



Information Paper

An Introduction to Socio-Economic Indexes for Areas (SEIFA)

Australia

2006 (Preliminary)

Information Paper

An Introduction to Socio-Economic Indexes for Areas (SEIFA)

Australia

2006 (Preliminary)

Brian Pink
Australian Statistician

AUSTRALIAN BUREAU OF STATISTICS

EMBARGO: 11.30AM (CANBERRA TIME) MON 18 FEB 2008

ABS Catalogue No. 2039.0

ISBN 0 642 47936 4

© Commonwealth of Australia 2008

This work is copyright. Apart from any use as permitted under the *Copyright Act 1968*, no part may be reproduced by any process without prior written permission from the Commonwealth. Requests and inquiries concerning reproduction and rights in this publication should be addressed to The Manager, Intermediary Management, Australian Bureau of Statistics, Locked Bag 10, Belconnen ACT 2616, by telephone (02) 6252 6998, fax (02) 6252 7102, or email: <intermediary.management@abs.gov.au>.

In all cases the ABS must be acknowledged as the source when reproducing or quoting any part of an ABS publication or other product.

Produced by the Australian Bureau of Statistics

INQUIRIES

- For further information about these and related statistics, contact the National Information and Referral Service on 1300 135 070 or Jonathon Khoo on Canberra (02) 6252 5506.

PRELIMINARY RELEASE

The full release of the suite of four Socio-economic Indexes for Areas (SEIFA), together with a full information paper and a detailed technical manual, will be made available on Wednesday 26 March 2008. However, in view of strong demand for earlier access to the Index of Relative Socio-economic Disadvantage (IRSD) to enable important and time-critical uses of the IRSD, the ABS has decided to release a preliminary version of the IRSD only.

This preliminary release will be superseded by the full release on 26 March, which will contain more detailed information about SEIFA, its compilation and its uses. We do not expect the actual index values for the IRSD to change between this preliminary release and the final, but this may occur as a consequence of the final validation process that will precede the full release.

The ABS advises users, especially those with no prior experience of SEIFA indexes, to wait for the full release with the complete accompanying documentation. Users also need to note that the IRSD may not necessarily be the most suitable index for all applications.

Recommended reading

Users of the data contained in this preliminary release are advised to carefully read the information which accompanied the 2001 version of SEIFA in conjunction with this paper. However, it is important to note that there are a number of differences between SEIFA 2001 and the preliminary release. These changes are summarised in Section 2.2 of this paper.

The *Technical Paper* (ABS cat. no. 2039.0.55.001) associated with the 2001 release provides a full description of the technical issues, methods used and results for the creation of SEIFA 2001. The 2001 *Information Paper* (ABS cat. no. 2039.0) describes the 2001 indexes and provides some examples of how to use SEIFA.

In 2006, the ABS released a *Methodology Research Paper*: “Socio-economic Indexes for Areas: Introduction, Use and Future Directions” (ABS cat. no. 1351.0.55.015). This paper discusses a number of important features of SEIFA and provides examples of how to use SEIFA effectively in analysis.

1. WHAT ARE THE SOCIO-ECONOMIC INDEXES FOR AREAS?

The Socio-economic Indexes for Areas (SEIFA) are a set of four summary measures of relative socio-economic status at a small area level. The indexes are created from a wide range of variables collected in the 2006 Census of Population and Housing.¹ These variables summarise a range of characteristics of all the people living in the area. The indexes attempt to identify and rank areas where a high proportion of people are relatively more, or less, disadvantaged. This means that the indexes can provide contextual information about the area in which a person lives.

There are a number of factors related to socio-economic status which the indexes do not represent well. First, the indexes contain only limited information about accumulated wealth and health status. Second, an area's infrastructure such as schools, community services, shops and transport is not represented by the indexes. Third, the indexes do not capture the difference in cost of living across different areas. The Census of Population and Housing does not collect information about these three factors, and so they cannot be included in the construction of the SEIFA indexes.

This preliminary release contains only one of the four SEIFA indexes: the Index of Relative Socio-economic Disadvantage. This index focuses on low income earners, relatively low educational attainment, high unemployment and other variables reflecting disadvantage.

The full release of SEIFA will also contain the following indexes:

- Index of Relative Socio-economic Advantage/Disadvantage – looks at the whole continuum of advantage to disadvantage;
- Index of Economic Resources – focuses specifically on financial aspects of advantage and disadvantage; and the
- Index of Education and Occupation – includes education and occupation variables only.

SEIFA is an area level measure – in other words, it is a summary measure of *all* people living in the area. The socio-economic status of individuals and families can be very diverse within an area. This means that there is a risk of making incorrect conclusions if SEIFA indexes are used as a proxy for individual or family level disadvantage, rather than as a measure of area level disadvantage.

¹ SEIFA 2006 is created using a method called Principal Components Analysis. Further details on these methods can be found in *ABS Technical Paper: Census of Population and Housing: Socio-economic Indexes for Areas, Australia, 2001* (ABS cat. no. 2039.0.55.001). For this Preliminary IRSD, the only methodological difference to the 2001 IRSD is the use of a loading cutoff of ± 0.3 . This is consistent with the construction of the other three SEIFA indexes.

1.1 Geographic areas available

This preliminary release of SEIFA only contains information for Census Collection Districts (CDs), of which there are almost 40,000 across Australia. CDs are the unit of analysis used to create SEIFA. In the full release, SEIFA will also be available for:

- Postal Areas (POAs);
- Statistical Local Areas (SLAs); and
- Local Government Areas (LGAs).

For more information on these geographical areas, please see the *Australian Standard Geographic Classification* (ABS cat. no. 1216.0) and the *Census Geographic Areas* (ABS cat. no. 2905.0).

1.2 Numbers used in SEIFA

For each index, SEIFA gives every geographic area:

- scores – a lower score indicates that an area is relatively disadvantaged compared to an area with a higher score. Scores should ideally be used in distributive analysis. To enable easy recognition of high and low scores, the CD index scores have been standardised to have a mean of 1,000 and a standard deviation of 100 across all CDs in Australia.
- ranks – all areas are ordered from the lowest to highest score, then the area with the lowest score is given a rank of 1, the area with the second lowest score is given a rank of 2 and so on, up to the area with the highest score is given the highest rank (the highest rank is 37,457 for CDs).
- deciles – again all areas are ordered from lowest to highest score. The lowest 10 per cent of areas are given a decile number of 1 and so on, up to the highest 10 per cent of areas which are given a decile number of 10. This means that areas are divided up into ten groups, depending on their score.

The indexes have been constructed so that areas with high proportions of relatively disadvantaged people have lower index scores, ranks and deciles. When using the indexes, we recommend grouping areas into quantiles (e.g. deciles or quintiles), then using these quantiles as the basis for analysis.

The index scores, ranks and deciles are all ordinal. They can only be used rank areas and to indicate whether people living in one area tend to be more, or less, relatively disadvantaged than people in another area. However, the difference between scores has no real meaning, so we cannot say that an area with a score of 500 is twice as disadvantaged as an area with a score of 1,000.

1.3 Areas without a SEIFA score

We cannot always give an area a score. For SEIFA 2006, around 3% of CDs could not be given a score. There are a number of reasons for this. For example, the area could be an airport or a large office block, where either no-one, or very few people, usually live in the area. If only a few people responded to Census questions it becomes difficult to calculate a reliable score for the area, since those who did respond may not be representative of the area as a whole. Areas were not given a score if they fell into one of the following categories:

- Number of people usually living in the area was ten or fewer;
- Number of employed people was five or fewer;
- Proportion of applicable people not responding to the following Census questions was 70% or more: occupation, labour force status, type of educational institution attending, or non-school qualifications;
- Proportion of households where equivalised household income could not be calculated was 70% or more;
- Number of occupied private dwellings was five or fewer;
- Proportion of people usually living in non-private dwellings was 80% or more; or
- The area was classified as off-shore, shipping or migratory.

2. PRELIMINARY INDEX OF RELATIVE SOCIO-ECONOMIC DISADVANTAGE

The Preliminary Index of Relative Socio-economic Disadvantage (Preliminary IRSD) summarises a wide range of information on disadvantage, including low income, little education, high unemployment and unskilled occupations. An area could have a low score if there are many households in the area with low income, many people with no qualifications, or many people employed in unskilled occupations.

The index is designed to only focus on aspects of disadvantage. A low score on this index indicates a high proportion of disadvantaged people in an area. We cannot say that an area with a very high score has a large proportion of advantaged people, as there are no variables in the index to indicate this. We can only say that such an area has a relatively low incidence of disadvantage.

The preliminary IRSD is constructed as a weighted average of selected Census variables. Variables used in the index had to be available in the 2006 Census, and the index is dependent on the set of variables chosen for the analysis. If we had chosen a different set of variables, we would have created a different index. At the same time, because of the large number of variables in the index, removing or altering one variable will not usually have a large effect. The choice of different input variables leads to the creation of four SEIFA indexes which capture slightly different aspects of relative socio-economic advantage and disadvantage.

Before using the preliminary IRSD, you should consider the aspect of socio-economic status you are interested in, and examine the underlying set of variables included in the index. This will allow you to make an informed decision as to whether the preliminary IRSD is appropriate for your particular needs.

Table 2.1 describes each of the variables included in the Preliminary IRSD. The associated weights for each variable are also included in the table. In general, variables with higher loadings will contribute more to the index.

2.1 List of variables used for the Preliminary Index of Relative Socio-economic Disadvantage and their weights

<i>Variable description</i>	<i>Variable weight</i>
% Occupied private dwellings with no internet connection	–0.33
% Employed people classified as Labourers	–0.30
% People aged 15 years and over with no post-school qualifications	–0.30
% People with stated “ low ” household equivalised income (annual) of between \$13,000 and \$20,799 (approx. 2nd and 3rd deciles)*	–0.30
% Households renting from a Government or Community organisation	–0.27
% People (in the labour force) unemployed	–0.27
% Families that are one parent families with dependent children only	–0.26
% Households paying a “low” rent of less than \$120 per week (excluding \$0 per week)	–0.26
% People aged under 70 who have a long-term health condition or disability and need assistance with core activities	–0.24
% Occupied private dwellings with no car	–0.22
% People who identified themselves as being of Aboriginal and/or Torres Strait Islander origin	–0.20
% Occupied private dwellings requiring one or more extra bedrooms (based on Canadian National Occupancy Standard)	–0.20
% People aged 15 years and over who are separated or divorced	–0.20
% Employed people classified as Machinery Operators and Drivers	–0.20
% People aged 15 years and over who did not go to school	–0.17
% Employed people classified as Low Skill Community and Personal Service Workers	–0.17
% People who do not speak English well	–0.13

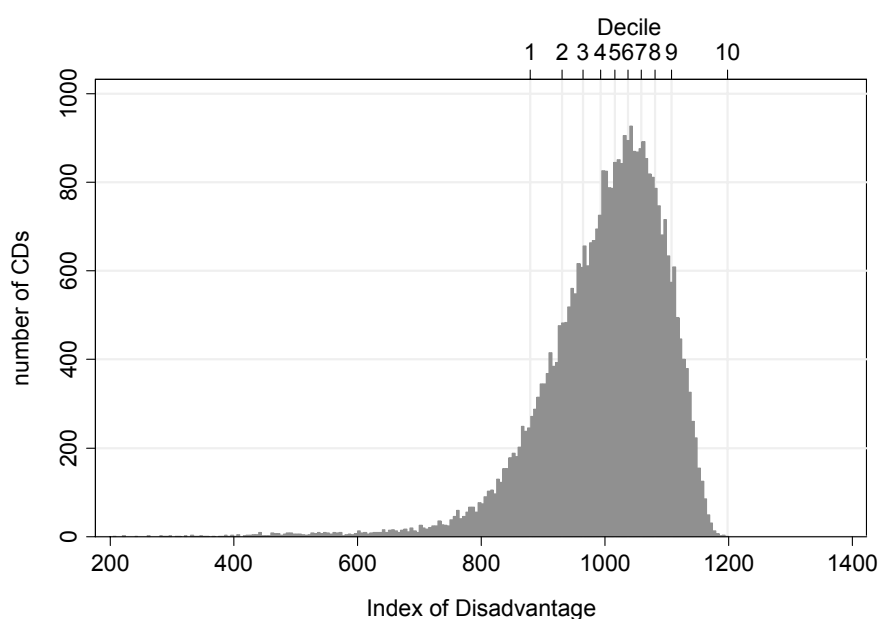
* The second and third equivalised income deciles are used in line with ABS standards. For more information see Explanatory notes 24-27 from *Household Income and Income Distribution, Australia, 2005-06* (ABS cat. no. 6523.0).

2.1 Distribution of scores

Figure 2.2 shows a frequency distribution for the preliminary IRSD. Each vertical bar represents the number of CDs within a range of five index points. The values range from around 200 to around 1200. However, the distribution has a very long left tail, and is left-skewed (i.e. the mean is lower than the median). This is because the index contains only disadvantage indicators, so there is more scope to distinguish between disadvantaged areas than between areas with relatively low levels of disadvantage.

Users should consider the features of this distribution when deciding how to use SEIFA in any form of analysis.

2.2 Preliminary Index of Relative Socio-economic Disadvantage distribution



2.2 Changes between 2001 and 2006

As a general rule, every effort is made to keep SEIFA the same as the previous release. However, some changes are important, or unavoidable.

New variables

Some new variables were introduced into the 2006 Census. One new Census variable included in the preliminary IRSD is:

- % People aged under 70 who have a long-term health condition or disability and need assistance with core activities.

Other variables included in the preliminary IRSD for the first time are:

- % Occupied private dwellings requiring one or more extra bedrooms (based on Canadian National Occupancy Standards); and
- % Households paying rent who pay less than \$120 per week.

Changed Census variables

Some variables collected in the 2006 Census were different to the corresponding variables collected in the 2001 Census. Some of these changes are minor, others are more important for SEIFA.

One of the largest changes was the change in classification standards for occupations. Occupation variables are now classified using *ANZSCO – Australian and New Zealand Standard Classification of Occupations, First Edition, 2006* (ABS cat. no. 1220.0).

Under this classification, each occupation is assigned a skill level. We used a combination of the ANZSCO major groups and skill levels to construct the 2006 SEIFA occupation variables.

Other major changes

A significant change to SEIFA in 2006 is the use of equivalised household income instead of (unequivalised) family income. Equivalised income takes into account the family structure of the household, such as the number of adults and children. A more detailed explanation of equivalised income is given in appendix 3 of *Household Income and Income Distribution, Australia* (ABS cat. no. 6523.0).

Previous SEIFA were created using Census information about people who were in an area on Census night. For 2006, SEIFA instead uses information about people who usually reside in the area. This ensures that the indexes reflect the characteristics of people living in the area, rather the characteristics of visitors to the area. This change removes the effects of unusual population movements at Census time, such as the large numbers of holiday makers visiting the New South Wales and Victorian ski fields.

The 2006 SEIFA indexes will be different to previous versions. We do not recommend comparisons over time of ranks or scores using different versions of SEIFA. A great deal of care needs to be taken when undertaking this kind of analysis. Conclusions need to take account of area boundary changes, changes to the variables used, and the way the relationship between variables might have changed.

2.3 Aggregating to higher geographical areas

Often, researchers want SEIFA scores at more aggregated geographical areas, either because their own data do not contain CD level information, or because the number of cases in the analysis at the CD level is small. To construct indexes for geographies higher than CD level, we use a population weighted average of the constituent CDs. This method can be applied using the following formula:

$$INDEX_{AREA} = \frac{\sum_{i=1}^n (INDEX_{CDi} \times POP_{CDi})}{POP_{AREA}}$$

where:

INDEX = index score for each CD or higher level area

POP = population for each CD or higher level area

n = total number of CDs (with SEIFA scores) in the higher level area

As the size of an area increases, it becomes correspondingly more heterogeneous and the socio-economic index becomes less meaningful. Using a simple average as described often masks differences in disadvantage that are present in smaller areas. To analyse the socio-economic differences between large areas, we recommend observing the distribution of CD scores within each area.

Alternative aggregation methods were investigated for SEIFA 2006. Due to a number of conceptual and technical issues, population weighted aggregation will be used for SEIFA 2006 (as described above).

FOR MORE INFORMATION . . .

<i>INTERNET</i>	www.abs.gov.au the ABS website is the best place for data from our publications and information about the ABS.
<i>LIBRARY</i>	A range of ABS publications are available from public and tertiary libraries Australia wide. Contact your nearest library to determine whether it has the ABS statistics you require, or visit our website for a list of libraries.

INFORMATION AND REFERRAL SERVICE

Our consultants can help you access the full range of information published by the ABS that is available free of charge from our website, or purchase a hard copy publication. Information tailored to your needs can also be requested as a 'user pays' service. Specialists are on hand to help you with analytical or methodological advice.

<i>PHONE</i>	1300 135 070
<i>EMAIL</i>	client.services@abs.gov.au
<i>FAX</i>	1300 135 211
<i>POST</i>	Client Services, ABS, GPO Box 796, Sydney NSW 2001

FREE ACCESS TO STATISTICS

All statistics on the ABS website can be downloaded free of charge.

<i>WEB ADDRESS</i>	www.abs.gov.au
--------------------	-----------------------



2203900013066

ISBN 0 642 47936 4

RRP \$11.00

© Commonwealth of Australia 2008
Produced by the Australian Bureau of Statistics

.....