discovered when we pursued it at Statistics Canada several years ago with the help of Martin Browning (Wolfson and Evans, 1991; Browning, 1988).

To be brief, we found first that the expenditure functions themselves were quite heterogeneous across different household types (e.g. blue or white collar occupation, family size, presence of children). Second, the more general functional forms which were **not** rejected by the data were not integrable, so no well-formed underlying utility functions could be recovered. Finally, even had the expenditure functions been integrable, they would have been unique only up to a monotone transformation. This is insufficient to set an EAU; it requires cardinal rather than just ordinal utility functions.

Furthermore, when we restricted the expenditure systems to forms that were integrable, we always found that the resulting utility functions were of a form that any implied EAU scales would vary by income as well as family size and composition.

As a result, the problem Bradbury is addressing is of fundamental importance. He is asking how in practice income distributions that do cover heterogeneous families can be compared. I strongly agree with his premises that some sort of account ought to be taken of family composition, and that the most practical approach is a form of sensitivity analysis. His paper is also valuable in looking not only at headcount approaches to measuring the extent of poverty, but also in trying to take account of the depth of poverty faced by various families.

Finally, the paper is to be commended for being explicit about sampling error; far too many published income distribution studies fail to acknowledge questions of statistical significance when comparing income distribution attributes like inequality (see Wolfson elsewhere in this volume), or, in this case, poverty measures. Nor is the paper shy regarding the obvious corollary - that the results should be expressed as ranges rather than point estimates.

The balance of my discussion is concerned not so much with what the paper is doing than with 'how'. Since the question of how one might adjust for family size and composition in poverty measures is so basic, it is highly desirable to have robust and accessible methods to accomplish the kinds of sensitivity analysis Bradbury has developed. Bradbury has proposed a method that is principally analytic. My sense is that there may be alternatives based on more computer intensive methods that will prove more intuitive, and easier to implement. Let me sketch three such alternatives.

To begin, let me re-express the poverty rankings developed in Bradbury's paper in a somewhat different, and I think more parsimonious manner.

Let $F_i(y)$ = proportion of individuals in family type i with actual income $\leq y$

m_i = number of equivalent adult units (EAUs) in family type i, where $m_1 = 1$ by definition, and $m = \langle m_i, m_2, \ldots, m_k \rangle = \{m_i\}$

z = y/m_i = equivalent income

so that $F_i(y) = F_i(m_i z)$.

We then define a class of poverty indices

$$P_F(t;m) = \sum_i \int_0^t p(z,t) f_i(m_i z)\,dz$$

where $m = \{m_i\}$ is an EAU scale,

t is a poverty line defined in terms of equivalent income

and $f_i(y) = d\,F_i(y)/dy$.

As noted in the paper, two classes of poverty index can be defined. The first, containing conventional headcount measures, requires:

(I) $\partial p(z,t)/\partial z \leq 0.$

The second class, containing poverty gap measures, is the subset of class I where

(II) $\partial^2 p(z,t)/\partial z^2 \geq 0.$

Let $F(y) = \Sigma_i F_i(y)$ and $G(y) = \Sigma_i G_i(y)$ be two income distribution cumulative density functions
(i.e. $\lim y \rightarrow \infty$ $F(y) = G(y) = 1$; also with $F(0) = G(0) = 0$).

Then define two orderings: FDPD (SDPD) is 'first (second) degree poverty dominance'.

Again, as noted in the paper (using somewhat different notation) where t* is some upper bound poverty line

$F \geq_{FDPD} G$ $\Leftrightarrow$ F has a smaller poverty headcount $\forall\, t \in [0, t^*]$

$\Leftrightarrow$ $P_F(t^*; m) \leq P_G(t^*; m)$ and p class I

$\Leftrightarrow$ $\Sigma_i F_i(m_i z) - G_i(mi) \leq 0.$ $\forall\, z \in [0, t^*].$

Similarly,

$F \geq_{SDPD} G$ $\Leftrightarrow$ F has a smaller poverty gap $\forall\, t \in [0, t^*]$

$\Leftrightarrow$ $P_F(t^*; m) \leq P_G(t^*; m)$ and p in class II

$\Leftrightarrow$ $\Sigma_i \int_0^z F_i(m_i z) - G_i(m_i z) \leq 0$ $\forall\, z \in [0, t^*].$

The essence of Bradbury's paper is to structure a set of sensitivity analyses for a variety of equivalence scales

$m \in \mathcal{M} = \{m \mid m \text{ is "reasonable"}\}.$

Bradbury defines $\mathcal{M}$ lexicographically or in terms of a series of 'nested constraints' where

$$m_1 = 1$$

$$1 \le \underline{m}_2 \le m_2 \le \overline{m}_2$$

$$m_2 \le \underline{m}_3 \le m_3 \le \overline{m}_3 \le 2m_2 - m_1$$

... etc.

and $\underline{m}_i$ ($\overline{m}_i$) is an externally imposed lower (upper) bound.

We are now in a position to sketch several different methods for comparing the extent of poverty in two income distributions F and G. These methods all trade simpler mathematics for greater computational intensity compared to the method developed by Bradbury. However, with the continuing dramatic fall in the cost of computing, this may be a reasonable trade-off.

The first method is not really rigorous in the sense of assuring first or second degree poverty dominance as just defined. But it does have the benefit of simplicity. We simply define a one-parameter class of equivalence scales - for example the power class in Buhman et al. (1988): $m_i = i^\alpha$ where $\alpha \in (0,1)$. We also define a range of poverty lines z, say in terms of median family income adjusted with an equivalence scale in turn based on a value of α in the middle of its range. We can then calculate headcount poverty incidences for both F and G stepping systematically through the ranges of α and z. The results of these 'brute force' computations can be displayed simply as a set of counts for F – G in a rectangular grid of cells where the rows correspond to different values of z, and the columns to different values of α. Alternatively, and a bit fancier, a contour plot could be interpolated from the rectangular grid of poverty headcount differences on the same (z, α) two-dimensional space showing curves where F – G was constant.

This approach, at least for now, likely involves batch processing of the underlying microdata. A more interactive graphical user interface (GUI) version of this procedure is also readily imagined. The F and G cdfs are each simplified and represented by two dimensional tables, where each column corresponds to a family type i, and each of 100 rows gives the average income of one per cent of the families in the column (after ordering by income). It would then be possible to have a GUI window with two 'verniers' or 'sliders' and one output box. One vernier would correspond to the range of z, while the other denoted the range for α. The output box would give the resulting headcount poverty difference F – G. The user of this software could, after loading the percentile data from F and G by family type, use his or her point device to slide the z or α verniers up or down, and then watch in real time the effect on the comparative headcounts. In other words, the user could explore manually the (z, α) space to see in which regions F had more or less headcount poverty than G.

The approach just sketched is rather restrictive in the space of equivalence scales because they are confined to the one-parameter 'power law' class. Building on the 'nested constraints' idea in the Bradbury paper, we can easily imagine a second and somewhat more general approach. This time, we focus on $P_F - P_G$ for arbitrary $p(z, t)$. For convenience, the bivariate income/family type distributions F and G can be approximated by percentile matrices just as in the previous paragraph. The idea here is to use a monte carlo procedure to draw repeatedly vectors m at random from the set of nested constraints that have been used to define a 'reasonable class' of EAU scales. (An algorithm for such a stochastic process is readily specified).

Next for each randomly drawn m, we compute $P_F - P_G$ for a series of zs in discrete steps up to a maximum z*. Finally, the results can be summarised in a set of 'box and whisker' plots as sketched in Figure 1 - where the boxes show the mean, median and quartile values of $P_F - P_G$ at each z, and the whiskers the 5th and 95th percentiles over, say, 100 monte carlo draws of vectors m from the set defined by the nested constraints. If, as in Figure 1, the whiskers all lie above the horizontal axis, we can be reasonably confident that F has more poverty than G (conditional on the poverty index function $p(z, t)$ and the upper limit z*).

This second approach is still not sufficiently general to test whether or not there is first or second degree 'poverty dominance'. For this purpose, we can imagine a third approach. Again, let us start with a pair of bivariate percentile income/family type matrices to represent F and G, and a 'reasonable class' of EAU scales that have been defined by some set of nested constraints. From the discrete percentile approximates of F and G (plus other information from the underlying microdata if we wish), we construct continuous piecewise linear approximations of F and G. Given these data, the next step is a monte carlo process as follows:

1 draw a vector m from the set of nested constraints;

2 given m, step along z from $z_{min} > 0$ to the first value z^1 or an upper bound z_{max} where $\Sigma_i\, G_i\,(m_i z) - F_i\,(m_i z)$ changes sign; and

3 start again and step along z from $z_{min} > 0$ to the first value z^2 or an upper bound z_{max} where $\sum_i \int_0^z G_i\,(m_i z) - F_i\,(m_i z)$ changes sign.

Figure 1

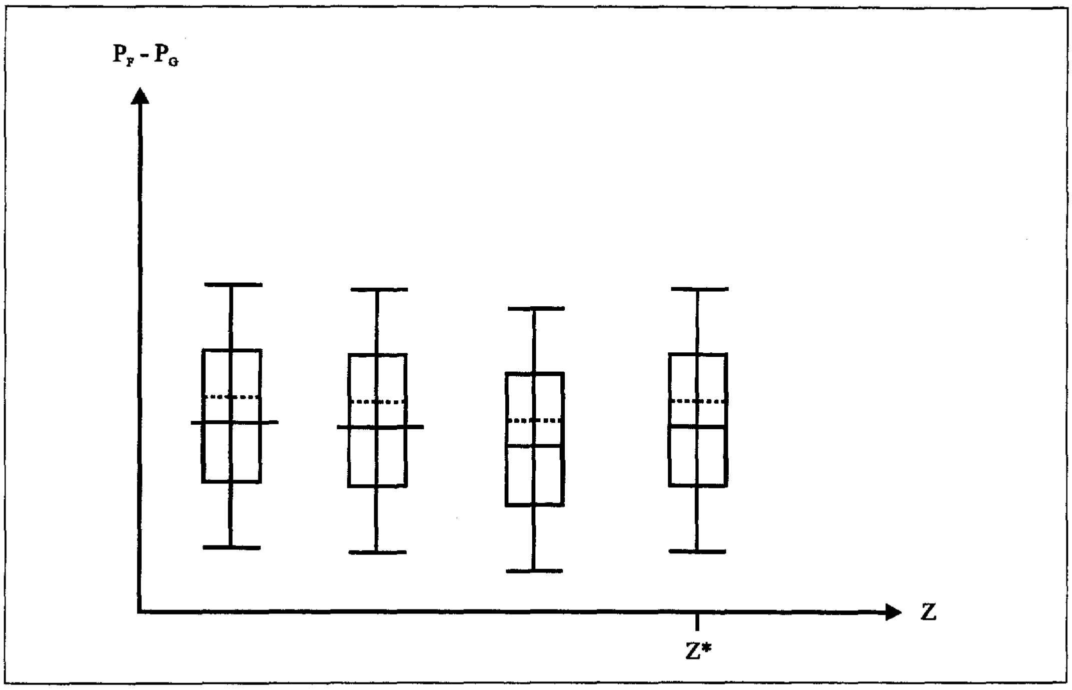
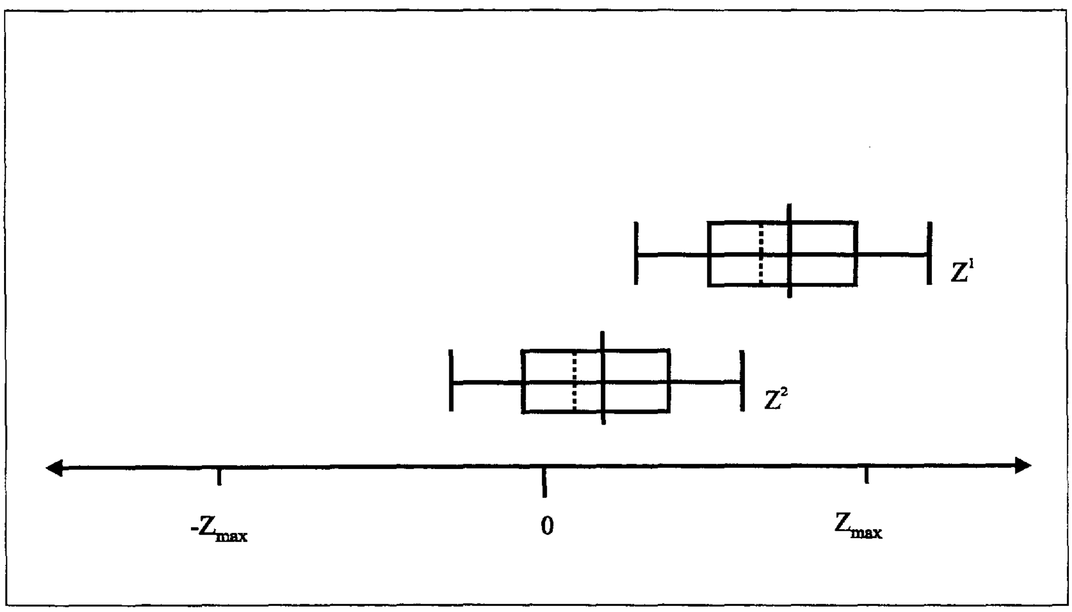


Figure 2



The values z^1 and z^2 then represent the maximum equivalent incomes where F first and second degree poverty dominates G respectively (assuming F and G have been labelled such that F poverty dominates G at z_{min}). Finally we collect these values z^1 and z^2 over 100, say, monte carol replicates and display the results in another box and whisker plot such as that sketched in Figure 2. This figure clearly shows a situation where F first degree poverty dominates G (conditional on the nested constraints set for the EAU scales and z_{max}), but the second degree poverty dominance is ambiguous.

Finally, there is the question of sampling error in these rankings. The mathematically straightforward but computationally intensive solution is simply to nest the procedures outlined above within a larger sample re-use monte carlo loop - a bootstrapping method of variance estimation. For example, twenty-five or more pairs of bivariate percentile income/family type matrices can be drawn (same size, drawn with replacement) to represent the distributions F and G from the full underlying microdata samples. Then either of the procedures above can be repeated for each of the twenty-five versions of the F and G distribution data in order to give an indication of the sampling error. If these sub-samples were drawn in a way that took account of the complexities of the underlying sample design, then the method would also avoid the usual simplifying assumption, as noted by Bradbury, that the data came from a simple random sample.

References

Browning, Martin (1988), 'The effects of household characteristics on behaviour and welfare', Report prepared for the Social and Economic Studies Division, Statistics Canada, July, mimeo.

Buhman, B., L. Rainwater, G. Schmaus and T. M. Smeeding (1988), 'Equivalence scales, well-being, inequality and poverty: sensitivity estimates across ten countries using the Luxembourg Income Study (LIS) database', *Review of Income and Wealth*, 34(2), June, 115-42.

Wolfson, M. and J. Evans (1991), *Statistics Canada's Low Income Cut-Offs: Methodological Concerns and Possibilities*, Analytical Studies Branch Research Paper, Statistics Canada, Ottawa.

Comparing Low Incomes in France and the United Kingdom: Evidence from Household Expenditure Surveys

A. B. Atkinson, Karen Gardiner,
Valérie Lechene and Holly Sutherland
Microsimulation Unit
University of Cambridge[1]

1 Introduction

Are there more people with low incomes, relative to the average, in the United Kingdom (UK) than in France? Is the aggregate poverty gap larger in France? To the extent that there are more people in poverty, are they primarily living in large households? To answer this kind of question, we need data on the resources of individual households in the two countries. This paper is concerned with the use of household survey data to make low income comparisons in France and the UK.

Recently cross-country comparisons of poverty have been given prominence in the report of Eurostat (1990) on 'Inequality and Poverty in Europe (1980-1985)'. This showed that the proportion of the population below 50 per cent of the national average in 1985 was 15.7 per cent in France, which was about the average for the European Community, but rather higher (18.2 per cent) in the UK. This ranking of France and the UK - with more poverty recorded in the UK - is different from that in the earlier study for the same year by O'Higgins and Jenkins (1989), according to which 17.5 per cent of the population in France lived in households with income below 50 per cent of the average, compared with a corresponding figure for the UK of 12.0 per cent. On this basis, there was less poverty in the

1 We are grateful to the Nuffield Foundation and to STICERD for their support of the project on poverty comparisons. The main part of the work for this paper was done during a visit to STICERD by Valérie Lechêne of DELTA, Paris, with the support of INRA (Institut National de Recherche Agronomique). At this time A. B. Atkinson was Tooke Professor of Economic Science and Statistics and Holly Sutherland was a research fellow with the Welfare State Programme at the London School of Economics. Since October 1992 they have been respectively Professor of Political Economy and Director of the Microsimulation Unit at the University of Cambridge. Karen Gardiner is a Research Officer at the LSE. A summary of the main results is given in Atkinson, Gardiner, Lechêne and Sutherland (1993).

UK. The same ranking emerges from the **national** studies which have been carried out. Comparing the *Households Below Average Income* (HBAI) study in the UK (Department of Social Security, 1990) with that carried out by Assémat and Glaude (1989) for France, we find 9.2 per cent of the population in households with incomes below 50 per cent of the mean in the UK, compared with 10.9 per cent in France in households with incomes below 50 per cent of the median. Allowing for the difference between mean and median (Atkinson and Cazes, 1990: 117), the proportion with low incomes again appears higher in France.

The comparison of poverty rates across countries is however a matter of some subtlety and we need to look at what lies behind figures such as those just cited. To begin with, the Eurostat figures related to **expenditure** as an indicator of resources, whereas the other figures cited all refer to **income**. Here we concentrate on those with low **incomes**. Secondly, we have to ask whether the income data are drawn from the same kind of source. In fact, the study by O'Higgins and Jenkins (1989) was based on sources which were different for the two countries. Whereas the UK data were drawn from a household budget survey, the French data used by O'Higgins and Jenkins, drawing on the work of Canceill, were derived from fiscal sources. Here we concentrate on what can be said using evidence from the same kind of source: household budget surveys.

The main point of the paper however is that even where the estimates of the low income population are all based on the same kind of source (household budget surveys) there are a number of reasons why estimates of the low income population may not be comparable. We have just noted the difference between the mean and the median, but there are other important differences. An apparently tight specification of the low income criterion may still allow for considerable variation. In this paper, we assume that 'low income' is defined as having a household equivalent disposable income below a specified percentage (such as 50 per cent) of the national average. Yet, as emphasised in Atkinson and Cazes (1990) and Atkinson (1991), there can still be differences in such matters as the period over which income is measured, the equivalence scale, the weighting of units, and the choice of summary measure. And these variations can affect the conclusions drawn.

Moreover, the adoption of a common criterion does not in itself eliminate sources of ambiguity in making comparisons, since the choice of the common method to be applied in both countries is likely to affect the comparison. Two different bases for standardisation may lead to diametrically opposed conclusions. The results of Mitchell (1991) based on the Luxembourg Income Study show, for example, that the ranking of France and the UK depends on the level of the low income cut-off. She estimates that with a cut-off of 40 per cent of the median, France has a higher poverty rate, that at 50 per cent of the median the rates are similar, and that at 60 per cent of the median the UK has the higher rate (1991, Table 4.1, OECD equivalence scale). The cumulative frequency appears from her findings to cross at around 50 per cent of the median. The choice of cut-off crucially affects the results of the comparison.

In order to investigate the sensitivity of the low income estimates, we need access to the raw micro-data. This approach has been pioneered in the Luxembourg Income Study (LIS), which has collected micro-data sets and has put cross-country comparisons on a much more systematic basis: for example, Buhmann et al. (1988) and Smeeding et al. (1990). The LIS data set covers a wide range of countries; here we concentrate on France and the UK. The LIS project contains data from a range of types of source. In the case of the UK, the data are from the household budget survey; in the case of France, the data are those from the fiscal returns. Since we are concerned with what can learned from data from the same type of source - household budget surveys - we have not used the LIS data, but have had direct access to the micro-data tapes for each country.

To sum up, in this paper we consider what can be learned regarding the extent and composition of the low income population in the mid-1980s from a detailed examination of the household survey data for France and the UK. The comparability of the sources is the subject of Section 2. The problems of ensuring consistent definitions, and the range of possible bases for comparison, are the subject of Section 3. The results obtained from our comparisons of the budget surveys in France and the UK are presented in Section 4, and in Section 5 we explore more fully what lies behind the findings. The conclusions are summarised in Section 6.

2 Problems in Making Comparisons: Data Sources

In this section, we consider first the data sources in each country separately; we then make a comparison.

2.1 United Kingdom

The principal source of evidence about low incomes in the United Kingdom at the present time is the Family Expenditure Survey (FES). The primary purpose of this continuous household budget survey is to collect the expenditure information necessary to construct the weights for the Retail Prices Index, but it also collects a substantial quantity of income data. There are extensive questions covering the income of each household member from a wide range of sources, together with other information on household composition, housing costs, etc.

The FES is a representative sample of private households in the UK. The sample size is about 11,000, equivalent to around a one in 2,000 sample of all private households. The response rate in 1985 was 67 per cent, giving a total of 7,012 households. In only about two per cent of cases was the interviewer unable to contact anyone at the selected address; the primary reason for non-response (31 per cent) was refusal to co-operate. There is evidence that the characteristics of non-respondents differ from those of respondents. A special study in 1981 (Redpath, 1986) found a lower response among households without children and where the head was self-employed; there was a fall in response with the age of the head of household. In order to adjust for differential non-response, differential grossing-up weights are applied in a number of studies (see Atkinson et al., 1988).

The procedure applied here for grossing-up is based on the weights devised by the Department of Social Security (DSS). These vary with family composition and age of the family head.[2] There are however three problems in applying their weights: (i) they relate to Great Britain, excluding Northern Ireland, (ii) the DSS exclude certain cases (such as those with absent spouses) and their weights apply to the sample excluding these cases, and (iii) the weight applies to family units rather than to households. We have re-calculated the weights to cover the UK (assuming the same family composition as in Great Britain) and to allow for the exclusions. When calculating results on an **individual** basis, the weights for a family unit are applied to each individual in that unit. When calculating results on a **household** basis, the weight is the mean of those for the individuals in the household.

The sample studied here is that for the United Kingdom, and all households interviewed are included. Unlike the DSS, we have not excluded families where there is recorded to be an absent spouse. We have however re-coded a number of cases where a detailed examination shows there to be inconsistencies (e.g. where the woman is coded as married with spouse absent and the man is coded as married with spouse present). We have re-coded foster children to show them as dependants, rather than as separate family units; and we have re-coded certain students living at home as separate units. These re-codings do not affect total household income, but do affect the grossing-up weights applied and, in some circumstances, the equivalence ratios.

2.2 France

In France, two major sources have been used to assess the extent of low incomes. The first is the 'Enquête sur les Revenus Fiscaux', referred to as ERF, based on income tax declarations. The inquiry has been carried out periodically since 1956, and among the most recent are those for the years 1979 and 1984. The ERF proceeds by drawing a representative sample of the population for which the income tax declarations are supplied by the authorities. Where no declaration is found, this 'non-response' is allowed for by re-weighting of respondents. In the ERF for 1984 there are some 47,000 households. The majority of non-taxable income (family benefits, *minimum vieillesse*) is incorporated by an imputation procedure. This source is used in studies of income inequality in France (see

2 In the case of a couple, the head of family is defined to be the husband for the purposes of analysis.

Canceill and Villeneuve, 1990) and it is these data that are lodged in the Luxembourg Income Study.

The second source in France is the household budget survey, the 'Enquête sur les Budgets Familiaux', referred to as EBF. The EBF is conducted periodically and here we refer to the 1984-5 survey carried out between mid-1984 and mid-1985.[3] Information is obtained by interview on expenditure, income and other variables. The sample of approximately 20,000 is about twice the size of that in the UK FES, representing approximately 1 in 1,000 households (in 1984-5 there were 20,326,000 households in the total population of Metropolitan France). Non-response reduced the effective sample is set out in Table 1. This led to 11,977 cases being included in the INSEE data base from which we started.

In our analysis below, we use only those returns where the income questions were answered, and discarded those for which simply a range of income was given. The size of the effective sample is further reduced by the fact that there are households which answered the income questions but for which there are missing values for certain items of income. We use only those returns where there is a full answer to the income questions. We adopted this strict definition in order to ensure comparability with the United Kingdom FES data tape, where there are no missing values. It should be noted that this absence of missing values reflects the use of imputation procedures in the case of item non-response. The French EBF data are in this sense less 'processed'. In our analysis of the EBF data, we have also excluded cases of zero income which did not appear to be coherent with the other information supplied. This led to a final sample of 9,837. This may be compared with the figure of 11,076 used by Assémat and Glaude (1989) in their work on the low income population.

To adjust for differential non-response between different types of household in the French EBF, a grossing-up procedure is applied to yield results representative of the population. The weights are calculated for households grouped into 115 categories based on the type of commune (local authority district), socio-economic category of the household head[4], and number of persons in the household

3 The survey was carried out between 18 June 1984 and 19 July 1985, excluding the first two weeks of August and the second two weeks of December. A few interviews (35) were carried out before 18 June 1984, and a few (23) after 19 July 1985.

4 The definition of head of household (HoH) in multi-person households in the EBF is as follows (it must be an adult aged 18 or over):

i) if there is a couple, then man = HoH,
ii) if more than one couple, oldest working man = HoH, if no working man then oldest man = HoH,
iii) if no couples but one parent with children, then parent = HoH,
iv) if no couples and several parents with children, oldest working parent = HoH, if no working parent, then oldest parent = HoH,

Table 1: Reasons for Non-Response

	%
Absent	6.7
Refusal to participate	11.1
Failure to complete	13.0
Refusal to income questions	0.9
Income questions only answered in ranges	4.3

Source: Assémat, 1989, Table 1.

(Moutardier, 1988: 12). In our analysis, we have applied these weights, but it is not clear that they apply to the particular sample that we are using. The exclusion of households with item non-response for income may not be adequately allowed for.

As already noted, the French data have been less extensively processed than those in the FES. This necessitated re-coding of certain observations. The need may be illustrated by the case of family benefits paid as part of the wage. Respondents were first asked (Question 26) about their wage, and then about family benefits, including whether (Question 29) these are included in the wage reported in response to the earlier question. Where the benefit was reported as included, it is then subtracted from the wage, the benefit appearing as a separate item. In certain cases this procedure generates a negative net income, and we have assumed that the second question was not answered correctly. This example is mentioned, not because it is quantitatively particularly important, but because it illustrates the difficulty in replicating the calculations of others. Quite a lot has to be done to the data tape before it can be used to prepare tabulations.[5]

The existence of both fiscal and budget survey data in France is analogous to the relation between the Survey of Personal Incomes (SPI), based on income tax returns, and the FES in the United Kingdom. The position is, however, rather different in that the SPI does not provide adequate coverage of non-taxpayers and that the overall income distribution statistics published by the Central Statistical Office (CSO) combine information from the SPI with that from the FES. These statistics have not been used to measure the extent of low incomes. In France, on

v) if no couples and no children, oldest working person = HoH, if no working person, then oldest person = HoH.

5 A second example is that the original data showed 37 households with more than one person apparently married to the head of household. By examining the individual observations it was possible (using the marital status codes) to identify the true spouse and correct this error.

the other hand, the ERF has been used to measure the extent of low incomes. Suesser (1988) has used the ERF from 1979 to compare the position of France with that of other countries in the LIS data set, including the United Kingdom.[6]

The two French sources have been compared by Assémat and Glaude (1989) and Assémat (1989). They have made use of the fact that part of the sample from the 1984 ERF was made up of the households in the sample for the 1984-5 EBF, which allowed a matching of households in the two surveys. This in turn permitted an examination of differential non-response and of the comparability of the income figures reported. There is an inverse-U shaped relationship with income: the response rate was around 2/3 in the middle income ranges but fell to 53.7 per cent in the bottom decile and to 51.0 per cent in the top five per cent (Assémat, 1989, Table 1). This suggests that, to the extent that this is not corrected by the re-weighting, the proportion of low and high incomes would be under-stated in the EBF. The calculations of Assémat and Glaude show 8.9 per cent of households below 50 per cent of the median in the ERF in 1984 (a figure similar to that found by Suesser for 1979), but that when attention is restricted to the matched sample of respondents to the EBF, the proportion falls by 1 percentage point.

Assémat goes on to compare the responses to the two surveys and concludes that, if we exclude those answering only in terms of ranges, that on average total income is very close: the mean response in the EBF is 99.6 per cent of that in the ERF and the median 98.8 per cent (Assémat,1989, Table 4).[7]

One important difference concerns the reporting of income-related social security benefits. The ERF imputes receipt of minimum pension (*minimum vieillesse*) in all cases of apparent eligibility, whereas the EBF relies on reported information. The latter records many fewer recipients (656,000) than are imputed in the former (1,540,000). Under-reporting in the EBF almost certainly leads to the size of the low income population being over-stated. Assémat notes that conversely, non-take-up may mean that the ERF understates the low income population:

> L'imputation automatique du minimum vieillesse dans notre source fiscale alors qu'il existe des bénéficiaires potentiels ne faisant pas valoir leurs droits, risque quant à elle de remonter artificiellement le bas de la distribution pour les inactifs. (Assémat,1989: 145)

6 He finds (Suesser, 1988, Table 2) that 8.4 per cent of the population in France had an equivalent household disposable income less than 50 per cent of the median, although he notes that the proportion in France would rise to 9.86 per cent (Table 3) if the unit of analysis were the tax unit.

7 The comparison also shows that the answers of those who only responded in terms of the range of total income had a strong downward bias (of the order of 30 per cent). We did not make use of these responses in our work.

On the other hand, the number of recipients of the *minimum vieillesse* reported in the ERF, after imputation, of 1,540,000 (Assémat and Glaude, 1989: 13) is close to that recorded in the administrative statistics: 1,553,000 in 1984 and 1,487,000 in 1985 (Annuaire des Statistiques Sanitaires et Sociales, 1990: 182). The comparison of the results for the matched EBF-ERF sample by Assémat and Glaude shows that overall the percentage of households with equivalent disposable incomes less than 50 per cent of the median was 7.9 per cent in the ERF, compared with 10.1 per cent in the EBF (Assémat and Glaude, 1989, Table 1). The proportions less than 60 per cent of the median were 13.3 per cent in the ERF and 17.5 per cent in the EBF.

To sum up, the two sources of data in France have their relative advantages and disadvantages. The fiscal data (ERF) understates the low income population, but probably to a lesser degree than it is overstated in the budget survey (EBF). On the other hand, there can be little doubt that the EBF is closer to the FES in the UK, and it is the EBF data that we use in this paper. The fact that both the EBF and FES are household budget surveys should not however be allowed to obscure the fact that there may be significant respects in which they are not comparable, as is explained below.

2.3 Comparison of the Budget Surveys in France and the UK

We list below the main differences between the two surveys which are likely to affect the conclusions drawn with regard to the size of the low income population.

Choice of Sample

As already noted, in our analysis of the French EBF, we use only those returns where there is a full answer to the income questions, and we exclude all households where there is a missing value for any item of income. This strict definition is adopted in order to ensure comparability with the UK FES data tape, where there are no missing values. However, it should be noted that the FES uses imputation procedures in the case of item non-response, and it is possible that the exclusion from the FES of households where there had been imputation would affect the conclusions drawn in our comparison.

Grossing-up

In both cases we apply differential grossing-up factors to allow for differential non-response, with the intention of arriving at estimates representative of the total population. This adjustment can however only be approximate. In the UK, the weights vary with family composition and age of the family head, but their application may cause the sample to become less representative in other respects. In the case of the French data, these weights are those applied to the original set of respondents and do not allow for our exclusion of those with incomplete income information. It is not evident how this affects the results.

Coverage of Income

The aim is to measure the total net flow of resources accruing to the household in a specified period. However, this ideal is unattainable and the only feasible procedure is to form a sum of those items available from the survey information which form a reasonably consistent aggregate. Thus, the accretion of resources in principle includes capital gains and losses on asset holdings, but information on accrued gains is not available, and we make no attempt to include this item, leaving out of consideration information such as that on the disposal of assets (the French EBF includes data on the sales of durables and land). Similarly, no attempt is made to impute rent on assets such as owner-occupied houses or consumer durables.

The main components of income are those from earnings, self-employment income, state social security benefits (including housing benefit), private transfers (including occupational pensions and maintenance payments), investment income, and certain forms of income in kind. From these are deducted income tax and social security contributions. Components corresponding to these may be identified in the two surveys, but there are undoubtedly differences in the way that they are implemented. The UK questionnaire seeks information about income in considerable detail: the Income Schedule is 39 pages long and contains 91 questions. The French survey contains 10 questions and is less detailed, although there is scope for adding supplementary information (for example describing sources of income), and this is incorporated in the coding. This difference in the degree of detail is likely to affect more seriously certain types of income. In the case of self-employment income, for example, respondents to the French EBF were simply asked about 'the resources received in the previous 12 months' (Question 33), whereas in the UK FES there are 3½ pages of questions for the self-employed which gave alternative methods of estimating their income. It is for this reason, amongst others, that in the Appendix we present results excluding households where the head is self-employed.

There are also differences in the time period over which income is measured. Both surveys seek to measure 'current' resources, but to eliminate transitory variations. In the case of earnings, the French question asks 'if earnings are not regular, give a monthly average'; in the UK, the calculations are based on the 'normal' earnings variable where different from the last pay-period. The surveys accept also that certain forms of income are reported over a longer period. Thus, in the case of interest income, the FES figures in the UK relate to that received over the past 12 months. This applies to rather more types of income in France, and it may be said that the French definition is closer to an annual basis.

Comparison with External Information

In both countries, certain forms of income tend to be under-stated in the budget surveys: interest and dividend income, and the income of the self-employed. The relative importance of these differs however across the two countries. This applies in particular to the income of farmers in France, reckoned to be under-reported by about 50 per cent in the EBF (Assémat and Glaude, 1989: 12),[8] where home production may also be more important. This re-enforces the interest in looking at results excluding the self-employed.

Top- and Bottom- Coding

In the French data, there is an upper limit to the income that may be recorded under certain headings, but this does not apply to any variables for the cases considered here. In the French data, no negative values are recorded, and for comparability we have set negative incomes to zero in the UK data.

To sum up, access to the micro-data from the surveys in the two countries allow us to go a considerable way towards harmonising the measures of the low income population, as is discussed below, but there remain differences between the sources which have to be borne in mind when considering the results. In some cases these differences could be eliminated by changes in the surveys in the two countries, but in other cases they reflect differences in structure between the countries (such as the relative importance of farm income).

3 Problems in Making Comparisons: Definitions

In this paper, low income is assumed to be defined in terms of a household having an equivalent disposable income of less than a specified percentage of the average. This is a restrictive definition. It accepts that the standard is a **relative** one, rather than related to some external standard. It is based on **income** rather than the expenditure concept now espoused by Eurostat. It accepts the household as the unit of analysis, assuming a degree of income-pooling which may not in fact take place (see for example Atkinson, 1991).

Yet even this restrictive definition allows considerable room to manoeuvre. The definitional issues which remain include:

- choice of mean or median,
- choice of equivalence scale,
- weighting of different units,
- before or after housing costs, and
- headcount, poverty gap or other poverty indicator.

---

8 In part, the difference may be one of timing, in that self-employment income is reported for an earlier period and hence needs to be up-rated.

These are discussed in turn below, as is the issue of sampling error and statistical significance.

3.1 Choice of Mean or Median

The first European Poverty Programme adopted as its measure of poverty a cut-off equal to 50 per cent of the **mean** equivalent disposable income, and this was followed in O'Higgins and Jenkins (1989). In contrast, the majority of studies using the Luxembourg Income Study data, such as Smeeding et al. (1990), Buhmann et al. (1988) and Mitchell (1991), have employed the **median** in place of the mean. This may make a significant difference, in view of the fact that the median is typically considerably less than the mean. The calculations of Assémat and Glaude (1989, Table 1) using the EBF in France show for example that the median is about 88 per cent of the mean. A cut-off of 50 per cent of the median is like taking 44 per cent of the mean.

The choice between different measures of central tendency is discussed in this context by Hagenaars (1991), who notes that the median is less sensitive to sampling fluctuations. The mean is more affected by the occurrence of a few very high incomes in the sample. This argument depends however on the form of the tails of the distribution, and on the behaviour of the density function around the median.[9] In the UK, the Department of Social Security reports empirical investigation of the stability over time in the medians for decile groups that suggests that 'the median figures over the years are much more stable than those based on means' (Department of Social Security, 1990: 70).

Concern for sensitivity may be related not to sampling fluctuations but to errors of measurement. Here, we may note that the median is not affected by top-coding or bottom-coding of observations: for example, by setting negative entries to zero. The mean would be affected by such re-codings. On the other hand, the median is affected by the deletion of observations.

The choice may favour the median on statistical grounds, but a cut-off based on the mean income is more easily related to aggregate statistics, particularly when comparing across countries. Many people have in mind the differences in per capita national income when making such comparisons, whereas the relative median incomes may be different on account of the different distribution within the countries, as is indeed the case for France and the UK. In what follows, we consider both mean and median.

9 If equivalent income were normally distributed, then the standard error of the median would be 25 per cent **larger** than that of the mean (see Kendall and Stuart, 1969: 237). The standard error of the median is inversely proportional to the density at the median, so that it depends sensitively on the shape of the distribution (see Maritz, 1981: 28).

3.2 Choice of Equivalence Scale

The equivalence scale applied in the analysis of the French EBF by Assémat and Glaude (1989) is simple: 1 for the first adult, 0.7 for other adults and 0.5 for children aged less than 14. This scale is that recommended by the OECD (1982), and for shorthand we refer to it as the OECD scale; it was also used by O'Higgins and Jenkins (1989), along with two variations (raising the child allowance to 0.7 and reducing that for other adults to 0.5).

An approximate, but valuable, way of summarising differences in equivalence scales has been proposed by Buhmann et al. (1988), who parameterised the equivalence scale as n^s, where n is the number of household members. The exponent *s* summarises the differences in scales. If we take, for example, the case of a couple plus 2 children, the OECD scale is 2.7 (1+ 0.7+ 2 x 0.5). This is equal to $n^{0.72}$. Buhmann et al. (1988) show that a value for the exponent of 0.72 is towards the high end of the range of those used. Scales based on benefit parameters, and the official US poverty line, tend to have values around 0.55; estimates based on observed consumption patterns and identifying restrictions tend to be lower (Buhmann et al. take a value of 0.36 as representative); scales based on subjective evaluations (of what is needed 'to get along') tend to be lower still (around 0.25 is representative). The OECD scale tends to be relatively generous to large families; it can also be criticised for not being sufficiently finely graduated. The McClements scale applied in the HBAI study in the UK varies the amount per additional adult according to the number in the household (and is less for the spouse of the household head) and the amount per child is graded with age, as set out in Table 2.

The before housing costs scale for a couple with two children aged under 14 ranges between 1.94 and 2.52. The HBAI scale makes less allowance for this family than the OECD scale, particularly if the children are young.

In what follows, we examine the sensitivity of the results to the choice of equivalence scale, comparing the findings with the OECD and HBAI scales, and showing the effect of taking different values of the parameters.

3.3 Weighting of Different Units

As has been brought out by O'Higgins and Jenkins (1989), the mean or median can be calculated in different ways depending on how the units are weighted,[10] and the same applies to the calculation of the proportion with low incomes. For each household we calculate the income per equivalent adult. If the total income of household h is y_h, and the number of equivalent adults e_h, then the equivalent

10 For earlier discussions of this issue, see Atkinson and Harrison (1978: 244) and Danziger and Taussig (1979).

Table 2: McClements (HBAI) Equivalence Scale

		Before housing costs	After housing costs
Single adult		1.00	1.00
Couple		1.64	1.82
2nd adult		0.79	0.82
3rd adult		0.69	0.82
4th adult		0.59	0.73
Child age	16-17	0.59	0.69
	13-15	0.44	0.51
	11-12	0.41	0.47
	8-10	0.38	0.42
	5-7	0.34	0.38
	2-4	0.30	0.33
	0-1	0.15	0.13

Source: Calculated from Department of Social Security, 1992: 125 (the scale is quoted there in terms of a couple = 1.00).

income is y_h/e_h, and it is according to this that the households are ranked. The question is now: how do we weight these households when adding them up (to find the median or to calculate the proportion below the cut-off)? The same question applies when we add up the incomes of the households to form the mean.

There are at least three possibilities: a weight of unity to all households, a weight of e_h to household h, and a weight equal to the number of individuals in household h, denoted by n_h. As is pointed out in Atkinson and Cazes (1990) and Atkinson (1991), the first of these methods has been applied in French studies; the second in estimates for Germany; the third is that applied in the HBAI study in the UK. In terms of national aggregates, the second method has the attraction that the mean equals total income divided by total equivalent adults, and does not depend on the distribution of those individuals among the income units. On the other hand, if the individual is the basic unit for concern then there are attractions in giving everyone an equal weight, rather than a weight which depends on their status in the household (and age).

The sensitivity of the findings to the weighting is examined below, where we compare the household weights (first method) and the individual weights (third method).

3.4 Before or After Housing Costs

The HBAI calculations carried out by the UK Department of Social Security are made on two bases: before and after housing costs. In the former case, the figures relate to the distribution of equivalent net income, which includes housing benefit but makes no deduction for housing costs (including the interest paid on loans for house purchase). In the latter case, what is measured is net resources, defined as net income minus housing expenditure, again expressed per equivalent adult (with a different equivalence scale). The housing costs deducted are, for owner-occupiers, mortgage interest (but not repayment of principal), insurance of the structure, ground rent, water rates and rates (local taxes); for tenants, they are rent (not in principle including service charges), ground rent, water rates and rates.[11]

The second of these calculations - that of net resources after housing costs - may appear rather strange to observers from outside the UK. It has however to be borne in mind that the official series which preceded the HBAI tables - the estimates of Low Income Families (Department of Health and Social Security, 1988a) - took as its standard the Supplementary Benefit (SB) level, which consisted of a fixed scale plus an allowance for individual housing costs. The net resources available after deduction of housing costs was therefore the natural variable to compare with the SB scale. It may also be noted that the calculation of net income in distributional analyses based on UK income tax returns used to allow for the deduction of interest paid on mortgages.

The argument for using net resources rather than net income is not just tradition. It may be held that housing expenditure is a relatively exogenous element of a household's outgoings and one which varies across households in a way which reflects accidents of geographical location and tenure rather than the quality of the accommodation occupied. Such a view does not necessarily mean that we should concentrate **solely** on net resources. (For a clear discussion of the case for and against, see Johnson and Webb, 1988: 57-60.) We agree with the Technical Review of the Department of Health and Social Security that

> both measures - income before and after housing costs - have value and may throw light on the position of those

11 The definition of housing costs in terms of FES codes for tenants is variable 010 (net rent) if coded, otherwise 020 (which may include services if these cannot be distinguished, which applies to 80 households). For owner-occupiers, mortgage interest is taken as 130 where coded. Where this variable is not coded, 150 is used providing that it is sufficiently recent. Otherwise a formula based on length that mortgage has run (A133) is applied to the total mortgage payment (200). For owner-occupiers, add other regular expenses (060) and buildings insurance (110). Housing benefits and rebates are added back. For all households, add rates (030) and water/sewerage rates (050) and subtract heating and other service costs included in rent (211 and 212).

For the French calculations, we have made use of the following EBF variables. For loan repayments, MON081 - 84 (i.e. up to 4 loans) and the corresponding period variables PER081 - 84. For tenants, rent is based on MONLOY (amount of rent) and PERLOY (number of months covered). In both cases, if housing benefit has been subtracted (AIDED = 1), then it is added back based on MALLOC and PERALOC. Local taxes are based on MONTAX. If water paid direct (EAUDIR=1) then MONEAU added.

> in the lower income ranges. (Department of Health and Social Security, 1988: 23)

We have therefore made calculations of net resources after housing costs for France. It should, however, be noted that these calculations are not fully comparable because the mortgage costs deducted in the French case are the total costs (including repayment of principal).

Viewed in terms of comparisons across countries, the case for considering both results for net income and for net resources is that differences in net incomes could represent differences between the countries in their housing policies. One country may pursue a policy of low rents (either via low rents in the public sector or via rent control legislation in the private sector); the other country may provide increased income transfers. The proportion in poverty may be lower in the latter country when measured in terms of net income, even though the situation after housing cost is identical in the two countries. This is particularly obvious where the increased transfer takes the form of housing benefit, and this suggests that a third possible calculation is that of net income **excluding housing benefit**. Such a measure has been proposed by Johnson and Webb (1990) and is explored in Section 5.

3.5 Headcount, Poverty Gap and Other Poverty Indicators

We have so far talked in terms of the proportion of the population with low incomes, or the **headcount** measure of poverty. As was spelled out by Watts (1968) and Sen (1976), this indicator is insensitive to the severity of poverty; it does not distinguish situations where the poor are close to the income cut-off from those where there is a large gap. The simplest alternative is indeed the **poverty gap**, or the amount of income required to raise all households to the low income cut-off. This is expressed here as a percentage of the total income of all households:

$$\Sigma[e_h z - y_h] \quad / \ \Sigma \ y_h$$

sum taken over those below cut-off	sum taken over whole population

where z is the cut-off defined for a single person (e.g. 50 per cent of mean equivalent income), e_h is the number of equivalent adults, and y_h denotes total household income.

The poverty gap measures the total income shortfall; it does not take account of its **distribution**. A wide variety of alternative measures have been proposed, giving differential weight to the income short-fall at different distances from the cut-off. As argued in Atkinson (1987), it may be more fruitful to consider the poverty deficit curve. The deficit curve is analogous to the Lorenz curve and shows the cumulative shortfall from different cut-offs. For example, we first calculate the shortfall from a poverty line of 10 per cent of mean income, then 20 per cent, then 30 per cent, and so on up to the maximum poverty line of interest (100 per cent of the mean seems a reasonable upper bound).

3.6 Sampling Errors

In addition to the systematic sources of bias in the comparison of the two data sources, we have also to allow for sampling error. In the case of a pure random sample of n_i observations drawn from population i, the sampling error, s_i, for a proportion P_i estimated from that sample is

$$\sqrt{\{P_i(1-P_i)/n_i\}}$$

So that with a sample of 7,000 the sampling error for a headcount of 15 per cent is about 0.4 per cent. In order for the difference between the proportions drawn from two samples of size n_1 and n_2 to be significant at the 5 per cent (1 per cent) level, we require that the difference be greater than 1.96 (2.58) times[12]

$$\sqrt{\{P(1-P)[1/n_1 + 1/n_2]\}}$$

where P is the overall proportion from both samples combined. If the samples are of size 7,000 and 9,000, this requires a difference, with a proportion of 15 per cent, of 1.12 per cent to be significant at the five per cent level and 1.47 per cent to be significant at the one per cent level. These standard errors have, furthermore, to be adjusted for the fact that the surveys are not simple random samples but involve a multi-stage design. In the UK, the Central Statistical Office calculate a design factor, which in the case of mean gross normal weekly income in 1988 is 1.36 (1991: 135). Applying this factor, we may conclude that we are looking for differences of 1½ per cent for significance at the five per cent level and two per cent at the higher one per cent significance level.

4 Extent of Low Incomes in France and the UK

The results presented below are primarily concerned with the comparison of France and the UK, but we begin by explaining their relation to the evidence from national studies.

12 The difference is approximately normally distributed with variance equal to $s_1^2 + s_2^2$. Under the null hypothesis, the values of P_1 and P_2 are equal. See, for example, Hoel, 1962: 149-50.

4.1 Point of Departure

In the case of France, the natural starting point is the work of Assémat and Glaude (1989) which uses the EBF to arrive at estimates of the size of the low income population, defined relative to the median household income per equivalent adult. Their results are summarised in the upper part of Table 3.

Our own calculations are shown in the bottom line of the table. Although the method of calculation is the same (for example, we apply the same equivalence scale), the sample is different, consisting of the 9,837 cases described in Section 2. This sample excludes those where income is reported only in ranges, and hence is more closely comparable to the second line in Table 3 than to the full sample results in the first line.

Comparing the results from our sample with those of Assémat and Glaude excluding cases where income is only reported in ranges, we see that the median is 2.7 per cent higher with our sample, and that the mean is 1.2 per cent higher. The proportions with less than 50 per cent of the median are slightly lower with our sample: 9.6 per cent compared with 10.1 per cent below 50 per cent of the median. To the extent that we have been more restrictive in the criteria for inclusion (the sample size is smaller), our calculations may be more reliable; on the other hand, we have not been able to adjust the grossing-up factors to allow for this stricter exclusion principle. There are several other reasons why our figures may differ from those of Assémat and Glaude, including the treatment of missing values for those cases which are included. The main conclusion to be drawn, however, is that the differences between the results are re-assuringly small.

In the case of the UK, the natural starting point is the official Department of Social Security study of *Households Below Average Income*. For 1985 this gives the results shown in Table 4 for the mean income and the proportions of individuals in households below different percentages of the mean. There are again a number of reasons why our calculations differ from those made in the earlier study. These include the facts that our estimates relate to the whole of the UK, rather than Great Britain, that we have not excluded any households from the calculations (e.g. those with absent spouses), and our different treatment of negative self-employment income.[13] It would not therefore be surprising to find differences in the results. Our results for the mean income (see lower part of Table 4) are however very close. The proportion below 50 per cent of the mean is estimated by us to be 8.6 per cent before housing costs, compared with 9.2 per cent in the HBAI estimates, but this is the largest discrepancy. Again the comparison with the earlier study is re-assuring.

---

13 Where the self-employed report negative income, the DSS has replaced this by the amount reported to be drawn from the business for personal use (Department of Social Security, 1991, para 2.14). We have set to zero total income of **individuals** that are negative (before any deduction of housing costs).

Table 3: Comparison of Assémat and Glaude Estimates with Our Calculations: France, 1984/85

	Median (FF per year)	Mean (FF per year)	Less than 50% median	Less than 60% median
Assémat and Glaude				
Full sample (11,823 cases)	47,478	53,869	10.9	18.1
Excluding cases where income only reported in ranges (11,076 cases)	47,628	53,964	10.1	17.2
Matched sample with ERF,including cases where income only reported in ranges (7,251 cases)	47,881	54,557	10.1	17.5
Our calculations	48,937	54,604	9.6	16.8

Source: Assémat and Glaude, 1989: 5.

Table 4: Comparison of Official HBAI Estimates with Our Calculations: United Kingdom, 1985

	Mean (£ per week)	Less than 50% mean	Less than 50% mean
HBAI	86.55	9.2	20.1
Before housing costs	67.05	13.4	25.4
After housing costs			
Our calculations			
Before housing costs	86.58	8.6	19.9
After housing costs	67.13	13.6	25.0

Source: HBAI from Department of Social Security (1990), *Households Below Average Income 1981-87*, Tables C1 and F1 (means supplied by S. Webb). It should be noted that the means are expressed in the original source **per couple.** The means given in Table 3 are **per person.**

4.2 Different Definitions

We now turn to the comparison of France and the UK on the basis of our calculations. We consider the effect of putting them on a comparable basis, working from the French to the UK definitions, proceeding by stages. As may be seen from Table 5, adopting the same definitions as Assémat and Glaude but applying them to the UK yields an estimate of some 4.1 per cent of the population below 50 per cent of the median. The difference is comfortably larger than the one per cent confidence interval of plus or minus two per cent. On this basis, the low income population is much larger in France than in the UK. In broad terms, there are about 10 per cent of the French population in poverty, a figure which may be contrasted with the much smaller proportion of four per cent in the UK when measured in this way.

This conclusion does, however, depend on the choice of definition. The first difference in definition lies in the use of the **mean** in place of the **median**. As may be seen from Table 5, the median is 90 per cent of the mean in France and 85 per cent in the United Kingdom. Taking 50 per cent of the median is like taking a cut-off of 45 per cent or 42.5 per cent of the mean. The figures are therefore higher with 50 per cent of the mean, and they are higher to a different extent in the two countries, as may be seen from line B in Table 5. The proportion below 50 per cent in France rises from 9.6 per cent to 13.5 per cent, or by slightly more than the figure of one-third suggested in Atkinson and Cazes (1990). In contrast, the proportion in the UK more than doubles. As a result, poverty is now around half as much again in France, rather than twice as much.

The second difference in definition lies in the replacement of the **household** by the **individual** as the unit of analysis. It may be seen from a comparison of lines B and C in Table 5 that this makes only a difference of about one percentage point to the estimates for each country individually, but that they move the two countries in opposite directions. The poverty count is reduced in France and increased in the United Kingdom, so that it is now only around a quarter higher in France.

The third issue of definition concerns the choice of equivalence scale, which affects both the mean and the cumulative distribution.

In Table 5 we show the effect of moving from the OECD scale used in the French study (line C) to the HBAI scale used in the official UK study (line D). As already noted, the latter makes less allowance for the needs of younger children, and it may be observed that the mean equivalent income rises by 11 per cent in France and eight per cent in the UK in moving from line C to line D. The effect is to reduce the numbers below 50 per cent of the mean, with the reduction being larger in the UK.

Line D of Table 5 provides a comparison of France and the UK on definitions comparable to those in the official UK study. On this basis, with a cut-off of 50

Table 5: Estimated Size of Low Income Population on Different Definitions: France 1984/85 and the United Kingdom 1985

	France Proportion less than			United Kingdom Proportion less than		
	40%	50%	60%	40%	50%	60%
A: median households OECD scale before housing	5.3	9.6	16.8	1.7	4.1	9.9
median	48,937 FF per year			£70.44 a week		
B: **mean** households OECD scale before housing	7.0	13.5	22.5	3.1	9.2	20.9
mean	54,604 FF per year			£83.09 a week		
C: mean **individuals** OECD scale before housing	6.4	12.5	22.0	3.8	10.3	21.0
mean	51,356 FF per year			£80.01 a week		
D: mean individuals **HBAI** before housing	6.5	11.9	20.1	2.6	8.6	19.9
mean	57,188 FF per year			£86.58 a week		
E: mean individuals HBAI **after housing**	7.4	13.0	21.2	5.3	13.6	25.0
mean	44,739 FF per year			£67.13 a week		

per cent, the poverty rate in France appears to be 12 per cent in round figures compared with 8½ per cent in the UK. The position is, however, quite different with the alternative DSS definition, based on resources available after deducting housing costs, shown in line E in Table 5. Although the reduction in mean income arising from the deduction of housing costs is virtually the same in the two countries (21.8 per cent in France and 22.5 per cent in the UK), the proportion below 50 per cent of the mean rises much more sharply in the UK. On the basis of line E, there is little difference in the extent of the low income population in the two countries. The difference of 0.6 per cent is not significant at the five per cent level.

From these results it is evident that the choice of definition **does** affect the conclusions drawn. The impact of a change in definition is not the same in the two countries. While taking a cut-off of 50 per cent of the 'average', we can vary the conclusion from

> the low income population is more than twice as large in France than in the UK (line A)

or

> the low income population is about 50 per cent larger in France than in the UK (line B)

to

> the low income population is about the same size in France and the UK (line E).

5 What Lies Behind the Differences?

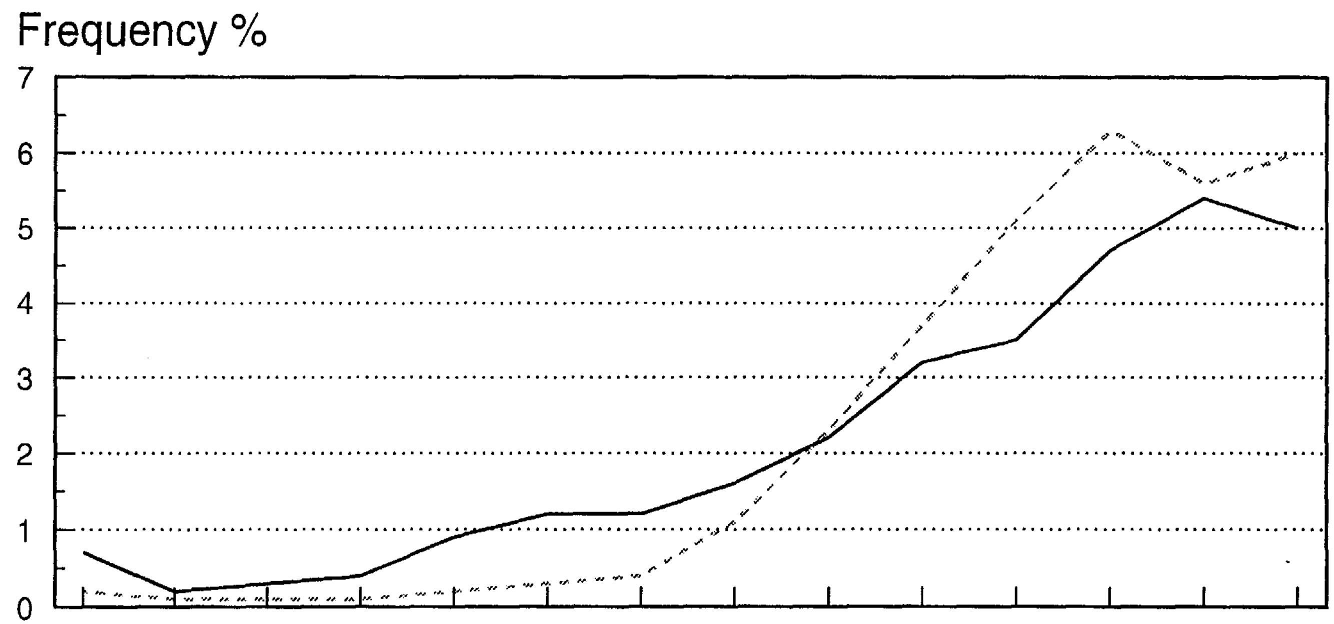
A full explanation of the differences between France and the UK would go far beyond the scope of this paper, but in this section we explore in more detail four aspects which seem to us important in understanding the statistics presented in the previous section: the shape of the lower part of the income distribution, the sensitivity to choice of equivalence scale, the role of household size, and the implications of housing benefit.

5.1 Shape of the Income Distribution

From the results in Table 5, it is evident that the comparison of France and the UK depends on the cut-off chosen. If we take the estimates using the HBAI scale (line D), then we see that the proportion in France below 40 per cent of the mean is more than twice that in the UK, whereas the proportion below 60 per cent of the mean is virtually identical. This suggests that the shape of the income distribution is different in the two countries.

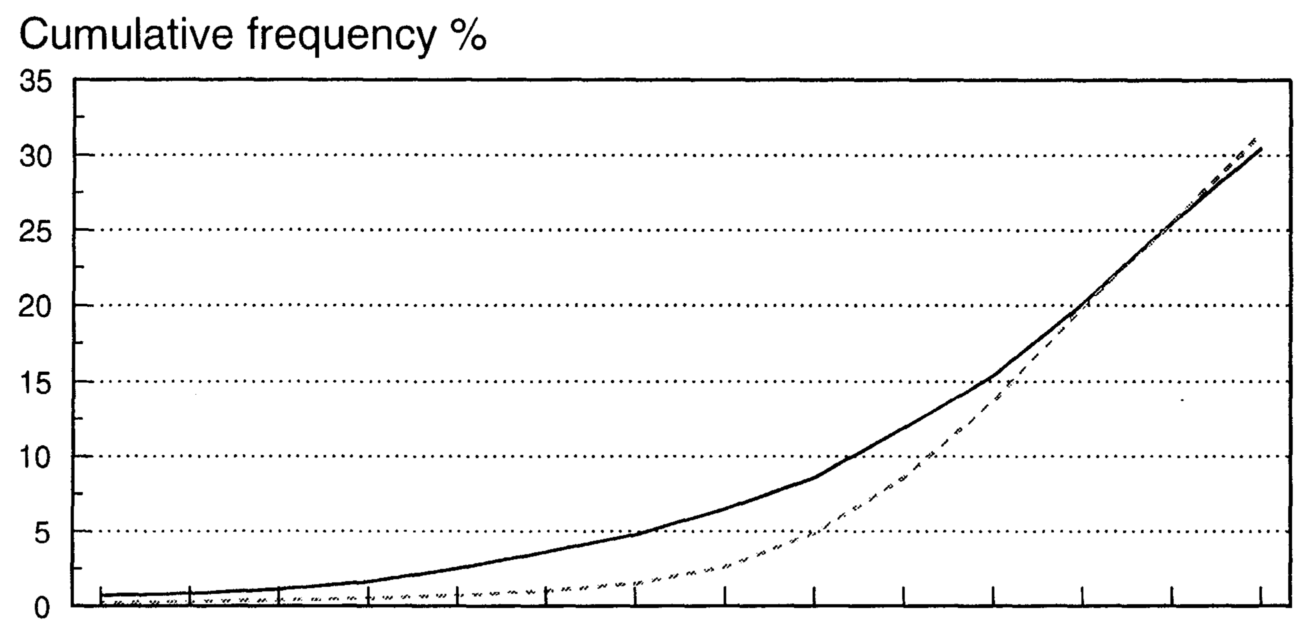
The shape of the distribution of income is illustrated in Figure 1a, which plots the frequency distribution (by ranges of five per cent of the mean) of individuals, and in Figure 1b which shows the cumulative distribution, for equivalent income, using the HBAI scale, before housing costs. (These figures correspond to line D in Table 5.) The solid line relates to France and it may be seen that the frequencies are in general higher until we reach 45 per cent of mean income. There appear to be more very low incomes in France.

Figure 1a: Frequency Distribution in France and the UK (Definition D)



Percent of mean	5	10	15	20	25	30	35	40	45	50	55	60	65	70
France	0.7	0.2	0.3	0.4	0.9	1.2	1.2	1.6	2.2	3.2	3.5	4.7	5.4	5.0
UK	0.2	0.1	0.1	0.1	0.2	0.3	0.4	1.1	2.3	3.7	5.1	6.3	5.6	6.0

Figure 1b: Cumulative Frequency Distribution in France and the UK (Definition D)



Percent of mean income	5	10	15	20	25	30	35	40	45	50	55	60	65	70
France	0.7	0.8	1.1	1.6	2.5	3.6	4.8	6.5	8.6	11.9	15.4	20.1	25.5	30.5
UK	0.2	0.3	0.4	0.5	0.7	1.0	1.5	2.6	4.9	8.6	13.7	19.9	25.5	31.5

In contrast, in the United Kingdom frequencies are more concentrated in the range 45-70 per cent of the mean, so that the cumulative distribution for the UK 'catches up' and intersects that for France at 65 per cent of the mean.

The difference in the shape of the distribution is apparent with other definitions, as is shown by Figures 2a and 2b which show the cumulative distributions corresponding to lines B (households, equivalent incomes adjusted using the OECD scale) and E (individuals, equivalent resources adjusted using the HBAI scale). There is the same pattern of the UK starting below and catching up, although the point of intersection comes slightly later in the case of definition B and much earlier (below 50 per cent of the mean) in the case of definition E.

The shape of the distribution affects the conclusions drawn with different poverty indicators. Suppose that we consider individual equivalent net resources (definition E). If we apply the 50 per cent cut-off, the headcount is virtually the same in France and the UK. However, the fact that more of those in poverty are to be found at low levels of income in France than in the UK causes the poverty gap to be larger in France. In Table 6 we show the poverty gaps expressed as a percentage of total income.

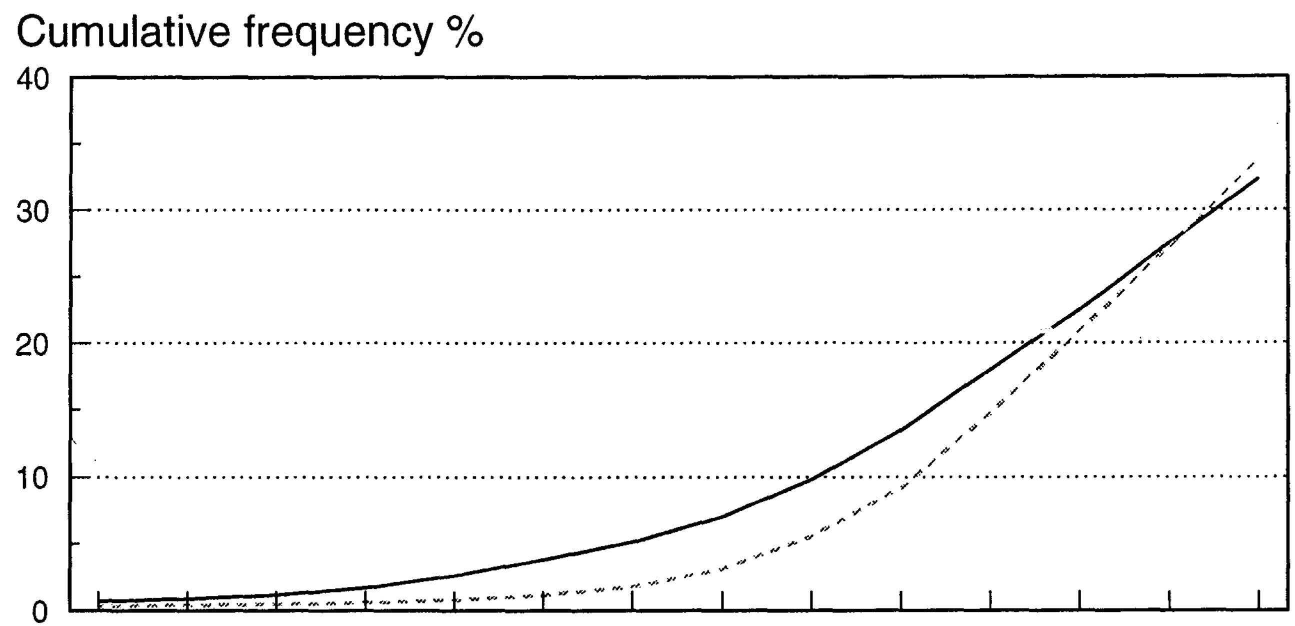
On definition E, the gap at 50 per cent of the mean is 2.28 per cent in France, compared with 1.47 per cent in the UK. This means that some 2¼ per cent of total income would be needed to bring all French households up to 50 per cent of the mean, compared with 1½ per cent in the UK. This is a sizeable difference. Put another way, the average gap per person below the 50 per cent cut-off is 3,920 FF per year[14], whereas the average gap in the UK is £3.98 a week, or £206.90 a year.

The severity of the poverty problem is, on this basis, greater in France, and the same is borne out for other definitions and cut-offs: see Table 6. Taking the cut-off as 50 per cent of the mean, the poverty gap in the UK is approximately half that in France for all measures which take income before deduction of housing costs.

By calculating the poverty gaps at different cut-offs, we are in effect constructing the poverty deficit curve over that range. This allows us to extend the results to other poverty indicators (Atkinson, 1987). Where the poverty deficit curve is higher at all cut-offs in the range, we can deduce that we will arrive at the same ranking with a poverty indicator which gives more weight to large than to small poverty deficits. For the definitions involving income before deduction of housing costs (lines B - D in Table 6), the French poverty deficit curve lies above that for the UK all the way up to mean income. For definition E, the French deficit curve lies above until an income cut-off of 80 per cent of the mean - see Figure 3. Since the poverty line is unlikely to be taken above this level, this suggests that the choice of poverty indicator is not critical to the present comparison.

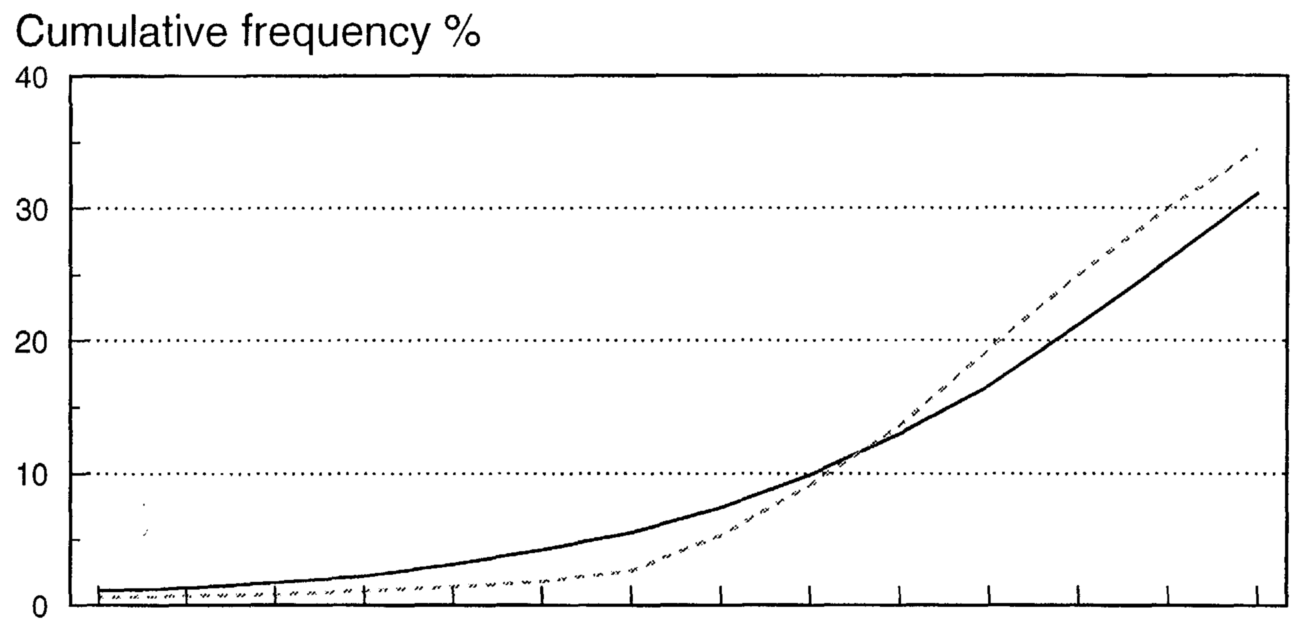
14 3,920 is 17.5 per cent of the 50 per cent cut-off (22,370), and 17.5 per cent times the percentage below 50 per cent (13.0 per cent) gives 2.28 per cent.

Figure 2a: Cumulative Frequency Distribution in France and the UK (Definition B)



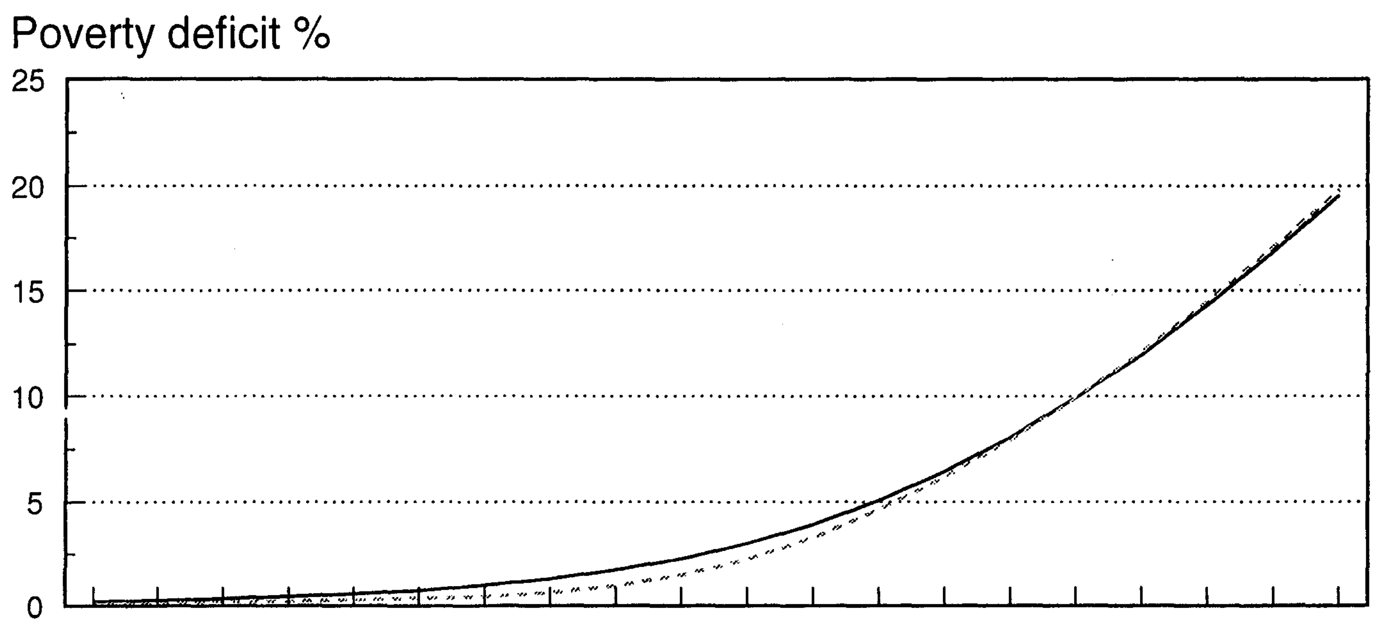
Percent of mean income	5	10	15	20	25	30	35	40	45	50	55	60	65	70
France ———	0.7	0.9	1.2	1.7	2.6	3.8	5.1	7.0	9.8	13.5	18.0	22.5	27.5	32.3
UK - - - -	0.3	0.4	0.5	0.6	0.8	1.2	1.8	3.1	5.5	9.2	14.7	20.9	27.2	33.8

Figure 2b: Cumulative Frequency Distribution in France and the UK (Definition E)



Percent of mean income	5	10	15	20	25	30	35	40	45	50	55	60	65	70
France ———	1.1	1.3	1.7	2.2	3.1	4.2	5.5	7.4	9.9	13.0	16.6	21.2	26.1	31.1
UK - - - -	0.6	0.7	0.8	1.1	1.4	1.8	2.6	5.3	9.1	13.6	19.4	25.0	30.0	34.4

Figure 3: Poverty Deficit Curves for France and the UK (Definition E)



Percent of mean income	5	10	15	20	25	30	35	40	45	50	55	60	65	70	75	80	85	90	95	100
France ———	0.22	0.28	0.35	0.45	0.58	0.75	0.99	1.30	1.72	2.28	3.00	3.91	5.06	6.44	8.06	9.92	12.00	14.31	16.83	19.57
UK - - - -	0.12	0.15	0.19	0.23	0.30	0.37	0.47	0.64	0.96	1.47	2.24	3.30	4.63	6.19	7.94	9.93	12.12	14.49	17.06	19.81

Table 6: Estimated Size of Poverty Gap (Expressed as per cent of Total Income) on Different Definitions: France 1984/5 and the UK 1985

	France			United Kingdom		
		Cut-off			Cut-off	
Definition	40%	50%	60%	40%	50%	60%
A: median households OECD scale before housing	0.80	1.53	2.90	0.21	0.50	1.22
B: **mean** households OECD scale before housing	1.08	2.23	4.33	0.39	1.12	2.81
C: mean **individuals** OECD scale before housing	0.91	1.80	3.46	0.33	0.93	2.33
D: mean individuals **HBAI scale** before housing	0.92	1.82	3.38	0.28	0.74	2.05
E: mean individuals HBAI scale **after housing**	1.30	2.28	3.91	0.64	1.47	3.31

Note: French results from XWRKGAP1 dated 14 August 1992; UK results from WORKGAP3 dated 25 June 1992.

5.2 Sensitivity to Equivalence Scale

We have seen that the change in equivalence scales from the OECD to HBAI scale affected not just the level of poverty but also the relative poverty rates in the two countries. This is not perhaps surprising in view of what we know about the differences in policy towards families of different sizes in the two countries. People tend to think in terms of French policy being more generous to families, whether in the form of income tax allowances (the *quotient familial*) or child benefits. In particular, the French child benefit system is more tilted in the direction of larger families.

In order to explore the sensitivity of the comparison to the equivalence scale, we take a scale of n^s, where s varies between 0 and 1. The results for the proportions below 40 per cent, 50 per cent, and 60 per cent of the mean are shown in Table 7 for definitions which otherwise correspond to lines B (household before housing costs), C and D (individual before housing costs), and E (individual after housing costs). The exponent s is a valuable method of summarising differences in scales, but it must be stressed that the parameterisation is only very approximate and does not capture the variation by age or other characteristics that one finds within a household of a specified size. Even, for example, the simple OECD scale for a household of five people ranges from 3.0 to 3.8, which would mean that the implied value of s could lie anywhere between 0.68 and 0.83.

Figure 4 shows how the poverty measures vary with s, taking the equivalence scale as given by n^s, for the cases corresponding to lines C and D in Table 5. (This is the middle block of figures in Table 7.) If we take the 50 per cent cut-off lines, which are both dashed, then we can see that in both countries they have a U-shape: the percentage below 50 per cent of the mean in France falls from 16 per cent with $s = 0$ to around 12 per cent and then rises again as s approaches 1. The trough in France appears to be around 0.55, that in the United Kingdom rather higher, around 0.65. The existence of such a U-shape has been examined by Coulter et al. (1992).

It is the difference between the countries which is of particular interest here. The two curves intersect at a value of s around 0.6. Poverty is higher in the United Kingdom if we take values below this: for instance, if one were to take a value of 0.25, as with the scales based on subjective evaluation, then the United Kingdom would have a poverty rate of 15.4 per cent compared with 13.0 per cent in France. On the other hand, with values of s higher than 0.6, poverty appears lower in the UK. On a per capita basis, the poverty rate in France is 16.5 per cent compared with 14.3 per cent in the United Kingdom.

Why do we get this intersection? As s changes, there are two effects which have to be taken into account when considering the position of an individual household. The first is that, for all households except those consisting of just a single person, there is a change in their calculated equivalent income. Suppose that one considers the extreme cases of $s = 0$ (no adjustment for household size) and $s = 1$ (a per capita calculation). Then, if the household total income is Y, then its equivalent income is Y on the former basis and Y/n on the latter basis, where n is the number of household members. The second effect is that the poverty line changes, since average equivalent income changes. In both countries the average equivalent income is reduced by broadly a factor of 3 in moving from $s = 0$ to $s = 1$.

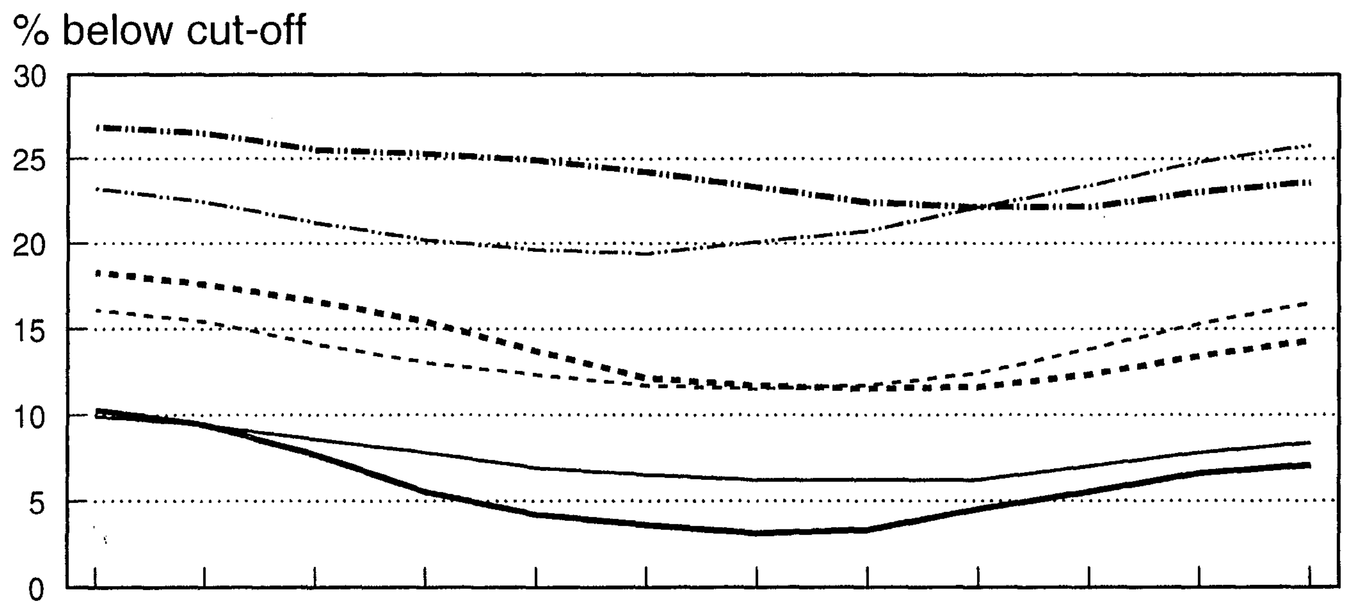
We can therefore deduce how households of different sizes are affected. For single person households, only the second effect operates. The rise by a factor of 3 in average equivalent income in moving from $s = 0$ to $s = 1$ means that for a single

Table 7: Sensitivity of Low Income Headcount to Different Equivalence Scales: France 1984/5 and the UK 1985(a)

Value of s	France Cut-off 40%	France Cut-off 50%	France Cut-off 60%	France Mean ,000FF	United Kingdom Cut-off 40%	United Kingdom Cut-off 50%	United Kingdom Cut-off 60%	United Kingdom Mean £week
Definition B:mean/households/before housing costs								
0.0	13.1	20.0	28.4	106.7	14.1	23.2	31.8	160.1
0.05	12.4	18.9	27.2	101.5	12.4	22.2	31.3	152.7
0.15	11.0	17.2	25.3	92.2	9.0	20.0	30.2	139.0
0.25	9.5	15.7	23.5	84.0	6.5	17.3	29.2	126.9
0.35	8.5	14.8	22.5	76.7	5.0	14.4	28.1	116.2
0.45	7.8	13.7	21.6	70.3	3.7	12.0	26.2	106.7
0.55	7.1	13.2	21.8	64.5	3.0	10.9	24.2	98.2
0.65	6.9	13.1	21.9	59.4	2.9	10.4	22.5	90.7
0.75	7.0	13.3	22.5	54.9	3.6	10.3	21.8	84.0
0.85	7.3	14.5	23.5	50.9	4.6	10.8	22.3	78.0
0.95	8.3	15.9	24.9	47.3	5.5	11.6	23.0	72.7
1.0	8.9	16.5	25.5	45.7	6.0	12.3	23.4	70.2
Definitions C and D: mean/individuals/before housing costs								
0.00	9.9	16.1	23.2	122.0	10.3	18.3	26.9	188.2
0.05	9.4	15.4	22.4	114.8	9.4	17.6	26.5	177.3
0.15	8.6	14.1	21.2	101.8	7.7	16.6	25.5	157.5
0.25	7.8	13.0	20.2	90.5	5.5	15.4	25.3	140.3
0.35	6.9	12.3	19.6	80.6	4.2	13.7	24.9	125.3
0.45	6.5	11.7	19.4	71.9	3.6	12.1	24.2	112.1
0.55	6.2	11.5	20.1	64.3	3.1	11.7	23.3	100.5
0.65	6.2	11.7	20.7	57.7	3.3	11.5	22.4	90.3
0.75	6.2	12.4	22.1	51.8	4.5	11.6	22.1	81.4
0.85	7.0	13.8	23.4	46.7	5.5	12.3	22.1	73.5
0.95	7.8	15.3	24.8	42.2	6.6	13.4	23.0	66.6
1.0	8.4	16.5	25.8	40.2	7.1	14.3	23.6	63.5
Definition E: mean/individuals/after housing costs								
0.00	11.1	17.4	24.5	107.7	15.0	22.6	29.4	162.5
0.05	10.6	16.7	23.9	101.3	14.1	22.2	28.9	153.0
0.15	9.8	15.4	22.9	89.6	12.7	21.4	28.2	135.9
0.25	8.9	14.4	22.4	79.5	11.7	20.5	28.3	120.9
0.35	8.0	13.9	21.9	70.7	10.9	20.3	27.8	107.7
0.45	7.4	13.4	21.6	63.0	9.6	19.9	27.9	96.3
0.55	7.2	13.1	21.6	56.2	8.3	19.0	28.0	86.2
0.65	7.3	13.2	22.2	50.3	8.5	17.4	28.3	77.4
0.75	7.6	13.8	23.1	45.1	8.3	15.8	28.3	69.6
0.85	8.1	15.0	24.1	40.5	8.9	16.2	27.7	62.8
0.95	8.8	16.6	25.7	36.5	9.8	16.8	27.1	56.8
1.0	9.2	17.8	26.5	34.7	10.5	17.2	27.2	54.0

Note: a) French results from XWRKEBFX dated 13 August 1992, pages 35-38, 57-60, and 79-82; UK results from FRCOMP6A dated 10 June 1992, pages 18-21, 40-43, and 84-87.

Figure 4: Sensitivity of Headcounts to Equivalence Scales, France and the UK (Definitions C and D)



s	0	0.05	0.15	0.25	0.35	0.45	0.55	0.65	0.75	0.85	0.95	1.00
UK40%	10.3	9.4	7.7	5.5	4.2	3.6	3.1	3.3	4.5	5.5	6.6	7.1
UK50%	18.3	17.6	16.6	15.4	13.7	12.1	11.7	11.5	11.6	12.3	13.4	14.3
UK60%	26.9	26.5	25.5	25.3	24.9	24.2	23.3	22.4	22.1	22.1	23.0	23.6
FR40%	9.9	9.4	8.6	7.8	6.9	6.5	6.2	6.2	6.2	7.0	7.8	8.4
FR50%	16.1	15.4	14.1	13.0	12.3	11.7	11.5	11.7	12.4	13.8	15.3	16.5
FR60%	23.2	22.4	21.2	20.2	19.6	19.4	20.1	20.7	22.1	23.4	24.8	25.8

person household the poverty line is reduced to a third, relative to their incomes. This reduces measured poverty in both countries, but the effect is larger in the United Kingdom, since the United Kingdom distribution is steeper - as we have seen in aggregate. In the case of two-person households, both effects operate, but the second is larger. In moving from $s = 0$ to $s = 1$, their income is halved but the poverty line falls to a third. It is only for households of 4 persons or more that the first effect is larger. The reason why the curves intersect for the two countries appears therefore to be that the United Kingdom estimates are more sensitive to variations in the poverty line, and that the effects for small households outweigh those for larger households.

Consistent with this explanation is the fact that the crossing-over point occurs at a lower value of s when the unit of analysis is the household (definition B), since larger households then get no more weight than smaller households. Figure 5 shows the diagram corresponding to definition B (the upper part of Table 7). The intersection of the curves for the proportion below 50 per cent of the mean (the dashed curves) now takes place around 0.3.

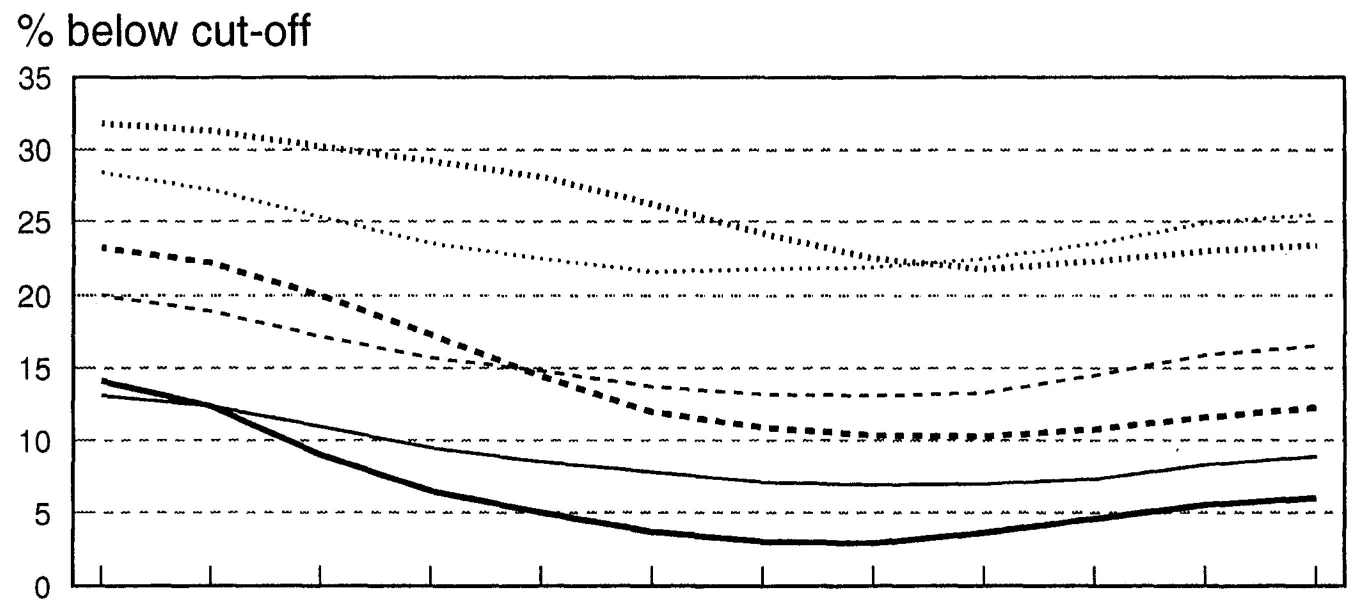
The comparisons made above assume that the same equivalence scale should be applied in both countries, but the appropriate equivalence scale may vary from country to country. Where for instance the fixed costs of a household are relatively low, then standard of living considerations may point to a scale which is close to per capita. In another country, where fixed costs, such as those for housing, heating and property taxes, are relatively larger, then the costs of additional household members may be less.

5.3 Role of Household Size

The analysis of different equivalence scales has highlighted the role of household size. Differences in household size between the two countries may be one of the factors driving the results. In fact the distribution by household size is relatively similar, but there are slightly more small households in the UK. Single person households account for 26.6 per cent, compared with 24.1 per cent in France, and there are in all 58.7 per cent with only one or two persons, compared with 54.0 per cent in France.

Table 8 shows the estimated size of the low income population broken down by household size, on the basis of definitions B, C and D, together with the distribution of all households (definition B) or all individuals (definitions C and D) by household size. If we begin with definition B, we can see that, taking the 50 per cent cut-off, the differences in poverty rates between countries are particularly marked for one and two person households. The poverty rate for the UK for these groups is about half that in France. From this figure, we can calculate the effect of the difference in distribution by household size. If France had the same distribution across household size as the UK (i.e. assuming that single person

Figure 5: Sensitivity of Headcounts to Equivalence Scales, France and the UK (Definition B)



s	0	0.05	0.15	0.25	0.35	0.45	0.55	0.65	0.75	0.85	0.95	1.00
UK40%	14.1	12.4	9.0	6.5	5.0	3.7	3.0	2.9	3.6	4.6	5.5	6.0
UK50%	23.2	22.2	20.0	17.3	14.4	12.0	10.9	10.4	10.3	10.8	11.6	12.3
UK60%	31.8	31.3	30.2	29.2	28.1	26.2	24.2	22.5	21.8	22.3	23.0	23.4
FR40%	13.1	12.4	11.0	9.5	8.5	7.8	7.1	6.9	7.0	7.3	8.3	8.9
FR50%	20.0	18.9	17.2	15.7	14.8	13.7	13.2	13.1	13.3	14.5	15.9	16.5
FR60%	28.4	27.2	25.3	23.5	22.5	21.6	21.8	21.9	22.5	23.5	24.9	25.5

Table 8: Estimated Size of Low Income Population by Size of Household: France 1984/5 and the UK 1985[(a)]

Household size	France				United Kingdom			
	Proportion in total population	% below cut-off			Proportion in total population	% below cut-off		
		40	50	60		40	50	60
Definition B: mean/households/OECD scale/before housing								
1	24.1	7.0	12.6	20.9	26.6	1.9	6.5	17.3
2	29.9	6.9	12.2	19.8	32.1	1.9	6.2	20.0
3	18.5	5.5	10.6	17.4	16.3	2.6	8.1	18.8
4	17.1	5.4	11.4	21.0	16.6	3.4	11.9	21.7
5	6.9	9.4	20.1	36.1	5.8	8.6	21.4	33.1
6+	3.6	19.9	42.6	64.5	2.7	19.0	36.2	49.9
Total	100	7.0	13.5	22.5	100	3.1	9.2	20.9
Definition C: mean/individuals/OECD scale/before housing								
1	9.1	5.9	10.6	18.1	10.5	1.9	5.1	14.9
2	22.5	5.5	10.5	16.5	25.4	1.7	4.9	16.3
3	20.8	4.7	8.8	14.6	19.4	2.2	6.8	16.1
4	25.7	4.7	9.0	17.2	26.3	2.7	10.0	19.7
5	12.9	8.1	15.2	29.2	11.4	6.5	19.6	31.0
6+	8.9	15.9	34.6	60.5	7.0	17.9	34.2	49.6
Total	100	6.4	12.5	22.0	100	3.8	10.3	21.0
Definition D: mean/individuals/HBAI scale/before housing								
1	9.1	7.7	15.0	23.2	10.5	2.1	8.6	21.1
2	22.5	7.1	12.7	20.7	25.4	1.9	6.7	20.7
3	20.8	4.8	9.0	14.9	19.4	1.8	5.3	14.4
4	25.7	4.2	7.8	14.1	26.3	2.2	5.6	15.8
5	12.9	7.2	11.5	20.1	11.4	2.9	13.7	23.4
6+	8.9	13.0	25.6	44.5	7.0	8.5	27.4	40.9
Total	100	6.5	11.9	20.1	100	2.6	8.6	19.9

Note: a) French results from XWRKEBFX dated 13 August 1992 pages 12, 15 and 16, UK results from FRCOMP6A dated 10 June 1992, pages 6, 10 and 11.

households with a poverty rate of 12.6 per cent made up 26.6 per cent of all households, as opposed to 24.1 per cent, etc.), this would reduce the overall poverty rate from 13.4 per cent to 13.2 per cent. In other words, the difference in distribution by household size contributes relatively little to explaining the overall difference.

From the first part of Table 8 there does appear to be an interesting difference between France and the UK. Whereas for France the poverty rates are similar for households with one, two, three or four members, and are only substantially higher for households with five or more members, in the UK the poverty rate rises more steadily with household size for the 40 per cent and 50 per cent cut-offs. In the UK the proportion of four person households below 50 per cent of the mean is nearly double that for one or two person households.

The second and third sets of figures in Table 6 shows the variation by household size on an **individual** basis. It is evident that this gives much more weight to larger households. In France, households with six or more members account for only 3.6 per cent of all **households,** but the individuals who live in these households constitute 8.9 per cent of all **individuals** in the population. Around one in 10 households in France has five or more members, but around one person in five lives in such a household. The move from definition B to definition C also involves a change in the poverty line, since the mean income is lower with the latter definition. It is this which accounts for the lower poverty rates for households of a specified size. The difference between France and the UK noted in the previous paragraph continues, however, to be observed. The poverty rate for a four person household is double that of a single person household in the UK, but in France it is actually lower.

The final part of Table 8 shows the effect of a change in the equivalence scale. It is evident that this affects the conclusions drawn about the variation in poverty rates by household size. The move to definition D, with the HBAI scale leads to a reduction in the poverty rates for larger households, and an increase for one and two person households (as the mean equivalent income, and hence the poverty cut-off, is higher). It remains nonetheless the case that four person households in France do relatively better than in the UK, the French poverty rate now being some half of that for single persons.

5.4 Role of Housing Benefit

The treatment of housing is particularly important, in view of differences across countries in relative prices and in the arrangements for finance. We have already seen the effect of subtracting housing costs, but now consider the treatment of income-related subsidies to housing, known in the United Kingdom as housing benefit. These are related to housing costs, and indeed may be paid direct to the landlord where the accommodation is rented.

This raises the question of the appropriate treatment of housing benefit. Is it an income supplement, or is it a subsidy to spending? The latter would be parallel for example to transport subsidies, which simply show up in lower prices. In our estimates it has been treated as income, but this makes quite a difference to the figures, as may be seen from Table 9, which compares the estimates, on

Table 9: Effect of Excluding Housing Benefit from Income on Estimated Size of Low Income Population on Different Definitions: France 1984/5 and the United Kingdom 1985
(Figures in brackets include housing benefit)

Definition	France			United Kingdom		
	Proportion less than			Proportion less than		
	40%	50%	60%	40%	50%	60%
B: mean households OECD scale before housing	7.7	14.7	24.0	7.0	16.9	29.3
	(7.0)	(13.5)	(22.5)	(3.1)	(9.2)	(20.9)
mean	53,912 FF per year (54,604 FF per year)			£79.92 a week (£83.09 a week)		
D: mean **individuals HBAI scale** before housing	6.9	12.8	22.0	6.8	16.0	24.6
	(6.5)	(11.9)	(20.1)	(2.6)	(8.6)	(19.9)
mean	56,315 FF per year (57,188 FF per year)			£84.33 a week (£86.58 a week)		

Note: French results excluding housing benefit from output XWRKNHBX dated 14 August 1992; UK results from output FRCOMP6A dated 10 June 1992.

definitions B and D, with and without housing benefit. The reduction in mean income, on definition B, is 1.3 per cent in France and more than twice (3.2 per cent) in the UK. Housing benefit is not only more important in the UK figures, but it is more concentrated on those with low incomes. Subtracting housing benefit causes the estimated proportion below 50 per cent of the mean in the UK to rise from around nine per cent to some 16-17 per cent. The increase in France is only around one percentage point, and the effect of this adjustment is to reverse the ranking of the two countries. The exclusion of housing benefit from income makes little difference to the French estimates of the headcount, but increases those for the United Kingdom very significantly. The cumulative distribution now intersects between 40 and 45 per cent of the mean, and the difference at 50 per cent is large enough to satisfy the two per cent criterion for significance at the one per cent level. At the same time, the poverty gap remains larger in the United Kingdom until about 65 per cent of the mean. This is a further illustration of the potential sensitivity of the conclusions.

6 Conclusions

In comparing poverty in France and the UK in this paper, we have assumed that the object is to measure the extent of low income defined as having a household equivalent disposable income of less than a specified percentage of the average. This is a highly restrictive definition. Yet even this restrictive definition allows considerable room to manoeuvre, and the choice of definition can materially affect the conclusions drawn. While taking a cut-off of 50 per cent of the 'average', we can vary the conclusion from

> the low income population is more than twice as large in France than in the UK,

or

> the low income population is about 50 per cent larger in France than in the UK,

to

> the low income population is about the same size in France and the UK,

or

> the low income population is some 3 percentage points larger in the UK.

There are many qualifications which have to be entered concerning the results reported here. Indeed, one of the principal lessons from the research is the need for caution when making comparisons across countries. To the qualifications which emerge in the course of the paper should be added the fact that we focus on one particular period - the mid-1980s - and comparisons at a different date may lead to different conclusions. The official estimates for the UK show that the proportion of the population with less than 50 per cent of average income more than doubled between 1979 and 1988/89. A comparison with France at the end of the 1980s may therefore give a different picture if there has not been the same growth in the proportion with low incomes.

Appendix: Effect of Excluding Households with Self-Employed Heads

The income data collected from the self-employed in sample surveys is typically of lower quality than that collected from employees or pensioners. As has been described in the text, the French EBF questions are less detailed than those in the UK FES. In view of this, we show in this appendix the effects of eliminating those households where the head is self-employed. It should be noted that this does not exclude all households with self-employment income. It should also be noted that the definition of head of household differs between the two countries.

Table A is in the same format as Table 5, but relates to the sample excluding households where the head is self-employed (the full sample figures are shown in brackets for comparison). In the UK, the omission of self-employed households reduces the sample size by 7.8 per cent (households) and 9.7 per cent (individuals). The numbers involved in France are rather smaller: 6.4 per cent (households) and 7.8 per cent (individuals). This may reflect the fact that we have omitted from the French data those households providing income data in terms of ranges, or where there are incomplete items. A second interesting difference is that the omission of the self-employed **reduces** the mean income in the UK but *raises* the mean income in France. This suggests that the self-employed are rather different in the two countries.

Even though the mean incomes move differently in the two countries with the exclusion of households with self-employed heads, in both the proportions with low incomes are reduced. The proportions with less than 50 per cent of the mean are in general reduced by about one percentage point. This suggests that the self-employed were over-represented among those with low incomes, but that their overall contribution is not sufficient to change the basic conclusions concerning the comparison of France and the UK in 1985.

Table A: Effect of Excluding Households with Self-Employed Heads on Estimated Size of Low Income Population on Different Definitions: France 1984/5 and the United Kingdom 1985 (Figures in brackets for full sample)

Definition	France Proportion less than 40%	50%	60%	United Kingdom Proportion less than 40%	50%	60%
A: median households OECD scale before housing	4.7 (5.3)	8.8 (9.6)	16.1 (16.8)	1.3 (1.7)	3.6 (4.1)	9.2 (9.9)
median	49,400 FF per year (48,937 FF per year)			£69.82 a week (£70.44 a week)		
B: **mean** households OECD scale before housing	6.2 (7.0)	12.5 (13.5)	21.8 22.5)	2.6 (3.1)	8.2 (9.2)	19.7 (20.9)
mean	54.850 FF per year (54,604 FF per year)			£82.01 a week (£83.09 a week)		
C: mean **individuals** OECD scale before housing	5.6 (6.4	11.6 12.5	21.0 22.0)	3.0 (3.8	9.2 10.3	19.9 21.0)
mean	51,820 FF per year (51,356 FF per year)			£79.13 a week (£80.01 a week)		
D: mean individuals **HBAI scale** before housing	5.6 (6.5)	10.8 (11.9)	18.9 (20.1)	1.9 (2.6)	7.6 (8.6)	18.6 (19.9)
mean	57,645 FF per year (57,188 FF per year)			£85.42 a week (£86.58 a week)		
E: mean individuals HBAI scale **after housing**	6.1 (7.4)	11.8 (13.0)	20.2 (21.2)	4.3 (5.3)	12.4 (13.6)	24.0 (25.0)
mean	45,359 FF per year (44,739 FF per year)			£66.16 a week (£67.13 a week)		

Note: French results excluding self-employed HoHs from output XNSEEBFX dated 13 August 1992; UK results from output FRCOMP6B dated 29 May 1992.

References

Assémat, J. (1989), 'La qualité de la mesure des revenus au milieu des années quatre-vingt', in *Mélanges en l'honneur de Jacques Desabie*, INSEE, Paris.

Assémat, J. and M. Glaude (1989), 'Source fiscale et/ou enquête par interview: l'experience française en matière de mesure des bas revenus', paper presented at Eurostat Conference, Noordwijk.

Atkinson, A. B.(1987), 'On the measurement of poverty', *Econometrica*, 55(4), 749-64.

Atkinson, A. B. (1991), 'Comparing poverty rates internationally: lessons from recent studies in developed countries', *World Bank Economic Review*, 5, 3-21.

Atkinson, A. B. and S. Cazes (1990), 'Mesures de la pauvreté et politiques sociales', *Observations et Diagnostics Economiques*, Revue de l'OFCE, 33, 105-30.

Atkinson, A. B., K. Gardiner, V. Lechene and H. Sutherland (1993), *Comparing Poverty in France and the United Kingdom*, Welfare State Programme Discussion Paper No WSP/84, London School of Economics.

Atkinson, A. B., J. Gomulka and H. Sutherland (1988), 'Grossing-up FES data for tax-benefit models', in A. B. Atkinson and H. Sutherland, eds, *Tax-Benefit Models*, STICERD Occasional Paper 10, London.

Atkinson, A. B. and A. J. Harrison (1978), *Distribution of Personal Wealth in Britain*, Cambridge University Press, Cambridge.

Buhmann, B., L. Rainwater, G. Schmaus and T. M. Smeeding (1988), 'Equivalence scales, well-being, inequality, and poverty', *Review of Income and Wealth*, 34(2), 115-42.

Canceill, G. and A. Villeneuve (1990), 'Les inégalités de revenu: quasi statu quo entre 1979 et 1984 pour les salariés et les inactifs', *Economie et Statistique*, 230, 65-74.

Central Statistical Office (1991), 'The effects of taxes and benefits on household income 1988', *Economic Trends*, March, 107-49.

Coulter, F., F. A. Cowell and S. P. Jenkins (1992), 'Equivalence scale relativities and the extent of inequality and poverty', *Economic Journal*, 102, 1067-82.

Danziger, S. and M. K. Taussig (1979), 'The income unit and the anatomy of income distribution', *Review of Income and Wealth*, 25(4), December, 365-75.

Department of Health and Social Security (1988), *Low Income Statistics: Report of a Technical Review*, Government Statistical Service, London.

Department of Health and Social Security (1988a), *Low Income Families, 1985*, Government Statistical Service, London.

Department of Social Security (1990), *Households Below Average Income 1981-87*, Government Statistical Service, London.

Department of Social Security (1991), *Households Below Average Income: Stocktaking*, Department of Social Security, London.

Department of Social Security (1992), *Households Below Average Income: A Statistical Analysis 1979 - 1988/89*, HMSO, London.

Eurostat (1990), 'Inequality and poverty in Europe (1980-1985)', *Rapid Reports*, Population and Social Conditions, No 7.

Hagenaars, A. (1991), 'The choice of a measure of central tendency', note prepared for Eurostat, EURDP91.03.

Hoel, P. G. (1962), *Introduction to Mathematical Statistics*, third edition, Wiley, New York.

Johnson, P. and S. Webb (1990), *Poverty in Official Statistics: Two Reports*, IFS Commentary No. 24, London.

Kendall, M. G. and A. Stuart (1969), *The Advanced Theory of Statistics*, Volume I, Griffin, London.

Maritz, J. S. (1981), *Distribution-Free Statistical Methods*, Chapman and Hall, London.

Mitchell, D. (1991), *Income Transfers in Ten Welfare States*, Avebury, Aldershot.

Moutardier, M. (1988), *Les Budgets des Menages en 1984-1985*, INSEE, Paris.

O'Higgins, M. and S. Jenkins (1989), 'Poverty in Europe: Estimates for 1975, 1980 and 1985', paper presented at Eurostat Conference, Noordwijk.

Organisation for Economic Co-operation and Development (1982), *Social Indicators*, OECD, Paris.

Redpath, R. U. (1986), 'A second study of differential response comparing Census characteristics of FES respondents and non-respondents', *Statistical News*, 72.

Sen, A .K. (1976), 'Poverty: an ordinal approach to measurement', *Econometrica*, 44, 219-31.

Smeeding, T. M., M. O'Higgins and L. Rainwater (1990), *Poverty, Inequality and Income Distribution in Comparative Perspective*, Harvester Wheatsheaf, Hemel Hempstead.

Suesser, J. R. (1988), 'Quelques éléments de comparaison internationale sur les revenus: la position de la France au debut de la décennie', LIS Working Paper 30, Luxembourg.

Watts, H. W. (1968), 'An economic definition of poverty', in D. P. Moynihan, *On Understanding Poverty*, Basic Books, New York, 316-29.

Comments on paper by A. B. Atkinson, Karen Gardiner, Valérie Lechene and Holly Sutherland

Deborah Mitchell
Research School of Social Sciences
Australian National University

The motivation for this study arises from what appears to be highly contradictory findings on poverty rates for the UK and France published over the past few years. The authors set themselves two major tasks: first to examine the differences which lie behind the published studies; and second to conduct a thorough investigation of poverty rates in the UK and France which exposes the range of methodological choices to be made in such research and the consequent differences for measured poverty rates.

In the first instance we find that there are quite striking differences in the previously published studies that go a long way toward explaining the variations in findings e.g.: Eurostat used an **expenditure** based set of measures while the other studies used **income** based measures. We should therefore not be surprised by the differences reported. More subtly, however, are questions concerning the use of a median rather than a mean income standard to set a relative poverty line To get a handle on such differences the authors have carefully constructed and made comparable household expenditure surveys carried out in each country around the mid-1980s. The discussion in Section 2 of the paper, of the issues which arose in the course of rendering their surveys comparable is extremely useful and should act as a guide for those attempting similar tasks.

In Section 3 we get to the core of the methodological choices which face researchers in the field. The authors begin by restricting their field of methodological choices to a relative income measure rather than using an expenditure based measure or some other external standard. Having adopted this baseline, they then turn their attention to a series of further choices facing researchers: choice of mean or median; equivalence scale; weighting of units; before or after housing costs; and headcount, poverty gap or other indicator.

The implications of these are discussed in turn.

From this discussion, Table 5 sets out five methods to give a spread of possible measures (calculated from their data) variously combining choices of mean or median; different equivalence scales; different weighting methods; before or after housing costs; and providing both headcounts and poverty gap measures for each alternative.

At first sight the spread of the findings reported in Table 5 may seem somewhat alarming. The authors note that using a 50 per cent cut-off of each of their averages, the results vary from the low income population being: twice as large in France (A); 50 per cent larger in France (B); or about the same size as in the UK (E).

My reaction to these findings is that they are not as dramatic as the authors suggest. To begin with, I would argue that the calculations in line E (income after housing costs) are not strictly comparable with the other measures presented in the table. While the concept may be useful **within** a particular national setting, its comparability **between** countries is of limited value. As the authors point out themselves this measure may say more about a housing policy than it does about the income distribution or the effects of income support programs. Moreover, the authors also note that the French calculations are not fully comparable to the English calculations. In the remainder of my comments I therefore leave this measure to one side and consider only measures A to D.

The considerable variation of the within country estimates between the median and the mean require the greatest attention in this research field. The authors have spelt out some of the issues here regarding the pros and cons of choosing one or the other of these measures e.g.: the mean is much more affected by the occurrence of a few high incomes while the median is sensitive to the shape of the distribution. It would seem however that the median has much more going for it than the mean on statistical grounds yet the authors have used the mean for four out of five of their measures. They justify this on the grounds that it is 'more easily related to aggregate statistics'.

Once we move beyond the median vs mean standard, we are left with measures B, C and D. Here the gap narrows considerably between the measures and we find a fairly consistent picture arising both within and between the two countries. At each level of the poverty line (40, 50 and 60 per cent) and for each measure B, C and D we find that the headcount and the poverty gap are lower in the UK than in France. This is also true for the headcount and gap based on measure A.

Accordingly, if we return to the questions raised in the opening paragraph of the paper, the authors findings suggest that we **can** claim that there are more people on lower incomes 'relative to the average' in France than in the UK and that the aggregate poverty gap is also larger in France than in the UK. What we **cannot** say with any great certainly is exactly what the percentage differences are or the exact size of the poverty gap. Even if it were possible to make such statements, it seems unlikely that we would need to know these exact differences; the overall trend is all we need to see for most purposes. I note that if a specific question was being asked, for example in relation to a policy change which might affect both countries, then the nature of the policy issue itself would suggest the type of measure to be used e.g.: expenditure vs income, or even median versus mean. This is a point which cannot be stressed too often in this field: what may constitute a low income for purposes of access to health services might be very different to that adopted for access to child care services and different again from that used to target income support.

Once we leave such concerns aside we can being to search for far more interesting issues which arise from this research. For example, Figures 1a and 1b tell an interesting story about different levels of basic income support in these two countries. It would appear for example, that UK income support programs offer better **coverage** of the low income groups.

To summarise: this paper clearly sets out the methodological issues to be dealt with in cross country comparisons. Unlike the authors, I feel less concerned about apparently large differences in the headcount poverty rates and look mainly to the overall trends and the reassuring back up of the poverty gap data best summarised in the poverty deficit curve discussion arising out of Figure 3.

The paper leaves me with two questions to be puzzled over. First, the stark contrast between the findings reported here based on income measures and those of Eurostat based on expenditure measures. Why is there such a divergence between an income and expenditure based standard? What is the poverty concept behind each measure? Second, the paper once again leaves us to consider the pros and cons of the median vs the mean as a relative standard.

Contributions from Gender and Unions to Earnings Differences Among Young Australians: The Analysis of a Panel

Gábor Kõrösi
La Trobe University
Russell J Rimmer
Deakin University
Sheila M Rimmer
La Trobe University[1]

> You mean the youngsters are - unfortunate? No, they're only, like all the modern young, ... terrible little baffling mysteries. (Henry James, *The Awkward Age*)

1 Introduction

Many may believe that the young generation of the eighties was much luckier than its parents. Female labour force participation and educational participation for all had expanded quickly. At least officially, much discrimination was ruled out by 1984. The aim in this research is to test such beliefs and identify some sources of good fortune for the young.

For some years now one wave of the Australian Longitudinal Survey (ALS) has been used to study various influences on young people's educational experiences, their transitions from schooling to work, and their lives as workers and family members. Two issues have recently become the focus of intensive study using other cross-sectional data as well as the ALS. Mostly these issues, gender and union membership, have been studied separately. Here the approach is to exploit all four years of the ALS to investigate gender and union influences on earnings, and on inequality among the young. Panel methods as well as cross-sectional analysis yield several insights on the interactions between, gender, union

1 The authors are pleased to acknowledge the interesting comments made by Elizabeth Savage and others at the Conference on Contemporary Issues in Income Distribution, University of New South Wales, 3 December 1993.

memberships, and other variables on pay gaps. The different developments in inequality through time are explored for unionists and others.

1.1 Some Recent Australian Research

Several papers based on Australian data have analysed the effects on earnings of gender and of trade union membership. In two papers (Rummery, 1992, and Christie, 1992) the analysis was for people aged at least 18, based on the 1984 Australian National Social Science Survey (ANSSS). Other papers (Miller and Mulvey, 1992, and Miller and Rummery, 1989) reported work on data for males in the 1985 wave of the Australian Longitudinal Survey (ALS), while Kõrösi et al. (1993) also included the 1988 wave of the ALS. Variations of a human capital cross-section regression model form the basis of the empirical work in each of these papers. In a recent survey paper Miller and Mulvey (1993) review Australian empirical work on why people join unions, trade union strike activity, and pay differentials attributable to unionisation.

Two of the papers about unionisation were restricted to males. The paper by Miller and Rummery, included a model to predict trade union membership. They also estimated earnings equations, corrected for selectivity bias arising from non-random effects of young people's attachment, or otherwise, to unions. For their sample data, they concluded that the matter of non-random samples was not too important (Miller and Rummery, 1989: 197). They established from their regressions that a union premium of 13.12 per cent was evident for the young men (1989: 207). Using the same 1985 data base, the second of these papers investigated the incidence of fringe benefits for young men aged 19 to 25 years. They found that 'In respect of fringe benefits generally, Australian union members enjoy a significant advantage over their non-union counterparts' (Miller and Rummery, 1989: 139). Allowance for fringe benefits raised the union premium by 1.4 percentage points from 12.5 per cent. Both of the papers concerned with young Australian men confirm the importance of unions in obtaining financial benefits for their members.

The gender earnings gap was the major concern for Rummery (1992). She found that '... the earnings of females in the Australian labour market are 15.12 per cent lower than that of males, other things being equal' (Rummery, 1992: 359). This estimate was obtained using the conventional measures of experience defined as age minus school-leaving age. She adjusted the conventional measure by estimating actual years worked to allow for women's interrupted working life and their tendency to part-time employment on return to the labour market. Her adjustment brought down the estimate of pay discrimination for women to 10.33 per cent (Rummery, 1992: 360-1). When experience is measured similarly for men and women, it gives the appearance of a larger discrimination factor than when women's workforce interruptions are allowed for. Rummery (1992: 360-1) also found that the largest factor contributing to discrimination arises from differences in the returns to education for men and women.

Like Miller and Rummery, the issues of the wage premium and why men and women join unions were of interest to Christie (1992). She found that it is the characteristics of jobs, rather than of individuals, which are of importance as indicators of the trade union membership decision. She found the correction for selectivity bias helpful for the union wage equation. After the correction, it was found that the standardised wage differential in favour of unionists is 17.22 per cent (Christie, 1992: 52). However she suggested that not correcting OLS earnings equations for selectivity '... has probably not seriously distorted the overall picture of union wage effects' (Christie, 1992: 52). So far as gender is concerned, Christie found little difference in the coefficients for unionists and the others. This was so whether or not the selectivity correction is made.

Of the papers reviewed so far, only Christie considered both gender and union effects. But this was done too by Kõrösi et al. (1993). They estimated earnings equations with ALS data for 1985 and 1988. Thus they were able to comment on differences in the outcomes for the same young men and women across time. In their regressions (not corrected for selectivity) they found a gender wage gap of 5.9 per cent in 1985, and it was significantly greater, being 8.4 per cent in 1988. Trade unionists enjoyed a premium of 9 per cent in 1985. This fell to 6.5 per cent three years later, but the difference was not significant. In terms of weekly earnings the trends were the same, but as proportions of the relevant means, the gender and union effects were generally larger.

An important difference in the sample data used by Christie (1992) and Kõrösi et al. (1993) is that for the former the data related to the whole workforce aged 18 and over, while for the latter it was for people aged 16 to 25 years in 1985 (and the same group aged by three years in 1988). The age earnings profiles in the latter study showed a growing gender gap in raw earnings over the survey years. The widening gap was confirmed in the regressions after some human capital, industry, and occupation variables were allowed for. This gap appeared before work experience interruptions were widespread among the young women in the sample. It points rather more towards discriminatory pay practices, than to the decay of human capital due to truncated experience, as the source of the gap. This is consistent with Rummery (1992: 362) who remarked that the insignificance of her 'home time' variable '... suggests that human capital is not subject to depreciation during time out of the labour market'. That is, she suggested that discrimination, not interrupted experience, dominates the gender wage gap.

The results discussed above are summarised in Table 1.

1.2 New Directions

The issue of inequality among young Australians was the major concern for Kõrösi et al. They looked for a relationship between gender, unions, and inequality. The inequality within the cohort aged 23 to 25 in 1985 was examined. Generally they found that inequality within the groups was similar for men and

Table 1 : Some Recent Australian Research on Gender and Trade Union Pay Differences

Author	Data	Period	Sample Size N	Regression	Mean Y $	Feature	Percentage Group Female	Percentage Group Union	Other independent variables[a]
Miller and	ALS 19-	1985	1904	Pooled	7.30			13.6	Education, experience, region, occupation, industry, F/T, duration of job.
Rummery	25 year		904	Union	8.02				
(1989)	old males		1000	Non-union	6.71				
Christie	ANSSS	1984	1316	Pooled	8.41		-17.3	15.4	Education, experience, marital status, location, occupation.
(1992)	18 and		487	Union	9.45		-20.0		
	over		829	Non-union	7.77		-17.1		
Miller and	ALS 19-	1985	1833	Pooled	8.12	Adjusted for		13.9	Education, region, occupation, industry, experience, F/T, job duration.
Mulvey	25 year		884	Union	8.97	fringe benefits			
(1992)	old males		949	Non-union	7.40	and overtime pay			
Rummery	ANSSS	1984	1386	Pooled	8.41	Conventional	-15.1		Education, experience, region, marital status.
(1992)	18 and		531	Female	7.69	experience			
	over		855	Male	8.85	measure			
			1386	Pooled	8.41	Experience adjusted for career breaks	-10.3		As above but experience modified for period actually worked.
Kőrösi,	ALS	1985	4638	Pooled	6.75	Corrected for	-5.9	9.0	Age, experience, occupation, industry, F/T, hours, student, multiple jobs, government programs, income unit, marital status, household size.
Parkinson,	cohorts					heteroscedasticity			
Rimmer and	16-25								
Rimmer	years	1988	3899	Pooled	7.53		-8.4	6.6	
(1993)[b]									

Notes:
a) Although broad categories of variables are indicated here, it should be noted that the measures used differ throughout the literature.
b) Kőrösi et al. use the level of hourly earnings as the dependent variable, whereas the other authors reported in the table use the logarithm of hourly wage rates. Care should therefore be taken in comparing the estimates of the gender and union effects in the table.

women in 1985 - whether unionised or not. However, three years later it was only among unionised men that inequality fell. For the cohort, inequality within a group was greatest for the women not unionised. But in contrast to the case of men, the unions were (apparently) unable to prevent inequality from increasing among their female members. Using two cross-sections from the ALS it was apparent that the trade union effect was not static.

This result on the gender difference in inequality among unionists, along with the sequence of papers to which reference has been made already, is suggestive of further work. It is agreed in the research dealing with trade unions that there is a substantial pay premium for union members. From Christie's work it is unclear whether the unionisation effect has a gender bias. Her (selectivity-corrected) gender coefficients for unionists and non-unionists are not significantly different. But there is evidence of rising inequality among young female unionists, while inequality was falling among young unionised men. This apparently greater dispersion in female unionists' earnings occurred while the tendency through time was for falling inequality among Australia's young people (Kõrösi et al. 1993: 15-17). Comparable findings for the US are reported in Flaherty and Caniglia (1992). Their key result was that trade union density tends to promote equality among all men but '... union density does not appear to equalise and may even contribute to a disequalisation of the earnings distribution for all women' (Flaherty and Caniglia, 1992: 391-2). An approach which would allow insights on developments through time of gender and union pay gaps, and the dispersion of pay, is indicated.

A comprehensive approach to study the time-dependent processes arising in unionised labour markets was devised by Freeman (1989). He stated that the selectivity bias could be dealt with by using longitudinal methods (Freeman, 1989: 294). Modest errors of measurement in assigning workers to unionised or other categories would, he noted, bias downwards the estimates of the union premium to a greater extent than cross-sectional estimates (1989: 299-302). Freeman deduced that under reasonable assumptions on the incidence of errors of measurement and of selectivity '... cross-section estimates of union effects provide an upper bound and longitudinal estimates provide a lower bound on the "true" union impact in the model under study' (Freeman, 1989: 311). In his longitudinal framework Freeman also investigated the effect of unions on the dispersion of hourly pay. He found that pay dispersion for the group who joined unions tended to fall, whereas for the group of workers who left unions, pay dispersion tended to rise. He took this as evidence confirming the finding from cross-section studies of lower wage dispersion for unionists (Freeman, 1989: 308). On this, Miller and Mulvey (1993) noted that while Kornfeld's (1990) work did not conform with Freeman's results, the age range of the (ALS) sample is probably too young for it to be a satisfactory test.

On dispersion Elliott (1991) summarised some evidence to conclude that '... the otherwise least well paid do better as union members and the effect is ... to narrow the dispersion of earnings' (Elliott, 1991: 440). He also remarked on the tendency for earnings dispersion to be low in strongly unionised economies, and that pay dispersion is least in the unionised parts of labour markets where unions are less dominant (Elliot, 1991: 439). A linkage between pay dispersion, labour market institutions, and the gender pay gap was suggested by Blau and Kahn (1992: 1). They commented that '... the wage determination process in the US is more decentralised than elsewhere, quite likely contributing to its higher level of wage inequality' (1992: 28). Their argument is that if the US pay structure resembled those of highly unionised labour markets like Sweden and Australia, then that pay compression would act to lower the gender pay gap.

2 Analysis of a Panel

To focus on our principal interests of gender and union pay discrimination, a balanced sample was selected of individuals for whom earnings were reported in each survey year. These data were used to estimate a series of regressions: from pooled data for all four years; for annual cross-sections; and for the random effects panel estimator. The coefficients from the regressions are used in two ways in this paper. There is first the conventional discussion of the models and their explanatory variables. This gives an indication of the magnitude of the pay gaps and their development through time. Second, pay discrimination between sub-sets of the panel is investigated. Finally, an analysis is made of inequality within similar labour force groups.

The explanatory variables in the regressions fall into several groups. There are human capital variables, industry and occupation variables, and trade union membership and gender. The set of human capital variables together specify the usual quadratic terms for the influences of education and experience. The greater the number of years of education and work experience, the greater earnings are expected to be. However their influence is thought to diminish over time. Hours worked is modelled as a continuous variable. Working part-time might raise the average hourly wage in some jobs if a premium is paid on hourly rates. It may be that individuals work to a target level of income, accepting some jobs at lower rates to achieve the target. There is sufficient variation in mean earnings across the industrial and occupational structures of the youth labour market to expect different influences from them on hourly earnings.

The expected effect of trade union membership for the young is ambiguous. Although trade unions may aim to secure higher wages for their members, it does not necessarily follow that they are successful. On evidence that pay dispersion is less in highly unionised labour markets, Elliott (1991: 439-40) argues that unions narrow the dispersion of pay and are thus a force for lessening inequality. Against this it can be argued that, all else equal, a union premium for members would involve more overall inequality than in a non-unionised market. On such dispersion, Kornfeld (1990) used ALS data for 1988 and found slightly greater dispersion in the wages of people who did not belong to trade unions. The degree of union density and the chance of spillovers to non-union workers however, are among the factors which make the effect of unionisation on inequality uncertain. Here, in this section there are three empirical issues: first the extent of the union mark-ups, second to what extent there is a gender bias in any union premium, and third whether the dispersion of earnings is different among the unionists and among the others.

The age-earnings profiles given in Kõrösi et al. (1993, Chart 3) show a gender gap in real hourly earnings for 1985. By 1988 this gap had expanded so that, for example, the women who were 21 years old in 1985 earned 96 per cent of the earnings of 21 year old men. Only three years later the gap for these **same individuals** had widened to eight percentage points. Few of the respondents to the survey had formed family units or had taken career breaks. Thus evidence of a widening gender pay gap among these young Australians is unexpected. One explanation is that employers' expectations of interrupted workforce attachment for women could result in pay discrimination, less training opportunity, and relatively fewer chances of promotion for women. In combination, these factors could produce a widening gap through time. Employers' expectations are proxied by gender.

2.1 Earnings Gaps

Who then are the lucky young Australians? The coefficients and test statistics for the first set of regressions are in Appendix Table A1. The sample size is 1453 individuals. Almost all of the estimated coefficients are significant at the one per cent level or better. The diagnostic testing is reported at the foot of the table. From estimations not reported here, it was found that a number of industries, occupations, and variables reflecting household social and financial structures, studying, marital status, and trade union size contributed little to the models. For the pooled regression reported in column one, there is strong evidence of mispecification. Tests for structural breaks were significant. In each regression heteroscedasticity was indicated and the standard errors were corrected. There is clear evidence that the residuals are not normally distributed.

In general the estimated coefficients are broadly similar across the columns. The pooled and random effects models provide comparable evidence on the importance of all the explanatory variables, save two of the industry dummies. The prior expectation of the diminishing importance of human capital influences is confirmed. After 1986 the estimated role of education - years of schooling - is slight. As anticipated, working more hours was associated with lower hourly earnings, on average.

Some lucky individuals, at least in terms of their pay, worked in mining, but others who were less fortunate worked in agriculture, finance, and wholesale and retail trade. By 1987 the premium for mining workers was about three times larger than was the pay penalty for agricultural workers. These inter-industry pay differences for the young reflect an industrial pay ranking as it appears in many economies. Agricultural workers are poorly paid while miners may receive up to double average earnings. Elliott (1991: 345) gives the example of the hourly pay differences among unskilled male workers in Britain having a top rate two-thirds greater than in the lowest paying industry. He records that industry 'fixed effects' have been shown to account for between seven and 30 per cent of differences in individual earnings in the US. Similarly, King (1990: 133) is impressed by longitudinal evidence which shows that ' ... as workers move from industry to industry, their earnings change, while their individual endowments and personality traits do not'. It is not the purpose here to explain why particular inter-industry differences occur. Rather, the regressions show that working in a particular sector may bring a young worker a special benefit or a cost. As in the evidence referred to by Elliott and King, controlling for variations in education, experience, occupation and gender does not eliminate the dispersion of pay across industries.

Managerial and supervisory workers, the professionals, and para-professionals all enjoyed a pay premium, on average of about ten per cent. The trade unions are certainly of importance to young Australians. After controlling for human capital, industry, occupation, working hours, and gender, it is estimated that the union pay premium is in the range of six to nine per cent. This accords with the estimates in Kōrösi et al. (1993). This premium is however about half that reported in the research summarised in Appendix Table A1. A pooled regression for men suggests a union premium of nine per cent, whereas in the comparable regression the women's union premium is about five per cent.[2]

The gender bias in the union premium remains to be discussed later in more detail in this paper. After allowance for the other explanatory factors listed in the table, gender appears to be becoming more important through time. Between 1985 and 1988, occupational segregation by gender was falling, and would have been an equalising force. The raw pay gap in the sample was nine per cent. The pooled regression suggests a premium of eight per cent, whereas the random effects model suggests it to be nine per cent. Little, if any, of the raw gap would appear to be explained by the other regressors.

Is it that current earnings of the young depend on what they earned previously? In Appendix Table A2 results are reported for similar models to those already discussed, except for the inclusion of a one-year lag on the dependent variable, the log of hourly earnings. For unionists there remains a premium above the earnings of non-unionists, but in the latter table the coefficients are smaller (but significant still). For gender, the pooled and random effects estimates are comparable in the

2 These regressions are not reported here. With the balanced data for additional regressions were estimated - for men and women unionised and non-unionised. Differences in co-efficient for almost all of the variables were not sufficiently great to warrant including another four regression tables. They are available on request from the authors.

Table 2: Age at which the Return to Experience is Maximised[(a)]

	Educational Attainment		
	Early School Leaver	Completed Secondary School	Three years Post-secondary
	Age	Age	Age
Men	26	28	30
Unionised	25	27	28
N/unionised	26	28	30
Women	26	27	29
Unionised	24	26	27
N/unionised	26	27	29

Note: a) The ages were found from the earnings/education profiles in the pooled regression by maximising the quadratic in experience for each educational level.

two tables. Whereas in Table A1 the gender pay gap was larger in 1988 than three years earlier, here the coefficients on gender are not significantly different in each of the cross-section models.

Over the survey years ALS respondents reached various educational targets. For some, leaving school at age 15 was the target, for others it was acquiring post-school credentials. Age-earnings profiles typically show faster earnings growth and higher levels of earnings for the better educated. Commonly, earnings peak sooner for the less-educated workers. From the human capital elements of the regressions, the age, all else equal, at which earnings cease growing with further experience, can be calculated for differently educated people, in various labour force groups. Such an exercise is reported in Table 2.

The entries in the table show surprising uniformity, given that the years of schooling range from nine for early school leavers to sixteen for those with more than a secondary school level of education. The biggest age difference in reaching the peak return to experience is between men who were early school leavers, and men with post-secondary schooling. Women in unions attain the peak sooner than all other groups.

It would be expected that early school leavers, in jobs for which experience counts for little, would attain the peak return to growing experience at a young age. Against other evidence (Berndt, 1990: 173-4), it would be expected that the returns to experience are maximised at similar ages for people with very different educational attainments. Berndt reports that for US men, it was found that the returns to experience are maximised for elementary school graduates, high school graduates and college graduates at about 45 years of age. The problem in interpreting the results in Table 2, is that the underlying data refer to a small range of ages (16 to 28) over a small number of years. These factors may bring a downward bias to the ages estimated. Also, the truncation of the sample with respect to age, further suggests that the quadratic earnings function, which is the usual (convenient) way to model human capital, is inappropriate.

2.2 Pay Discrimination

Earnings differences can arise because workers of similar characteristics (or endowments) receive different rewards for those characteristics, as well as because workers have differing endowments. A common method used to identify these two effects is to write the mean difference in the predicted log of earnings from the regressions for two groups of workers as follows[3]. Write

$$\log y_{g,i} = \alpha_g \beta_g{}' X_{g,i} + u_{g,i};\ g = 1,2 \tag{1}$$

and form the difference

$$\overline{\log y}_1 - \overline{\log y}_2 = \Delta\alpha + \overline{X}'_1 \Delta\beta + \beta'_2 \Delta\overline{X}_1\ , \tag{2}$$

where $\overline{\log y}_g$ denotes the mean of log earnings for workers in group g, $\overline{X}_2$ denotes the mean endowments of group 2 workers, β_1 denotes the coefficients (estimated for group 1 by OLS regression) of the form of equation (1) and Δ denotes change. In (2) the mean difference in log earnings has been decomposed into an effect we call 'Discrimination' (due to a difference in coefficients) and the other 'Endowment' effect (due to different means). While the decomposition is helpful in sorting out the effects of discrimination from differing endowments, Berndt (1990: 184) warns that if relevant endowments are not included, and if something other than discrimination could influence the coefficients, the 'procedure would be inappropriate'. Using the results from the pooled regressions for all males and all females, the decomposition (2) of the overall change in the means of the dependent variables for men and women is set out in Table 3.

When decomposed, changes in 'rewards' rather than 'basic stocks' dominate the gender difference in earnings.[4] The discrimination against women (in Table A1) has its source in the returns to experience, mining jobs, and managerial and supervisory work. When lagged earnings are modelled, it is past earnings, as well as the returns to education and experience which dominate the discrimination

3 The method is attributable to Blinder and Oaxaca and is fully described in Berndt (1990, Chapter 5).

4 This was true whatever order was taken with the decomposition.

Table 3: Decomposition of the Male/Female Earnings Difference

	Overall difference in the mean of log earnings between males and females	Part due to Discrimination[a] (coefficients)	Endowment (means)
Without lagged dependent variable	0.08	0.10	-0.02
(%)		(119.0)	(-19.0)
With lagged dependent variable[b] (%)			
	0.09	0.08	0.01
		(91.8)	(8.2)

Notes: a) Chow tests for the coefficient differences between males and females are significant at the one per cent level in both models.
b) This estimate is for 1986 to 1988 only.

effect (reported in column two). In terms of their relative endowments, men have lower educational attainment, but they are able to compensate for that through their greater experience, and the composite variable of education and experience. Their strong tendency to work longer hours than women -about four hours a week - represents another endowment favouring men. Given the very small ß coefficients on hours worked, this will provide only a little of the explanation for the difference in the mean of log earnings by gender.

Recall evidence cited earlier of the widening gender pay gap in the age earnings profiles (Kõrösi et al., 1993), and also the gender coefficients in the cross-section regressions in Table A1. The gender pay gap is decomposed for each survey year in Table 4.

This decomposition shows that the discrimination effect may have accounted for more of the pay difference as time passed. Underlying sources of this discrimination include working in any of the included industries, in managerial and supervisory occupations, and in 1986 trade union membership had a small role.

Similarly to gender-based discrimination, it is possible to decompose the overall change in the means of the dependent variables for unionists and others. The result is set out in Table 5.

Table 4: Gender Earnings Gasp. 1985 to 1988

	Overall difference in the mean of log earnings between males and females	Part due to Discrimination[a] (coefficients)	Part due to Endowment (means)
1985	0.062	0.067	-0.004
(%)		(107.1)	(-7.1)
1986	0.094	0.105	-0.010
(%)		(111.2)	(-11.2)
1987	0.091	0.104	-0.013
(%)		(113.9)	(-13.9)
1988	0.091	0.127	-0.036
(%)		(140.1)	(-4.01)

Note: a) Chow tests for the coefficient differences between males and females were not significant in any year at the one per cent level.

Table 5: Change in Earnings by Gender and Union Membership

	Overall difference in the mean of log earnings between unionists and non-unionists	Part due to Discrimination[a] (coefficients)	Part due to Endowment (means)
Females			
U - NU	0.11	0.05	0.06
(%)		(42.6)	(57.4)
Male			
U - NU	0.20	0.08	0.12
(%)		(41.2)	(58.8)

Note: a) Chow tests for the coefficient differences between unionists and others are significant at the one per cent level for females and males.

There appears to be a smaller union pay gap for the young women than the young men. But in this case it is obvious that the pay gap for both men and women is due more to differences in endowments, than to the union premium indicated by discrimination. For the young women the sources of discrimination are agriculture, mining, para-professional jobs, and working hours. Professional, agricultural, and finance sector jobs, and working hours are the sources of the discrimination among men. Further, the strength of the endowment contribution means that unionists are different to other workers (of the same gender) with respect to the variables modelled in the regressions. This is not to ignore the extent of the discrimination factor. What the decomposition shows is that among women (and among men), about forty per cent of the pay difference between the unionists and the rest, is due to bigger rewards (and/or smaller penalties) for unionised workers with respect to the independent variables.

2.3 Inequality

From the regressions it has been possible to examine mean earnings for various groups of ALS respondents. Now the focus is on the dispersion of earnings. Earlier work indicated (for a different sample from the ALS) that inequality was generally lessening over the survey years (Kõrösi et al., 1993: 15). Rising (or falling) inequality has two sources which can be identified: that which emerges 'between' groups of respondents, I_B, and that which arises 'within' groups, I_W.[5] Changes in the mean earnings of parts of the sample relative to the overall mean, give rise to 'between-group' inequality. It is also influenced by demographic change as groups come to account for larger (or smaller) shares of the sample. However this is ruled out as an influence on gender-based inequality where the calculations are for the balanced sample. There, all of the 'between' gender inequality must be attributable to changes in the relative earnings of men and women. The 'within' component of inequality is due to influences which cause earnings to be more (or less) dispersed around mean earnings for the relevant group. The calculations for gender-based inequality are set out in Table 6.

Overall, inequality for the balanced sample was falling, as shown by I_0. Most of the inequality is attributable to I_W : that within the group of male respondents and within the group of women. However, the small growth in the gender pay gap by 1988 is confirmed again as I_B shows that gender-based inequality was rising, especially after 1985. Although its (percentage) contribution to overall inequality was small, it was larger at the end of the survey period.

Observed deunionisation in many labour markets has been thought of as a possible reason for growing inequality. Here among young Australians inequality was falling over the survey years. The dominance of I_W was evident for gender, here the issue is to examine inequality within the unionised and within the un-unionised workers. The relevant within-group indicators of inequality are set out in Table 7, for each survey period.

Inspection of each row shows that inequality was falling. Through time, inequality fell fastest among non-unionists. In all years, the level of inequality among unionised male workers was less than among the other men. But the difference in I_W for the men for these groups was much greater than for women. Except for 1985, inequality was greater for females in unions than the other

5 Shorrocks (1980) identifies the class of inequality indices which are additively decomposable into within and between effects. Here his I_0 index is used.

Table 6: Gender-based Inequality, 1985-1988

	1985	1986	1987	1988
I_W	69.26	60.94	47.81	42.38
I_B	0.72	1.52	1.06	1.06
I_0	69.98	62.46	48.87	43.44

Table 7: Inequality by Gender and Union Membership, 1985-1988

		1985	1986	1987	1988
	Males				
I_W	Non-union	82.6	77.8	47.2	45.1
	Union	55.1	50.6	45.1	37.6
	Females				
I_W	Non-union	58.5	47.3	42.1	37.7
	Union	57.0	48.3	46.2	39.5

females. It could be argued on this basis that although the unions seem to achieve a significant pay premium, on the face of it, for women, this appears to have made little differential impact on the dispersion of their earnings. The big dispersion in pay (relative to all the others) is for men who do not belong to unions.

Among the balanced sample of 1453 cases, the share of young people who belonged to unions grew. In 1985, 47.1 per cent of the women were unionised and this grew slightly to 48.1 per cent by 1988. For men unionisation rose from 45.7 per cent to 53.4 per cent. Given these shifts and the evidence of falling union pay premia, the between-group inequality for unionists and others cannot be predicted in advance. The between components can be partitioned into the effects of increasing unionisation and the falling union pay relativity. Such calculations show that the falling relativity for unionists accounts for most of the between component, with the shifts in union shares having a trivial impact. Although the sample is balanced with respect to earnings, these individuals need not have had the same union status in each year. We have not attempted work similar to Freeman (1989) or Kornfeld (1990) who compared changes in the dispersion of pay between union entrants and leavers.

3 Concluding Observations

The young people in the ALS were born in the 1960s, and much of their education took place in the 1970s. All of them would have begun working after Australia's equal pay legislation was in place. The oldest members of the panel could have begun working in 1975, while the youngest could have begun in 1984 when the Sex Discrimination Act was passed by the Federal Parliament. Leading up to this, many issues of relevance to women, such as maternity leave, the appointment of ministerial advisers on women's matters, and the Federal Child Care Act, brought to prominence the expectation of equity for women in the labour market. Given this background, it could be expected that the ALS panel members would have experienced less gender bias at the work place than their parents.

For the parents, much of their working life occurred when (after 1950) the legislated basic wage for women was 75 per cent of that for men. There has been a considerable closure of the gender pay gap since the fifties. Over the ALS survey period, the pay gap for all non-managerial adult Australians was around 12 per cent. For the ALS respondents, the raw gap in the pooled sample was nine per cent. It could speculated that this young generation did benefit from the recent enlightened period. Against that, several pieces of evidence reported here suggest a widening gender pay gap. Also, in spite of legislation covering award rates[6], it is very obvious that Australian employers continued to provide different rewards to young people of similar characteristics, save gender. While others have established the size of the gender pay gap to lie between ten and twenty per cent, here the gap is shown to be smaller in all of the regressions. It is not possible to know here whether this is due to different background events, or whether the different ages included in the ALS and the survey data (from ANSSS) where the whole workforce aged 18 and over is included, account for the different estimates of the gap. The extension provided here is the unfolding through time of rising pay discrimination for this panel.

As with gender, it has been possible to use the panel to study the time path of the falling trade union premium. The decomposition of union non-union pay differences produced several new insights. Whereas in the case of gender, discrimination dominated the earnings difference, in the union decomposition the endowment effect was dominant. However there was almost no difference in the union discrimination effect for women and men, despite the larger union pay premium for men.

On inequality it was found that the dispersion of earnings was approximately equal for three groups : unionised males, and females, whether union members or

6 The award is the (legally binding) minimum rate of pay set by one of a range of wage setting tribunals. Award coverage is extensive, especially in public sector employment, but is declining.

not. The greatest improvement in inequality was where the dispersion of earnings was greatest in 1985, that is for men who were not unionists.

At several points it has been noted that errors of specification have probably occurred. Two obvious problems are the small number of years of data and the restricted age range in the panel. It may not be reasonable to judge the relative returns to different educational attainments with this data set using the quadratic specification of education and experience. While the **level** of earnings will differ with people's years of education, what was observed was that the peak return to experience was reached at very similar, and very young, ages. On the evidence here it would be hard to decide how fast any bonus from more education is dissipated. For some fields of tertiary education, the returns to the additional years of study may not be maximised within the age range of the ALS respondents.

Some members of the panel easily qualify as lucky young Australians. They are men, especially those in unions, miners, managers and supervisors, and professional workers. In the regression analysis it was possible to dismiss as unimportant several 'family-related' variables. On the one hand it could be thought that the further maturation of the panel and the growing incidence of family formation would lead to an even greater gender pay gap beyond 1988. On the other hand, the evidence on discrimination in this paper suggests that the ALS women may have already been penalised for what employers expected about their continued labour market participation.

Appendix

Table A1: Regressions for Log Earnings, 1985-1988[a][b][c]

Variable	All years	1985	1986	1987	1988	Random Effects
Education	0.33	0.48	0.38	0.18	0.12	0.34
	(14.04)**	(10.56)**	(7.87)**	(2.64)**	(1.53)	(17.99)**
Education^2	-0.008	-0.01	-0.01	-0.004	-0.002	-0.008
	(-9.03)**	(-7.58)**	(-5.12)**	(1.51)	(-0.52)	(-11.18)**
Experience	0.31	0.47	0.38	0.21	0.17	0.32
	(22.32)**	(12.72)**	(11.55)**	(6.17)**	(4.49)**	(22.13)**
Experience^2	-0.009	-0.01	-0.01	-0.007	-0.005	-0.009
	(-23.18)**	(-12.08)**	(-11.1)**	(-7.23)**	(-5.44)**	(-22.61)**
Education x Experience	-0.01	-0.02	-0.02	-0.007	-0.006	-0.01
	(-13.95)**	(-8.75)**	(-7.73)**	(-3.27)**	(-2.42)*	(-13.72)**
Hours	-0.008	-0.009	-0.009	-0.007	-0.007	-0.008
	(-15.14)**	(-9.74)**	(-8.26)**	(-6.39)**	(-5.87)**	(-22.35)**
Industry						
Agriculture	-0.22	-0.28	-0.19	-0.13	-0.30	-0.13
	(-5.05)**	(-2.87)**	(-2.31)*	(-2.04)*	(-3.15)**	(-4.30)**
Mining	0.40	0.37	0.45	0.45	0.30	0.31
	(8.82)**	(2.81)**	(5.40)**	(5.03)*	(4.51)**	(9.14)**
Wholesale/Retail	-0.10	-0.12	-0.09	-0.10	-0.09	-0.07
	(-11.77)**	(-6.47)**	(-5.42)**	(-6.05)**	(-5.64)**	(-7.96)**
Finance	-0.06	-0.08	-0.06	-0.06	-0.04	-0.05
	(-6.51)**	(-4.40)**	(-3.37)**	(-3.19)**	(-2.20)*	(-3.80)**
Occupation						
Managerial	0.13	0.10	0.11	0.14	0.15	0.11
	(10.31)**	(4.09)**	(4.12)**	(5.66)**	(6.26)**	(8.12)**
Professional	0.16	0.10	0.16	0.18	0.17	0.13
	(11.23)**	(2.96)**	(4.91)**	(6.62)**	(7.81)**	(9.31)**
Para-professional	0.08	0.06	0.04	0.13	0.08	0.07
	(4.99)**	(1.91)	(1.19)	(4.27)*	(2.80)**	(4.44)**
Union	0.07	0.08	0.09	0.07	0.06	0.07
	(10.92)**	(6.16)**	(6.17)**	(4.77)**	(4.92)**	(8.94)**
Gender	-0.08	-0.05	-0.09	-0.09	-0.11	-0.09
	(-11.84)**	(-3.56)**	(-6.31)**	(-6.51)**	(-7.54)**	(-8.30)**
Constant	-1.00	-2.13	-1.46	0.21	0.62	-1.07
	(-6.45)**	(-6.78)**	(-4.57)**	(0.45)	(1.10)	(-8.28)**
N	5812	1453	1453	1453	1453	1453 x 4
Mean of log earnings	1.94	1.85	1.92	1.98	2.03	1.02
Standard deviation of log earnings	0.35	0.38	0.35	0.31	0.30	0.25
R^2	0.48	0.54	0.50	0.40	0.35	0.38

Table A1: Regressions for Log Earnings, 1985-1988[(a)(b)(c)]
(Continued)

Variable	All years	1985	1986	1987	1988	Random Effects
Breusch-Pagan test for heteroskedasticity	88.95**	90.47**	59.93**	42.98**	111.84**	
Jarque-Bera test for normality	6784.3**	1364.4**	1648.3**	2672.0**	991.6**	
F test for zero slopes	357.7**	114.7	99.64**	64.62**	53.94**	
Chow test for parameter constancy	3.53**					
Reset text $(\log w)^2$	14.28**	0.08	0.26	0.02	4.52*	
Reset test $(\log w)^2$ and $(\log w)^3$	7.44**	3.08*	0.18	0.18	2.35	
Linear vs log linear specification (PE test) H_0: linear	4.54**	0.90	-0.67	0.32	2.38*	
H_0: log linear	4.70**	4.38**	5.44**	3.07**	0.74	
Lagrange multiplier test for random effects	1281.1**					

Notes:
a) T values are shown in brackets. Except for the random effects estimation heteroskedastic-consistent standard errors were used to calculate the test statistics.
b) ** denotes significance at the one per cent level, and * denotes significance at least at the 5 per cent level.
c) In addition to the reported tests other checks were conducted for omitted regressors and to check for structural breaks involving gender.

Table A2: Regressions for Log Earnings with Lagged Earnings, 1985-1988[(a)(b)(c)]

Variable	All years	1986	1987	1988	Random Effects
Log ($earnings_{-1}$)	0.42	0.37	0.45	0.45	0.20
	(22.41)**	(12.16)**	(12.15)**	(15.97)**	(16.13)**
Education	0.14	0.22	0.05	0.05	0.24
	(4.91)**	(4.68)**	(0.88)	(0.71)	(9.64)**
$Education^2$	-0.004	-0.006	-0.002	-0.000	-0.006
	(-3.46)**	(-3.34)**	(-0.78)	(-0.16)	(-6.51)**
Experience	0.12	0.21	0.04	0.07	0.22
	(6.94)**	(5.95)**	(1.18)	(1.97)*	(12.16)**
$Experience^2$	-0.004	-0.007	-0.002	-0.002	-0.006
	(-8.37)**	(-6.53)**	(-2.28)*	(-2.55)*	(-13.25)**
Education x Experience	-0.004	-0.008	0.000	-0.002	-0.009
	(-3.89)**	(-3.83)**	(0.17)	(-0.98)	(-7.53)**
Hours	-0.006	-0.007	-0.005	-0.005	-0.008
	(-9.55)**	(6.79)**	(-5.13)**	(-4.55)**	(-18.61)**
Industry					
Agriculture	-0.13	-0.10	-0.09	-0.22	-0.11
	(-3.61)**	(-1.56)	(-2.03)*	(-3.03)**	(-3.44)**
Mining	0.26	0.37	0.26	0.15	0.29
	(7.71)**	(5.72)**	(4.31)*	(4.18)**	(8.99)**
Wholesale/Retail	-0.05	-0.06	-0.05	-0.05	-0.05
	(-6.61)**	(-3.81)**	(-3.95)**	(-3.74)**	(-5.64)**
Occupation					
Managerial	0.10	0.09	0.09	0.10	0.10
	(7.81)**	(3.67)**	(4.33)**	(5.29)**	(7.35)**
Professional	0.13	0.14	0.13	0.11	0.14
	(8.45)**	(4.09)**	(5.07)**	(5.46)**	(10.11)**
Para-professional	0.08	0.05	0.12	0.05	0.07
	(4.85)**	(1.77)	(4.93)**	(1.93)	(4.58)**
Union	0.04	0.06	0.03	0.03	0.06
	(5.88)**	(4.73)**	(2.61)**	(2.90)**	(7.53)**
Gender	-0.07	-0.08	-0.06	-0.07	-0.09
	(-9.58)**	(-6.04)**	(-4.99)**	(-5.53)**	(-9.57)**
Constant	0.04	-0.60	0.77	0.65	-0.50
	(0.22)	(-1.96)*	(1.86)	(1.37)	(2.89)**
N	4359	1453	1453	1453	1453 x 3
Mean of log earnings	1.98	1.92	1.98	2.03	1.23
Standard deviation of log earnings	0.33	0.35	0.31	0.30	0.24
R^2	0.55	0.59	0.53	0.50	0.39
Breusch-Pagan test for heteroskedasticity	151.34**	48.92**	48.36**	105.55**	

Table A2: Regressions for Log Earnings with Lagged Earnings, 1985-1988(a)(b)(c)
(Continued)

Variable	All years	1986	1987	1988	Random Effects
Jarque-Bera test for normality	10637.9**	2924.3**	5620.0**	2614.5**	
F test for zero slopes	357.7**	137.8**	110.8**	99.13**	
Chow test for parameter constancy	2.09**				
Reset text $(\log w)^2$	1.13	0.02	0.06	0.64	
Reset test $(\log w)^2$ and $(\log w)^3$	0.63	0.17	0.85	0.30	
Lagrange multiplier test for random effects	2.88				

Notes:
a) T values are shown in brackets. Except for the random effects estimation heteroskedastic-consistent standard errors were used to calculate the test statistics.
b) ** denotes significance at the one per cent level, and * denotes significance at least at the 5 per cent level.
c) In addition to the reported tests other checks were conducted for omitted regressors and to check for structural breaks involving gender.

References

Berndt, E. (1990), *The Practice of Econometrics : Classic and Contemporary*, Addison Wesley, Reading.

Blau, F. and L. Kahn (1992), *The Gender Earnings Gap: Some International Evidence*, NBER Working Paper, 4224, National Bureau of Economic Research, Cambridge, Massachussets.

Christie, V. (1992), 'Union wage effects and the probability of union membership', *Economic Record*, 68(200), 43-56.

Elliott, R. (1991), *Labour Economics: A Comparative Text*, McGraw-Hill, Maidenhead.

Flaherty, S. and A. Caniglia (1992), 'The relative effects of unionism on the earnings distributions of women and men', *Industrial Relations*, 31(2), 382-93.

Freeman, R. (1989), *Labour Markets in Action, Essays in Empirical Economics*, Harvester Wheatsheaf, New York.

King, J. E. (1990), *Labour Economics : An Australian Perspective*, Macmillan, South Melbourne.

Kornfeld, R. (1990), 'Australia's trade unions: wage, fringe and job tenure effects', mimeo, National Bureau of Economic Research, Cambridge, Massachusetts.

Kõrösi, G., G. Parkinson, R. Rimmer. and S. Rimmer (1993), *Rising Inequality? Shifts in the Distributions of Earnings and Income*, La Trobe University Discussion Paper 12/93.

Miller, P. and C. Mulvey (1992), 'Trade unions, collective voice and fringe benefits', *Economic Record*, 68(201), 125-41.

Miller, P. and C. Mulvey. (1993), 'What do Australian unions do?' *Economic Record*, 69(206), 315-42.

Miller, P. and S. Rummery (1989), 'Unionism and the structure of male wages for youth', *Journal of Industrial Relations*, 2,185-211.

Rummery, S. (1992), 'The contribution of intermittent labour force participation to the gender wage differential', *Economic Record*, 68(202) 351-64.

Shorrocks, A. (1980), 'The additively decomposable inequality measures', *Econometrica*, 48(3), 613-25.

Comments on paper by Gábor Kõrösi, Russell J. Rimmer and Sheila Rimmer

Elizabeth Savage
Department of Economics

The recent and continuing period of industrial relations reform in Australia make this an appropriate time to consider the effects of unions, and indeed gender, on earnings. Detailed econometric analysis of earnings differentials provides an important input to discussions of the likely impact of changes to the system placing greater emphasis on bargaining at the enterprise level. One potential impact is on the extent of inequality in Australian society. If unions have been successful in obtaining wage mark-ups then a change which reduced their power to do so will impact on the earnings of a large portion of the populations. The work by Kõrösi, Rimmer and Rimmer provides an important contribution to this area. As well as reporting their own work, they provide an extremely useful survey of Australian work on union and gender earnings gaps.

Of course, the future implications of industrial relations reform depend on the facts that have been necessary for gaining a union wage premium, and identifying industrial relations activities which have led to a union premium is an important focus of the literature on union mark-ups. There has been quite a lot of recent work on union wage differentials both in Australia and overseas. The research tends to suggest that, ceteris paribus, unionised workers tend to gain more than those who do not belong to a union. However there has been considerable variation in the estimates of union mark-ups. More recent studies have used micro-data, in Australia primarily data from the Australian National Social Science Survey and the Australian Longitudinal Survey. These have tended to find smaller wage mark-ups than studies using data at the industry level. The work by Kõrösi et al. find relatively small wage mark-ups of less than 10 per cent, similar to that obtained in recent work by Kornfeld (1993).

The quite considerable variation in the size of estimated mark-ups means that it is important to consider the impact of the control variables available in the data, and this is where I focus my comments. The work by Kõrösi et al. and the papers which they survey, use data at the level of the individual. This provides good data on personal training and experience variables, but it has no data on workplace characteristics. There are two reasons why this might be of some concern. Firstly, the workplace industrial relations environment may provide important determinants of union mark-ups and omitting these important variables may bias the estimated earnings differentials. Secondly, if we are to draw implications for industrial relations reform, workplace industrial relations activities are crucial since this is where change is focused. The Australian Workplace Industry Relations Survey (AWIRS) provides such data but it comes at the cost of human capital data. For estimation based on the AWIRS data we can tell whether workplace industrial relations variables are likely to be important factors determining earnings differentials and hence whether omitting them might be of concern. The AWIRS results are drawn from recent work in Apps, Dearden, Jones and Savage (1994).

Before I discuss the results I have some general comments on methodology. The estimations in Kõrösi et al. on which the union and gender earnings differentials are based include dummy variables for gender, union status, occupation and industry, as well as human capital variables .

This approach does not allow spillover effects between these dummies and other explanatory variables. It is often found in the literature of union mark-ups that the impact of explanatory variables on earnings differs by union status and that full interactions should be allowed between union status and other explanatory variables by estimating separate equations for the union and non-unionised workers (see work by Stewart, 1987 and 1990). I would also expect that union and gender earnings differentials would vary by occupation so I have some misgivings about the reliance on dummy variables. In the work by Apps et al. we found that separate equations for fully unionised and less unionised workplaces were preferred, indicating significant spillover effects. Pooled estimates may conceal important features of the data. Ideally if both gender and union earnings differentials are being analysed, splitting the sample into four subgroups would provided a very detailed description of the problem.

Whether industrial relations characteristics of the workplace are important for earnings differentials depends on the system of wage setting and how wage determination procedures might impact on earnings for union and non-union members. In Australia, Federal and State Tribunals adjust wages for all workers covered by awards. In addition industry cases adjust minimum wages and conditions for particular awards. Strong unions can achieve favourable outcomes from these cases and these awards then apply to all workers, union or non-union. Unions may negotiate higher award wages where industry and occupation unionisation is high. This effect may be overlooked unless variables for aggregate union density are included. Otherwise the effect will be in the coefficients on occupation and industry dummies and may not be identified as a union effect. Unions may also negotiate higher over award payments at the workplace level and workplace industrial relations variables such as presence of bargaining and union delegate activity would be indicators of such activity. Union membership of individuals fails to capture such activity. If the workplace industrial relations environment tends to vary by occupation and industry,. again the coefficients on these variables rather than the coefficient on the union dummy may capture these effects. Unions may also negotiate for promotion of union members to higher paying positions. This effect would be capture by the union dummy; however, again this effect may vary across occupations and industries.

In Apps et al., we find that both industrial relations characteristics of the workplace and of the industry are significant factors in determining earnings. This suggests the omission of workplace industrial relations variables in Kõrösi et al. may be of some concern especially if these are related to included variables such as occupation and industry. We also found that the implied ceteris paribus union mark-ups were very sensitive to the workplace environment. The largest workplace union mark-ups arise in workplaces with bargaining and active union delegates, but bargaining effects usually dominate. There are no mark-ups from workplace unionisation for plant and machine operators, and for this occupation and for labourers the extent of industry unionisation is crucial for achieving mark-ups in the unionised sector. These results indicate the importance of undertaking separate analyses for major occupation categories and allowing full interactions, rather than pooling and relying on occupation dummy variables.

Unless this is done it is difficult to be confident about conclusions concerning, for example, discrimination versus endowment effects. In summary, my major concern is that the individual data available from the Longitudinal Survey of young people does not permit the detailed exploration of union and gender earnings effects because of the omission of the workplace variables which appear to be important determinants of earnings differentials.

References

Apps, P. F., L. Dearden, G. S. Jones and E. J. Savage (1993), *Union Wage Differentials for Manual Workers in Australia*, Working Paper, Economics, Research School of Social Sciences, Australian National University, Canberra.

Kornfeld,. R. (1993), 'The effects of union membership on wages and employee benefits: the case of Australia', *Industrial Relations and Labour Review*, 47, 114-28.

Stewart, M. (1987), 'Collective bargaining arrangements, closed shops and relative pay', *Economic Journal*, 97, 140-56.

Stewart, M. (1990), 'Union wage differentials, product market influences and the division of rents', *Economic Journal*, 100, 1122-37.

Inequalities of Income, Health and Happiness: The Stratification Paradigm and Alternatives

Bruce Headey
University of Melbourne
Peter Krause
Deutsches Institut für Australia
Wirtschaftsforschung, Berlin[1]

1 Introduction

Survey researchers are always asking people what they want from life. The three things most people say they want are a high standard of living, good health and happiness (Abrams, Gerard and Timms, 1985; Campbell, Converse and Rogers, 1976; Cantril, 1965; Rokeach, 1973). The purpose of this paper is to examine relationships between material standards of living, health and happiness both cross-sectionally and over time, and then to test alternative theories which might account for these relationships. The main data sources are the German Socio-Economic Panel (1984-91; N=9114) and the Australian (Victorian) Quality of Life Panel Survey (1981-89; N=942).

Many sociologists and other social scientists who have worked in the fields of social stratification and social mobility appear broadly to endorse a theory of inequality which might be termed **stratification theory** or the **stratification paradigm**. There are many variations within this paradigm but there seems broad agreement that all Western societies, and perhaps all industrial societies, are characterised by inequalities which are (1) cumulative (mutually reinforcing) (2) long term and (3) 'inherited' in the sense that they are substantially due to family social background. The assertion that inequalities are cumulative means that the same people tend to have high levels of all socially valued resources (income, status, power, health, happiness etc.), whereas other people have low levels. Asserting that inequalities are long term means that the same people remain at the top or the bottom of the heap for long periods. Reference to the importance of family social background as a determinant of inequalities, means that if one knows the social status of a person's parents, one can predict his/her probable level of resources as an adult. It seems

1 We would like to thank Mariah Evans and Jonathan Kelley of Australian National University for many helpful comments. Craig Lonsdale of the University of Melbourne gave much appreciated assistance with data management and statistical issues.

plausible to suppose that most social scientists and educated laypeople, as well as professional sociologists, have a picture of society not unlike the one conveyed by the stratification paradigm.

The first part of the paper presents German panel results to indicate that, as applied to the outcomes people care most about - material living standards, health and happiness - this theory or paradigm is only weakly true; it lacks explanatory power. It should be conceded immediately that stratification theory is not normally used to account for differences in living standards, health and happiness. The most common dependent variables are individual gross earnings, occupational status or prestige, and power. However, stratification theorists (if one may use that term) state or imply that all other socially valued resources are similarly distributed and some explicitly state that inequalities of health and happiness mirror inequalities of income, status and power to generate an alleged pattern of cumulative, long term and inherited stratification (see quotations below).

The persistence of the stratification paradigm may be illustrated by four quotations from leading sociologists, one for each of the last four decades:

> The social class system ... operates, largely through the inheritance of property, to ensure that each individual maintains a certain social position, determined by birth and irrespective of his particular abilities. (Bottomore, 1965: 16)
>
> It is the highly patterned nature of the inequalities we have so far examined which enables us to portray the reward system in terms of a dichotomous or two class model. (Parkin, 1972: 26)
>
> The further down one is in the social hierarchy, the fewer chances one has of living a long and healthy life...
>
> Sociological data show again and again that the chances for a successful 'pursuit of happiness' vary a great deal from class to class. (Coser et al., 1987: 201)
>
> ....all forms of stratification are characterised by powerful self-maintaining properties...Or, Lieberson concludes: Those who write the rules write rules that enable them to continue to write the rules. (Erikson and Goldthorpe, 1992: 394)

We suspect that in attacking the stratification paradigm we may be accused of attacking a straw man. We point out that T. B. Bottomore, Frank Parkin, Lewis A. Coser, Robert Erikson and John Goldthorpe are all alive, well and influential in sociology. Nevertheless, there have been some dissenting voices and much data that fails to fit the stratification paradigm. The 1968 International Encyclopaedia of the Social Sciences stated bluntly that, 'It is fundamental that social stratification is multidimensional'. Economists who have collected panel data on household income dynamics have found quite startling volatility and not a stable pattern of stratification

(Duncan et al., 1984). There are quite low correlations between the incomes of fathers and sons and even between the incomes of male siblings (Jencks, 1972 but see his revised account in Jencks, 1979). A recent analysis of living standards in Australia found weak-to-moderate relationships between measures of income and health, happiness and social participation (Travers and Richardson, 1993). Indeed different measures of living standards were only moderately correlated. Sociologists have documented high rates of absolute occupational mobility, and have sometimes questioned whether these are compatible with the stratification paradigm, even after allowing for the fact that the relative occupational prospects of middle class children remain considerably better than the prospects of working class children (e.g. Broom et al., 1980; Jones, 1979; Miller, 1960). Sociological research on second generation immigrants also appears to show that barriers to occupational mobility and high income are often overcome (e.g. Evans and Kelley, 1991). Finally, students of happiness or subjective well-being, contrary to Coser et al.'s (1987) claim, have generally remarked on the very weak association between social status and happiness (e.g. Andrews and Withey, 1976; Campbell et al., 1976; Headey and Wearing, 1992; Veenhoven, 1984).

In view of these many discordant findings, it may be true that the persistence of the stratification paradigm is due to (a) lack of agreement on precisely what would constitute a valid test and (b) inability to replace the paradigm with anything better.

What would constitute a fair and valid test? In testing a scientific theory one should look at its weak links not its strongest links. Almost any theory has evidence in its favour, including, for example Aristotle's flat earth theory: the issue is whether there is crucial evidence against it. Many tests of stratification theory concentrate on the area where it is probably strongest; that is, in showing linkages between education, occupational status and income. But stratification theory is quite clearly intended to be and is usually expressed as a theory about the distribution of the **most valued outcomes or resources** in society. These are, it seems, a high standard of living, good health and happiness. So it is reasonable to test the theory by seeing whether it accounts for cross-sectional and longitudinal relationships among these outcomes.

This is what we do in the first part of the paper. In the later parts we make an attempt to develop models which do a better job of accounting for short and medium term variance in living standards, health and happiness. German panel data are used to develop models of the determinants of living standards and health, and Australian panel data are used for a model of happiness or subjective well-being.

Each of these models could be termed **dynamic equilibrium models**. They have three components: stocks, flows, and utilities (satisfactions). Stocks (the capital account) are conceived of as stable or only slowly changing entities. Flows and utilities (the current account) are expected to change more rapidly, at least from year to year. In the language of social indicators research, flows are 'objective' indicators (e.g. annual income), while utilities are subjective indicators (e.g. income satisfaction). In this paper more or less stable stocks, measured at time 1, are hypothesised to constrain subsequent flows and utilities to fall within fairly broad ranges, but ranges which will not increase as time passes (i.e. the degree of constraint imposed by initial stocks levels will not diminish over time). Stable stocks, it is hypothesised, maintain flows and utilities in dynamic equilibrium. However the stocks which affect flows of income and income utility (satisfaction with income and standard of living) are **not** the same as those which affect health flows and utility, and different again from those which affect happiness.

It may be useful at the outset to indicate which stocks appear to constrain flows and utilities within the three domains. This is set out in Figure 1.

In the standard of living domain flows of income and satisfaction with income and living standards are viewed as dependent on initial stocks of human capital (skills conferred by education and job training) and material capital (wealth, net assets). In the health domain the relevant stocks are the physical constitution one was born with, including hereditary risk factors, and one's current physical condition which is a result of previous lifestyle (diet, exercise, smoking, drinking etc.) as well as constitutional factors. These stocks are viewed as determining subsequent health flows (illnesses and death: morbidity and mortality) and satisfaction with health. In the happiness domain the relevant stocks are hypothesised to be the personality traits of extraversion and neuroticism (Argyle, 1987; Eysenck and Eysenck, 1969; Eysenck, 1990; Headey and Wearing, 1989, 1992). These affect subsequent flows of favourable and adverse life events and a person's life satisfaction or subjective well-being.

The stocks-flows-utilities framework is of course borrowed from the field of accounting. It is envisaged that accounts would be audited annually and that the 'auditor' would review both the capital account (stocks) and the current account (flows and utilities). Calculations would be made of returns to capital in each domain and the consequences of decisions to consume or invest current flows would be assessed (Headey, 1993). An accounting framework is well suited to longitudinal analysis of inequalities and other important social phenomena. Indeed, it is hoped that the framework will prove valuable whether or not the particular dynamic equilibrium hypotheses tested in this paper are confirmed or falsified. These hypotheses, it should be stressed, cover just one type of accountancy calculation suggested by the framework.

2 Methods

2.1 Panel Studies: The German and Australian Panels

Both stratification theory and an alternative dynamic equilibrium theory propose hypotheses about the development of inequalities over time. It follows that in order to test them adequately it is necessary to have access to panel data, rather than just cross-sectional surveys. Only a limited number of panel surveys which contain data on stocks, flows and utilities are available. In this paper we rely on two such panels, one German and one Australian. Desirable as it would be to have panels for more

Figure 1: Stocks, Flows and Utilities in Three Domains: Living Standards, Health and Happiness

Domain	Stocks	Flows	Utilities (satisfactions)
Standard of Living	• human capital: education • material capital: wealth	income standards.	satisfaction with income and living
Health	• hereditary and constitutional factors • physical and mental condition/impairment	illnesses, death: morbidity and mortality	satisfaction with health
Happiness	personality: • extraversion • neuroticism	life events and experiences	life satisfaction

countries, it can be argued that, at least for providing a critique of stratification theory, data from one Western country are sufficient. Stratification theory is intended to apply to all Western countries (Blau and Duncan, 1967; Erikson and Goldthorpe, 1987a and b; Featherman, Jones and Hauser, 1975; Goldthorpe, 1987; Lipset and Zetterberg, 1959; Zagorski, 1985; but see Kelley, 1990).

A. **The German Socio-Economic Panel (GSOEP)** began in 1984 with a sample of 9114 West German respondents in 4528 households; everyone aged 16 and over in sample households is interviewed. In addition the GSOEP includes a panel of 4085 foreigners (guest workers) living in 1,393 households. In 1990, following the revolution in the East, a panel of 4453 East Germans in 2197 households was added.

In the present paper analysis is restricted to West Germans. Foreigners (guest workers) are excluded because it did not seem a 'fair' test of stratification theory to include a large sample of first generation immigrants whose lives were almost certain to change a lot during the years in question, making it less rather than more likely than inequalities would appear cumulative and long term. The East German panel was also excluded (a) because only two years of data were available and (b) because the lives of East Germans also were likely to be changing at atypical rates.

Most of the analyses include all West German respondents who agreed to be interviewed eight times between 1984 and 1991. However economists who do research on income dynamics usually prefer to work with samples of males of prime working age (30-55 years) who were in employment throughout the period in question (e.g. Creedy, 1985; Lydall, 1976, but see Saunders et al., 1992). In deference to this practice (although not really agreeing with it, when the aim is to assess inequality in society as a whole), the income/standard of living section of the paper contains analyses drawing on (a) the entire West German panel and (b) only prime age males in continuous employment.

It should be noted that all panel results are weighted, using longitudinal weights calculated to compensate for differential drop-out rates among social and economic groups (Rendtel, 1990).

B. The Victorian Quality of Life Panel Study (VQOL) began in 1981 with a sample of 942 respondents located in both Melbourne and non-metropolitan areas. Respondents were interviewed five times (1981-89) at two-yearly intervals. By 1989 502 panel members remained and only data from them are used in this paper. The aim of the study was to assess the determinants of subjective well-being (or happiness), including the impact of life events and experiences occurring during the period of the study.

2.2 Measures

This section describes the main measures of stocks, flows and utilities used in the income/standard of living domain, the happiness domain and the health domain.

The Income/Standard of Living Domain

Stocks of human and material capital are measured as follows:

Human capital. Years of school education and job training were combined to construct a single measure (Years of Education 84, Years of Education 85 etc). The exact method of calculating these measures is described in Schwarze et al. (1993).

Material capital. Each year respondents are asked whether or not their family owns five assets: a bank savings account, a homebuilding savings account, life insurance, stocks and shares, and business assets. This 0-5 measure of assets predicts income just as well as a more detailed of wealth (net worth) obtained just once in the 1988 survey.

Income. The main income measure used here is the log of household **equivalent income.** This is the household's income net of taxes and including benefits, adjusted for household size, using the OECD equivalence weights which are 1.0 for the first adult, 0.7 for other adults, and 0.5 for children under 18 (Saunders et al., 1992). Plainly if one is concerned to measure family living standards as validly as possible one must measure disposable not gross income, and plainly too one must adjust for household size. The obvious adjustment is to calculate household per capita income but this makes no allowance for economies of scale in larger households or for the fact that children are cheaper to maintain than adults. It can plausibly be argued that equivalent income is the best available measure of material standard of living, or, to be more exact, of a household's potential consumption level (Ringen, 1987).

Utility/satisfaction. Income satisfaction is measured each year in the German panel with a question asking respondents to rate satisfaction with their household income on a 0-10 scale where 0 means totally dissatisfied and 10 means totally satisfied. A question about standard of living rather than income would perhaps be conceptually more appropriate as a measure of utility. However such a question is only asked intermittently in the GSOEP. In fact results correlate very highly with answers to the income satisfaction question. Conceptually, it would be preferable to ask about satisfaction during the last year (rather than now) but answers would probably be much the same.

The Happiness Domain

Two personality traits - extraversion and neuroticism - are generally believed to be largely stable over a lifetime and to a considerable degree hereditary (Costa and McCrae, 1984, 1985; Eysenck and Eysenck, 1969). In this paper they are viewed as stocks hypothesised to affect happiness (Argyle, 1987; Eysenck, 1990; Headey and Wearing, 1989, 1992). Extraversion is hypothesised to be positively related to happiness and neuroticism negatively.

The following measures were included in the VQOL Panel Survey.

Stocks

Extraversion. 24-item scale measuring factors of sociability and impulsiveness (Eysenck and Eysenck, 1969).

Neuroticism. 24-item scale measuring anxiety, hypochondriasis and psychosomatic symptoms (Eysenck and Eysenck, 1969).

Flows: Life Events

List of Recent Experiences. At each interview VQOL respondents completed an inventory listing 93 events which might have happened to them since their last interview two years previously. This life events inventory was a modified form of inventories developed by Henderson et al. (1981) and Holmes and Rahe (1967).

For present purposes the inventory was scaled by simply assigning +1 to all favourable or presumably satisfying events (e.g. job promotion, marriage) and -1 to all adverse or presumably distressing events (e.g. job loss, death of spouse). So for each individual in each time period two life events scores were calculated - a favourable events score and an adverse events score. It has been shown that this straightforward scoring system yields measures which account for as much variance in outcomes (e.g. measures of happiness, or measures of subsequent illness) as more elaborate scoring systems using predetermined weights for more or less serious events (Dawes and Corrigan, 1974; Rahe and Arthur, 1978; Schroeder and Costa, 1984).

Utility: Life Satisfaction. In the VQOL survey happiness was measured with two identical questions about twenty minutes apart in which panel members were asked, 'How do you feel about your life as a whole?' The response scale was as shown in Figure 2.

Responses to the two questions were averaged to give the Life-as-a-whole Index. This index has been found to be one of the most reliable and valid measures of happiness or subjective well-being available (Argyle, 1987; Headey, Kelley and Wearing, 1993; Larsen, Diener and Emmons, 1985). In the present paper, for ease of interpretation, the index has been transformed to run from 0 to 10.

In the German panel study a single item measure of **life satisfaction** was used (0-10 scale : 0 = totally dissatisfied, 10 = totally satisfied).

The Health Domain

To date it has not been possible to obtain a health panel survey with at least 4-5 waves of data which contains all variables required to implement the stocks-flows-utility framework. The GSOEP has much valuable data but lacks measures of some important health stocks. In particular, the only indicators of hereditary or constitutional risk factors are the age of death of one's father and mother, and there are no measures of lifestyle variables (smoking, drinking, exercise etc.) which damage or promote health. Also, all measures are self-reports not medical observations. Consequently, the analyses reported below should be regarded as merely illustrative of the application of the stocks-flows-utilities framework to the health field.

Stocks

Health impairment. A 1-3 scale on which respondents rate the extent to which health problems reduce their ability to engage in normal daily activities (1 = no impairment, 3 = great impairment).

Chronic Illness. Respondents reported whether they had a chronic illness which had lasted more than a year (1 = no, 2 = yes).

Flows

Doctor visits. At each interview respondents are asked how many times in the last year they have attended various types of doctor (general practitioner, gynaecologist etc.). The measure used here is the sum of all doctor visits. It is conceded that doctor visits are by no means a perfectly valid measure of current

Figure 2: Delighted-Terrible Scale

1	2	3	4	5	6	7	8	9
terrible	very unhappy	unhappy	mostly dissatisfied	mixed feelings	mostly satisified	pleased	very pleased	delighted

health, since they vary by sex, age, social class and available free time, as well as the need to attend.

Utility

Respondents rate satisfaction with their health on a 0-10 scale (0 = totally dissatisfied, 10 = totally satisfied).

Additional Measures

In assessing stratification theory it is necessary to use measures of family social background. The two conventional variables are father's occupational prestige and father's education. The former is measured by the Treiman score of father's occupation (Treiman, 1977) and the latter by years of school education, plus years of job training, calculated as for respondents (see above).

Finally, it should be noted that in all equations (except in the section on 'Happiness')[2] results are net of gender and age. Gender and age were used as 'control' variables so that their influence on flows and utilities was not confounded with other stocks included in analyses.

2.3 Note on Use of Standardised Estimates

Most of the results in this paper are in the form of **standardised** (rather than metric) estimates: Pearson product moment correlations and ordinary least squares regression Betas. Standardised estimates were preferred because they assess people's **relative position** in an income distribution, a happiness distribution, or a health distribution at one or more points in time. It is worth emphasising that stratification theory is a theory about relative positions. It says that people who have relatively high incomes will also rate relatively highly on health and happiness measures, and that relatively high incomes at time 1 will be associated with relatively high incomes at time 2, time 3 etc.

2 Gender and age are almost uncorrelated with life satisfaction so there was no point in including them as 'controls'.

Standardised estimates could be misleading if the variances of the variables used in the paper changed much. None does.

3 Results

3.1 Contra Stratification Theory

In order to make a preliminary assessment of the extent to which inequalities of income, health and happiness are cumulative, long term and due to family background, the following model (Figure 3) was estimated, using the GSOEP data, 1984-91.[3]

These results offer only minimal support for stratification theory. Inequalities of income, health and happiness appear only slightly cumulative. The observed correlation between equivalent income and ill-health (number of doctor visits) in 1984 was -0.02, the correlation between income and life satisfaction was 0.14, and the correlation between ill-health and happiness was -0.14. Plainly, even if one disattenuates these correlations to adjust for measurement error, the inference still has to be that the degree of cumulativeness is slight (see the upper right triangle of the correlation matrix beneath Figure 3).

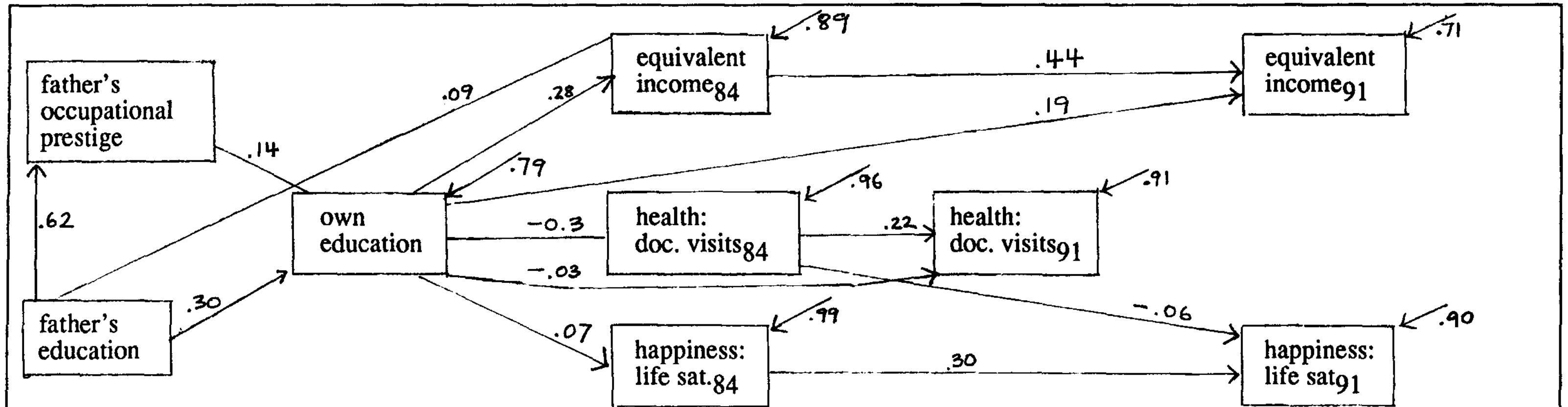
It might also be an indication of cumulative inequality if a person's standing in one domain of life at an earlier point in time affected their standing in another domain at a later point. It transpired, however, that the only statistically significant link of this kind was between ill-health$_{84}$ and equivalent income$_{91}$ (B= -0.06).

The German panel evidence also suggests that inequality within domains is only moderately persistent over time. The observed correlation between disposable income$_{84}$ and disposable income$_{91}$ was 0.50 (disattenuated r = 0.62), while the regression coefficient was 0.44. The observed correlation between health (doctor visits)$_{84}$ and health (doctor visits)$_{91}$ was 0.26 (disattenuated r = 0.41), while the correlation between life satisfaction$_{84}$ and life satisfaction$_{91}$ was 0.30 (disattenuated r = 0.47).

The third central proposition of stratification theory, that family social background substantially affects life chances, receives modest support for income but little or none for health and happiness. Our measures of family social background are father's occupational prestige and father's years of education. Some sociologists would also include here a respondent's own level of education on the grounds that it

3 It may seem confusing to include in the model flow measures of income and health alongside a utility measure in the happiness domain (life satisfaction). The reason is that, when stratification theorists speak of the cumulativeness of inequalities in these three domains, they presumably have in mind actual ('objective') income and health (rather than subjective income satisfaction and health satisfaction) but subjective happiness.

Figure 3: Preliminary Assessment of Stratification Theory: German Panel Data[a]



Note: a) Coefficients are standardised (Betas) and are net of gender and age. Estimates not significant at the .05 level have been omitted.

Correlations[a]

		(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
Father's occupational prestige	(1)	1.00	.76	.40	.20	-.07	.01	.21	-.06	.01	-.04	-.10
Father's education	(2)	.62	1.00	.48	.26	-.03	.01	.30	.04	.01	-.03	-.14
Own education	(3)	.33	.39	1.00	.38	-.10	.08	.43	-.10	.07	-.24	-.12
Equiv. income 84	(4)	.17	.21	.31	1.00	-.03	.19	.62	-.01	-.11	-.06	.08
Health 84	(5)	-.05	-.02	-.07	-.02	1.00	-.22	-.06	.41	-.17	.15	.23
Happiness 84	(6)	.01	.01	.06	.14	-.14	1.00	.08	-.08	.47	.01	.05
Equiv. income 91	(7)	.17	.24	.35	.50	-.04	.06	1.00	-.07	.17	-.12	-.07
Health 91	(8)	-.04	-.03	-.07	-.01	.26	-.06	-.05	1.00	-.33	.13	.24
Happiness 91	(9)	.01	.01	.05	.08	-.11	.30	.12	-.21	1.00	-.01	-.03
Gender	(10)	-.04	-.03	-.22	-.05	.12	.01	-.11	.10	-.01	1.00	.10
Age	(11)	-.09	-.13	-.11	.07	.18	.04	-.06	.19	-.02	.10	1.00

a) Observed correlations are below the diagonal; disattenuated (estimated 'true') correlations are above. Disattenuated correlations were obtained by assuming validity coefficients of 0.9 for occupational prestige, years of education and income, and 0.8 for health and happiness (Jencks, 1979). The control variables, gender and age, were assumed to have validities of 1.00.

Table 1: Total Effect of Family Social Background on Income, Health and Happiness[a]

	Effects of father's occupation and education	Effects of father's occupation and education and own education
$Income_{84}$	.21	.33
$Income_{91}$	.24	.36
$Health_{84}$	.05	.08
$Health_{91}$	.04	.08
$Happiness_{84}$	.01	.07
$Happiness_{91}$	.01	.05

Note: a) For the method of calculating total effects see Alwin and Hauser (1975)

is a consequence of family background rather than individual achievement. This appears a somewhat tendentious viewpoint but, as Table 1 shows, it does not greatly affect our conclusions.

Depending on whether or not respondent's own education is included, family background accounts for four to 12 per cent of the variance in equivalent income, at most one per cent of the variance in health and virtually no variance in happiness (life satisfaction).

An additional piece of analysis in the health field involved examining relationships between indicators of parents' status and mortality. It is known that lower status people are at considerably greater risk of dying from particular causes (e.g. respiratory problems and road trauma), but there are 'diseases of affluence' as well as 'diseases of poverty', so that overall there is only a weak-to-moderate relationship between social status and mortality (DHSS, 1980; McMichael, 1985; Travers and Richardson, 1993). In the German panel the correlation between father's occupational status and age of death was only 0.04 and between father's years of education and year of death the correlation was 0.02. The correlations between these indicators of status and mother's age of death were, respectively, 0.03 and 0.04.[4]

In examining relationships between a person's own status (as distinct from parental status) and health, one must allow for the possibility that health could contribute to status, as well as vice-versa. In particular, it has been shown that good health is associated with upward social mobility (Fogelman, Fox and Power, 1989).

4 Only respondents aged 50 and over, most of whose parents had died, were included in these analyses. Living parents were given an expected age of death, based on their sex and current age.

A further point of interest in the results in Figure 3, which bears on the validity of stratification theory, is how little variance in income, health and happiness in 1991 is accounted for by the combination of family background and prior (ie. 1984) measures of these outcomes. The adjusted R^2s for the 1991 measures were just 29.4 per cent, 8.9 per cent and 9.6 per cent. Plainly, individual attributes and life experiences which do not rate a mention in stratification theory account for most of the variance.

3.2 The Income/Standard of Living Domain

The aims of this section are first to provide more detailed evidence of the volatility of household standards of living; the extent to which inequality of equivalent incomes is or is not persistent over time. We then test alternative theories of income dynamics, concluding that a dynamic equilibrium theory which views current income flows and utility (income satisfaction) as being determined by human and material capital best fits the data.

Further Evidence of Income Volatility

To what extent, then, is income stratification long term? In fact, since we only have eight years of panel data, we can only answer the question, 'Are incomes stratified in the medium term?' However, in so far as it transpires that incomes are not highly stratified in the medium term, the a fortiori argument can be made that there is unlikely to be rigid stratification in the longer term. Plainly this argument is not foolproof, but it will be bolstered by showing that even among males of prime working age, there is considerable income volatility. Surely, if members of this group have incomes which are fairly unstable relative to each other, then a 'strong' theory of income stratification - one which implies that people are stratified into narrow income bands - begins to look decidedly shaky.

Table 2 gives a cross-tabulation of equivalent income quintiles in 1984 and 1991.

This matrix answers the question, 'What had happened by 1991 to people starting in different income quintiles in 1984?' The bold figures on the top left to bottom right diagonal show the percentages who remained in the same quintile as they started. Of those who started in the poorest quintile, 43.6 per cent remained there, and 55.1 per cent of those who started in the top quintile stayed put. In the middle three quintiles only 26 to 31 per cent stayed put. The appearance of greater stability in the top and bottom quintiles is, however, misleading. People in these quintiles can only move in one direction, while people in the middle quintiles can shift either way. In fact, if one looks at the percentage change in 1984-91 incomes of people in each quintile, one finds the biggest change and the highest standard deviation of change among people who started in the poorest quintile. The people starting in the top quintile

Table 2: Transition Matrix of Equivalent Income Quintiles 1984-91 (N = 5000)

	Quintiles 1991					
Quintiles 1984	Q1 %	Q2 %	Q3 %	Q4 %	Q5 %	%
Q1	**43.6**	25.5	15.1	10.8	5.1	20
Q2	28.0	**28.9**	19.4	14.0	9.8	20
Q3	12.0	27.0	**26.4**	22.1	12.5	20
Q4	10.2	12.8	27.9	**31.3**	17.8	20
Q5	5.2	7.6	10.6	21.6	**55.1**	20

recorded the smallest income increase with a similar standard deviation to those in the third and fourth quintiles.[5]

The results show a considerable degree of income volatility, perhaps no surprise to economists who specialise in income dynamics and who expect considerable regression to the mean, but quite strongly counter to mainstream sociological theory. If we focus on households starting in the poorest quintile, we find that 31.0 per cent had moved up the distribution by two quintiles or more after seven years. At the other end of the distribution, 23.4 per cent of those starting in the top quintile moved down by two quintiles or more.

Very similar quintile transition matrices have been produced for the United States, based on the University of Michigan Panel Study of Income Dynamics (Duncan et al., 1984). So it can be concluded that for two Western societies with very different political, economic and welfare state frameworks, income stratification is far from rigid.

The Michigan and German research both show that the main reasons for income changes are major family events: marital break-up and repartnering, birth and departure of children, and changes in family labour force participation (Burkhauser et al., 1990; Duncan et al., 1984; Hauser and Semrau, 1989). However, changes in the head of household's earnings also have considerable impact (Bane and Ellwood, 1986).

5 The means and standard deviations of change in income, after adjustment for inflation, were

	Mean	**Standard Deviation**
Q1	+90.5%	105.9%
Q2	+42.5%	57.4%
Q3	+27.2%	44.9%
Q4	+10.5%	39.0%
Q5	+ 2.8%	45.1%

We have suggested that equivalent income provides the best measure of material standard of living. However, if we use gross earnings as our measure and restrict the sample to males aged 30-55 in continuous work, we get a picture of less volatility. The correlation between the log of gross earnings in 1984 and 1991 was 0.72, compared with 0.50 for equivalent incomes.[6] There was also, as expected, less movement among quintiles (see Table 3).

Stability of quintile positions ranged from 65.2 per cent in the top quintile to 38.6 per cent in the middle. Even so, it is worth noting that 22.8 per cent of men starting in the lowest quintile moved up by two quintiles or more, and 11.1 per cent of those starting in the top quintile moved down by at least two quintiles.

Medium Term Income Inequality is Considerably Less than Annual Inequality

Partly because of movement up and down the income distribution by individuals and households from year to year, income inequality in the medium term is considerably less than inequality at one moment in time. This result is well known to economists who have undertaken simulations showing that lifetime household disposable income inequality is very much lower than annual figures might suggest. In Britain, for instance, the ratio of the top decile's to the bottom decile's annual disposable income is about 10:1, whereas the ratio of lifetime disposable income is estimated at about 3:1 (Falkingham, Hills and Lessof, 1993). Results like these suggest that, in a broad sense, the welfare state achieves redistributive objectives (Kakwani, 1986; Ringen, 1987), contrary to what many sociologists, basing their accounts on annual not long term data, appear to believe.

The German panel data tell the same story as the simulation results. Column (2) of Table 4 shows Gini coefficients of household equivalent income calculated for one year (1984), two years (1984-85) and up to eight years (1984-91). Column (1) gives results for **household** gross earnings, which may be regarded as a proxy for what incomes and standards of living would have been in the absence of government taxes and benefits.[7] The figures were calculated by summing the earnings of each household's members and dividing by the household's equivalence score.[8] By comparing columns (1) and (2) we can see the impact of government ('the welfare state') on income distribution. Column (3) provides a percentage estimate of this impact.

6 People reporting zero or negative incomes were excluded from all calculations here as elsewhere in the paper.

7 But see Ringen (1987) Appendix F for discussion of the weakness of this counterfactual.

8 This is the only calculation in the paper in which people with zero income are included. This makes sense when estimating the redistributive impact of government, since many households genuinely have zero earnings and are supported by the state.

Table 3: Transition Matrix of Gross Earnings: Males Aged 30-55 Continuously in Work 1984-91 (N=1054)

	Quintiles 1991					
Quintiles 1984	Q1 %	Q2 %	Q3 %	Q4 %	Q5 %	%
Q1	**60.8**	16.5	16.3	2.8	3.7	20
Q2	31.7	**41.1**	12.1	5.8	3.3	20
Q3	4.8	30.6	**38.6**	23.2	2.8	20
Q4	2.0	5.4	27.2	**45.9**	19.5	20
Q5	0.7	1.9	8.5	23.7	**65.2**	20

It can be seen that, partly because people's relative incomes fluctuate from year to year, income inequality diminishes the longer the period considered. In Germany it appears that inequality of household equivalent earnings reaches an asymptote after 6 years (Gini = 0.426)

The impact of the German state in reducing income inequality is very considerable, especially in the medium term. In a one year period (1984) household equivalent disposable incomes are 43.5 per cent more equal than household equivalent gross earnings. Summing over 8 years (1984-91), the estimated reduction in inequality due to the state is 48.8 per cent. It is interesting that there is no clear indication after 8 years that the increasing reduction in inequality approaches an asymptote.

It is not clear whether the results given here are in line with microsimulation estimates, which also show increasing equality over longer time periods, but which have in some cases indicated that the redistributive impact of the state is less not greater on a lifetime basis than an annual basis (Falkingham et al., 1993; Galler and Wagner, 1986; Harding, 1993; Merz, 1988; Orcutt, 1957).

It must be conceded, however, that although the estimated trend is probably correct, these figures somewhat **overstate** gross income equality. The gross income figures relate only to earnings and thus exclude income from capital, self-employment and private transfers.[9] However, it can be argued that the inclusion of superannuation payments as transfers, when they are partly returns on private investment, overstates the redistributive impact of the state.

9 The GSOEP collects data on income from capital and private transfers, so more accurate estimates of household gross incomes will eventually be made.

Table 4: Income Inequality Diminishes as Time Passes and this is Largely Due to the State: German Panel Results 1984-91

	Gini Coefficients		
Time period	Household equivalent gross earnings[b] (1)	Household equivalent disposable incomes (2)	Reduction in inequality due to the State[a] (3)
			%
1 year (1984)	.474	.268	43.5
2 years (1984-85)	.450	.247	45.1
3 years (1984-86)	.436	.234	46.3
4 years (1984-87)	.430	.231	46.3
5 years (1984-88)	.431	.227	47.3
6 years (1984-89)	.426	.224	47.4
7 years (1984-90)	.426	.222	47.9
8 years (1984-91)	.426	.218	48.8

Note: a) Estimated effect of State action = 100 x (household gross Gini-household net Gini)/household gross Gini (Kakwani, 1986; Ringen, 1987).

b) $ incomes rather than log incomes were used for calculations in this table.

Are There Long Term Poor and Long Term Rich?

Despite considerable fluctuations of gross and disposable incomes in the population as a whole, it is often claimed that a fairly substantial percentage is locked into poverty or near-poverty. Claims about the 'new poverty', 'the two-thirds society' and 'the culture of contentment' rest on the belief that persistently high unemployment in the 1980s and 1990s has left many people in more or less continuous poverty (Balsen et al., 1984). The 'two-thirds society' thesis claims that up to a third of people in West European countries may be regarded as persistently poor (Glotz, 1984).

Detailed research based on the Michigan panel data has given only slight support to these claims, although long term poverty may have increased somewhat in the 1980s and 1990s (Duncan et al., 1984; Duncan et al., 1991). The German panel data indicate that in the 1984-89 period about 10 per cent of people were below 60 per cent of median equivalent household income for more than half the time, and thus may be regarded as having been persistently poor (Headey, Krause and Habich 1993). Furthermore, spell analysis showed that 60 per cent of people who became poor ceased to be so within two years.

Less research appears to have been done on the long term rich (or long term high income people) than the poor. Preliminary results from the German panel study indicate that, generally speaking, high income is nearly as transient as poverty income. Few people manage to stay rich for long periods; typical 'spells' seem to last two to three years. Such conclusions, however, need to be tempered by recognition of less-than-adequate data on wealth.

Theories of Income Dynamics

We have seen that in an 8 year period there is considerable movement up and down the income distribution. Does this mean that household income is more or less a random walk, so that if a long enough period of data were available, there would be no systematic relationship between the incomes of different households? Surprisingly, some students of income distribution appear to think so. There appear to be three main sets of theories (Lydall, 1976).

- Stochastic theories which postulate that in the long run income distribution is a random walk.
- Human capital theories which postulate that income depends on education and job training (Becker, 1964; Friedman, 1957; Mincer, 1974).
- Multivariate theories which attribute income differences to a range of variables usually including human capital, material capital (wealth or net assets), gender, intelligence and so forth.

An obvious objection to stochastic theories is that they are not really theories at all. To say (for example) that a large number of small multiplicative random effects could account for the fact that income distributions tend to be approximately lognormal is not really an explanation. However, at least in the case of the German panel data, initial inspection suggest some empirical support for the most obvious and common prediction of stochastic theory, namely that correlations between income at time 1 and later points in time (time 2, time 3 etc.) will progressively and evenly diminish in a sequence suggesting a Markov chain. Table 5 gives correlations between income at time 1 (1984) and subsequent years for household equivalent income (all households) and individual earnings (men aged 30-55 in continuous employment).

The equivalent income column in Table 5 suggests a first order Markov chain, while the results in the earnings column are somewhat ambiguous; the correlations generally tend to decline but by no means evenly. However, there is no need to flirt with a stochastic 'theory' if a more coherent explanation can be found.

In assessing the relative merits of human capital theory and a less parsimonious multifactorial theory, it immediately became clear that we could account for more variance in both flows of income and utility (income satisfaction) if measures of

Table 5: Correlations of Income Over Time

	Equivalent income	Individual earnings
1984-85	0.70	0.79
1984-86	0.66	0.75
1984-87	0.64	0.78
1984-88	0.59	0.75
1984-89	0.55	0.71
1984-90	0.53	0.65
1984-91	0.50	0.72

material capital (and also gender and age) were included. In what follows we propose a dynamic equilibrium model in which income flows and utility in 1985-91 are viewed as being constrained by stocks measured in 1984.

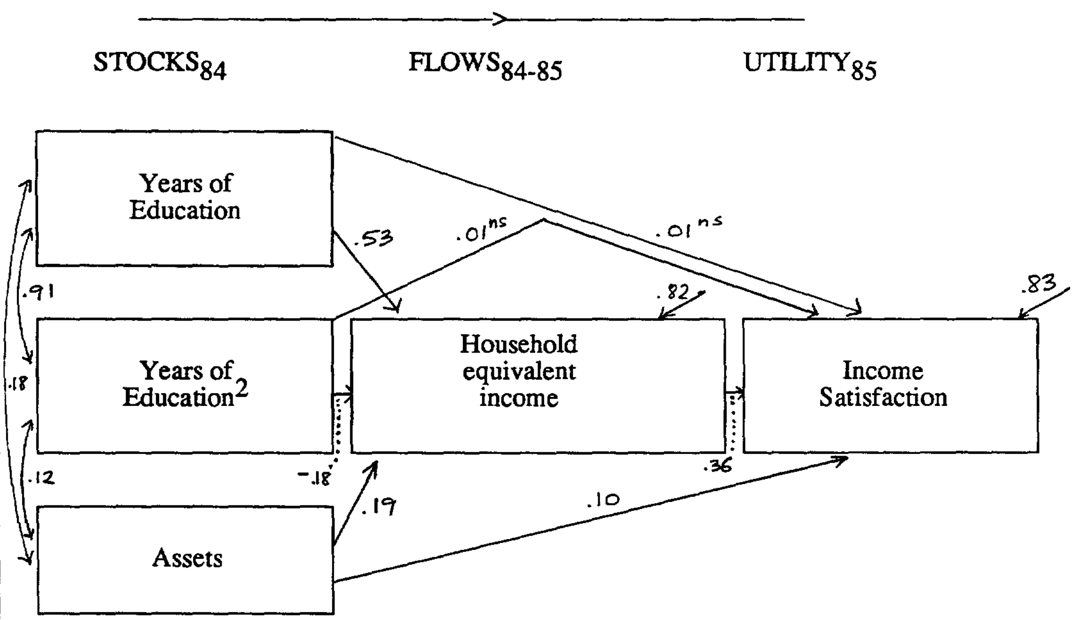
The reasoning underlying the model is as follows. It is assumed that in his/her twenties a person is engaged in stock-building. In Germany education and job training usually continue during this period of life and people may also acquire assets like a savings account, a home, shares or a business. By the age of thirty it is hypothesised that initial stock-building is complete. The model then proposes that one's subsequent income and utility depend on these initial stocks. It should be remembered that equivalent income comprises both market income and goverment benefits (minus taxes). So in proposing that initial stocks constrain future income, we hypothesise that **people with the highest stock levels are best at playing both the market and government systems to maximise their equivalent incomes (real standard of living). People with low stock levels lack the resources to play these systems effectively**.

In line with this reasoning, the sample used to test the dynamic equilibrium model (Figure 4) comprises Germans aged 30 and over in 1984. People under 30 are excluded.

The key test of the model involves assessing whether (a) 1984 stocks account for substantial variance in income flows and utility 1985-91 and (b) more importantly, whether 1984 stocks predict a **constant range** within which subsequent incomes and utility fall. Thus if 1984 stocks predict incomes and utility in later years (i.e. 1990-91) just as well as they predict these outcomes in earlier years (i.e. 1985-86), we can infer that, at least in the medium term, outcomes are in dynamic equilibrium due to initial stock levels. If, on the other hand, outcomes at later time points are less and less strongly correlated with initial stocks, the inference would be that the dynamic equilibrium model is incorrect (or at least that the wrong stocks were identified).

Before giving longitudinal results, it may be useful to give a static model. Figure 4 gives OLS regression equations and Pearson correlations indicating relationships between stocks in 1984 and flows and utility (income satisfaction) in 1985.

Figure 4: Living Standards: Stocks, Flows and Utility (N = 3700)[(a)]



Notes: a) Estimates are standardised (see next page for metric estimates) and net of gender, age and age squared. Years of education squared and age squared are included to take account of non-linearities in the relationships of both education and age with income.

ns Not significant at the .05 level.

There are moderate correlations between equivalent income and both years of education (r = 0.32) and assets (r = 0.24). An equation including these variables plus sex and age accounts for 20.2 per cent of the variance in income. Income satisfaction is largely accounted for by actual income (r = 0.38), education (r = 0.15) and assets (r = 0.15). An equation including these latter three variables, plus gender and age, accounts for 16.9 per cent of variance in income satisfaction.

We now come to the main test of the dynamic equilibrium model. Do initial stocks (1984) predict (a) income and (b) utility in later years as well as they predict these outcomes in 1985, or do the predictions get worse over time? To answer these questions we first calculated the 95 per cent confidence interval for the multiple correlation obtained by regressing Equivalent $Income_{85}$ on $Stocks_{84}$, and also the

The following are the OLS metric equations related to the results in Figure 4.

Metric Equations

Equation 1.1

$$\text{Equivalent Income} = 5.35 + .10\text{YrsEduc} - .01\text{YrsEduc}^2 + .08\text{Assets} + .05\text{Gender} + .01\text{Age}^{ns} - .00\text{Age}^2$$

(.12) (.01) (.00) (.01) (.02) (.00) (.00)

R = 0.43

Equation 1.2

$$\text{Income Satisfaction} = -7.46 + 1.85\text{Equiv.Inc.} + .02^{ns}\text{YrsEduc} - .00^{ns}\text{YrsEduc}^2 + .22\text{Assets} + .25\ \text{Gender}$$

(1.00) (.12) (.05) (.01) (.05) (.11)

$$-.01^{ns}\text{Age} + .00\text{Age}^2$$

(1.00) (.00)

R = .41

ns = not significant at the .05 level

Correlations[(a)]

		(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Equiv income (*ln*)	(1)	1.00	.53	.40	.30	.30	-.08	-.01	-.03
Inc. satis	(2)	.38	1.00	.21	.18	.21	.03	.09	.11
Yrs educ.	(3)	.32	.15	1.00	1.00	.20	-.31	-.20	-.13
Yrs educ^2	(4)	.24	.13	.91	1.00	.15	-.20	-.14	-.09
Assets	(5)	.24	.15	.18	.12	1.00	-.11	-.32	-.33
Gender(M = 1, F = 2)	(6)	-.07	.02	-.28	-.18	-.10	1.00	.12	.11
Age	(7)	.01	.07	-.18	-.13	-.29	.12	1.00	.89
Age^2	(8)	-.03	.09	-.12	-.08	-.30	.11	.89	1.00

Note: a) Observed correlations are below the diagonal and disattenuated correlations above. The valid variances assumed for disattenuating were: 0.9 for equivalent income, education and assets, 0.8 for income satisfaction, and 1.0 for gender and age.

confidence interval for Income Satisfaction_{85} regressed on Stocks_{84}.[10] The first of these intervals was 0.43 (±.03) and the second was 0.25 (±.03). Next, we obtained the (six) regressions of Income_{86} to Income_{91} on Stocks_{84}, and the regressions of Income Satisfaction_{86} to Income Satisfaction_{91} on Stocks_{84}. The key test was to see whether the multiple correlations for these later years fell within the confidence intervals calculated for 1985.

The results in Table 6 (column 1) indicate that initial stocks do constrain (relative) incomes within a fixed range. All six predicted correlations for the years 1986-91 fall within the required confidence interval of 0.43 (±0.03).[11] The results relating to utility (column 2) are less decisive. There appears to be a slight decline over the years in the relationship between Stocks_{84} and $\text{Utility}_{86\text{-}91}$. However, three of six predicted correlations (1986-91) fall within the 95 per cent confidence interval of 0.25 (±0.03).

It is also worth noting that the magnitudes of the regression coefficients of the 1984 stock variables scarcely change when they are used to predict income and income satisfaction in later years. In particular, the coefficients for Years of Educations_{84} and Assets_{84} were much the same in equations for predicting outcomes in 1986-91 as they had been in 1985.

Another method of expressing the results relating to income is first to divide people into quintiles on the basis of their stocks in 1984[12] and also divide them into income quintiles for 1985-91. Table 7 shows the percentage of people who were in the bottom quintile of stocks in 1984 and who were also in the bottom two quintiles of income in 1985-91. Similarly, Table 8 shows the percentages in the top quintile of stocks in 1984 who were in the top two quintiles of income in subsequent years.

It should be clear what is not claimed as well as what is claimed here. Initial stock levels are **not** highly predictive of subsequent income but they do constrain relative incomes within the same range throughout 1985-91. Table 7 indicates that, of the people in the bottom quintile of stocks in 1984, 34.6 per cent were in the bottom quintile of income in 1985 and 60.2 per cent were in the bottom two quintiles combined. By 1991, 57.7 per cent of those in the lowest quintile of Stocks_{84} were still in the bottom two income quintiles. At the top of the income distribution the picture is similar; 63.6 per cent of those in the top quintile of Stocks_{84} were in the top two income quintiles in 1985 and 59.9 per cent were in these quintiles in 1991.

10 Confidence intervals were obtained by calculating sheaf coefficients for the stock variables as predictors of (a) Equivalent Income_{85} and (b) Income Satisfaction_{85} (Heise, 1972). It was then straightforward to regress the outcome variables on their respective sheafs and get confidence intervals.

11 The first of the seven correlations in each column of Table 6 is of course a benchmark and not a 'prediction' in the sense used here.

12 This was done by means of sheaf coefficients (see note 10).

Table 6: Impact of Stocks in 1984 on Equivalent Income Flows and Utility 1985-91
(N = 3700)

$Flows_{85\text{-}91}$	(1) Multiple Rs for Income	$Utility_{85\text{-}91}$	(2) Mutliple Rs for Income Satis
$Income_{85}$	.43	Income $Satis_{85}$	.25
$Income_{86}$	.42	Income $Satis_{86}$	.23
$Income_{87}$	.43	Income $Satis_{87}$	.22
$Income_{88}$	.41	Income $Satis_{88}$	.19
$Income_{89}$	.43	Income $Satis_{89}$	.23
$Income_{90}$	.41	Income $Satis_{90}$	.21
$Income_{91}$	.44	Income $Satis_{91}$	.20
Confidence Interval	= .43 (±.03)	Confidence Interval =	.25 (±.03)

Table 7: Percentage of Respondents in Bottom Quintile of $Stocks_{84}$ who were in Lowest 2 Quintiles of Equivalent Income 1985-91

	% who were in		
	1st Quintile of income	2nd Quintile of income	1st and 2nd Quintiles of income
1985	34.6	25.6	60.2
1986	31.9	25.3	57.2
1987	33.8	22.4	56.2
1988	32.5	24.8	57.3
1989	32.4	24.6	57.0
1990	32.7	23.4	56.1
1991	34.0	23.7	57.7

Exactly the same analyses as those just described were also undertaken using the more homogeneous sample of German males of prime working age (30-55 years) who were in continuous employment throughout the period. The dependent variable here is individual gross earnings (Table 9).

Clearly, there is no tendency for these multiple correlations to diminish over time. All correlations fall well within the 95 per cent confidence interval of 0.59 (±0.06).

Tables 10 and 11 confirm these results by showing that a more or less constant proportion of men starting in the bottom and top quintiles of $Stocks_{84}$ were, respectively, in the lowest and highest two quintiles of earnings in 1985-91.

Table 8: Percentage of Respondents in Top Quintile of $Stocks_{84}$ Who Were in Top 2 Quintiles of Equivalent Income 1985-91

	% who were in ..... Top Quintile of income	4th Quintile of income	4th and 5th Quintiles of income
1985	42.4	21.4	63.6
1986	39.8	22.8	62.6
1987	40.7	22.3	63.0
1988	38.8	22.7	61.5
1989	40.1	22.0	62.1
1990	49.1	20.9	60.0
1991	39.6	20.3	59.9

Table 9: Impact of Stocks in 1984 on Earnings 1985-91 (N=742)

Flows		Multiple Rs for Earnings
$Earnings_{85}$		0.59
$Earnings_{86}$		0.60
$Earnings_{87}$		0.61
$Earnings_{88}$		0.60
$Earnings_{89}$		0.61
$Earnings_{90}$		0.57
$Earnings_{91}$		0.59
Confidence Interval	=	0.59 (±0.06)

Table 10: Percentage of Individuals in Bottom Quintile of $Stocks_{84}$ Who Were in Lowest 2 Quintiles of Earnings, 1985-91

	% who were in ..... Bottom Quintile of earnings	2nd Quintile of earnings	1st and 2nd Quintiles of earnings
1985	33.5	24.2	57.7
1986	33.1	24.8	57.9
1987	35.4	25.3	60.7
1988	32.7	28.2	60.9
1989	35.6	23.1	58.7
1990	33.5	27.8	61.3
1991	37.3	20.9	58.2

Table 11: Percentage of Individuals in Top Quintile of Stocks$_{84}$ Who Were in Top 2 Quintiles of Earnings, 1985-91

	% who were in ..... Top Quintile	4th Quintile of earnings	4th and 5th Quintiles of earnings
1985	45.6%	24.2%	69.8%
1986	44.8%	26.0%	70.8%
1987	47.7%	26.5%	74.2%
1988	46.7%	26.7%	73.4%
1989	48.1%	23.4%	71.5%
1990	44.2%	26.3%	70.5%
1991	45.1%	26.1%	71.2%

Taken together, the results relating to household equivalent income and individual earnings provide fairly strong confirmation of a dynamic equilibrium model in which income flows and (less certainly) utility in subsequent years are predicted on the basis of initial stocks of human and material capital. The key test of the predictive capacity of initial stocks is that they correlate approximately as highly with income and utility (income satisfaction) in 1991 as in earlier years.

The results given in this section on incomes/standard of living may be summarised as follows: (1) household and individual incomes are not rigidly stratified over time but nor are they a random walk; (2) the degree of volatility in household incomes is mainly due to family events, especially changes in family labour force participation, the birth and departure of children, and marital break-up and repartnering; (3) the degree of stability in correlations of income over time is largely due to stocks of human and material capital and only to a small extent to family social background.

3.3 The Happiness Domain

We now assess a dynamic equilibrium model in the happiness domain. The relevant stocks in this domain are the personality traits of extraversion and neuroticism, the flows are favourable (satisfying) and adverse (distressing) life events, and utility is measured by self-reports of life satisfaction. In this domain (in contrast to the income/standard of living domain), the main interest lies in the relationship between initial stocks and subsequent utility. This is what we will focus on; only limited attention will be given to the relationship between stocks and life events (but see Headey and Wearing, 1989).

We begin by showing that, whereas the personality traits of extraversion and neuroticism are very stable over time, there is a moderate degree of change in reports of life satisfaction. We then present static and dynamic models linking stocks to flows and utility, again constructing sheaf coefficients of initial stocks to see whether people's subsequent flows and utility are constrained within predictable bounds. The data source is the 5-wave Victorian Quality of Life Panel (1981-89).

Happiness is Just Moderately Stable: Somewhat less than Income

Table 12 gives observed and disattenuated (estimated 'true') correlations between life satisfaction in 1981 and life satisfaction in later waves of the Australian panel.

It can be seen that these correlations tend to diminish over time, suggesting a moderate degree of change in levels of life satisfaction. There is, however, a hint that the relationship between the time 1 (1981) measure and subsequent measures may be reaching an asymptote after about five years, but with only five waves of data, collected over eight years, one cannot be sure. It is of some interest that the estimated 'true' correlation of life satisfaction (or happiness) over 8 years is not a great deal less than the correlation of income over the same period (0.49 compared with 0.62). Critics of attempts to measure happiness sometimes claim that such measures merely indicate current mood - not so (Diener et al., 1984; Headey and Wearing, 1992; Schwarz and Strack, 1991).

Table 13 confirms moderate shifts in life satisfaction 1981-89 with a quintile transition matrix similar to the more familiar income transition matrix given earlier.

As with the income transition matrix, stability appears to be higher in the top and bottom quintiles of life satisfaction, but this can be shown to be misleading by comparing mean changes in the scores of 1981 quintiles; that is, changes occurring by 1989. For the sample as a whole, mean life satisfaction scores on the 0-10 scale changed only slightly in the decade , moving in the range between 7.0 and 7.5. The least satisfied quintile in 1981 showed a 0.74 increase in satisfaction, the second quintile reported a 0.19 decline, the third quintile a 0.46 decline, the fourth quintile a 1.01 decline and the top quintile in 1981 showed a decline of 1.39 by 1989.

Readers unfamiliar with life satisfaction data may like to note that in Australia, as in all Western countries, a substantial majority of people record satisfaction levels above the mid-point of the scale, indicating that they are more satisfied than dissatisfied with their lives. Naturally, there has been considerably controversy about the validity of this result, but mood data reported at random intervals in response to a random buzzer appear to confirm it (Diener and Larsen, 1984).

Personality 'Stocks' Constrain Subsequent Happiness

The personality traits (stocks) of extraversion and neuroticism have been shown to be highly stable over lifetimes and, indeed, are substantially hereditary (Costa and McCrae, 1984; Costa et al., 1983). In the Australian panel the estimated stability of extraversion 1981-89 was 0.99 and of neuroticism 0.82.

Table 12: Correlations Among Life Satisfaction Measures Over Time

	Correlations Life Satisfaction$_{81}$ with...	
	Observed	Disattenuated[a]
Life Sat$_{83}$	0.64	0.70
Life Sat$_{85}$	0.47	0.49
Life Sat$_{87}$	0.44	0.45
Life Sat$_{89}$	0.47	0.49

Note: a) Disattenuated r = Observed r/Cronbach alpha - Correlated error (Andrews and McKennell, 1980; Headey and Wearing, 1992) i.e. Disattenuated r = observed r/0.80-0.1

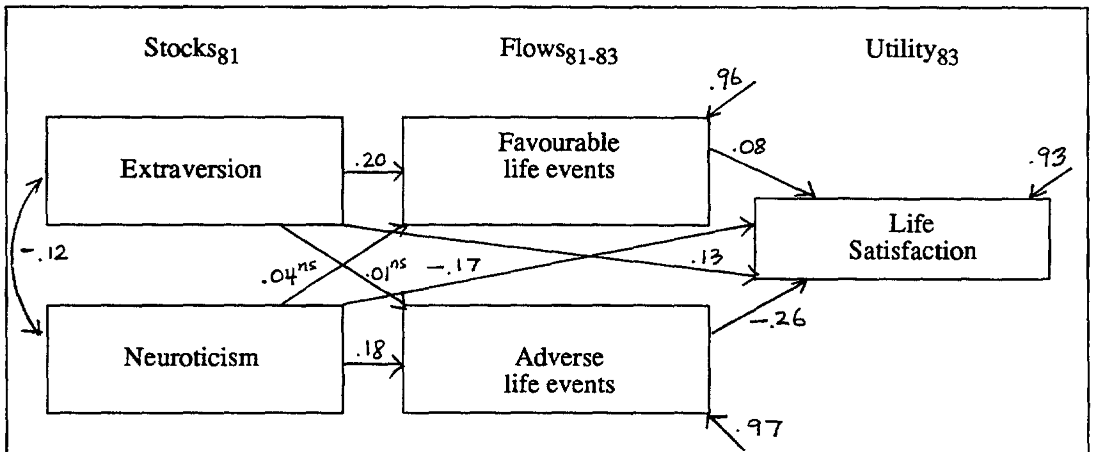
Table 13: Life Satisfaction: Quintile Transition Matrix 1981-89

Life Satisfaction 1981	Life Satisfaction 1989 Q1 %	Q2 %	Q3 %	Q4 %	Q5 %	%
Q1	**50.0**	27.3	12.1	7.6	3.0	20
Q2	28.8	**26.3**	23.8	12.5	8.8	20
Q3	10.5	31.4	**20.9**	29.1	8.1	20
Q4	12.0	14.7	26.7	**22.7**	24.0	20
Q5	10.8	8.4	12.0	16.9	**51.8**	20

Figure 5 shows OLS estimates of relationships between extraversion and neuroticism measured in 1981, favourable and adverse life events 1981-83, and life satisfaction 1983.

It can be seen that extraversion and neuroticism measured in 1981 account for a modest amount of variance in favourable and adverse life events 1981-83 (R=0.20 and R=0.18 respectively), and that personality traits and life events combined account for rather more variance in 1983 levels of life satisfaction (R=0.38). However, even though only small amounts of variance in life events are statistically accounted for, it is of considerable interest that initial personality traits partly explain the kinds of life events and experiences that subsequently happen to people. People who were assessed as relatively extraverted in 1981 persistently reported lots of favourable friendship and also job related events in the period 1981-89. They reported job promotions, increases in job related skills, praise from their boss, and also claimed to have made lots of new friends and joined lots of clubs and

Figure 5: Happiness: Stocks, Flows and Utility[(a)] (N=502)



Notes: a) Standardided OLS estimates (Betas)
n.s. Not significant at the .05 level.

Metric equations

Eq. 2.1 Fav. $\text{Events}_{81\text{-}83} = 1.00 + .14\ \text{Extraversion}_{81} + .02^{ns}\ \text{Neuroticism}_{81}$
(.44) (.03) (.02)
R=0.20

Eq. 2.2 Adv. $\text{Events}_{81\text{-}83} = 1.49 + .01^{ns}\ \text{Extraversion}_{81} + .08\ \text{Neuroticism}$
(.41) (.02) (.02)
R=0.18

Eq. 2.3 Life $\text{Sat}_{83} = 7.58 + .05\ \text{Extraversion} - .05\ \text{Neuroticism} + .05\ \text{Fav. Events}_{81\text{-}83}$
(.26)(.02) (.01) (.02)

$-.17$ Adv. Events
(.02)
R=0.38

Correlations[(a)]

		(1)	(2)	(3)	(4)	(5)
Extraversion_{81}	(1)	1.00	-.16	.27	$-.01^{ns}$	.25
Neuroticism_{81}	(2)	-.12	1.00	$.02^{ns}$	.20	-.28
Fav. $\text{Events}_{81\text{-}83}$	(3)	.20	$.02^{ns}$	1.00	.17	.08
Adv. $\text{Events}_{81\text{-}83}$	(4)	$-.01^{ns}$	.18	.16	1.00	0.33
Life Sat_{83}	(5)	.18	-.23	.07	-.28	1.00

Note: a) Observed correlations are below the diagonal; disattenuated (estimated 'true') correlations are above. Estimated validity coefficients: $\text{Extraversion}_{81} = 0.63$, $\text{Neuroticism}_{81} = 0.82$, Fav. Events81-83 and Adv. Events81-83 = 0.90, Life Satisfaction = 0.80.

organisations (Headey and Wearing, 1989, 1992). Similarly, but less fortunately, people who were assessed as relatively neurotic in 1981 reported lots of adverse events and experiences in 1981-89. In particular, they reported adverse experiences relating to their standard of living and jobs: continuous financial worries and crises, unemployment, and disputes with colleagues at work.

It is important to note that at least some of the experiences which appear to be personality-driven are sufficiently concrete or 'objective' (e.g. job promotions and unemployment) that people's reports that they occurred could scarcely be misperceptions or distortions coloured by personality. Thus it is not difficult to believe that an extraverted person might, due to personality, perceive that he/she had made lots of new friends in the same actual circumstances as a less extraverted person might have no such perception. But when an extravert persistently reports job promotions, this seems likely to be true. Similarly, a neurotic person might be biassed towards reporting financial worries, but getting the sack and being unemployed are more or less objective experiences (Headey and Wearing, 1989).

We now assess a dynamic equilibrium model of happiness in exactly the same way as the income model was tested (Table 14). Confidence Intervals (95 per cent) were calculated for the multiple correlations obtained by regressing (a) Favourable Events_{81-83} (b) Adverse Events_{81-83} and (c) Life Satisfaction_{83}, on Extraversion_{81} and Neuroticism_{81}. Extraversion_{81} and Neuroticism_{81} were then used to predict Life Satisfaction in later years and the test was to see whether later multiple correlations fell within the initial year's confidence intervals.

It can be seen that most of the multiple correlations are consistent with predictions generated by a dynamic equilibrium model. All predictions of life satisfaction, based on personality in 1981, fall within a 95 per cent confidence interval. So do all predictions of favourable life events 1983-1989 and two of three predictions of adverse life events 1983-1989. It seems reasonable to claim that if one knows a person's personality, at least with regard to extraversion and neuroticism, one can, within a broad range, predict their medium term future levels of happiness and even the types of life events that will happen to them.

As a further check on the dynamic equilibrium model for life satisfaction, Tables 15 and 16 show the quintile positions for Life Satisfaction 1983-89 of people who started in (a) the lowest quintile of sheaf scores for 1981 and (b) the highest quintile.

These results confirm that the capacity of the personality stocks of extraversion and neuroticism to constrain subsequent life satisfaction within a fixed (though fairly wide) range does not diminish over time.

In concluding this section we recall Coser et al.'s (1987) statement that, 'Sociological data show again and again that the chances of a successful "pursuit of happiness" vary a great deal from class to class'. We saw earlier that family social background (father's occupational prestige and education) has zero predictive power in accounting for happiness. A person's own status position also accounts for little

Table 14: Assessing a Dynamic Equilibrium Model: Do Personality Stocks$_{81}$ Constrain Subsequent Life Events and Happiness?

	Multiple Rs: Life Sat.	Multiple Rs: Fav. Events	Multiple Rs: Adv. Events
Life Sat$_{83}$	0.28	Fav.Events$_{81-83}$ 0.20	Adv.Events$_{81-83}$ 0.19
Life Sat$_{85}$	0.22	Fav.Events$_{83-85}$ 0.20	Adv.Events$_{83-85}$ 0.19
Life Sat$_{87}$	0.24	Fav.Events$_{85-87}$ 0.16	Adv.Events$_{85-87}$ 0.11
Life Sat$_{89}$	0.30	Fav.Events$_{87-89}$ 0.18	Adv.Events$_{87-89}$ 0.20

Confidence Interval (C.I)= 0.28 ($\pm$.07) C.I.=0.20 ($\pm$.07) C.I.=0.18 ($\pm$.07)

Table 15: Percentage of People in the Lowest Quintile of Stock Scores in 1981 Who Were in the Lowest 2 Quintiles of Life Satisfaction in Later Years

Life Satisfaction	Lowest Sheaf Quintile	% Who Were In 2nd Sheaf Quintile	1st and 2nd Quintiles
Life Sat$_{83}$	38.8	20.2	54.0
Life Sat$_{85}$	36.6	16.7	53.5
Life Sat$_{87}$	28.3	23.9	52.2
Life Sat$_{89}$	35.5	30.3	65.8

Table 16: Percentage of People in the Top Quintile of Stock Scores in 1981 who were in Top 2 Quintiles of Life Satisfaction in Later Years

Life Satisfaction	Top Sheaf Quintile	4th Sheaf Quintile	4th and 5th Quintiles
Life Sat$_{83}$	22.6	26.3	48.9
Life Sat$_{85}$	20.8	27.5	48.3
Life Sat$_{87}$	25.0	22.4	47.4
Life Sat$_{89}$	26.1	28.4	54.5

variance. In the Australian panel an index of socio-economic status (occupational status of the main breadwinner, gross household income and respondent's educational level) correlated only 0.10 with life satisfaction (Headey and Wearing, 1992). Similar results have emerged from all Western countries in which quality of life surveys have been conducted (Argyle, 1987; Veenhoven, 1984). There is simply no basis for the claim that happiness is substantially related to social class.

A model which views flows of life events and utility (life satisfaction) as being constrained by initial personality stocks has only modest capacity to account for variance (quite low R^2s) but does correctly predict the range within which later life satisfaction falls.

3.4 The Health Domain

In the health domain we have three measures of stocks, namely age of death of parents (a hereditary risk factor), a health impairment which interferes with daily activities (1-3 scale, where 3 represents a high degree of impairment) and a measure of chronic illness which has lasted more than a year (2 = chronic illness, 1 = no illness). Our flow measure, intended as a proxy for frequency and severity of health problems, is annual doctor visits (1984-91), and the measure of utility is health satisfaction (0-10 scale: 0 = totally dissatisfied, 10 = totally satisfied). The data source is the GSOEP.

The section begins by documenting the degree of stability/change in health flows (1984-91). We then give static and dynamic results, assessing the fit of a dynamic equilibrium model.

Considerable Change in Health Flows (doctor visits)

Table 17 gives observed and disattenuated correlations for annual doctor visits.

Clearly, even using the disattenuated results, there is considerable annual change in people's propensity or need to attend doctors. The correlations tend to diminish over time but not in any very consistent way, suggesting that there could be long term factors affecting current health.

Table 18 confirms this picture of just moderate stability with a quintile transition matrix similar to the income and life satisfaction matrices given earlier. The base year in 1984 and 1991 the concluding year.

It is clear that, although there is some tendency for people to remain in or close to the health quintiles they started in, there are also many people who become much healthier, and others who become much less healthy in a seven year period. For example, of those who visited doctors most in 1984, 21.1 per cent were in the two

Table 17: Annual Doctor Visits 1984-91[a]

	Correlations Observed	Disattenuated[b]
	Doctor Visits 1984 with...	
Doctor Visits$_{85}$	0.30	0.47
Doctor Visits$_{86}$	0.28	0.44
Doctor Visits$_{87}$	0.28	0.44
Doctor Visits$_{88}$	0.29	0.45
Doctor Visits$_{89}$	0.22	0.34
Doctor Visits$_{91}$	0.26	0.41

Notes: a) Data on doctor visits were not collected in 1990.
b) Disattenuated r = Observed r/Validity squared.

Table 18: Health (Doctor Visits) Quintile Transition Matrix 1984-91

Doctor Visits 1984	Doctor Visits 1991 Q1	Q2	Q3	Q4	Q5	
Q1	**38.8**	20.0	18.3	13.1	9.8	20
Q2	18.1	**29.3**	23.8	15.2	13.4	20
Q3	18.2	22.7	**23.8**	18.7	16.6	20
Q4	15.8	15.7	16.5	**27.1**	24.9	20
Q5	9.0	12.2	17.6	25.9	**35.2**	20

healthiest (or least doctor visiting) quintiles by 1991. In parallel fashion, of those who were in the healthiest quintile in 1984, 22.9 per cent were heavy doctor attenders (highest two quintiles) in 1991.

As expected, there was considerable regression-to-the mean in doctor visits 1984 to 1991. People who were in the lowest quintile of doctor visiting in 1984 recorded an average increase of 2.16 visits by 1991. People in the second quintile in 1984 recorded 1.83 more visits in 1991. In the third quintile the increase was 1.15 visits, and in the fourth 0.48 visits. People who most frequently attended doctors in 1984 recorded an average reduction of 6.57 visits by 1991.

Static Model

Preliminary analysis indicated that the stock variables 'age of death of father' and 'age of death of mother' were only very weakly correlated with health flows and utility; no correlation exceeded 0.05. These stocks were therefore excluded from analyses given below. Figure 6 shows a static model linking health stocks, flows and utility.

Health stocks (impairment of daily activities and chronic illness) account for a moderate amount of variance in doctor visits (R = 0.32), and stocks and flows combined account for rather more variance in health satisfaction (R = 0.53).

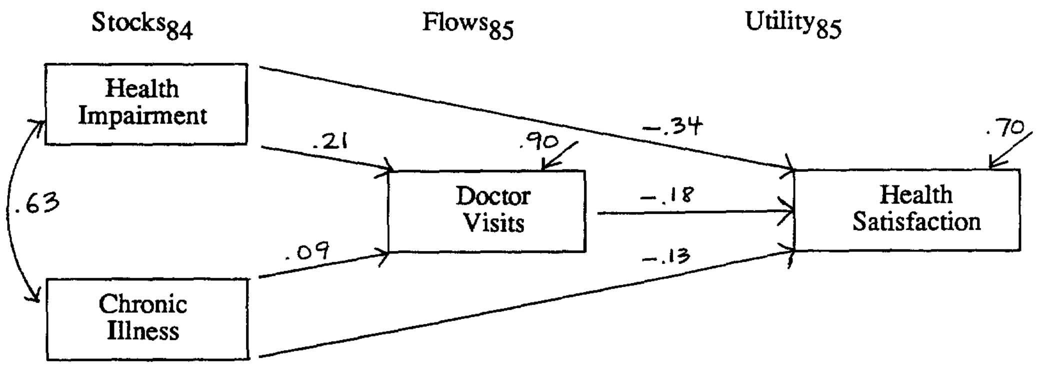
A dynamic equilibrium model is now assessed, using identical procedures to those used in the income and happiness sections. Confidence intervals (95 per cent) were calculated for the multiple correlations obtained from regressing (a) Doctor Visits$_{85}$ and (b) Health Satisfaction$_{85}$ on the two stock measures. Then a dynamic equilibrium model was tested by seeing whether predictions of flows and utilities in later years fell within the initial confidence intervals.

The results in Table 19 do not support the proposed dynamic equilibruim model, although it is possible that a slightly revised model would fit the data. Most correlations fall outside - but in the doctor visits column only just outside - the 95 per cent confidence interval. However, the relationship between Stocks$_{84}$ and subsequent flows and utilities appears to stabilise a year or two down the track. Thus Stocks$_{84}$ predicts annual doctor visits in the years 1986-91 with about equal accuracy, the correlations all falling within the range 0.27-0.29. Similarly, Stocks$_{84}$ predicts annual health satisfaction in the years 1987-91 with correlations in the range 0.45-0.46. The fact that both sets of correlations soon reach an asymptote suggests a possible amendment to the dynamic equilibrium model. It may be that there are temporary (as well as permanent stock-like) elements in our measures of health stocks and that these temporary elements produce the higher correlations between stocks in 1984 and flows and utilities in 1985-86 than are found for the years 1987-91. Thus in 1984, when stocks were measured, people may have had temporary health problems which led them to give higher 'readings' on the health impairment and chronic illness stock measures than they normally would. Conversely, some people may have been exceptionally fit in 1984 and thus given more favourable stock 'readings' than normal. These abnormal 'readings' would have the effect of temporarily inflating correlations between stocks and the indicators of flows and utility. Correlations would then stabilise (reach an asymptote) at a lower level in later years after these temporary health problems (or temporary abnormal fitness) had been resolved.

There is some evidence that our self-report stock measures are not as stable as perhaps they should be, or would be if medical reports were available. The correlations between Health Impairment$_{84}$ and Health Impairment$_{92}$ is only 0.53, while the correlation between Chronic Illness$_{84}$ and Chronic Illness$_{89}$ is 0.45.[13]

13. Correlations between 1984 and 1991 measures (given for other variables in this paper) are not available. Neither measure was included in the 1991 survey.

Figure 6: Health: Stocks, Flows and Utility[a] (N = 5000)



Note: a) Standardised OLS estimates. Estimates are net of gender and age.

Metric equations

Eq. 3.1 Doc. Visits85 = -1.67 + 1.03 Health Impairment
(.24) (.11)
+0.72 Chronic Illness +0.35 Gender + .01 Age
(.15) (.11) (.00)

R = 0.32

Eq. 3.2 Health Sat.85 = 10.37 -1.13 Health Impairment
(.14) (.06)
-.67 Chronic Illness -.12 Doctor Visits -.11^{ns} Gender -.01 Age
(.09) (.01) (.06) (.00)

R = 0.53

ns = not significant at the .05 level.

Correlations[a]

		(1)	(2)	(3)	(4)	(5)	(6)
Health Impairment	(1)	1.00	.78	.42	-.71	.09	.49
Chronic Illness	(2)	.63	1.00	-.35	-.58	.07	.36
Doctor Visits	(3)	.30	.25	1.00	-.52	.09	.23
Health Sat.	(4)	-.51	-.42	-.33	1.00	-.10	-.39
Gender	(5)	.08	.06	.07	-.08	1.00	.13
Age	(6)	.44	.32	.18	-.31	.13	1.00

Note: a) Observed correlations are below the diagonal; disattenuated correlations are above. Estimated validity coefficients: Health Impairment = 0.9, Chronic Illness = 0.9, Doctor Visits = 0.8, Health Satisfaction = 0.8, Gender = 1.0, Sex = 1.0.

Table 19: A Dynamic Equilibrium Model: Do Health Stocks Constrain Flows and Utility?

	Multiple Rs for Doctor Visits[a]		Multiple Rs for Health Satisfaction
Doc. $Visits_{85}$	0.32	Health Sat_{85}	-0.53
Doc. $Visits_{86}$	0.29	Health Sat_{86}	-0.49
Doc. $Visits_{87}$	0.29	Health Sat_{87}	-0.45
Doc. $Visits_{88}$	0.28	Health Sat_{88}	-0.46
Doc. $Visits_{89}$	0.27	Health Sat_{89}	-0.46
Doc. $Visits_{91}$	0.27	Health Sat_{90}	-0.45
		Health Sat_{91}	-0.45
Confidence Interval:			
	CI = 0.32 (± .03)	CI =-0.53 (± .02)	

Note: a. No data collected for 1990

It must be conceded that, whether or not the results in this section are taken as indicating some potential support for a revised dynamic equilibrium model, the tests that can be made with the GSOEP data are somewhat limited. All data are based on self-reports rather than medical observation, and there is an element of tautology (or at least contamination among variables) in showing linkages between self-reported health impairment, chronic illness, doctor visits and health satisfaction. Ideally, the dynamic equilibrium model and alternatives should be assessed with data which include medical reports and also better evidence of hereditary risk factors. It may be noted that British studies have shown moderate links between hereditary (and constitutional) factors, some types of morbidity and some causes of mortality (DHSS, 1980; Wadsworth, 1987).

4 Discussion

The stratification paradigm has some explanatory power in accounting for inequalities of income, but less for health and happiness. One reason is that the stocks which account for variance in income are **not** highly correlated with the stocks accounting for variance in health, nor with the stocks accounting for happiness. Neither of our panel studies has ideal stock measures in all three domains but the Australian panel data illustrate how low the correlations probably are. Educational level, a key stock in the income/standard of living domain, correlates very weakly with happiness stocks; -0.02 with extraversion and -0.21 with neuroticism. It correlates -0.12 with health impairment. This health stock, in turn, correlates just -0.13 with extraversion and +0.13 with neuroticism.

So the reason why inequalities of income, health and happiness are only slightly cumulative is that the stocks which determine these outcomes are themselves only weakly correlated. It would be a rare human being who was in the top quintile of income, health and happiness stocks, or for that matter in the bottom quintile.

Contrary to the implied claims of stratification theory, it is rare for people to be either 'rich' or 'poor' in all important stocks.

Why is the paradigm of stratification theory so widely if rather vaguely endorsed? Tentatively, we would suggest that it looks plausible because (a) it is usually tested with data relating to domains of life which matter less rather than more to people (b) some of the variables included in tests are highly correlated virtually by definition and (c) many tests use only cross-sectional data not panel data.

As noted earlier, tests of stratification theory usually relate not to the domains of standard of living, health and happiness - which people care most about - but to gross incomes, occupational prestige, educational attainment and power, which they care rather less about. Furthermore, the impression that inequalities in these latter four domains are cumulative is partly due to the fact that the most widely used occupational status and prestige scales (including the Duncan and Treiman scales) either explicitly or implicitly define occupational status/prestige as a function mainly of gross incomes and educational qualifications. So it follows that high correlations between occupational status (or prestige) and income and education are true by definition. It also seems fair comment that the claim that power (political power?) is correlated with the other variables is only weakly supported by evidence. Power is extremely hard to measure and it is of little value to show (as is often done) that holders of elite positions in the political and economic system are mostly from moderately high status social backgrounds. It does not follow that their policy attitudes or policy decisions (their exercise of power) generally favour higher status groups at the expense of lower status groups (Polsby, 1963).

Finally, in so far as tests of stratification theory often rely on cross-sectional rather than panel data, they are plainly unable to test the hypothesis that inequalities are persistent or long term.

As an alternative to the stratification paradigm, this paper has tested dynamic equilibrium models of standard of living, health and happiness. Whereas stratification theory proposes that outcomes depend upon a single stock - family social background - we have tested separate models for the three domains. The tests have yielded mixed but somewhat promising results. In the standard of living domain, initial stocks of human and material capital came close to constraining subsequent flows of income and income satisfaction within predictable ranges, although the satisfaction correlations showed some tendency to diminish over time. In the happiness domain the model fitted well: the personality stocks of extraversion and neuroticism contrained subsequent flows of life events and life satisfaction within predictable ranges. In the health domain the model fitted least well. There were diminishing correlations for two or three years between initial health stocks and subsequent indicators of current health and health utility, but after that the correlations reached an asymptote. This perhaps indicated that stocks do constrain subsequent outcomes and that alternative specifications of the model might yield better results.

Clearly, this paper only scratches the surface of complex issues which require much more investigation. It is readily conceded that the dynamic equilibrium models offered here have quite limited predictive power, and that additional variables need to be included (especially in the health domain) to improve predictions. It is hoped, however, that the proposed framework - a stocks, flows and utility framework - will prove useful in future research. The framework seems well suited to analysis of panel data: and panel data are essential to test theories of inequality.

References

Abrams, M., V. Gerard and N. Timms (1985), *Values And Social Change In Britain*, Macmillan, London.

Andrews, F.M. and S. B. Withey (1976), *Social Indicators Of Well-Being*, Plenum, New York.

Andrews, F. M. and A. C. McKennell (1980), 'Measures of self-reported well-being: their affective, cognitive and other components', *Social Indicators Research*, 8, 127-55.

Argyle, M. (1987), *The Psychology Of Happiness*, Methuen, London.

Balsen, W. H., H. Nakeilski, K. Rosler and R. Winkel (1984), *Die Neue Armut*, University Press, Koln.

Bane, M. J. and D. T. Ellwood (1986), 'Slipping into and out of poverty: the dynamics of spells', *Journal of Human Resources*, 21, 1-23.

Becker, G. S. (1964), *Human Capital*, Bureau of Economic Research, New York.

Blau, P. M. and O. D. Duncan (1967), *The American Occupational System*, Wiley, New York.

Bottomore, T. B. (1965), *Classes in Modern Society*, Allen and Unwin, London.

Broom, L., F. L. Jones, P. McDonnell and T. Williams (1980), *The Inheritance of Inequality*, Routledge and Kegan Paul, London.

Burkhauser, R. V., G. J. Duncan, R. Hauser and R. Bernsten (1990), 'Economic burdens of marital disruption: a comparison of the US and Germany', *The Review of Income and Wealth*, 36, 319-34.

Campbell, A., P. E. Converse and W. L. Rodgers (1976), *The Quality of American Life*, Sage, New York.

Cantril, H. (1965), *The Pattern Of Human Concerns*, Rutgers University Press, New Brunswick, NJ.

Coser, L. A., S. L. Nock, P. A. Steffan and B. Rhea (1987), *Introduction To Sociology*, Harcourt, Brace, Jovanovich, New York.

Costa, P. T., R. R. McCrae and P. Arenberg (1983), 'Recent longitudinal research on personality and ageing', in K. W. Schaie, ed., *Longitudinal Studies of Adult Psychology Development*, Guildford, New York.

Costa, P. T. and R. R. McCrae (1984), 'Personality as a lifelong determinant of well-being', in C. Malatesta and C. Izard, eds, *Affective Processes in Adult Development And Ageing*, Sage, Beverley Hills.

Creedy, J. (1985), *Dynamics Of Income Distribution*, Blackwell, Oxford.

Dawes, R. M. and Corrigan, B. (1974), 'Linear models in decision-making', *Psychology Bulletin*, 81, 95-106.

Department of Health and Social Security (Great Britain)(1980), *Inequalities in Health*, (Black Report), HMS0, London.
Diener, E., R. J. Larsen, S. Levine and R. A. Emmons (1984), 'Bias in mood recall in happy and unhappy persons', Annual Meeting of the American Psychology Association, Toronto.
Diener, E. and R. J. Larsen (1984), 'Temporal stability and cross-situational consistency of affective, behavioral and cognitive responses', *Journal of Personality and Social Psychology*, 47, 871-83.
Duncan, G. J. et al. (1984), *Years Of Poverty, Years Of Plenty*, Institute for Social Research, Ann Arbor.
Duncan, G. J., T. Smeeding and W. L. Rodgers (1991), 'W(h)ither the middle class?', Levy Institute Conference on Income Inequality, Bard College, June 18-20.
Erikson, R. and J. H. Goldthorpe (1987a), 'Commonality and variation in social fluidity in industrial nations. Part I: A model for evaluating the "FJH Hypothesis"', *European Sociological Review*, 3, 54-77.
Erikson, R. and J. H. Goldthorpe (1987b), 'Commonality and variation in social fluidity in industrial nations: The model of core social fluidity applied', *European Sociological Review*, 3, 145-66.
Erikson, R. and J. H. Goldthorpe (1992), *The Constant Flux: A Study Of Class Mobility In Industrial Societies*, Clarendon, Oxford.
Evans, M. D. R. and J. Kelley (1991), 'Prejudice, discrimination and the labour market: attainments of immigrants in Australia', *American Journal of Sociology*, 97, 721-59.
Eysenck, H. J. and S. G. B. Eysenck (1969), *Personality Structure And Measurement*, Routledge and Kegan Paul, London.
Eysenck, M. (1990), *Happiness: Facts And Myths*, Lawrence and Erlbaum, New York.
Falkingham, J., J. Hills, and C. Lessof (1993), *William Beveridge versus Robin Hood: Social Security and Redistribution over the Lifecycle*, Discussion Paper WSP/88, Suntory-Toyota International Centre for Economics and Related Disciplines, London School of Economics.
Featherman, D., F. Jones and R. Hauser (1975), 'Assumptions of social mobility research in the US: the case of occupational status', *Social Science Research*, 4, 329-60.
Fogelman, K., Fox, A.J. and Power, C. (1989), 'Class and tenure mobility: do they explain social inequalities in health among young adults in Britain?', in J. Fox, ed., *Health Inequalities in European Countries*, Gower, Aldershot.
Friedman, M. (1957), *A Theory Of The Consumption Function*, Princeton University Press.
Galler, H. P. and G. Wagner (1986), 'The microsimulation model of the Sfb3 for the analysis of economic and social policies', in R. Orcutt, ed., *Microanalytic Simulation Models To Support Social And Financial Policy*, New Holland, New York.
Glotz, P. (1984), *Die Arbeit der Zuspitzung*, Deutsche Verlag Anstalt, Berlin.
Goldthorpe, J. H. (1987), *Social Mobility And Class Structure In Modern Britain*, 2nd ed., Clarendon, Oxford.

Harding, A. (1993), 'Lifetime vs annual tax-transfer incidence: how much less progressive?', *Economic Record*, 69, 179-91.

Hauser, R. and P. Semrau (1989), *Trends in Poverty and Low Income in the Federal Republic of Germany*', Working Papers, No 306, Sfb3, University of Frankfurt.

Headey, B. W. and A. J. Wearing (1989), 'Personality, life events and subjective well-being: towards a dynamic equilibrium model', *Journal of Personality and Social Psychology*, 57, 731-39.

Headey, B. W. and A. J. Wearing (1992), *Understanding Happiness*, Longman Cheshire, Melbourne.

Headey, B. W., J. Kelley and A. J. Wearing (1993), 'Dimensions of mental health: life satisfaction, positive affect, anxiety and depression', *Social Indicators Research*, 29, 63-82.

Headey, B. W., P. Krause and R. Habich (1993), 'Long and short term poverty: is Germany a two-thirds society?', *Social Indicators Research*, 32, 39-64.

Headey, B. W. (1993), 'An economic model of subjective well-being: integrating economic and psychological theories', *Social Indicators Research*, 28, 97-116.

Heise, D. R. (1972), 'Employing nominal variables, induced variables and block variables in path analysis', *Sociological Methods and Research*, 1, 147-73.

Henderson, S., D. Byrne and P. Duncan-Jones (1981), *Neurosis And The Social Environment*, Academic, New York.

Holmes, T. H. and R. H. Rahe (1967), 'The Social Readjustment Rating Scale', *Journal of Psychosomatic Research*, 11, 213-18.

Jencks, C. (1972), *Inequality: A Reassessment Of The Effects Of Family And Schooling In America*, Basic, New York.

Jencks, C. (1979), *Who Gets Ahead?*, Basic, New York.

Jones, F. L. (1979), 'Inequality across generations: one hypothesis - half an answer', in Jean I. Martin, ed., *Counter-Predictive Research Outcomes*, Department of Sociology, Research Monograph No. 2, Australian National University, Canberra.

Kakwani, N. (1986), *Analysing Redistribution Policies*, Cambridge University Press.

Kelley, J. (1990), 'The failure of a paradigm: log-linear models of social mobility', in J. Clark, C. Modgil and S. Modgil, *John H. Goldthorpe: Consensus And Controversy*, Falmer, London.

Larsen, R. J., E. Diener and R. A. Emmons (1985), 'An evaluation of subjective well-being measures', *Social Indicators Research*, 17, 1-18.

Lipset, S. M. and H. Zetterberg (1959), 'Social mobility in industrial societies', in S. M. Lipset and R. Bendix, eds, *Social Mobility In Industrial Societies*, University of California, Berkeley.

Lydall, H. F. (1976), 'The theories of the distribution of earnings', in A. B. Atkinson, ed., *The Personal Distribution Of Incomes*, Westview, Boulder, Colorado.

McMichael, A. J. (1985), 'Social class (as estimated by occupational prestige) and mortality in Australian males in the 1970s', *Community Health Studies*, 9, 220-30.

Merz, J. (1988), 'Microsimulation', paper prepared for Eighth Symposium on Forecasting, ISF, 88, Amsterdam, June 12-15.

Miller, S. M. (1960), 'Comparative social mobility', *Current Sociology*, 9, 1-89.

Mincer, J. (1974), *Schooling, Experience And Earnings*, National Bureau of Economic Research, New York.
Orcutt, R. (1957), 'A new type of socio-economic system', *Review of Economics and Statistics*, 58, 773-97.
Parkin, F. (1972), *Class Inequality and Political Order*, Paladin, St. Albans.
Polsby, N. W. (1963), *Community Power And Political Theory*, Yale, New Haven.
Rahe, R. H. and R. H. Arthur (1978), 'Life changes and illness studies', *Journal of Human Stress*, 4, 3-15.
Rendtel, U. (1990), 'Teilnahmbereitschaft in Panelstudien: zwischen Beeinflussung, Vertrauen und sozialer Selektion', *Kolner Zeitschrift fur Soziologie une Sozialpsychologie*, 42, 280-99.
Ringen, S. (1987), *The Possibility Of Politics*, Clarendon, Oxford.
Rokeach, M. (1973), *The Nature Of Human Values*, Free Press, New York.
Saunders, P. et al. (1992), 'Noncash income, living standards, inequality and poverty: evidence from the Luxenberg Income Study', SPRC Discussion Paper, No. 35, Social Policy Research Centre University of New South Wales, Sydney.
Schroeder, D. H. and P. T. Costa (1984), 'Influence of life event stress on physical illness: substantive effects or methodological flows?', *Journal of Personality and Social Psychology*, 46, 853-63.
Schwarz, N. and F. Strack (1991), 'Evaluating one's life: a judgement model of subjective well-being', in F. Strack, M. Argyle and N. Schwarz, eds, *Subjective Well-Being: An International Perspective*, Pergamon, Oxford.
Schwarze, J. et al. (1993), 'Human capital and returns to education in East Germany 1989-92', Conference on the German Socio-Economic Panel, June 7-9, 1993.
Travers, P. and S. Richardson (1993), *Living Decently*, Oxford University Press, Melbourne.
Treiman, D. (1977), *Occupational Prestige In Comparative Perspective*, Academic, New York.
Veenhoven, R. (1984), *Conditions Of Happiness*, Reidel, Dordrecht.
Wadsworth, M. E. J. (1987), 'Follow-up of the first national birth cohort: findings from the Medical Research Council National Survey of health and development', *Paediatric and Perinatal Epidemiology*, 1, 95-117.
Zagorski, K. (1985), 'Class, prestige and status attainment in comparative perspective', *International Journal of Comparative Sociology*, 26, 167-80.

Comments on paper by Bruce Headey and Peter Krause

Peter Travers
Flinders University

Headey and Krause draw on a 1984 to 1991 German panel of 9114 people and a smaller 1981 to 1989 Victorian panel of 941 people to construct an ambitious model to explain inequalities of income, health and happiness. Their finding that these inequalities are only slightly cumulative, moderately persistent, and only in the case of income inequality are substantially due to family social background, is used to refute what they term stratification theory, a theory said to have wide support among social scientists, educated laypeople, as well as professional sociologists. The central tenets of stratification theory are that inequalities are mutually reinforcing, long term, and are substantially due to family background.

Headey and Krause use their German data on income and health and the Victorian data on happiness to test their own alternative thesis, which they term dynamic equilibrium theory. According to the theory, differing 1984 **stocks** such as human and material capital, physical impairment, and personality traits constrain in subsequent years both **flows** such as income, illnesses, and life events, and **utility** as measured by satisfaction with living standards, health and by life satisfaction. The data support the theory most strongly in the domain of living, with more modest indications that it holds also in the domains of health and happiness. The chief explanation of their finding that inequalities are not mutually reinforcing is that the stocks that prove to be so important in constraining flows and utilities are themselves only weakly correlated across individuals and households.

The refutation of the version of stratification theory presented is powerful. The paper shows clear evidence of the volatility of income over time, and of relatively low rates of persistence in very high and very low incomes. The effect of family background as measured by an index of prestige of father's occupation is non-existent in the case of health and happiness, and very modest in the case of standard of living. Above all, the paper points to what I regard as the reassuring conclusion that there is little cumulation of privilege or disadvantage in the three domains under study.

As a sociologist raised in one of the mainstream schools of stratification theory typified by such writers as John Goldthorpe and Robert Erikson, I would have to say that I am neither surprised nor troubled by the findings described in the paper. Sociologists in this tradition are well aware of the volatility of current income, and that is why some of us make use of alternative measures such as full income, a measure in dollar terms of all the major sources of material well being (Travers and Richardson, 1993: 26 ff). On theoretical grounds, we would expect full income to show much greater stability over time than current income.

We are also familiar with the tautology involved in defining prestige or status in terms of gross income and education, and then using status to predict income and education. This rival social stratification tradition (based above all in Europe rather than the United States) has always used class in its classical sense of a categorical variable based on relationship to the means of production in its descriptive and analytical accounts of modern industrial societies. Some of the conclusions of this tradition would sit well with the findings of this paper, despite the very different methodology employed by Headey and Krause.

To illustrate this, let me take the example of what I regard as the chief contribution of Headey and Krause, their model based on stocks, flows and utility. 'Stocks' refer to stable factors such as education, wealth, hereditary and constitutional factors, and personality traits such as extroversion and neuroticism, which in turn **constrain** flows and utility. In the case of income, for instance, though the stock levels are not highly predictive of subsequent income, they do constrain relative incomes within the same range throughout 1985-91. It is precisely at this point that I find myself on familiar territory. Take for instance, Erikson and Goldthorpe's recent study (1992) of mobility in 9 European countries, together with the US, Japan and Australia. Their class model was based not simply on prestige but on the relative desirability of different class positions considered as destinations, the relative advantages afforded to individuals by different class origins, and the relative barriers that face individuals in gaining access to the different class positions (1992: 391). What they found was that these constraints operate in remarkably similar ways, such that the chances of movement between different class positions fell within a similar range across all 12 countries. Despite the different questions they are asking, and the fact that the constraints are expressed in terms of class rather than of 'stocks', I would regard the findings of these mainstream stratification theorists as being in parallel to rather than in conflict with those of Headey and Krause.

I would, however, draw attention to one disturbing feature of the world we are entering that may be underestimated by Headey and Krause. As has been pointed out, they find a low correlation between parental status/prestige and income. If, however, we look at class rather than status, and examine the patterns of transition between the various class positions, we already get a rather different picture of what has been happening in the recent past. In all industrial societies, we find a low probability of movement from service class (professional and managerial) origins to working class destinations. We do find, however, quite high rates of movement from working class origins to service class destinations. One of the reasons for this is that whereas the working class positions have been shrinking in numbers, there has been a massive increase in service class positions. In other words, while families in privileged positions have been relatively successful in protecting themselves and their children from downward mobility, there has been literally more room at the top, and hence considerable scope for upward mobility, even in a relatively stratified society. This world I have just described is rapidly changing. Those of us who are attempting to make strategic plans in the tertiary education sector are painfully aware of how rapidly definitions of what are considered desirable occupations are changing. In some cases, the issue now is not so much access to this or that class position, but to **any** position involving full time employment. We know very little about how people are likely to react in such a world, and what strategies they will employ. My own guess is that it would be unlikely that family background in the richer sense I have described would decline in importance. In other words, in such a polarised world, I would not expect those with resources in the form of human or material capital and social networks to remain passive, but rather to strive to protect both themselves and their families, and to enhance what advantages they already enjoy.

The authors concede that their section on health is limited to owing to the data available to them. They also point out that a British panel study has shown strong links between hereditary and constitutional factors and subsequent morbidity and mortality. I would add that a stratification theorist would have been more likely to have used a different methodology, such as odds ratios to test the relative risks experienced by different social groupings.

To this point, I have said little about utility as measured by life satisfaction, and satisfaction with health and standard of living. It is an area where I myself feel a degree of ambivalence. As Headey and Krause note (and Aristotle and Plato before them), people are very interested in happiness! Since this is among the things people want most in life, surely any study of inequality should place inequality in happiness at centre stage? This is, for instance, the conclusion of one strand of the Scandinavian levels of living studies represented by authors such as Allart, who argue this case against their more resource-focused colleagues (Erikson et al., 1987). Against this, we have the warnings of Sen (1990) of the inherent conservative bias of such a focus. Precisely because levels of utility have so little to do with standards of living, with literacy, or even with basic freedoms, it is a small step to conclude that public intervention to secure these other goods is of little importance. Every liberation movement faces the dilemma of whether to take steps that will widen the horizons of previously contented people and perhaps in the short term increase, not their happiness, but their discontent. I have no doubt that the answer to this dilemma lies in facing the inescapability of making not just empirical observations, but also moral judgements on what constitutes a decent life. Thus, I would note with interest that people in the most appalling circumstances may be more or less as happy as affluent Australians or Germans. But I may still conclude that those appalling circumstances are not fit for human beings.

References

Erikson, R. and H. Uusitalo (1987), 'The Scandinavian approach to welfare research', in R. Erikson et al., eds, *The Scandinvaian Model: Welfare States and Welfare Research*, M. E. Sharp, New York, 177-93.

Erikson, R. and J. H. Goldthorpe (1992), *The Constant Flux. A Study of a Class Mobility in Industrial Societies*, Clarendon, Oxford.

Sen, A. (1990), 'Individual freedom as a social commitment', *The New York Review of Books*, 37, 14 June, 49-54.

Travers, P. and S. Richardson, (1993), *Living Decently*, Oxford University Press, Melbourne.