Session Number: 4C
Session Title: Improving Estimates from Survey Data
Session Organizer(s): Stephen Jenkins, University of Essex, Colchester, UK
Session Chair: Stephen Jenkins, University of Essex, Colchester, UK

Paper Prepared for the 29th General Conference of
The International Association for Research in Income and Wealth

Joensuu, Finland, August 20 – 26, 2006

THE SENSITIVITY OF INCOME DISTRIBUTION MEASURES TO CHANGES IN SURVEY COLLECTION TOOLS AND ESTIMATION TECHNIQUES IN AUSTRALIA

LEON PIETSCH and BOB McCOLL
Australian Bureau of Statistics
PETER SAUNDERS
Social Policy Research Centre, University of New South Wales

For additional information please contact:

Author Name(s) : Bob McColl
Author E-Mail(s) : bob.mccoll@abs.gov.au

This paper is posted on the following websites: http://www.iariw.org

THE SENSITIVITY OF INCOME DISTRIBUTION MEASURES TO CHANGES IN SURVEY COLLECTION TOOLS AND ESTIMATION TECHNIQUES IN AUSTRALIA

LEON PIETSCH and BOB McCOLL
Australian Bureau of Statistics

PETER SAUNDERS
Social Policy Research Centre, University of New South Wales

DRAFT

INTRODUCTION AND ABSTRACT

Time series of household income and income distribution have proved very sensitive to changes in the survey collection tools and estimation techniques used. Even when these have not changed, concerns about time series consistency can arise.

This paper presents three case studies that describe circumstances that have led to concerns about the validity of time series taken from Australian Bureau of Statistics household income surveys, and the actions that were taken to address those concerns.

First, despite no apparent change to the way that the survey was being conducted, the survey coverage of aggregate social assistance benefits paid by government declined markedly over a period of two years. After an intensive study of the possible causes of this divergence, two major changes were introduced into the estimation process - an additional government transfers calibration benchmark was introduced into the estimation phase of the survey, and results for all surveys in the current series from the mid-1990s were re-estimated using more consistent demographic calibration benchmarks.

The second case study compares income surveys conducted prior to the mid-1990s and those conducted in the years following the introduction of the Survey of Income and Housing in 1994-95. It is concluded that the survey changes introduced in 1994-95 were of such a magnitude that there is a distinct break in time series at that point. A degree of comparability over time could be achieved if the 1994-95 survey estimates were re-estimated in a way that better reflects the definitions and structures of the earlier surveys. However, this would only provide a linked series, and any time series analysis would have to use 1994-95 estimates on both the "old" and "new" basis.

The third case study examines uncertainties underlying comparisons of data from the most recent income survey with those for earlier periods. Several changes were introduced to the income survey in order to improve the accuracy of the survey results, increase the range of data collected, and improve the way the income survey complemented other ABS household surveys. These included: taking an independent sample for household income measurement to replace the previous practice of taking a sample from households that had already been included in and continued to respond to the monthly labour force survey for eight months; using computer assisted interviewing; introducing additional income questions; collecting for the first time a comprehensive range of wealth data (with asset related incomes edited against the reported wealth holdings to improve business and investment income reporting); and integrating the household expenditure survey with the income survey for about 60 per cent of the income survey sample. Analysis of time series data from the survey suggests that estimates of income levels remained reasonably consistent through the introduction of the changes. But not all of the movements in income distribution measures such as the Gini coefficient can be immediately explained in terms of known real world changes, such as increased levels of social assistance benefit payments. Therefore the prospect of some instrument effects remains.

CASE STUDY 1: MANAGEMENT OF APPARENT SPONTANEOUS CHANGE IN SURVEY COVERAGE OF SOCIAL ASSISTANCE BENEFITS

Introduction

In compiling the estimates for the 2000-01 Survey of Income and Housing (SIH)[1], an unexpected and significant decline was observed in the coverage ratio of social assistance benefits reported in the SIH. The coverage ratio was obtained by comparing the aggregate value reported in the SIH to the value of aggregate payments made by the two government agencies responsible for nearly all payments, the Department of Family and Community Services (FaCS) and the Department of Veterans' Affairs (DVA).

In the first four SIH surveys, from 1994-95 to 1997-1998, the coverage of social assistance benefits had been relatively stable, falling between 82.9% and 86.2%. No SIH was conducted in 1998-99. In the 1999-2000 SIH the coverage was a little lower at 81.2%, but this was not yet seen as a cause for particular concern. However, when the 2000-01 estimates were compiled, the benefit coverage had fallen further, to 78.2%. If the lower coverage of SIH benefits largely related to missing payments made predominantly to households represented in the lowest two income quintiles, the change in understatement would have impacted very significantly on several of the measures used to assess income inequality.

This case study describes the results of the investigations undertaken and the steps taken to resolve the problem.

Possible sources of undercoverage in social assistance benefit estimates

Population scope

The SIH does not attempt to capture all social assistance benefit payments.

The scope of the SIH is restricted to urban and rural areas of Australia, excluding remote and sparsely settled areas of the Northern Territory, and includes only the usual residents of private dwellings, such as houses, flats, units, caravans, tents and other private structures that are places of usual residence at the time of interview. People living in non-private dwellings, such as hotels, boarding schools, nursing homes and other institutions, are excluded.

Some Australian government social assistance benefits are paid to people residing abroad, and they are of course are also excluded from the scope of the SIH. It is estimated that approximately 95% of total social assistance benefit payments are made to people falling within the scope of the SIH.

---

[1] Until 2001-02, the survey was known as the Survey of Income and Housing Costs, commonly abbreviated to SIHC. However, for simplicity, the survey has been consistently labelled as the Survey of Income and Housing (SIH) throughout this paper.

Underreporting

The SIH fails to collect some benefit payments that are made to people in scope of the survey. In some cases, respondents fail to report all their income, including social assistance benefits. Respondents are asked to report the latest benefit amount received. This amount is likely to be reported in SIH, at least in part, as the net cash transfer usually received by the respondent. Amounts that are deducted at source, such as tax, rent or other regular commitments for which arrangements have been made for automatic deduction by the payment agency, may be excluded by some respondents. Amounts that are received less frequently than fortnightly, such as a quarterly telephone allowance, may also be excluded. Respondents may also fail to report all their income for a variety of other reasons, such as privacy concerns, difficulties in remembering income details, and unwillingness to reveal fraudulent or other illegal activity.

Non-response bias

Survey results are expanded to estimates for the whole population by applying weights to survey responses. In calculating the weight to be applied to each respondent, the weights are calibrated to external demographic benchmarks to ensure that the expanded estimates are consistent with the demographic characteristics of the population as established through population censuses and intercensal demographic estimates. It is then assumed that survey respondents are representative of all people in the population who are within scope of the survey.

While demographic benchmarking ensures consistency for certain demographic characteristics, this may not be the case for other characteristics being collected in the survey, such as income and source of income. The most problematic aspect is the extent to which survey respondents may differ from people who reside in dwellings selected in the survey but from whom responses were not obtained. Such non-response bias may result in undercoverage or overcoverage in final survey estimates. In the case of the SIH, aggregate estimates of total social assistance benefits may, in principle, therefore reflect some element of undercoverage or overcoverage because of non-response bias. In practice, it seems unlikely that there is any significant overcoverage from non-response bias, given the known undercoverage of the survey aggregate.

Undercoverage over time

The net effect of scope restrictions, incomplete reporting and the population benchmarking adopted was a substantial but stable difference from 1994-95 to 1997-98 between aggregate benefits estimated from the SIH and aggregate benefits paid by government agencies. Variations from year to year were within the range that could be expected due to sampling error.

There also appeared to be undercoverage in the survey estimates of income from other sources such as wages and salaries, when compared to other data sets such as national accounting aggregates. However, the degree of undercoverage cannot be estimated with the same precision as for social assistance benefits because of the more substantial conceptual differences and methodological shortcomings of the alternate

data sets. The extent of undercoverage of each income source will affect the estimation of household income levels and the measurement of income distribution at any point in time. As long as undercoverage for each income source is relatively stable over time, the impact on measuring changes in income and its distribution will be limited.

However, coverage of social assistance benefits declined significantly over the two SIH cycles after 1997-98, as can be seen in the first row of table 1. From a peak of 86.2% in 1997-98, the ratio declined to 81.2% in 1999-00 and to 78.2% in 2000-01. If the increased SIH undercoverage was due to reporting error by individuals, or processing error, or a real world change not captured through the SIH methodology, there was the potential for significant misrepresentation of the changes in income distribution in Australia. In addition, comparisons between various life cycle groups was likely to be affected if some major income sources were more affected than others.

Investigations into possible causes of coverage change

A number of different avenues were investigated in seeking to understand and correct for the decline in benefit payments coverage. These included possible errors in the processing system (including clerical procedures), appropriateness of the coverage comparison being made between aggregate SIH estimates and aggregate benefits paid by government agencies, changes in the way that benefit payments were made which might not have been captured in the SIH, changes in the quality of reporting by households, and options for and appropriateness of the weighting methods used to compile aggregate results.

Processing error

The SIH processing system had been relatively stable since its inception in 1994-95. A review of the system did not identify either any system changes that might only have impacted on 1999-2000 and 2000-01 benefit estimates, or errors that might have been reflected systematically only in estimates for 1999-2000 and 2000-01.

Coverage comparison between SIH estimates and aggregate benefits paid by government agencies

Because of the audit scrutiny associated with government outlays, there is little likelihood of significant error in the aggregate benefit amounts published by government agencies. It was possible that changes in the nature of accounting for the expenditures, changes in the population composition of benefit recipients, or changes in the way that recipients were provided with their benefits may have impacted on the validity of the coverage analysis being undertaken. However, ABS investigations showed that a stable relationship could be expected, over the period 1994-95 to 2000-01, between SIH measures of benefits and the aggregate values published by the Departments of Family and Community Services and Veterans' Affairs because:

- the proportion of benefit recipients in special dwellings or overseas had been stable over the period when SIH coverage declined

- while accrual accounting was introduced as the basis of compilation for published benefit aggregates from 1998-99, the nature of the changes were not such that they would have had an adverse impact on apparent SIH coverage
- aggregate transfer values published by the benefit payment agencies only relate to benefits paid, and so do not include, for example, administrative costs which may have increased in recent years
- analysis of movements in selected survey aggregates, such as age pension payments, tracked combined changes in benefit levels and eligibility criteria announced by payment agencies
- even if all affected respondents failed to include the value of the automatic deductions made on their behalf by payment agencies, such as income tax or rent, the scale of the increase in such deductions was not sufficient to have had a marked effect on the coverage ratios over the period under analysis.

There was one error identified incidentally through this analysis. Under the income concept used in the SIH, the survey had failed to collect information about a special once-only payment to seniors paid in 2000-01 to social assistance benefit recipients who had reached age pension age. Correcting this error accounted for 1.1 percentage points of the 3.5 percentage point deterioration in coverage in that year, but did not account for any of the deterioration in 1999-2000.

In summary, except for the once-only payment to seniors, no errors were found in the process of comparing the SIH benefit transfer estimates with published aggregates.

Misreporting by SIH respondents (measurement error)

Possible causes for respondent error contributing to the declining coverage of benefit transfers reported in SIH included:

- respondents increasingly understating the amount of benefit transfers that they receive
- respondents increasingly declining to acknowledge that they were recipients of social assistance benefits, whether from an increasing concern about privacy, from a desire to hide fraudulent activity, or otherwise.

To assess the accuracy of respondents' reporting, the benefits reported by individuals were compared to estimates of apparent benefit entitlements modelled on the basis of other reported information such as age, non-benefit income, and number of children. The analysis did not reveal any obvious decline in the average individual benefit level being reported relative to the apparent benefits entitlement. If a decline had been detected it might have suggested an increasing tendency to understate the individual amounts received. Nor did this analysis identify any increase in people not reporting the receipt of benefits when they had no other significant sources of income. For example, the number of persons reporting that they received the age pension in SIH was a constant proportion of the total number of persons in the SIH sample who were of age pension age and also had little other income.

It is possible that persons who are not entitled to, but nevertheless receive, social assistance benefits do not report the fraudulently claimed benefit income in ABS surveys. While this possibility is plausible for some benefit types, no evidence of an

increase in fraud was identified. And no plausible explanation was identified for fraud to be the cause of an across-the-board decline in coverage of all major benefit types in 1999-2000, including age pensions, disability pensions and war service pensions, nor why that level of fraud would accelerate in 2000-01.

In summary, although there may well have been some misreporting by SIH respondents, no evidence was found for any significant deterioration over the two years of concern.

Differential undercoverage and demographic benchmarks

As with other household surveys, the estimation and weighting of SIH includes a process of calibrating the weights to known demographic totals (i.e. population totals of people and households, classified by age, sex, state, etc.). One of the reasons to benchmark a survey in this way is to maximise the extent to which the survey results represent the full population being surveyed. Subgroups that responded less well to the survey are therefore given larger weights than subgroups that responded more fully. However, if non-respondents differ from respondents in characteristics other than those being benchmarked, survey estimates are still subject to non-response bias.

There were several indicators that the impact of non-response on the SIH was changing and the profile of survey respondents was becoming less representative of that of non-respondents. As a result, the SIH estimation methodology may not have been fully effective in accommodating changing non-response patterns, leaving the potential for bias in the coverage of incomes that might result. These indicators were:

- SIH response rates that had been relatively stable at about 90% over the period 1994-95 to 1997-98, but slipped to 85% from 1999-2000, the first year of the decline in benefit transfer coverage
- an apparent and significant over-representation of children under 15 years of age in the weighted SIH results, indicating that households with children were more likely to respond in the SIH than households without children (the demographic benchmarks used at that time only included persons aged 15 years and over)
- the across-the-board nature of the decline in coverage of benefits suggested that weighting to demographic benchmarks was not fully compensating for differential undercoverage in the sample responses.

The demographic benchmark data used for earlier surveys were updated to reflect the incorporation of 2001 Population Census information into the estimates of Australia's resident population. Various demographic benchmarking regimes were then analysed to see which best adjusted for the undercoverage of different demographic sub-populations. While the estimates showed some sensitivity to the benchmark options tested, none of the options offered a solution to the benefits coverage gap.

Benchmarking to standard set of demographic benchmarks

Since it became clear that the survey estimates were sensitive to the choice of demographic benchmarking regime, and since the children under 15 years of age

were consistently overrepresented in the survey samples, it was decided that a standard, improved benchmarking regime should be introduced.

The main change was to benchmark to the number of children in the age ranges of 0-4 years, and 5-14 years, by state/territory. However, introducing this important improvement in benchmarking, and a desire to have an estimation regime consistent across all years, required the following benchmarks that had been previously applied to be foregone:

- quarterly and half yearly benchmarking
- state/territory by household counts.

The removal of sub-annual benchmarking was not found to be significant to the quality of the SIH results. While state/territory household counts were removed from the benchmarking, a range of state/territory benchmarks remained (age groups by sex, state by part of state, state/territory by labour force status), the new state/territory by children age groups benchmark was introduced, and national household benchmarks remained.

Benchmarking to social assistance aggregates

Following the investigation of the range of issues, discussed above, that could potentially contribute to the decline in SIH coverage of social assistance benefits, ABS concluded that the increasing SIH undercoverage of benefits resulted from an increase in the differential undercoverage of benefit recipients that could not be accommodated by demographic benchmarks alone, although there was no obvious cause for such an increase. Nor has any cause emerged in the period since these investigations were undertaken. Furthermore, the benefit coverage ratio in SIH surveys undertaken since 2000-01 have returned to about 85%.

To directly address the undercoverage of benefits the ABS introduced explicit benefit benchmarks for the 1999-2000 and 2000-01 SIH estimates. This was consistent with the general approach of benchmarking to address differential response rates and coverage deficiencies, such as not collecting data from certain geographic areas for which the populations are nevertheless incorporated in demographic benchmarks.

Several issues were considered in deciding how to benchmark to benefit transfers.

- Should benchmarking be to numbers of benefit recipients or to value of benefits paid?
- Should benchmarking be done at an aggregate level or by benefit type?
- Should benchmarking be to 100% of the FaCS/DVA values or some lower amount?

Numbers of benefit recipients or value of benefits paid?

It was decided to benchmark to the value of benefits rather than to the number of recipients because the available data on value of benefits more precisely accounts for variations in numbers of recipients over the year and differences between recipients eligible to receive maximum payments and those eligible to only receive partial

payments. While the benchmarking process ensured consistency with respect to the value of benefits, the process achieved this by increasing the survey weights assigned to respondents reporting benefits and decreasing the weights of other respondents. In other words, the benchmarking process increased the estimated number of benefit recipients, and did not amend the values of individual respondents.

<u>Aggregate level or by benefit type?</u>

In theory, it would have been desirable to benchmark to income from individual benefits, or at least to income from broad groups of benefits, because the undercoverage had behaved differently for different benefit types over the years that SIH had been run.

However, it was known that there was some tendency by respondents to misclassify certain benefit types, especially where the rules defining the boundary between two benefit types have changed over time. Therefore attempting to benchmark to individual benefit types would have implied a greater sense of accuracy than could have been achieved. An analysis of the impacts of the two choices of benchmarks showed that there would be little difference between the two approaches in practice, and so it was decided to benchmark to the total income from benefits.

<u>To 100% of the value paid by government agencies or some lower amount?</u>

Options also existed on whether to benchmark to 100% of aggregate benefits that were within scope of SIH, or to some lesser amount. For the early, apparently stable part of the series, the survey was accounting for about 85% of aggregate benefits. Some part of the difference is attributable to the scope differences, as discussed earlier. It is believed that the scope difference amount to about 5% of total payments, but the exact amount is not known.

In theory, if there is no measurement error in the data, the remaining undercoverage could be removed by benchmarking the sample to the total amount of benefits. However, there is likely to be some measurement error in the data, even though it had been concluded that increasing measurement error did not seem to be the cause of the decline in survey coverage of benefits. Measurement error may well have been a significant contributor to the 'base' amount of undercoverage through the whole period, since there may have been significant differences between the benefit reported by respondents and the actual amount of benefit transfers paid to them by government agencies. Benchmarking would not be an appropriate means of addressing this problem because it would increase the number of recipients, not the average value received.

Furthermore, as noted earlier, there also appears to be undercoverage in SIH estimates of income from other sources, and therefore they also are likely to be affected by measurement error. Correcting just the benefit income for such deficiencies, by increasing the incomes of those at the lower end of the income distribution, would alter the apparent income distribution observed in the SIH. But it is not possible at this time to determine whether such a change would increase or decrease the accuracy of the distribution measures.

As it was not known how much of the 15% 'base' undercoverage was attributable to the impact of differential undercoverage, it was decided that the benefit value benchmark should only be applied from 1999-2000 and that it should only be used to remove the deterioration in the survey coverage of benefit transfers that occurred from that time, that is, the increase in undercoverage beyond the base amount of approximately 15%.

Solution implemented

Three distinct changes were made to the SIH estimates of income as a result of the investigations described above.

First, the estimates for all years prior to 2000-01 were recalculated using the most up to date demographic benchmark data available and a consistent estimation and weighting system was introduced for all years through to 2000-01.

Second, estimates for the once-only payment to seniors were modelled and added to respondent records for 2000-01.

Third, the additional social assistance benefit benchmark was introduced for 1999-2000 and 2000-01 to maintain the SIH coverage of transfer benefits at a consistent level over time. As noted above, the undercoverage of social assistance benefits returned to the more usual level of around 85% in SIH cycles after 2000-01, and so the social assistance benefit benchmark was not maintained for the later cycles.

Impact on government benefit transfers

The impact of the changes on the SIH coverage rates of social assistance benefits is shown in table 1. As can be seen, at the start of investigations, the 1999-2000 coverage ratio of 81.2% was substantially below that of 1997-98 but not very far below the previous lowest point of 82.9%. The 2000-01 ratio fell a further 3.0 percentage points, to 78.2%.

Table 1 Coverage rates of FaCS and DVA social assistance benefits(a)

	1994-95	1995-96	1996-97	1997-98	1999-00	2000-01
% COVERAGE						
At the start of investigations	84.0	82.9	85.9	86.2	81.2	78.2
After standardisation of estimation(b)	84.7	83.6	84.5	84.9	82.5	79.0
After imputation of values for once-only payment to seniors in 2000-01	84.7	83.6	84.5	84.9	82.5	80.1
After introduction of benefit transfer benchmark in 1999-2000 and 2000-01	84.7	83.6	84.5	84.9	84.7	84.7

(a) Social assistance benefits paid by Departments of Family and Community Services and Veterans' Affairs that fall within the definition of income used in this publication
(b) Includes revision of demographic benchmark data for 1999-2000 and earlier years, and introduction of a standard benchmarking regime and estimation and weighting system for all years

After revisions were made to the demographic benchmarks for the years up to 1999-2000, the introduction of identical estimation and weighting procedures for all years, and the introduction of imputed estimates for the once-only payment to seniors in 2000-01, the fall in the coverage ratio between 1997-98 and 1999-2000 was not as great as previously estimated. However, the coverage ratios still showed a clear downward trend in the two years to 2000-01. The fall was even more apparent insofar as the ratios for the first four years showed less variation, after the estimation and weighting system had been standardised, than had been apparent at the start of the investigations. The first four observations now fell within a range of 1.3 percentage points, but there was still a 2.4 percentage point decline from 1997-98 to 1999-2000 and a further 2.4 percentage point decline to 2000-01. Without the contribution of the imputed estimates for the once-only payment to seniors, there would have been a 3.5 percentage point decline to 2000-01.

By definition, the introduction of the government benefit transfer benchmark for the last two years lifted the overall coverage ratio for those years to the benchmark level, that is, 84.7%. (This is marginally higher than the average of the first four years (84.4%) because the values feeding into the benchmark calculation were derived before the estimation and weighting system had been finalised.) The benchmark was applied to total benefits excluding the once-only payment to seniors.

Impact on income distribution

The introduction of the social assistance benefit benchmark tended to increase the sample weights of households with relatively low income and therefore lower the weights of households with relatively high income. Consequently, the values of income at the percentile boundaries shown in table 2 were all slightly lower after the introduction of the new benchmark. There was no impact on the percentage share figures (to one decimal place). Some of the percentile ratios measured slightly less income inequality, although P80/P20 and P20/P50 measured slightly greater inequality in 1999-2000. The Gini coefficient would have been slightly higher in 2000-01 if a benefit benchmark had not been introduced (0.313 compare to 0.312). In all cases, the revisions to the measures were considerably smaller than one standard error, that is, they do not make a significant difference to the interpretation of the indicators.

Similarly, the correction to include imputed values for the once-only payment to seniors decreased the measures of inequality very slightly, and slightly increased the values of income at the percentile boundaries.

Table 2 Income distribution, equivalised disposable household income

Indicator		1999-2000		2000-01		
		Without benefits benchmark	With benefits benchmark	Without benefits benchmark or OOPS(a)	With benefits benchmark but not OOPS(a)	With benefits benchmark and OOPS(a)
Income per week at top of selected income percentiles, in 2000-01 dollars(b)						
20th [P20]	$	242	241	244	243	245
50th [P50]	$	407	405	416	413	414
80th [P80]	$	636	636	647	643	644
Share of total income received by persons with						
High incomes(c)	%	38.4	38.4	38.5	38.5	38.5
Low incomes(d)	%	10.5	10.5	10.4	10.4	10.5
Ratios of incomes at top of selected income percentiles						
P90/P10	ratio	3.90	3.89	4.03	4.01	3.97
P80/P20	ratio	2.63	2.64	2.66	2.65	2.63
P80/P50	ratio	1.57	1.57	1.56	1.56	1.56
P20/P50	ratio	0.60	0.59	0.59	0.59	0.59
Gini coefficient	no.	0.310	0.310	0.313	0.312	0.311

(a) Imputed value of once-only payment to seniors.
(b) Adjusted for changes in the Consumer Price Index.
(c) Persons in the top income quintile (9th and 10th deciles) after being ranked by their equivalised disposable household income.
(d) Persons in the 2nd and 3rd income deciles after being ranked by their equivalised disposable household income. The characteristics of people in the 2nd and 3rd deciles are regarded as being more representative of people with low economic resources because there is strong evidence that many households in the 1st decile with very low reported income have significant other economic resources available to them.

Conclusion to Case Study 1

The ABS made extensive investigations into the decline in the SIH coverage of aggregate social assistance benefit payments that occurred in 1999-2000 and 2000-01. Those investigations revealed a number of ways in which the survey estimates could be improved but did not provide any explicit explanation for the decline in coverage. It was therefore concluded by default that there had been an unexplained increase in non-response biases that particularly impacted on the coverage of social assistance benefits. With the passing of time since those investigations, no further information has emerged to better explain what had occurred in those years. Equally inexplicably, the coverage rate has returned to historic levels of around 85% in the surveys conducted since 2000-01.

Since social assistance benefits are payments made by government to support people with low income, ABS considered that it would be inappropriate to produce time series of income distribution measures that may have been significantly impacted by problems in the survey data collected. Therefore corrective action was taken to ensure that social assistance benefits made a consistent contribution to household income as measured in the survey.

CASE STUDY 2: IMPACT OF SIGNIFICANT METHODOLOGY CHANGE TO INCOME SURVEY, 1994-95

Introduction

The Survey of Income and Housing (SIH) described in the previous case study was introduced in 1994-95. It replaced an earlier family of surveys conducted in 1982, 1986 and 1990 which can collectively be labelled as the Income Distribution Surveys (IDSs). There were a number of significant changes between the IDSs and the SIH which potentially impacted on the comparability of time series across the two surveys. Examination of those differences was part of a joint research project undertaken by the Social Policy Research Centre of the University of New South Wales and the ABS. Key findings of the project are summarised here.

Major differences between the IDSs and the SIH

The IDSs were conducted independently of any other survey whereas the sample for the SIH was chosen from households that had just completed a period of inclusion in the monthly Labour Force Survey. Households selected for the Labour Force Survey remain in that survey for eight months and a subsample of those completing their eighth month in the survey were selected as the SIH sample. It is therefore likely that non-response to the SIH and the degree and form of any mis-reporting in the SIH was influenced by that previous eight month's experience of the selected households. It cannot be assumed that the errors were better or worse, but only that they are likely to be different. For example, a household that was asked to be part of the relatively sensitive SIH after eight months of inclusion in the Labour Force Survey might be more inclined to refuse because of the burden of previously responding to surveys, or they might be more cooperative because they have become used to being part of a statistical survey.

The IDSs were conducted in the months September to December whereas the SIH was conducted continuously through the Australian financial year of July to June. Therefore questions relating to annual income in the previous financial year were likely to be answered differently in the two surveys because respondent recall and the ability to access taxation records differed between the two situations. In addition, the estimates of current period income were subject to seasonal biases in the IDSs but not the SIH, since only the SIH was conducted throughout the year.

In both the IDSs and the SIH, some records were excluded from estimates of previous financial year income because changing circumstances in the household meant that the data collected were not likely to be representative of a full year's economic situation. However, the previous financial year exclusion flags were set in more circumstances in the IDSs than in the SIH, partly because more information about household circumstances, such as time spent overseas in the previous financial year, was collected in the IDSs.

There were changes to data item definitions and treatments between the surveys. Of particular relevance to income distribution analysis was the change to the treatment of negative net income reported with respect to business or property rental activity. In

the IDSs these were set to zero whereas in the SIH they were allowed to remain negative.

Calibration of survey weights to external demographic benchmarks was discussed as part of the first case study considered in this paper. In the IDSs weights were derived and calibrated for each person record in the survey. Only persons aged 15 and above were included in the survey. Weights for higher level units such as income units, families and households were calculated as the harmonic mean of the weights for each of the individuals comprising the unit. In the SIH, integrated weighting was introduced, whereby the weights for all individuals in the household and the household weight were all constrained to be equal and the household weights were also calibrated to external demographic benchmarks. As described in the first case study above, the SIH weights were re-calibrated at a later stage, and calibration to external demographic benchmarks for children less than 15 years of age was also introduced.

Impact of the differences

It was not possible to directly quantify the overall impact of the differences between the IDSs and the SIH because the two surveys were not conducted concurrently. Therefore quantitative analysis largely comprises comparisons between survey estimates and external aggregates taken from estimates of resident population and the Australian System of National Accounts.

Comparison with aggregate population estimates

Comparison with aggregate population estimates is separated into three parts, differentiating between persons aged 15 to 64 years, those aged 65 years and over, and those aged less than 15 years.

Chart 1 Population aged 15 to 64, ratio of survey estimate to estimated resident population of Australia
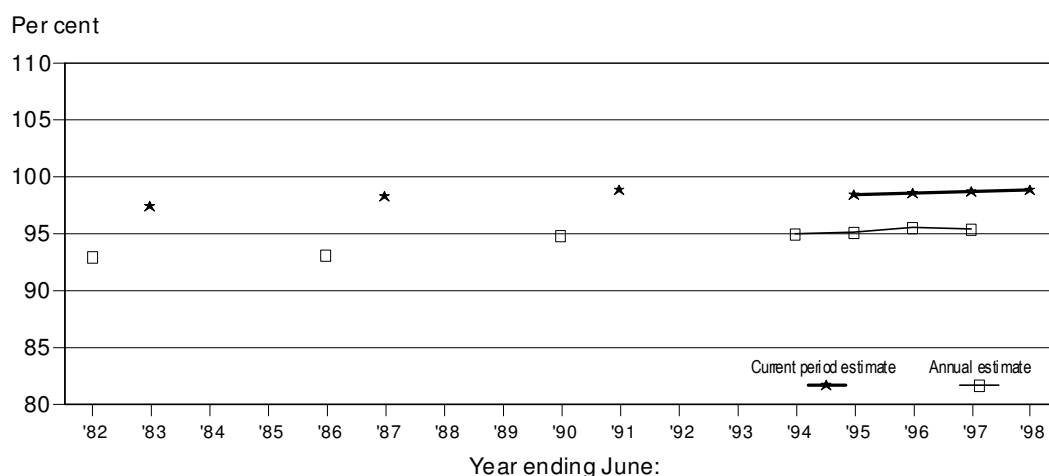

Chart 1 compares the number of people aged 15 to 64 years (approximately the working age population) as estimated from the surveys with the demographic

estimates of the total Australian resident population. In the chart, the ratios labelled 'current period' relate to the average population in the year that the survey was conducted.

The ratios for the current period estimates are close to 100 per cent for all years, since the estimates are calibrated to demographic benchmarks and only a relatively small proportion of the working age population are institutionalised and therefore out of scope of the survey.

The chart also contains estimates labelled as 'annual'. These are the ratios of the people in the surveys who are eligible to contribute to the estimates of previous year income compared with the average estimated resident population of the year prior to the survey. The annual ratios are somewhat lower than the current period ratios because a number of people are excluded from contributing to the annual estimates due to their changed circumstances. The slightly larger gap between the current period and annual ratios reflects the greater range of exclusions that were incorporated into the IDSs up until 1990.

Chart 2 shows the same ratios for the population aged 65 years and older. The ratios for the current period estimates are somewhat lower than for those aged 15 to 64 years, reflecting the higher proportion of people in the older age group that are institutionalised. The ratios for the annual estimates are similar to those aged 15 to 64. The impact of a higher proportion of the older age group being institutionalised is offset by the survey estimate (that is, the numerator) including those who were still aged 64 in the previous year.

Chart 2 Population aged 64 and over, ratio of survey estimate to estimated resident population of Australia
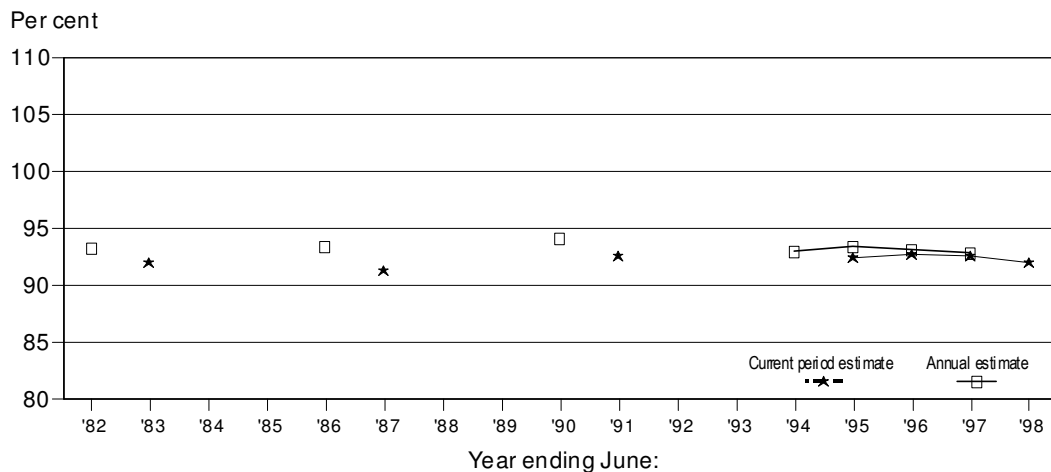

Chart 3 provides similar comparisons for children aged less than 15 years. It shows ratios for current period estimates close to 100 per cent from 1995 onwards, reflecting their calibration to external demographic benchmarks. Such calibration has not been undertaken for the IDSs. Annual estimates are somewhat lower than current period estimates because the children have been excluded if their parents were excluded, as discussed above.

Chart 3 Population aged under 15, ratio of survey estimate to estimated resident population of Australia

Per cent

Year ending June:
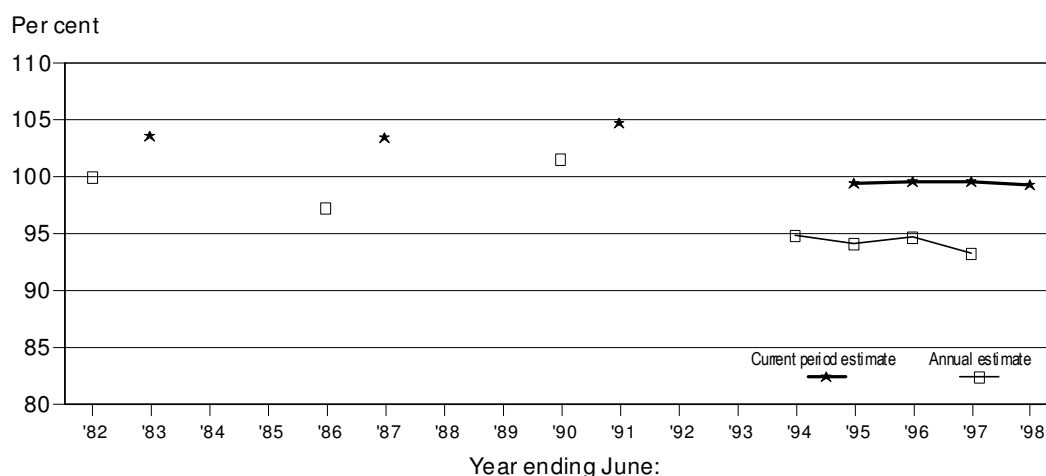
The higher ratios for children in the IDSs shows that they are overrepresented in the survey, with the implication that households with children are overrepresented. This is likely to have a significant impact on the comparability of results from the IDSs and the SIH. The first case study described in this paper showed that estimates of income distribution are sensitive to the benchmarking regime chosen. The IDSs did not include calibration to benchmarks of household numbers either, with additional likely impact on comparability between the IDSs and the SIH.

Comparison with national accounting aggregates

In practice, there are substantial differences between the scope and definitions of income used in compiling the Australian System of National Accounts (ASNA) and the scope and definitions implemented in the SIH. Therefore comparisons are only presented here for the two components for which the differences are the smallest, wage and salary income and social assistance benefit income. Between them, they account for over 80% of total household income.

Chart 4 provides ratios of SIH estimates of aggregate wage and salary income to ASNA estimates of wage and salary income and ASNA estimates of the compensation of employees. While less comparable in scope, the latter is available for a longer period of time in the ASNA. The relationship between the SIH and ASNA estimates appears relatively stable between the IDSs and the SIH, except for the lower ratios for the current period estimate for 1982-83 and the annual estimate for 1981-82. In the 1982 IDS only, the wages and salaries estimate did not include wages and salaries and directors' fees paid from own incorporated businesses. Instead, this income was grouped with the income of own unincorporated business. Therefore there is a comparability issue between the 1982 IDS and later surveys when examining wages and salaries, but this difference should not impact on the comparability of estimates of total income.

It might be noted that the annual estimates of wages and salaries tended to be higher than the current period estimates. This reflected the focus of the current period estimates on regular or usual wage and salary income, whereas the previous financial

year data item included all wage and salary income including irregular overtime, bonuses and the like.

Chart 4 Wage and salary income/compensation of employees, ratio of survey estimate to ASNA estimates
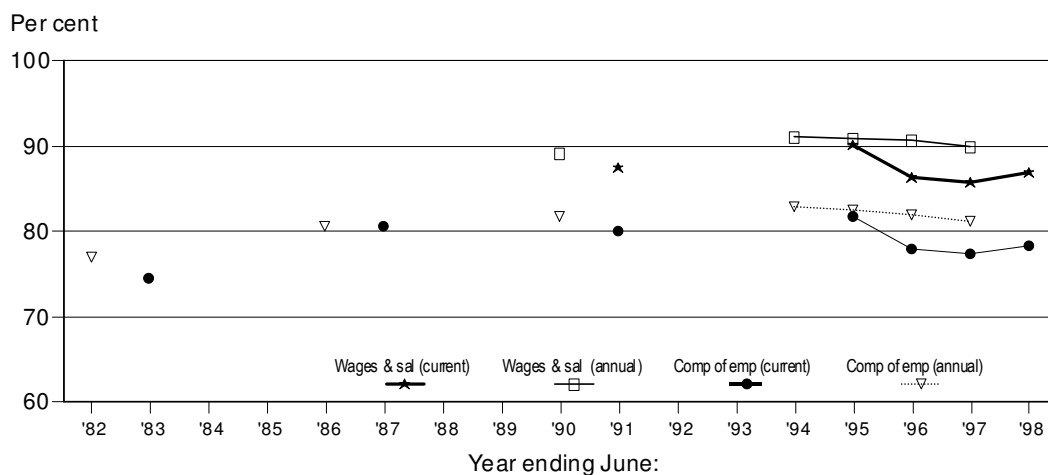
As discussed in Case Study 1, the SIH estimates of social assistance benefits are lower than the outlays made by the agencies responsible for payment of the benefits. This is reflected in the comparison between the survey estimates and the corresponding ASNA aggregate, since the ASNA aggregate is based on outlays by the paying agencies. Chart 5 shows that the ratios for the current period estimates from the 1986 and 1990 IDSs are similar, at about 85%. The situation for the 1982 IDS is substantially different however, with the ratio at 78%, and the annual ratio for that year is also substantially different from that of other years. The reason for this is not understood at this point in time.[2] The ratios for the annual estimates from the 1986 and 1990 IDSs are a little higher than those for the SIH. This may reflect the greater ease of recalling the value of benefits received in the previous financial year in the IDSs, since they were conducted in the months September to December, while the SIH was conducted throughout the financial year.

---

[2] Comparison of the current period and annual estimates for 1982 with those of later years might suggest that the 1982 ratios have been inadvertently transposed. However, checking of data has shown this is not the case, and the same phenomenon has been observed independently by other researchers. See Siminski, P., Saunders, P., Waseem, S. and Bradbury, B. 2003, 'Assessing the quality and inter-temporal comparability of ABS household income distribution survey data', Discussion Paper no. 123, Social Policy Research Centre, University of New South Wales.

Chart 5 Social assistance benefits, ratio of survey estimate to ASNA estimates

Per cent

Year ending June:
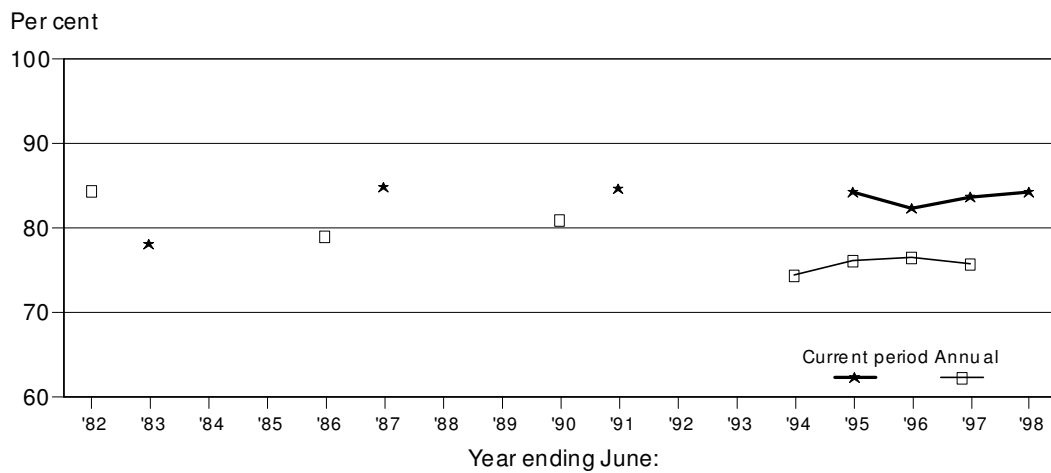
Conclusions to Case Study 2

There are sufficient differences between the IDSs and the SIH to be wary of any time series comparisons between the two, especially in the light of the results from Case Study 1 which showed sensitivity to changes in benchmarking regimes even within the cycles of the SIH. The study comparing IDS and SIH results concluded that the best way to maximise comparability between the IDSs and the SIH was to generate alternate estimates from the 1994-95 SIH, using the exclusion flags and benchmarking regimes of the IDSs to the extent possible. This would not address all the causes of discontinuity between the surveys, but would provide approximate overlapping estimates. However, resources have not been available to undertake this exercise.

CASE STUDY 3: IMPACT OF SIGNIFICANT METHODOLOGY CHANGE TO INCOME SURVEY, 2003-04

Introduction

The 2003-04 Survey of Income and Housing (SIH) introduced a number of major changes designed to improve survey quality and effectiveness. Some of the changes impacted on the comparability between the 2003-04 estimates and earlier data. This case study describes the changes to the survey, presents time series of interest, analyses available data, and summarises the approach taken in presenting the survey results.

Changes introduced in the 2003-04 SIH

A number of changes were introduced that were expected to have improved the quality of the income data produced from the survey, including:

- a larger sample of 11,361 households (comprising 22,315 persons aged 15 and over) for 2003-04 compared with 10,211 households (comprising 19,400 persons aged 15 and over) for 2002-03 (lower sample error)

- interviewer use of a laptop computer instead of a paper form to collect information from respondents (possible improvement in data capture)

- an expanded range of questions to collect details about income - in particular, information was collected about expected income in the current financial year from own unincorporated business and investments, whereas previous "current period" estimates for these components of income were based only on information about reported income for the previous financial year (a significant impact on the coverage of such income streams in current period income measures)

- a comprehensive range of questions to collect details about the assets and liabilities of the household (may have improved the quality of reporting of associated income streams)

- instrument wording changes to explicitly ask that reported dividends include the value of imputation credits - previously this direction was only included in interviewer instructions (clearer specification of data collected).

Other changes introduced in 2003-04 are also likely to have had an impact on the estimates, but not necessarily resulting in an improvement in quality. As explained at the beginning of Case Study 2, from 1994-95 the sample for the SIH was chosen from those households that had just completed eight months inclusion in the Labour Force Survey (LFS). From 2003-04, this approach to selecting the SIH sample has changed. The SIH sample is now taken as an independent sample of households not recently selected in ABS household surveys, that is, it is once again a stand-alone survey in the same way that the IDSs were prior to 1994-95.

A key incentive to making this change was the ability to integrate the SIH with the Household Expenditure Survey (HES). This integration was achieved by selecting a subsample of the households in the SIH and asking them the additional questions required for HES purposes. Total respondent burden was lower than if the two surveys were not integrated. Also, the resultant dataset is richer because HES and SIH results are more comparable than previously. However, response rates for the HES subsample were lower than achieved in the 2003-04 SIH-only subsample, reflecting the reluctance of some respondents to provide the extra information required in the HES part of the survey. There are therefore two non-response issues to be considered.

- Total non-response in the 2003-04 SIH was higher than in 2002-03 because of the impact of the high HES non-response. This is a negative impact on quality. However, encouragingly for the future, it should be noted that the SIH-only subsample of the 2003-04 SIH had a lower non-response rate than the 2002-03 SIH, which was still conducted as an extension of the labour force survey.

- Even if the 2003-04 SIH had not had a HES sub-sample, but was nevertheless still conducted as a stand-alone survey, the non-response would probably have been somewhat different in character to the non-response in the previous years, when the SIH sample was drawn from the LFS sample. The quality of the data is not necessarily better or worse for any given level of non-response, but the results may well be a little different.

Editing and imputation procedures changed in a number of ways in the 2003-04 SIH. There is always a question of the extent to which changes in processing staff and changes in the procedures they adopt in themselves impact on survey results. But in this instance there was a specific known forced change to procedures because LFS records were no longer available to use as input to imputation for SIH non-respondents who had been respondents in the LFS.

Impacts on estimates by source of income

This section examines estimates of income by source for indications of discontinuity in the 2003-04 results.

For wages and salaries, no obvious impacts were detected. Average wages and salaries in the 2003-04 results were 4.8% higher than in 2002-03, in line with the increase in average total weekly earnings reported in ABS business surveys. Calculation of Gini coefficients and quintile shares for wages and salaries showed little change between 2002-03 and 2003-04, that is, there appeared to be no unusual change in the distribution of wage and salary income.

No obvious impacts were detected for social assistance benefit income either. For 2003-04, the coverage of survey reported benefits compared to the aggregate benefits reported by the paying agencies was slightly above the longer term average in cycles from the mid to late 1990s, but within one standard error of that average. Therefore, while a benefit benchmark had been introduced for the 1999-00 and 2000-01 cycles (when coverage fell significantly, see Case Study 1), no benchmark was used in either 2002-03 or 2003-04.

There was a significant impact from the changes in 2003-04 to directly ask about current investment income and current own unincorporated business income, rather than imputing them on a "no change" assumption from reported income for the previous financial year. An indication of the extent of the impact was obtained by examining the current period estimates from the 2002-03 SIH, the 2003-04 previous financial year and current period estimates and the related income estimates from the Australian System of National Accounts (ASNA). Overall, it appeared that the previous methodology had resulted in an underestimate of current period income in 2002-03 for investment income and an overestimate for own unincorporated business income. The latter was due to a slow-down in the Australian economy in 2002-03.

Thus the impacts on current investment income and current own unincorporated business income largely offset each other, and therefore the movements in average total income between 2002-03 and 2003-04 were not affected as much as might have been expected. However, there was a more substantial effect on comparisons between 2003-04 and years prior to 2002-03. Also, as seen in the next section, the new methodology had a discernable impact on the overall income distribution measures.

Income from superannuation, annuities or allocated 'private' pensions (that is, pensions other than government paid social assistance benefits such as the age pension) was substantially higher in 2003-04 than in 2002-03. There were no explicit changes in methodology affecting the reporting of these values. However, it is possible that changes in non-response impacted on the series, or that the reporting of superannuation assets in conjunction with income improved the quality of reporting. The increase in gross weekly household incomes from superannuation etc. of about $7 between 2002-03 and 2003-04 is statistically significant (about 3 standard errors). It is also likely that the reported superannuation assets were an underestimate of the total value of these assets and it may be that the superannuation income series, although higher than previously reported, is still a somewhat conservative measure.

Impacts on measures of income distribution

Inspection of the Gini coefficient and other income distribution measures (table 3) draws attention to the question of whether all the changes have introduced significant discontinuities into the time series of the summary measures. The income distribution measures are the survey output that attract the most attention as time series.

The Gini coefficient declined from 0.309 in 2002-03 to 0.294 in 2003-04, a drop of 0.015. The previous largest annual movement was 0.011, between 1996-97 and 1997-98. The P90/P10 ratio declined by 0.30, with the previous largest annual movement being 0.11. Similarly for the other measures.

Table 3 Indicators of the distribution of equivalised disposable household income

	Gini	P90/P10 (a)	P80/P20 (a)	Share of Q1(b)	Share of D2/D3(c)	Share of Q3(b)	Share of Q5(b)
1994-95	.302	3.77	2.56	7.9	10.8	17.7	37.8
1995-96	.296	3.73	2.58	8.1	11.0	17.7	37.3
1996-97	.292	3.66	2.53	8.3	11.0	17.8	37.1
1997-98	.303	3.77	2.56	7.9	10.8	17.7	37.9
1999-2000	.310	3.89	2.64	7.7	10.5	17.7	38.4
2000-01	.311	3.98	2.63	7.7	10.5	17.6	38.5
2002-03	.309	4.00	2.63	7.7	10.6	17.6	38.3
2003-04	.294	3.70	2.49	8.2	10.9	17.9	37.4

(a) Ratio of percentile values (b) Quintile share (c) Share of 2nd and 3rd deciles. The characteristics of people in the 2nd and 3rd deciles are regarded as being more representative of people with low economic resources because there is strong evidence that many households in the 1st decile with very low reported income have significant other economic resources available to them.

Examination of the huge (in historical terms) decline in measured inequality as shown in table 3 can be approached in two directions. First, try to estimate or find pointers to the impact of any of the effects outlined in the early part of this paper. Second, try to estimate or find pointers to the impact of any specific forces that might have genuinely decreased inequality.

Table 4 illustrates the impacts of those factors for which the summary measures can be re-estimated for 2003-04, and they are discussed further in the following sections.

Table 4 Impact of various influences on indicators of EDHI(a) distribution, 2003-04

	Gini	P90/P10 (b)	P80/P20 (b)	Share of Q1(c)	Share of D2/D3(d)	Share of Q3(c)	Share of Q5(c)
Initial estimate							
1 Based on whole sample	0.294	3.70	2.49	8.2	10.9	17.9	37.4
2 Based on HES subsample	0.290	3.68	2.47	8.3	11.0	18.0	37.1
3 Based on SIH-only subsample	0.301	3.77	2.50	8.2	10.8	17.7	38.0
Impact of changes in survey							
4 Revert to old methodology for current period business and investment income	0.297	3.72	2.49	8.0	10.9	17.9	37.6
5 Assume imputation credit reported in value of dividends	0.298	3.72	2.50	8.0	10.9	17.9	37.6
Impact of real world changes (compare lines 6 and 7 to line 4)							
6 Remove once-only payments to families and carers	0.302	3.81	2.55	7.9	10.7	17.8	37.8
7 Impose personal tax regime of 2002-03	0.297	3.70	2.49	8.1	10.9	17.9	37.5

(a) Equivalised disposable household income (b) Ratio of percentile values (c) Quintile share (d) Share of 2nd and 3rd deciles. The characteristics of people in the 2nd and 3rd deciles are regarded as being more representative of people with low economic resources because there is strong evidence that many households in the 1st decile with very low reported income have significant other economic resources available to them.

Impact of changes in survey

New data and methodology for current period business and investment income

The impact of the introduction of reported current year income for own unincorporated business income and investment income could be quantified precisely because sufficient information was available to replicate the previously used methodology. It increased mean household income by $16.54 (or 1.5%) per week. The impact on the income distribution measures can be seen by comparing lines 1 and 4 of table 4, with the Gini coefficient increasing from 0.294 to 0.297 and P90/P10 increasing from 3.70 to 3.72. However, P80/P20 and the income share of D2/D3 do not change.

While reported current business and investment income may contain an overly optimistic assessment from some respondents (especially those interviewed early in the financial year), the results of the new methodology appeared significantly more credible than the estimates based the old methodology. In particular, it reduced the number of households with gross HH income below $20/wk by 55,000 (69%). Therefore it is believed that the new methodology is significantly superior.

But clearly the introduction of the new methodology leads to a series break. It is not known whether the difference between the two methodologies would be of the magnitude shown here for all years, or whether it would vary substantially with changes in the business cycle.

Integrating SIH and HES, and SIH sample no longer taken from the LFS sample

Lines 2 and 3 of table 4 decompose line 1 into results based on the HES subsample and results from the SIH-only subsample. The differences in the Gini coefficient and the income shares going to Q3 and Q5 were statistically significant at about the 90% confidence level. This implies that there is a high probability that including HES with the SIH had a noticeable impact on the income inequality measures, even though separate analysis found no significant difference on income estimates due to differential non-response between the two subsamples.

There is also likely to have been a substantial impact from moving from a LFS sample based survey to a stand-alone survey. Differences in non-response rates are likely to be part of the story. Non-response in the SIH-only sample of 2003-04 was significantly less than for the 2002-03 SIH and for the 2003-04 HES subsample. However, the high non-response in the HES subsample was not necessarily for the same reasons as the non-response experienced in the previous LFS-based SIHs, and therefore the non-response impact is likely to be different between the two.

Because all the estimates were benchmarked to demographic totals, it has not been possible to get an understanding of the composition and other characteristics of non-response through comparison to independent estimates of the population.

Changes in imputation and editing practices

Hot-decking is used to impute for missing data in the SIH, where sufficient information is available to identify appropriate donor records. Up until 2002-03 hot

decking could often be used to impute data for people that responded to the LFS but became non-respondents in the SIH, since the LFS has a much higher response rate than the SIH. Such imputation was particularly important for lone person and single parent households. Since the 2003-04 survey was conducted independently of the LFS, this particular imputation option was no longer available.

There were no other significant changes in imputation and editing practices.

<u>Assumptions about imputation credits</u>

Prior to 2003-04, the model used to estimate income tax liabilities assumed that reported dividend income excluded imputation credits on the expectation that most respondents would report in that way. In 2003-04, the survey instrument explicitly asked for imputation credits to be included in the reported value of dividend income.

A comparison of lines 4 and 5 in table 4 shows the result of amending the tax model in 2003-04 to be consistent with the 2003-04 survey instrument. In effect, the difference is the discontinuity in the series that would have arisen if respondents did not in practice change their reporting behaviour for this income item, while the change in tax modelling assumed reporting behaviour had changed. As can be seen, the differences were inconsequential for the income distribution measures (but they were more substantial for estimates of mean tax liability).

Impact of real world changes

<u>Once-only payments to families and carers</u>

In 2003-04, the Australian Government made a special once-only payment to people receiving regular social assistance benefits under families assistance and carer assistance programs. Because these payments were targeted as a boost to social assistance they were deemed to be income for SIH purposes. Due to their once-only nature, and since all recipients received the payments at the same time of the year, the value to SIH respondents was modelled rather than collected in the survey. A comparison of lines 4 and 6 in table 4 shows the impact. Without the once-only payments, the Gini coefficient would have been 0.005 higher, P90/P10 would have been 0.009 higher, etc.

<u>Tax changes</u>

In 2003-04, marginal rate thresholds were increased for personal income tax, and there were other changes to other parts of the tax system such as an aged person's rebate. A comparison of lines 4 and 7 in table 2 shows the results of retaining the 2002-03 tax regime for 2003-04 income. There is not a significant impact on income inequality.

Conclusion to Case Study 3

There were substantial changes in the income inequality measures between 2002-03 and 2003-04.

Adjusting the income estimates to retain constant methodology for current year business and investment income reduces the difference in the Gini coefficient to 0.012 and the difference in P90/P10 to 0.28, still greater than any historic change.

On their own, the once-only payments to families and carers have resulted in the Gini coefficient being 0.005 lower than it would otherwise have been, and P90/P10 being 0.09 higher.

However, bringing these two together still leaves a large gap of unexplained change in the inequality measures, and we cannot with confidence say how much of the gap is due to other sources of discontinuity and how much is due to genuine change in inequality.

The indicators now show less inequality than in 1994-95, the first year of the SIH. In presenting these results in survey output, the ABS has said "While it is difficult to assess changes in income distribution over time due to the methodological improvements introduced with the 2003-04 survey, it appears that there has been no significant change in income inequality from the mid 1990s to 2003-04."

When the results from future income surveys become available (2005-06 results will be available shortly), it should become clearer whether there has been a significant break in time series in 2003-04.