2039.0

Information Paper

Census of Population and Housing

Socio-Economic Indexes for Areas

Australia

2001

Information Paper

Census of Population and Housing

Socio-Economic Indexes for Areas

Australia

2001

Dennis Trewin
Australian Statistician

CONTENTS ....

ABBREVIATIONS ....

ABS	Australian Bureau of Statistics
ACT	Australian Capital Territory
ASGC	Australian Standard Geographical Classification
Aust.	Australia
CD	Collection District
CED	Commonwealth Electoral Division
HES	Household Expenditure Survey
IRDB	Integrated Regional Database
ISDR	indirect standardised death rate
LGA	Local Government Area
NSW	New South Wales
NT	Northern Territory
POA	Postal Area
Qld	Queensland
S/T	State or Territory
SA	South Australia
SD	statistical division
SED	State Electoral Division
SEIFA	Socio-Economic Indexes for Areas
SLA	Statistical Local Area
SSC	State Suburbs
SSD	Statistical Subdivision
Tas.	Tasmania
TAFE	Technical and Further Education
Vic.	Victoria
WA	Western Australia

INTRODUCTION ....

INTRODUCTION

This publication describes four summary measures, or indexes, derived from the 2001 Census of Population and Housing to measure different aspects of socio-economic conditions by geographic areas.

The 2001 Census of Population and Housing provides information on a broad range of social and economic aspects of the Australian population. Nearly fifty questions of social and economic interest are asked in the census. People using census data are often interested not just in these items taken one at a time, but in an overview or summary of a number of related items. Statistical techniques can be used to provide such summaries and the indexes presented in this publication are one type of summary measure.

A measure of socio-economic disadvantage was first produced by the Australian Bureau of Statistics (ABS) following the 1971 census. The Socio-Economic Indexes for Areas (SEIFA), in their present form, were first produced in 1990 and consisted of five indexes formed from 1986 census data. In 1994, five indexes were produced from 1991 census data using essentially the same methodology as in 1990. Indexes using the same methods were also created from the 1996 census.

For the 2001 indexes, a review of the indexes and methods used to construct the indexes was conducted. This involved extensive user consultation. One of the outcomes of this review was a new variable selection strategy. Previous SEIFA indexes used variables chosen in 1990. These variables were chosen after extensive consultation with academics and other potential users. The choice of variables was based largely on people's experience with other indexes, rather than referring to a theoretical model.

For the 2001 review, we decided to develop a new variable selection strategy for SEIFA. The strategy was based on a theoretical model of disadvantage taken from our reading of the relevant literature. This theoretical model starts with core variables that represent socio-economic status. These 'first level' variables are Education, Income and Occupation. These variables are always included in the SEIFA indexes because they are fundamental to measuring socio-economic status.

The 'second level' of variables were those that measured aspects of disadvantage. These aspects of disadvantage relate to things like wealth, living conditions, and access to services. Examples include number of bedrooms in a house, whether the house is owned or rented, and access to the Internet.

The 'third level' of variables are those that are often associated with disadvantage generally, rather than a specific aspect of disadvantage. Some components of the disadvantage may have already been captured by level one and level two variables. Level three variables have been included where it appeared that some additional aspect of disadvantage still remained to be measured over and above that from level one and two variables. An example of a level three variable is a high percentage of Indigenous people

INTRODUCTION *continued*

in an area. Level three variables should not be interpreted as causing disadvantage, they are better thought of as indicators which signal that an area has some disadvantage.

Using the variable selection strategy, we can sort the indexes in descending order of their coverage of aspects of advantage/disadvantage.

In 1996, there were five indexes; the Index of Disadvantage; the Urban Index of Advantage; the Rural Index of Advantage; the Index of Economic Resources and the Index of Education and Occupation. The Index of Disadvantage and the Indexes of Advantage had the broadest coverage of advantage/disadvantage, using variables that both measure and reflect advantage/disadvantage (i.e., level one to level three variables). The Index of Economic Resources and the Index of Education and Occupation were targeted towards specific aspects of advantage/disadvantage. The Index of Education and Occupation uses level one variables only. The Index of Economic Resources uses level one and level two variables to get a measure of economic disadvantage.

In 2001, there are four indexes. Again, the most general index is the Index of Disadvantage. This index includes all variables that either reflect or measure disadvantage (level one to level three variables). The inclusion of level three variables means that while it may reflect an area's disadvantage, it is not possible to identify all aspects of disadvantage being represented. Of all the 2001 indexes, this index is most comparable to its 1996 counterpart. It uses the same method, and the same variables as the 1996 Index of Disadvantage.

We have replaced the Urban and Rural Indexes of Advantage with an Index of Advantage/Disadvantage. This index is used to rank a Collection District (CD) in terms of both advantage and disadvantage. Any information on advantaged persons in an area will offset information on disadvantaged persons in the area.

The other two indexes—the Index of Economic Resources and the Index of Education and Occupation—fulfill the same purpose as their 1996 forerunners.

This Information Paper describes the indexes, illustrates their possible uses and describes the range of ABS geographic areas for which they are available. The approach for the construction of the indexes is outlined with a discussion of the limitations concerning their use. The socio-economic indexes which are available are described, and information is provided on how they can be obtained. A companion volume, the *Census of Population and Housing: Socio-Economic Indexes for Area's (SEIFA), Australia - Technical Paper* (cat. no. 2039.0.55.001), is also available. This manual outlines the method and results in detail, including detailed results from the principal component analysis; some results from the validation; and some analysis of the indexes for remote areas.

CHAPTER **2**   **THE INDEXES** ....

2.1 DESCRIPTION OF THE
INDEXES

There are four indexes described in this Information Paper. They relate to
socio-economic aspects of geographic areas. Each index summarises a different aspect of
the socio-economic conditions in an area. The indexes have been obtained by a
technique called principal components analysis. This technique summarises the
information from a variety of social and economic variables, calculating weights that will
give the best summary for the underlying variables. For the SEIFA indexes, each index
uses a different set of underlying variables.

The four indexes are:
- Index of Relative Socio-Economic Disadvantage
- Index of Relative Socio-Economic Advantage/Disadvantage
- Index of Economic Resources
- Index of Education and Occupation.

All the indexes (including the Index of Relative Socio-Economic Disadvantage) have been
constructed so that relatively disadvantaged areas (e.g. areas with many low income
earners) have low index values.

The Index of Relative Socio-Economic Disadvantage is derived from attributes such as
low income, low educational attainment, high unemployment, jobs in relatively unskilled
occupations and variables that reflect disadvantage rather than measure specific aspects
of disadvantage (e.g., Indigenous and Separated/Divorced).

High scores on the Index of Relative Socio-Economic Disadvantage occur when the area
has few families of low income and few people with little training and in unskilled
occupations. Low scores on the index occur when the area has many low income families
and people with little training and in unskilled occupations. It is important to understand
that a high score here reflects lack of disadvantage rather than high advantage, a subtly
different concept.

To maintain consistency with the other indexes, the higher an area's index value for the
Index of Relative Socio-Economic Disadvantage, the less disadvantaged that area is
compared with other areas. For example, an area that has a Relative Socio-Economic
Disadvantage Index value of 1200 is less disadvantaged than an area with an index value
of 900.

For the 1996 census, there were two indexes of advantage; the Urban and Rural Indexes
of Advantage. During the review of SEIFA, we found that these indexes were not widely
used; partly because different weights between urban and rural areas made them
incomparable; and partly because the same variables were being used for urban and rural
areas although with different weights (in reality, different variables measure advantage in
urban and rural areas). Our approach for the 2001 indexes has been to concentrate on an

2.1 DESCRIPTION OF THE INDEXES *continued*

Australia-wide Index of Relative Socio-Economic Advantage/Disadvantage, which we hope will be more useful.

A higher score on the Index of Relative Socio-Economic Advantage/Disadvantage indicates that an area has attributes such as a relatively high proportion of people with high incomes or a skilled workforce. It also means an area has a low proportion of people with low incomes and relatively few unskilled people in the workforce. Conversely, a low score on the index indicates that an area has a higher proportion of individuals with low incomes, more employees in unskilled occupations, etc.; and a low proportion of people with high incomes or in skilled occupations.

The Index of Economic Resources reflects the profile of the economic resources of families within the areas. The census variables which are summarised by this index reflect the income and expenditure of families, such as income and rent. Additionally, variables which reflect wealth, such as dwelling size, are also included. The income variables are specified by family structure, since this affects disposable income.

This index excludes Education and Occupation variables because they are not direct measures of economic resources. It also misses some assets such as savings or equities which, although relevant, could not be included because the information was not collected in the 2001 census.

A higher score on the Index of Economic Resources indicates that the area has a higher proportion of families on high income, a lower proportion of low income families, and more households living in large houses i.e. four or more bedrooms. A low score indicates the area has a relatively high proportion of households on low incomes and living in small dwellings.

The Index of Education and Occupation is designed to reflect the educational and occupational structure of communities. The education variables in this index show either the level of qualification achieved or whether further education is being undertaken. The occupation variables classify the workforce into the major groups of the Australian Standard Classification of Occupations (ASCO) and the unemployed. This index does not include any income variables.

An area with a high score on this index would have a high concentration of people with higher education qualifications or undergoing further education, with a high percentage of people employed in more skilled occupations. A low score indicates an area with concentrations of either people with low educational attainment, people employed in unskilled occupations, or the unemployed.

Appendix 1 lists the variables summarised by the four indexes, including the weights for each variable. The method for deriving the indexes is briefly described in Chapter 3. The *Census of Population and Housing: Socio-Economic Indexes for Area's (SEIFA), Australia - Technical Paper* (cat.no. 2039.0.55.001), contains a more detailed description of this method. Factors to be taken into account when interpreting the indexes are discussed in Chapter 4.

2.2 AVAILABLE GEOGRAPHIC AREAS

The *Australian Standard Geographical Classification (ASGC)* (cat. no. 1216.0) is the primary geographical classification used by the ABS for the collection and dissemination of geographically classified statistics. The Census Geographic Areas are designed to provide census data for administrative areas which are outside the scope of the ASGC.

The four SEIFA index scores are available for a range of ABS geographic areas from the ASGC 2001 and the Census Geographic Areas 2001.

Geographic areas used from the ASGC 2001 are:
- Collection District (CD)
- Statistical Local Area (SLA)
- Statistical Subdivision (SSD)
- Statistical Division (SD)
- State/Territory (S/T)
- Local Government Area (LGA).

For detailed information about the ASGC 2001 refer to the ABS publication *Statistical Geography: Volume 1 — Australian Standard Geographical Classification (ASGC), 2001* (cat. no. 1216.0).

Geographic Areas used from the Census Geographic Areas 2001 are:
- Postal Area (POA)
- State Suburbs (SSC)
- State Electoral Division (SED)
- Commonwealth Electoral Division (CED).

For detailed information about the Census Geographic Areas 2001 refer to the ABS publication *Statistical Geography: Volume 2 — Census Geographic Areas Australia, 2001* (cat. no. 2905.0).

The CD is the smallest geographic area of both the ASGC and Census Geographic Areas and is defined only in a census year. CDs aggregate to the higher level geographic areas of the ASGC and the Census Geographic Areas listed above. It is the smallest geographic area for which the SEIFA indexes are available. For the 2001 Census of Population and Housing, there were 37,209 CDs defined throughout Australia. In urban areas CDs comprise on average about 220 dwellings, while in rural areas they usually contain fewer.

SEIFA index scores for CDs have been used to calculate scores for the larger geographic areas by taking the weighted average, using population counts from the 2001 census, across all CDs which comprised the larger geographic area. A short description of each of the geographic areas follows:
- SLAs are based on the boundaries of incorporated bodies of local government where these exist. Where there is no incorporated body of local government, SLAs are defined to cover the unincorporated areas. An SLA consists of one or more whole CDs. SLAs cover, in aggregate, the whole of Australia without gaps or overlaps. SLAs do not cross state/territory borders.
- SSDs consist of one or more SLAs. They cover, in aggregate, the whole of Australia without gaps or overlaps and do not cross state/territory borders.
- SDs consist of one or more SSDs. They are the largest geographic area of the ASGC within each state/territory. They aggregate to form the states/territories.

2.2 AVAILABLE GEOGRAPHIC AREAS

continued

- State/Territory is the largest geographic area of the ASGC. For the purposes of the ASGC, Australia is divided into six states and three territories.
- LGAs are the cornerstone of the ASGC and are proclaimed by the various state/territory governments. These legally designated areas, in aggregate, cover only the incorporated part of Australia. There is a close relationship between the LGA and SLA in that LGAs comprise one or more SLAs.

Census Geographic Areas 2001 are approximations of official administrative areas defined by organisations other than the ABS. POAs are an approximation of the postcodes defined by Australia Post, Electoral Divisions (CEDs, SEDs) are an approximation of the divisions defined by the Australian Electoral Commission and the respective state Electoral Commissions. Census Geographic Areas are created by allocating each whole CD to one area in each Census Geographic Area classification. Should a CD share area with more than one official administrative area, a decision is made as to its final allocation. The allocation is based upon which administrative area is deemed to contain the majority of the population of the CD.

- POAs are groupings of whole CDs which have been assigned the same Australia Post postcode. Only the street delivery postcodes are used in the allocation.
- SSCs are groupings of whole CDs which have been assigned to localities gazetted by the Geographic Place Name authority of each state/territory. The classification does not cover the whole of Australia.
- SEDs are areas legally prescribed for the purpose of returning one or more members to the State or Territory Lower Houses of Parliament. The boundaries and census statistics produced for SEDs are CD derived.
- CEDs are areas legally prescribed for the purpose of returning one member to the Federal Lower House of Parliament. The boundaries and census statistics produced for CEDs are CD derived.

The index scores for the above ASGC and Census Geographic Areas were calculated by taking the weighted average, using population counts from the 2001 census, across all CDs which comprised the larger geographic area. The use of a weighted average to aggregate the indexes is discussed in more detail in the *Census of Population and Housing: Socio-Economic Indexes for Area's (SEIFA), Australia - Technical Paper* (cat.no. 2039.0.55.001).

Index values for other regions may also be derived. These values are based on the index score of the CDs which make up the region. Each CD score is multiplied by its census population count. The total score of all the CDs in the region is divided by the total regional population count. Population counts by CD and the CD index scores have been provided to enable weighted index scores to be calculated for user-defined regions.

The SEIFA indexes are not intended to be used for comparison of individual CDs. It is possible for index values at the CD level to be distorted by unusual characteristics. Therefore, indexes for CDs are provided in order to construct indexes for larger geographic areas, where the index values will be more stable.

2.3 DISTRIBUTION OF
THE INDEX VALUES

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. In practice, this means that around 95% of index scores are between 800 and 1,200. This has not been done for indexes aggregated to a larger geography.

Several tables of summary statistics have been provided in Appendix 2 to help give an intuitive understanding of what is a high value and what is a low value for the different indexes.

In the SEIFA 2001 product and in the tables in Appendix 2, the distribution of index values has been summarised by using quantiles or percentiles. Quantiles denote a point in the distribution of index values below which a specified percentage of index values fall. Quantiles which divide the distribution of index values into ten equal parts are commonly referred to as deciles; quantiles which divide the distribution of index values into four equal parts are commonly referred to as quartiles, with the 50% quartile also known as the median. Quantiles which divide the distribution of index values into five equal parts are known as quintiles.

For example, the 10% quantile for a state gives the value below which 10% of the index values for that state lie. To illustrate this, reproduced below is an excerpt from table 1 in Appendix 2. The 10% quantile for the Index of Relative Socio-Economic Advantage/Disadvantage for CDs in New South Wales, is 882. This means that one-tenth of CDs in New South Wales have a score on the Index of Relative Socio-Economic Advantage/Disadvantage at or below 882. Similarly, the 50% quantile or median for the Index of Relative Socio-Economic Advantage/Disadvantage for CDs in Western Australia is 994. This means that 50% of CDs in Western Australia have a score on the Index of Relative Socio-Economic Advantage/Disadvantage at or below 994.

SUMMARY DATA FOR CD LEVEL INDEX OF RELATIVE SOCIO-ECONOMIC ADVANTAGE/DISADVANTAGE

|		*Quantile*				
State/territory	Average	10%	25%	50%	75%	90%
NSW	1 011	882	932	999	1 085	1 165
Vic.	1 011	889	942	1 003	1 080	1 147
Qld	980	878	919	971	1 035	1 104
SA	973	859	907	966	1 036	1 108
WA	1 001	886	936	994	1 067	1 127
Tas.	943	831	881	935	1 007	1 073
NT	987	812	914	1 010	1 074	1 114
ACT	1 120	1 047	1 082	1 117	1 167	1 208
Aust.	**1 000**	**880**	**928**	**990**	**1 068**	**1 141**

2.3 DISTRIBUTION OF
THE INDEX VALUES
continued

The distributions of index scores are generally similar across the states. Most noticeable are the different distributions observed for the Northern Territory, the Australian Capital Territory and Tasmania. The graph below shows the comparison of Index of Relative Socio-Economic Advantage/Disadvantage values for the Northern Territory, the Australian Capital Territory, Tasmania and Australia. Interesting points to note are that the Northern Territory has a greater proportion of CDs with lower SEIFA values, but after a value of about 930, the distribution broadly follows the Australian distribution. This is similar to the 1996 distribution, see *Information Paper: Census of Population and Housing Socio-Economic Indexes for Areas, Australia, 1996* (cat. no. 2039.0). In 2001, Tasmania is more disadvantaged than Australia at all points. In 2001 it is the most disadvantaged state for three of the four indexes. The Australian Capital Territory has a higher proportion of CDs with high SEIFA values.

Another way to interpret this graph is to consider the 50% quantile or median point. This is the cutoff point below which 50% of index values lie. For CDs in Australia the 50% quantile for the Index of Relative Socio-Economic Advantage/Disadvantage is 990, for CDs in Tasmania it is lower at 935, for CDs in the Northern Territory it is also slightly higher at 1,010, and for CDs in Australian Capital Territory it is higher at 1,117. The median points for each state are shown in table 1, Appendix 2.

INDEX OF RELATIVE ADVANTAGE/DISADVANTAGE
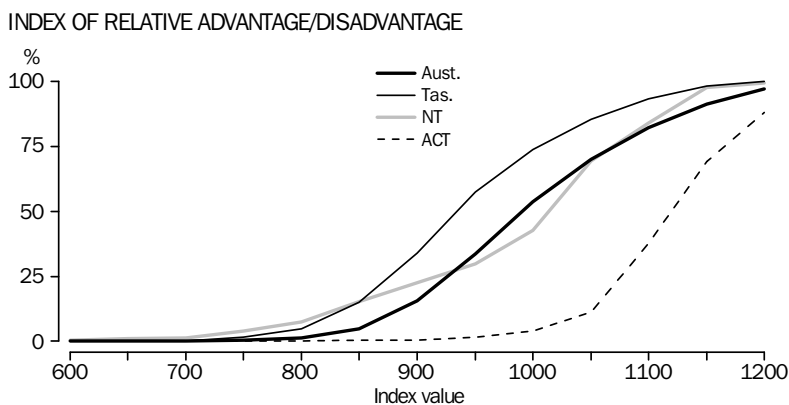
Note: % is cumulative percentage of people living in CDs below index value.

The indexes are 'ordinal measures' and not 'interval measures'. That is, the indexes can be used to order areas in terms of disadvantage; but any other arithmetic relationships between index values may not be meaningful. For example, a CD with an index value of 1,200 does not have twice the wellbeing of a CD with an index value of 600. Similarly, the socio-economic difference between two CDs with index values of 800 and 900, is not necessarily the same as the difference between two CDs with index values of 1,050 and 1,150. A technical explanation of this is given in the *Census of Population and Housing: Socio-Economic Indexes for Area's (SEIFA), Australia - Technical Paper* (cat.no. 2039.0.55.001).

2.4 APPLICATIONS OF
THE INDEXES

There are a number of ways the indexes can be used, such as targeting areas for business or services, 'demographic profiling', strategic planning, design of sample surveys, and social or economic research.

**2.4 APPLICATIONS OF
THE INDEXES** *continued*

The most appropriate index should be chosen for each application. The variables that contribute to each index should be considered when deciding which index to use. For example, if a user is interested in finding areas of disadvantage for allocation of services and they have used the 1996 Index of Disadvantage to do this before, they will probably want to use the Index of Relative Socio-Economic Disadvantage. On the other hand, if a user wanted to focus on finding areas containing relatively high proportions of people in unskilled jobs or with low levels of educational qualifications, the Index of Education and Occupation should be used. A full list of the variables included in each index is provided in Appendix 1.

Some examples of uses of the indexes are described below.

*2.4.1 Uses in
research/data analysis*

The indexes may be useful for modelling or explaining behaviour in other variables. In some studies it is desirable to determine if socio-economic factors are influencing a variable of interest. The researcher may also be interested in reducing the number of variables in the analysis. In such cases, one or more of the indexes can be used as a summary of a range of socio-economic factors.

Example:
A health researcher may be investigating death rates in different areas across Australia. He or she might choose to use an Indirect Standardised Death Rate (ISDR) for each SSD across Australia, (these are published by the ABS in *Mortality Atlas, Australia* (cat. no. 3318.0)). The researcher might wish to compare death rates in SSDs at each SSD's relative socio-economic standing.

In this example, the Index of Relative Socio-Economic Advantage/Disadvantage has been calculated for each SSD and plotted against the ISDR, to see if the ISDR is higher in disadvantaged areas. The result is shown in the graph below.

MORTALITY VARIATION AMONG STATISTICAL SUBDIVISIONS
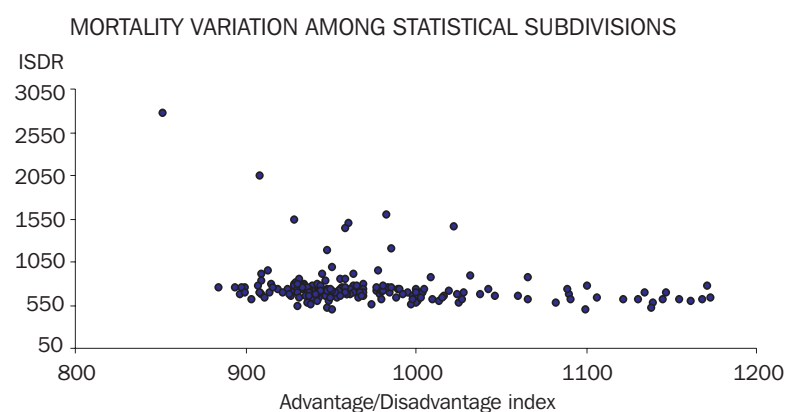
2.4.1 Uses in research/data analysis continued

The graph shows that more highly disadvantaged areas (those with lower index values) tend to have a higher ISDR than advantaged areas, although the relationship does not seem to be particularly strong. To gain a more complete understanding of the relationship, researchers might remove the two outlying SSDs, then investigate separately the relationship between SSDs with high ISDRs and SSDs with low ISDRs. For example, SSDs could be divided into those with index values above 950 and those with index values below 950 and separate graphs plotted, or a regression analysis done, to understand more fully relationships between socio-economic advantage/disadvantage and the ISDR.

2.4.2 Targeting areas for services

The indexes are of interest in their own right as summaries of area characteristics. Areas with different index values have different socio-economic characteristics. This information can be used by itself or in conjunction with other (more targeted) information to assist in determining the allocation of services.

Example:

A government agency responsible for funding of aged care facilities wants to ensure resources go to those localities which need them most. It decides to allocate funds to areas with low ratios of existing aged care places to population aged 70 years and over.

In reviewing the allocation of funding, the agency used the 1996 Disadvantage Index to check that relative socio-economic disadvantage localities have been given adequate funding. They want to recheck the distribution using the 2001 Index of Disadvantage, which is calculated along the same lines as the 1996 Disadvantage Index.

SLAs across Australia could be divided into five quintiles according to their ranking in the Index of Relative Socio-Economic Disadvantage. The ratio of existing aged care places to population aged 70 years and over could be calculated for each quintile. For example, the lowest quintile would contain the 20% of SLAs in Australia with the lowest disadvantage index values.

The average ratio of aged care places to population aged 70 years and over for each quintile could be graphed to look for systematic bias in aged care place funding with respect to socio-economic disadvantage.

Note that in this example SEIFA is being used to check that there is no systemic bias in the funding allocation, not to allocate the funds. Generally, SEIFA on its own is too broad to allocate funds for specific purposes. These are better allocated by identifying specific groups who need the service; and then SEIFA can be used to ensure there is no bias in the distribution.

2.4.3 Targeting areas for business

Businesses might use the indexes to assist with marketing and strategic planning. The indexes might be used simply as summaries of area characteristics. Information from the indexes will be useful for making business decisions, such as siting outlets and targeting promotion campaigns. The indexes are also useful for consumer research.

2.4.3 Targeting areas for business continued

Example:

A retail organisation wants to establish a chain of boutiques to sell designer-label women's clothing in Sydney and needs to know where to locate the shops. The Index of Relative Socio-Economic Advantage/Disadvantage could be used to identify Sydney suburbs with high advantage. This will help pinpoint suitable localities for the boutiques.

The map below shows the Index of Advantage/Disadvantage. The suburb level index is split into five quintiles, based on Australia-wide values. The map for Sydney is then extracted, showing some suburb and Statistical Subdivision names.

Index

1,114 to 1,267	(474)
1,050 to 1,114	(458)
995 to 1,050	(478)
934 to 995	(468)
597 to 934	(479)

Source: 2001 Census of Population and Housing

CHAPTER **3**   DERIVATION OF THE INDEXES . . . .
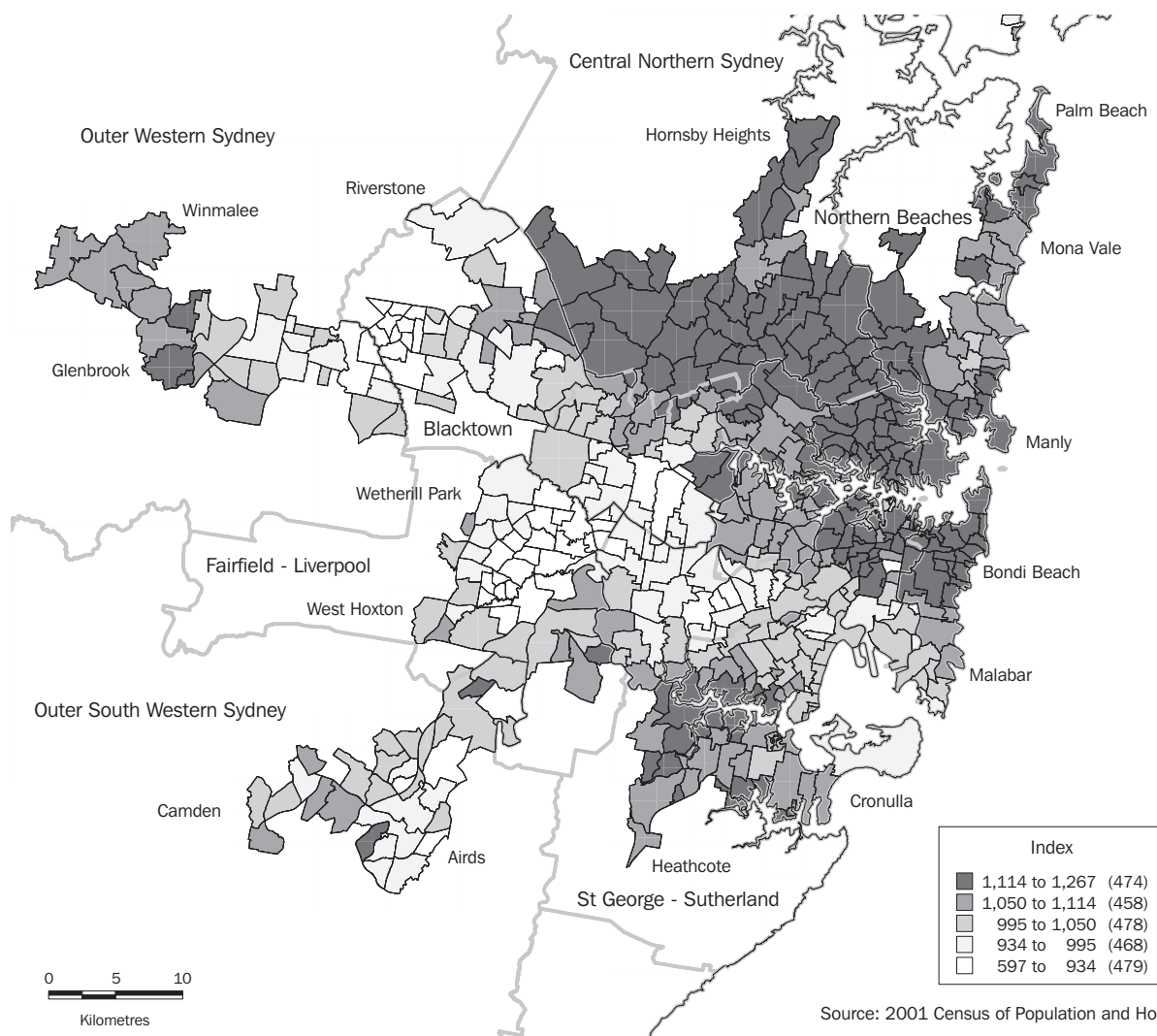
3.1 INTRODUCTION

Many aspects of the socio-economic profile of a community cannot be measured directly but there may be several variables that are recognised as contributing to a particular dimension. Often a single composite of these variables — an index — which reflects the population profile of these variables, is required to aid social and economic investigations.

Principal component analysis is a technique which is often used to summarise a large number of related variables. A socio-economic index can be derived using principal component analysis, and the resulting index measures what is common to the variables included in the analysis.

This section gives a brief description of the process that was followed to derive the indexes. A more detailed description is given in the *Census of Population and Housing: Socio-Economic Indexes for Area's (SEIFA), Australia - Technical Paper* (cat.no. 2039.0.55.001).

3.2 THE DATA

SEIFA is calculated using 2001 census data. All the data are calculated as ratios, so, the proportion of unemployed persons is calculated as the number unemployed divided by the total number of people in the labour force; and the proportion of people with a degree or higher is calculated as the number of people with a degree or higher divided by the number of people with a non-school qualification.

The distribution of each variable was examined to identify any skewed distributions (more details are given in the *Census of Population and Housing: Socio-Economic Indexes for Area's (SEIFA), Australia - Technical Paper* (cat.no. 2039.0.55.001)). Only two variables showed some skewness: % Households paying rent greater than $225 per week and % Households paying rent less than $88 per week. This skewness was not corrected, since it only affected two variables in one index; and the degree of skewness was not great.

Either a correlation or covariance matrix can be used to derive principal components. Using a covariance matrix will give higher weights to variables that have a large range of values. Using a correlation matrix ensures the range of values is the same for every input variable. The correlation matrix is normally used for principal components analysis, and is what we have used for the SEIFA indexes.

There are a number of features of the census data used to construct the indexes, that can affect the usefulness of the indexes. Users should be aware of the following:

3.2 THE DATA *continued*

- The variables included in analysis are limited to those for which data is collected by the census. Ideally, an indicator of a socio-economic factor should include all measures of relevance to that factor. However, the census does not obtain any information relating, for example, to wealth, or access to infrastructure such as schools, health and community services, transport and shops. The indexes cannot therefore purport to summarise these facets of socio-economic disadvantage.

- Missing data are a further impediment to index construction. Although non-response to individual census items is overall quite low, it does vary between CDs. It is possible that item non-response rates correlate directly with socio-economic disadvantage. Where possible, non-response for a variable has been dealt with by redefining the population associated with the variables, to include only those persons who answered the relevant questions. This approach implicitly assumes that non-respondents within a CD resemble respondents within that area, with respect to the characteristics measured by the variables.

- All variables pertaining to families and dwellings, in contrast to individuals, are based on data from occupied private dwellings. Families in non-private dwellings (e.g. motels, boarding houses, hospitals, refuges) are therefore 'under-represented'.

- The census data for social and economic aspects of the population are based on people's place of enumeration and not their usual residence, i.e. the population is classified to CDs according to where they were spending the night at the time of the census. Although the census is timed to capture the typical situation, holiday resort areas such as the Gold Coast may show a large enumeration count compared with the usual residence count.

- Some of the variables have a geographic context, so we would expect the weights from the principal component analysis to be different in different areas. Examples are percentage of households with no car, which in rural areas would be an indicator of disadvantage, but in urban areas it may mean the household is close to work or public transport; and rental or mortgage payments, which we expect to be higher in urban areas, so we would expect the monetary cutoffs we have used to be different. SEIFA 2001 considers disadvantage relative to an Australia-wide standard, so any monetary cutoffs (for Rent and Mortgage payments) or weights (for percentage of households with no car) were calculated for the whole of Australia.

- Some Indigenous people may be employed under the Community Development Employment Program (CDEP), so they are not registered as unemployed, but they are on a very low income. They will be counted in the low income variables but not as unemployed.

- The Internet use at home question only asked if the Internet was used at home in the week prior to census night. It gives no indication of the quality of the service.

CDs with small populations or a high proportion of families not responding to a census question may introduce some instability in the indexes. Consider for instance, a CD with eight people, all of whom are in the labour force, and two of those are unemployed, imagine if another two people become unemployed. The unemployment rate doubles from 25% to 50%. We therefore decided to exclude from the principal component analysis those CDs with one or more of the following characteristics:
- populations smaller than or equal to ten
- five people or fewer employed

3.2 THE DATA *continued*

- more than or equal to 70% of people not responding to any of the census questions on family income, occupation, labour force status, type of educational institution attended, and qualifications held
- more than 20% of dwellings are non-private
- off-shore and migratory CDs.

In total about 1,500, or 4% of Australian CDs fell into one of the above categories. This has increased from 1% in 1996 partly because the 2001 census tried to get better estimates for Indigenous people by creating extra CDs (within existing CDs), where Indigenous communities were known to exist. In the final census output, most of the population of these new CDs were added to the surrounding CD; and the populations of the smaller CDs were set to 0. In 2001, 909 of the 966 CDs with a population less than ten had a zero population. The *Census of Population and Housing: Socio-Economic Indexes for Area's (SEIFA), Australia - Technical Paper* (cat.no. 2039.0.55.001) shows details of the number of CDs excluded by criteria.

The Australian mean and standard deviation were calculated without using these CDs. When aggregating areas to a higher level the index values for these CDs are excluded from the weighted average.

Income variables in SEIFA experienced a major change between SEIFA 1996 and SEIFA 2001. In 2001, we decided to consider income by family type. We began looking at income by family type because what constitutes low income for a single person household is different to what constitutes low income for a family of four, say. The values for high and low incomes were calculated as the top and bottom quintiles of the income distribution for each family type.

The other cutoffs updated for SEIFA 2001 were Rent and Mortgage payments. Again, the top and bottom quintiles of the frequency distribution were used to calculate these cutoffs.

In the 2001 census, the Age Left School variable was replaced with Highest Level of Schooling Completed. For SEIFA 2001, those with Year 11 or below (i.e., those who had not completed Year 12) were considered disadvantaged.

3.3 CHOOSING THE VARIABLES

The variable selection process was described in Chapter 1. Variables measuring socio-economic status and aspects of disadvantage (level one and two variables) were included in the Index of Advantage/Disadvantage, the Index of Economic Resources and the Index of Education and Occupation. The Index of Disadvantage also included variables reflecting general disadvantage (level three variables).

The finest level at which complete census data are disseminated is the CD, which corresponds to the workload of one census collector. In 2001 there were 37,209 CDs defined across Australia. By calculating an index at the CD level, an index at any broader level can be derived using the constituent CD index scores. Thus the CD was chosen as the appropriate level for analysis.

3.3 CHOOSING THE VARIABLES *continued*

The correlations between all the input variables were analysed to ensure that particular socio-economic aspects were not over-represented in the analysis (as over-representation would lead to an unreasonably high index weighting for this aspect). Variables were reconsidered when they were correlated by more than a factor of 0.8. This criteria was included in the Variable Selection Decision Tree (see the *Census of Population and Housing: Socio-Economic Indexes for Area's (SEIFA), Australia - Technical Paper* (cat.no. 2039.0.55.001)).

It is also important that only variables which are well-related to the general thrust of the index are included. Variables which correlate poorly with the index do little but add to the variability of the index. These variables are not related to the main thrust of the index, and can make the index unnecessarily sensitive to small changes in the population over time. Therefore, for three of the four indexes, after the first principal component analysis, those variables which had low correlations (less than 0.3) with the index were excluded. The principal component analysis was then repeated to produce the final index. This cutoff is higher than it was in 1996 (0.2), because a literature search identified 0.3 as a better cutoff. However, the 2001 Disadvantage Index used the 0.2 cutoff to maintain consistency with the 1996 Index. The *Census of Population and Housing: Socio-Economic Indexes for Area's (SEIFA), Australia - Technical Paper* (cat.no. 2039.0.55.001) provides more detail on this.

3.4 THE METHOD

Principal component analysis was used to construct the Socio-Economic Indexes for Areas. Principal component analysis essentially achieves three things:

- it creates a 'raw index score' for each CD, based on the variables considered for each index
- it gives the contribution (called the 'weight') of each initial variable to the raw index score
- it allows the analyst to eliminate variables that have a low correlation with the index.

For the 2001 review of the indexes, we investigated using rotated components for the Index of Advantage/Disadvantage, the Index of Economic Resource and the Index of Education and Occupation. In a disadvantage index estimated using principal component analysis, rotation may associate particular aspects of disadvantage more strongly with each principal component; so we may find the first component represents income and the second component represents education and so on. Rotation usually aids in interpreting components, since the rotated components will often represent particular aspects of the underlying disadvantage in an area.

Unfortunately, we found that rotating the components of the SEIFA index was not helpful. The rotated components did not represent particular aspects of advantage/disadvantage, even though some variables had higher weights on the first component after rotation. So we used the first unrotated component in the 2001 index (this component gave a better general measure of advantage/disadvantage than its rotated counterpart). The full results from this analysis are shown in the *Census of Population and Housing: Socio-Economic Indexes for Area's (SEIFA), Australia - Technical Paper* (cat.no. 2039.0.55.001).

3.4 THE METHOD *continued*

Rotation was not attempted on the 2001 Index of Disadvantage because we wanted the method of calculating this index to be as close as possible to that used in the 1996 method.

To allow for easier recognition of high and low scores, the raw index scores produced by principal component analysis were standardised to have a mean of 1,000 and a standard deviation of 100 across all CDs in Australia. In practice, this means that around 95% of index scores are between 800 and 1,200.

Scores for areas larger than CDs were calculated by weighting together constituent CD scores, using the CD population size for weighting. These scores are CD weighted averages and are similar, but not quite the same as, those that would have been produced if the principal component analysis had been carried out separately on the larger geographic areas.

A more detailed description of the method is in the *Census of Population and Housing: Socio-Economic Indexes for Area's (SEIFA), Australia - Technical Paper* (cat.no. 2039.0.55.001).

3.5 VALIDATION

Socio-economic wellbeing is neither a simple nor well defined concept. Given the need to choose which variables to include and exclude, and the need to interpret the meaning of the summary variables resulting from the analysis, it was necessary to scrutinise the final indexes carefully to ensure their validity. The main validation exercises carried out on the final indexes were to:

- check that variables and their weights made sense
- compare the indexes with the 1996 indexes
- use local subjective knowledge to verify CD index results and to study CDs at the top and bottom end of the index
- compare the indexes with data from other sources
- subject the preliminary indexes to external scrutiny by an expert group.

These steps are summarised below; see the *Census of Population and Housing: Socio-Economic Indexes for Area's (SEIFA), Australia - Technical Paper* (cat.no. 2039.0.55.001) for a full description of the validation process.

One important check on the indexes was to ask whether the variables and their weights made sense. After each principal component analysis, the first principal component was examined to see if it was summarising the input variables adequately. In all cases the final indexes explained about 40% of the variability in the underlying input variables — a good indication that some common underlying factor was being identified and summarised. The weights of the variables in each index also made intuitive sense.

Where possible, the indexes were compared with the 1996 indexes by CD. Where there was a large change in a CDs position (defined as moving more than three quintiles), the change was analysed. We found that most change occurred because of the change in the income variables, in particular the change to income by family type.

3.5 VALIDATION *continued*

The indexes were 'ground truth' validated (using local subjective knowledge). In each state, maps of the indexes by CD were used to highlight areas of disadvantage. ABS offices in each state checked to see if the map reflected their understanding of the state (some of the questioned CDs were not comparable between 1996 and 2001, because CD boundaries had changed).

The top and bottom ranked CDs in each state were also examined using local knowledge. These CDs tended to be quite homogeneous (i.e. with levels of uniform characteristics). High values related to high income, home ownership and professional employment or trade qualification. CDs with low index values tended to be characterised by low car ownership, high unemployment or a relatively unskilled workforce and relatively low educational achievements and incomes.

Additional data sources were used to validate the index scores. Data from the current 1998–99 ABS Household Expenditure Survey (HES) was used to compare expenditure information with index scores from the Index of Economic Resources. Correlations were calculated for each broad expenditure group in the HES. The relationships identified appeared plausible, for instance housing costs and income tax payments were higher in more advantaged areas. The full results are given in the *Census of Population and Housing: Socio-Economic Indexes for Area's (SEIFA), Australia - Technical Paper* (cat.no. 2039.0.55.001).

An external source of data was also used to validate the indexes. The Public Health Information Development Unit (PHIDU) at the University of Adelaide have published the correlations between a number of variables and the 1996 Disadvantage Index. We replicated this analysis for the 2001 index. We found the correlations were similar, although correlations with some of the health measures were lower. Most of the changes could be explained.

A group of academics and expert users of the SEIFA indexes were also asked to analyse the indexes as part of the validation phase. This group received some preliminary indexes, and the full results from the principal components analysis. They also made some helpful suggestions, which were implemented in the final indexes; and made some longer-term suggestions, which we will be looking at for the 2006 indexes.

COMMENTS AND CAVEATS ON THE INTERPRETATION OF THE INDEXES ....

4.1 CHOICE OF VARIABLES FOR EACH INDEX

The indexes we produced depend upon the variables that were analysed using principal component analysis. A different choice of underlying variables would result in different final indexes; the indexes presented in this paper are just four of the many that could be derived from the census variables. The construction of an index depends on the socio-economic aspect of interest, and the underlying variables which represent those aspects most precisely.

Indexes produced using principal component analysis can be affected if some socio-economic aspects are over or under-represented in the underlying variables. Section 3.3. describes the way in which over-representation was dealt with. Under-representation occurs when the variables relating to an important aspect of a socio-economic dimension under consideration are absent from a particular index. The indexes described in this paper do not provide good measures for all social conditions. They are good overall indexes, but should be used in conjunction with other information that relates to the topic of interest. For example, the age structure of the population is not used directly in any of the indexes. Thus, if the topic of interest relates to infant or aged care, specific data on that segment of the population should be used in addition to the indexes.

4.2 DISADVANTAGE DIMENSIONS NOT REPRESENTED IN THE INDEXES

Users of the indexes should examine the constituents of the indexes (see Appendix 1) to ascertain whether they are appropriate to their problem or analysis. There are two factors in particular which the indexes do not represent well.

First, the indexes contain only limited information about wealth. While income and expenditure are included, aspects such as inherited wealth, savings, indebtedness, and property values are not (such data was not collected by the census). This shortcoming is most serious in the Index of Economic Resources.

Second, an area's infrastructure such as schools, community services, shops and transport is not represented by the indexes. Such information is considered to be important to the concept of advantage or disadvantage. For example, rapidly growing outer suburban areas may suffer from locational disadvantage rather than a socio-economic disadvantage.

4.3 UNDERSTANDING THE INDEXES

The indexes reflect the socio-economic wellbeing of an area, rather than that of individuals. They were calculated at the CD level, and reflect CD characteristics. Because all people within a CD are not identical, the index scores for a CD do not directly apply to individuals within that CD. It is possible for a relatively advantaged person to be resident in a CD which may have a low score on some or all of the indexes. Thus it is not appropriate to base inferences about a particular individual on index scores of his or her CD.

4.3 UNDERSTANDING THE INDEXES *continued*

The degree of heterogeneity within a CD influences its index score; the more homogeneous CDs tend towards the extreme index scores. This relationship is explored further in the *Census of Population and Housing: Socio-Economic Indexes for Area's (SEIFA), Australia - Technical Paper* (cat.no. 2039.0.55.001).

Partly because of this, the interpretation of the index values is more straightforward for areas which have extreme values (i.e. very high or very low index values). For example, it is usually easy to see why a CD which is in the top (or bottom) 5% of index values has that status. In contrast, areas with mid-range index values tend to contain a broader mix of people and households.

When the higher geographies are derived using CD level indexes, the extreme CD values are averaged out, so the indexes for larger areas are more stable than the CD level indexes. Aggregating by using a population weighted average can mean a loss of a lot of the information on disadvantage in the larger area if these more populous areas also tend to be areas of low disadvantage. Averaging is not the only method of summarising the indexes. If one's prime focus is on disadvantage say, one might choose to summarise the indexes by using the bottom decile of CD values in each SLA. The *Census of Population and Housing: Socio-Economic Indexes for Area's (SEIFA), Australia - Technical Paper* (cat.no. 2039.0.55.001) looks further at the relationship between population and the SEIFA 2001 indexes.

Appendix 2 shows indexes population weighted up to SLA and POA, then aggregated to state.

4.4 COMPARISON WITH THE 1996 CENSUS

Index scores from the 2001 census should not be compared directly to the indexes based on the 1996 census.

The index values for CDs are standardised to have a mean of 1,000 across Australia and so the difference between an area's score in 1996 and 2001 does not numerically represent any change in socio-economic conditions. Also, the indexes are not interval measures (see section 2.3). And a difference between the index scores of two areas in 2001 cannot be meaningfully compared to the difference in 1996 (that is the differences shed no light on whether the gap between the socio-economic conditions of the two areas is narrowing or widening).

SEIFA's weights depend on the constituent variable values, and will be different for each census. If a variable has a low weight, it will be dropped from the index. So SEIFA's variable weights and variable composition may also change between censuses.

Boundaries of some CDs have changed between censuses. The actual number of CDs in Australia has increased from 31,401 in 1991 to 37,209 in 2001. The boundaries of the higher geographic areas such as SLA and LGA may not be comparable. Approximately 86% of 2001 CDs in Australia are comparable within a 10% dwelling limit to the 1996 CDs.

DATA AVAILABILITY

SEIFA 2001, on CD-ROM, is available for each complete state/territory or Australia. The four indexes are provided for the ASGC 2001 areas of CD, SLA, SSD, SD, S/T and LGA, and for the Census Geographic 2001 areas of POA, SSC, SED and CED. SEIFA 2001 is available as either a 'Standalone' product or as a 'CDATA 2001 Add-On Datapack'.

Both the SEIFA 2001 'Standalone' and 'Add-On Datapack' are contained on a single CD-ROM, which offers the option to install one or both products.

The SEIFA 2001 'Standalone' product is a Windows 98SE, NT, 2000 or XP software package which provides a fully-documented user-friendly interface to the indexes. The 'Standalone' enables users to display selected indexes for any standard or customised geographic area, aggregate areas of interest and print a report. It provides online help.

The SEIFA 2001 'Add-On Datapack' for CDATA 2001 can only be accessed via the CDATA 2001 product. This option enables users to access and manipulate the indexes through the powerful functions within CDATA 2001 for area selection, mapping, graphing and display of information.

An Order Form and Conditions of Sale for both the SEIFA 2001 'Standalone' and 'Add-On Datapack' are provided in this publication. Alternatively, special index datasets for particular areas can be obtained on a range of media from ABS Information Consultancy or Statistical Consultancy Services in your state/territory ABS office. Our Statistical Consultants will also be able to provide assistance with using the indexes for various applications or, if necessary, designing other indexes to meet specific needs.

For users who are interested in data on specific census variables, CDATA 2001 contains both the Basic and Time Series Community Profiles for all applicable geographic areas and covering most topics on the census form. Further small area data are available from the Integrated Regional Database (IRDB) which provides access to ABS data from a wide range of social and economic data collections, including Population Census, Estimated Resident Population, Monthly Population Survey, Business Register, Agricultural Census, Manufacturing Census, Retail Census, Building Activity Survey, and many other collections. Two sets of data are available, one at SD level and above, the other at SLA level and above. The IRDB allows users to export data directly into CDATA and thematically display the data alongside census information.

If you are interested in the above products or services, your first point of contact in all circumstances should be the inquiry staff of your state/territory office of the ABS.

VARIABLES UNDERLYING SOCIO-ECONOMIC INDEXES ....

In this Appendix, we list the variables considered for inclusion in the various indexes. The variables are grouped by the value of their weight to indicate the contribution of each variable to the index.

The first group lists those variables which have been included in the final indexes whilst the second group lists those variables which were excluded from the indexes as a result of the analysis. The weights are to two decimal places, and are shown to four decimal places in the *Census of Population and Housing: Socio-Economic Indexes for Area's (SEIFA), Australia - Technical Paper* (cat.no. 2039.0.55.001).

VARIABLES UNDERLYING
SOCIO-ECONOMIC INDEXES
continued

INDEX OF ADVANTAGE/DISADVANTAGE

% Persons aged 15 years and over with degree or higher (0.24)

% Couple families with dependent child(ren) only with annual income greater than $77,999 (0.24)

% Couple families with no children with annual income greater than $77,999 (0.23)

% Employed Males classified as 'Professionals' (0.23)

% Persons aged 15 years or over having an advanced diploma or diploma qualification (0.21)

% Employed Females classified as 'Professionals' (0.21)

% Single person households with annual income greater than $36,399 (0.20)

% Persons using Internet at home (0.19)

% Couple families with dependents and non-dependents or with non-dependents only with annual income greater than $103,999 (0.18)

% Single parent families with dependent child(ren) only with annual income greater than $36,399 (0.17)

% Persons aged 15 years and over at university or other tertiary institution (0.15)

% Employed Males classified as 'Associate Professionals' (0.14)

% Single parent families with dependents and non-dependents or with non-dependents only with annual income greater than $62,399 (0.13)

% Employed Females classified as 'Advanced Clerical & Service Workers' (0.10)

% Dwellings with four or more bedrooms (0.08)

% Single parent families with dependents and non-dependents or with non-dependents only with annual income less than $26,000 (–0.10)

% Employed Females classified as 'Elementary Clerical, Sales & Service Workers' (–0.10)

% Employed Males classified as 'Tradespersons' (–0.13)

% Employed Females classified as 'Intermediate Production & Transport Workers' (–0.13)

% One parent families with dependent offspring only (–0.13)

% Couple families with dependents and non-dependents or with non-dependents only with annual income less than $52,000 (0.15)

% Females (in labour force) unemployed (–0.16)

% Males (in labour force) unemployed (–0.16)

% Single person households with annual income less than $15,600 (–0.18)

% Employed Males classified as 'Intermediate Production and Transport Workers' (–0.19)

% Employed Males classified as 'Labourers & Related Workers' (–0.19)

% Employed Females classified as 'Labourers & Related Workers' (–0.19)

% Couple families with dependent child(ren) only with annual income less than $36,400 (–0.20)

% Couple only families with annual income less than $20,800 (–0.20)

% Persons aged 15 years and over with highest level of schooling completed being Year 11 or below (–0.24)

% Persons aged 15 years and over with no qualifications (–0.25)

VARIABLES UNDERLYING
SOCIO-ECONOMIC INDEXES
continued

dropped initial variables

% Employed Males classified as 'Advanced Clerical & Service Workers' (0.08)

% Households who are group households (0.07)

% Employed Males classified as 'Intermediate Clerical, Sales & Service Workers' (0.06)

% Employed Males classified as 'Managers or Administrators' (0.06)

% Employed Females classified as 'Associate Professionals' (0.06)

Average number of bedrooms per person (0.05)

% Households purchasing dwelling (0.04)

% Dwellings with three or more motor vehicles (0.03)

% Households owning dwelling (0.03)

% Employed Females classified as 'Managers or Administrators' (0.03)

% Dwellings with one or no bedrooms (0.00)

% Persons aged 15 years and over who are still at school (–0.00)

% Persons aged 15 years and over at TAFE (–0.02)

% Employed Males classified as 'Elementary Clerical, Sales & Service Workers' (–0.02)

% Persons living in Caravan park (–0.03)

% Households living in improvised dwellings (–0.03)

% Persons aged 15 years and over with certificate qualification (–0.04)

% Lacking fluency in English (–0.05)

% Occupied private dwellings with two or more families (–0.05)

% Rental dwellings (–0.05)

% Employed Females classified as 'Intermediate Clerical, Sales & Service Workers' (–0.06)

% Dwellings with no motor vehicles at dwelling (–0.06)

% Employed Females classified as 'Tradespersons' (–0.07)

% Single parent families with dependent child(ren) only with annual income less than $15,600 (–0.08)

% Persons aged 15 years and over who did not go to school (–0.08)

INDEX OF ECONOMIC RESOURCES

% Couple families with dependent child(ren) only with annual income greater than $77,999 (0.33)

% Couple families with no children with annual income greater than $77,999 (0.32)

% Single person households with annual income greater than $36,399 (0.30)

% Households paying rent greater than $225 per week (0.30)

% Households paying mortgage greater than $1,360 per month (0.29)

% Couple families with dependents and non-dependents or with non-dependents only with annual income greater than $103,999 (0.27)

% Single parent families with dependent child(ren) only with annual income greater than $36,399 (0.24)

% Single parent families with dependents and non-dependents or with non-dependents only with annual income greater than $62,399 (0.20)

% Dwellings with four or more bedrooms (0.13)

% Single parent families with dependents and non-dependents or with non-dependents only with annual income less than $26,000 (–0.16)

% Households paying rent less than $88 per week (–0.19)

% Couple families with dependents and non-dependents or with non-dependents only with annual income less than $52,000 (–0.23)

% Single person households with annual income less than $15,600 (–0.27)

% Couple only families with annual income less than $20,800 (–0.28)

% Couple families with dependent child(ren) only with annual income less than $36,400 (–0.28)

VARIABLES UNDERLYING
SOCIO-ECONOMIC INDEXES
continued

dropped initial variables

% Households purchasing dwelling (0.11)

% Dwellings with three or more motor vehicles (0.07)

% Households who are group households (0.05)

Average number of bedrooms per person (0.03)

% Households owning dwelling (0.03)

% Dwellings with one or no bedrooms (–0.05)

% Households living in improvised dwellings (–0.06)

% Rental dwellings (–0.10)

% Dwellings with no motor vehicles at dwelling (–0.10)

% Single parent families with dependent child(ren) only with annual income less
than $15,600 (–0.11)

INDEX OF EDUCATION AND OCCUPATION

% Persons aged 15 years and over with degree or higher (0.33)

% Employed Males classified as 'Professionals' (0.31)

% Employed Females classified as 'Professionals' (0.29)

% Persons aged 15 years or over having an advanced diploma or diploma
qualification (0.28)

% Persons aged 15 years and over at university or other tertiary institution (0.21)

% Employed Males classified as 'Associate Professionals' (0.18)

% Employed Males classified as 'Advanced Clerical & Service Workers' (0.12)

% Employed Females classified as 'Elementary Clerical, Sales & Service
Workers' (–0.14)

% Males (in labour force) unemployed (–0.17)

% Females (in labour force) unemployed (–0.18)

% Employed Females classified as 'Intermediate Production & Transport
Workers' (–0.18)

% Employed Males classified as 'Tradespersons' (–0.19)

% Employed Males classified as 'Labourers & Related Workers' (–0.24)

% Employed Females classified as 'Labourers & Related Workers' (–0.25)

% Employed Males classified as 'Intermediate Production & Transport
Workers' (–0.26)

% Persons aged 15 years and over with highest level of schooling completed being
Year 11 or below (–0.32)

% Persons aged 15 years and over with no qualifications (–0.32)

dropped initial variables

% Employed Females classified as 'Advanced Clerical & Service Workers' (0.10)

% Employed Males classified as 'Intermediate Clerical, Sales & Service
Workers' (0.09)

% Employed Females classified as 'Associate Professionals' (0.08)

% Employed Males classified as 'Managers or Administrators' (0.08)

% Employed Females classified as 'Managers or Administrators' (0.05)

% Employed Males classified as 'Elementary Clerical, Sales & Service
Workers' (–0.01)

% Persons aged 15 years and over at TAFE (–0.02)

% Persons aged 15 years and over who are still at school (–0.03)

% Persons aged 15 years and over who did not go to school (–0.09)

% Employed Females classified as 'Intermediate Clerical, Sales & Service
Workers' (–0.09)

% Persons aged 15 years and over with certificate qualification (–0.10)

% Employed Females classified as 'Tradespersons' (–0.10)

VARIABLES UNDERLYING
SOCIO-ECONOMIC INDEXES
continued

INDEX OF RELATIVE SOCIO-ECONOMIC DISADVANTAGE

% Persons aged 15 years and over with no qualifications (0.31)

% Families with offspring having parental income less than $15,600 (0.29)

% Females (in labour force) unemployed (0.27)

% Males (in labour force) unemployed (0.27)

% Employed Males classified as 'Labourer & Related Workers' (0.27)

% Employed Females classified as 'Labourer & Related Workers' (0.27)

% One parent families with dependent offspring only (0.25)

% Persons aged 15 years and over who left school at or under 15 years of age (0.25)

% Employed Males classified as 'Intermediate Production and Transport Workers' (0.24)

% Families with income less than $15,600 (0.23)

% Households renting (government authority) (0.22)

% Persons aged 15 years and over separated or divorced (0.19)

% Dwellings with no motor cars at dwelling (0.19)

% Employed Females classified as 'Intermediate Production & Transport Workers' (0.19)

% Persons aged 15 years and over who did not go to school (0.18)

% Aboriginal or Torres Strait Islanders (0.18)

% Lacking fluency in English (0.15)

% Employed Females classified as 'Elementary Clerical, Sales & Service Workers' (0.13)

% Occupied private dwellings with two or more families (0.13)

% Employed Males classified as 'Tradespersons' (0.11)

dropped initial variables

% Households renting (non-government authority) (0.08)

% Employed Females classified as 'Intermediate Clerical, Sales & Service Workers' (0.08)

% Employed Males classified as 'Elementary Clerical, Sales & Service Workers' (0.08)

% Employed Females classified as 'Tradespersons' (0.07)

% Dwellings with one or no bedrooms (0.06)

% Recent migrants from non-English speaking countries (0.06)

% Households in improvised dwellings (0.04)

AVERAGE AND QUANTILE
INDEX VALUES

Tables 1–3 give the average index values and a range of quantiles for the geographic areas CD, SLA and POA in each state and in Australia. Quantiles denote a point in the distribution of index values below which a specified percentage of index values fall. Quantiles which divide the distribution of index values into ten equal parts are commonly referred to as deciles (the 10% and 90% deciles are given); quantiles which divide the distribution of index values into four equal parts are commonly referred to as quartiles, with the 50% quartile also known as the median (the 25%, 50% and 75% quartiles are given in this appendix).

The distribution of index values in tables 1–3 refer to different types of geographic area. Index scores of SLAs and POAs are formed by taking the population weighted average of index values of the CDs in the area. Their values depend on the distribution of population weights across the CDs.

Note that only the CD level indexes are rescaled to a mean of 1,000 and standard deviation of 100. The population weighted indexes are not rescaled.

TABLE 1 SUMMARY AREA DATA FOR CD LEVEL INDEXES

		Quantile				
	Average	10%	25%	50%	75%	90%

INDEX OF RELATIVE SOCIO-ECONOMIC ADVANTAGE/DISADVANTAGE

	Average	10%	25%	50%	75%	90%
NSW	1 011	882	932	999	1 085	1 165
Vic.	1 011	889	942	1 003	1 080	1 147
Qld	980	878	919	971	1 035	1 104
SA	973	859	907	966	1 036	1 108
WA	1 001	886	936	994	1 067	1 127
Tas.	943	831	881	935	1 007	1 073
NT	987	812	914	1 010	1 074	1 114
ACT	1 120	1 047	1 082	1 117	1 167	1 208
Aust.	*1 000*	*880*	*928*	*990*	*1 068*	*1 141*

INDEX OF ECONOMIC RESOURCES

	Average	10%	25%	50%	75%	90%
NSW	1 021	885	930	1 004	1 104	1 190
Vic.	1 006	893	937	997	1 067	1 134
Qld	980	880	919	971	1 033	1 097
SA	963	864	905	955	1 014	1 071
WA	997	891	934	988	1 056	1 119
Tas.	928	844	879	919	973	1 025
NT	1 002	823	909	1 033	1 092	1 143
ACT	1 107	1 031	1 062	1 106	1 151	1 201
Aust.	*1 000*	*883*	*925*	*986*	*1 065*	*1 142*

INDEX OF EDUCATION AND OCCUPATION

	Average	10%	25%	50%	75%	90%
NSW	1 009	889	936	996	1 078	1 162
Vic.	1 012	889	940	999	1 084	1 161
Qld	980	881	919	967	1 034	1 104
SA	978	859	910	969	1 037	1 122
WA	998	880	928	986	1 064	1 136
Tas.	959	840	888	945	1 026	1 102
NT	980	838	928	998	1 048	1 094
ACT	1 116	1 031	1 071	1 113	1 161	1 213
Aust.	*1 000*	*882*	*928*	*987*	*1 067*	*1 146*

INDEX OF RELATIVE SOCIO-ECONOMIC DISADVANTAGE

	Average	10%	25%	50%	75%	90%
NSW	1 000	876	945	1 011	1 073	1 119
Vic.	1 016	897	971	1 032	1 081	1 118
Qld	989	891	945	997	1 043	1 086
SA	994	861	945	1 013	1 064	1 104
WA	996	884	951	1 011	1 064	1 102
Tas.	969	846	919	983	1 038	1 079
NT	903	578	808	984	1 041	1 074
ACT	1 076	1 003	1 051	1 083	1 118	1 150
Aust.	*1 000*	*882*	*951*	*1 013*	*1 068*	*1 110*

TABLE 2 SUMMARY AREA DATA FOR SLA LEVEL INDEXES

| | Average | *Quantile* | | | | |
		10%	25%	50%	75%	90%
INDEX OF RELATIVE SOCIO-ECONOMIC ADVANTAGE/DISADVANTAGE						
NSW	979	909	924	948	1 009	1 108
Vic.	989	920	941	967	1 023	1 089
Qld	994	903	929	983	1 045	1 116
SA	960	887	916	942	994	1 068
WA	970	918	936	951	984	1 053
Tas.	928	877	892	911	944	986
NT	1 014	894	957	1 023	1 072	1 114
ACT	1 121	1 066	1 088	1 115	1 154	1 196
Aust.	*994*	*906*	*930*	*969*	*1 053*	*1 120*
INDEX OF ECONOMIC RESOURCES						
NSW	978	891	912	951	1 011	1 125
Vic.	979	906	927	958	1 022	1 092
Qld	994	903	938	989	1 044	1 099
SA	950	884	901	940	982	1 050
WA	960	884	909	941	994	1 070
Tas.	916	876	888	906	933	971
NT	1 027	890	979	1 047	1 102	1 142
ACT	1 111	1 040	1 071	1 102	1 141	1 195
Aust.	*990*	*893*	*924*	*971*	*1 047*	*1 111*
INDEX OF EDUCATION AND OCCUPATION						
NSW	981	915	930	952	1 003	1 105
Vic.	991	923	942	969	1 015	1 101
Qld	992	899	926	973	1 045	1 122
SA	962	887	918	944	996	1 073
WA	967	904	933	952	970	1 061
Tas.	940	885	903	921	949	1 022
NT	1 005	906	966	1 013	1 052	1 094
ACT	1 115	1 042	1 080	1 118	1 153	1 202
Aust.	*993*	*906*	*932*	*965*	*1 046*	*1 129*
INDEX OF RELATIVE SOCIO-ECONOMIC DISADVANTAGE						
NSW	992	937	957	981	1 022	1 080
Vic.	1 016	955	991	1 013	1 044	1 079
Qld	996	921	961	997	1 041	1 086
SA	994	911	966	1 005	1 041	1 074
WA	977	923	961	984	1 004	1 057
Tas.	966	925	936	958	996	1 025
NT	949	709	917	988	1 037	1 064
ACT	1 079	1 031	1 053	1 078	1 111	1 136
Aust.	*999*	*926*	*963*	*1 000*	*1 046*	*1 087*

TABLE 3 SUMMARY AREA DATA FOR POSTAL AREA LEVEL INDEXES

		Quantile				
	Average	10%	25%	50%	75%	90%

INDEX OF RELATIVE SOCIO-ECONOMIC ADVANTAGE/DISADVANTAGE

NSW	994	896	923	966	1 059	1 145
Vic.	987	908	936	967	1 029	1 097
Qld	961	886	914	947	996	1 062
SA	960	883	912	944	1 005	1 068
WA	976	909	932	957	1 009	1 077
Tas.	932	858	888	913	963	1 032
NT	991	882	938	998	1 040	1 098
ACT	1 126	1 082	1 100	1 137	1 149	1 188
Aust.	*978*	*892*	*923*	*957*	*1 023*	*1 105*

INDEX OF ECONOMIC RESOURCES

NSW	1 000	889	914	969	1 074	1 170
Vic.	974	886	916	957	1 025	1 083
Qld	962	880	911	951	1 008	1 068
SA	946	868	899	938	991	1 038
WA	967	882	912	952	1 010	1 083
Tas.	916	843	880	907	956	997
NT	1 004	876	938	999	1 071	1 126
ACT	1 120	1 075	1 100	1 112	1 150	1 189
Aust.	*973*	*880*	*910*	*954*	*1 023*	*1 101*

INDEX OF EDUCATION AND OCCUPATION

NSW	993	903	929	962	1 056	1 147
Vic.	990	908	940	974	1 022	1 109
Qld	958	886	911	942	988	1 058
SA	964	886	915	949	999	1 084
WA	971	899	930	956	994	1 086
Tas.	946	865	901	927	981	1 051
NT	980	893	943	987	1 018	1 059
ACT	1 116	1 049	1 074	1 122	1 155	1 178
Aust.	*979*	*896*	*925*	*959*	*1 017*	*1 104*

INDEX OF RELATIVE SOCIO-ECONOMIC DISADVANTAGE

NSW	996	916	950	990	1 044	1 103
Vic.	1 020	950	989	1 025	1 058	1 092
Qld	982	915	954	982	1 013	1 047
SA	1 002	925	970	1 009	1 049	1 086
WA	985	918	960	990	1 023	1 062
Tas.	971	901	931	965	1 012	1 059
NT	924	736	848	970	1 018	1 038
ACT	1 081	1 042	1 060	1 080	1 104	1 123
Aust.	*998*	*923*	*962*	*1 000*	*1 044*	*1 085*

SEIFA2001 Order Form

Socio-Economic Indexes for Areas are summary measures derived from the Population Census to measure different aspects of socio-economic conditions by geographic area. These indexes reveal where the affluent (as opposed to just high income earning) live; where the disadvantaged (as opposed to the unemployed) live and where the highly skilled and educated (as opposed to simply the degree-holding people) live. The 'stand-alone' SEIFA2001 contains the indexes in tabular form, and the names and codes of the geographic areas. The fully documented software is designed to enable you to calculate and analyse index values and produce reports on the areas defined. The indexes can be exported to spreadsheets and other packages. SEIFA2001 add-on module for CDATA2001 contains the same data as the stand-alone version but the software will be fully compatible with CDATA2001, the most up-to-date package of tabulation and mapping software and detailed census data on CD-ROM. You will be able to make use of the powerful area selection and display functionality in CDATA2001 to access and analyse the indexes, and present the results as tables, maps or graphs.

Note: for information on the functionality and pricing of the product please refer to the SEIFA2001 Product Brief. (www.abs.gov.au > Census > Products and Services Briefs)

Instructions:

1. Client Manager/Salesperson to complete Sections A & C.
2. Client to complete Sections B & D and sign Section C.
3. Client to read Section E.
4. When the order form is signed, please fax/post or deliver to your Client Manager/Salesperson.
5. MapInfo Sales Agents: please send to;

> Attn: Shelley Bartlett
> MapInfo Australia Pty Ltd
> Fax: 07 3844 0448
> Mail: PO Box 3055, South Brisbane QLD 4101

MapInfo Australia Pty Ltd: please fax to the Australian Bureau of Statistics - Publishing Section - Facsimile number 02 - 62528512.

SECTION A - *(Client Manager/Salesperson to complete)*

Client Manager / Salesperson :   **Contact No :** *(include area code)*

....……………

Clients ANZSIC Code : *(refer ANZSIC codes document)*

....

SECTION B - Client Details *(Client to complete)*

Title :   Mr   ❏   Ms   ❏   Other...

Client Name :
First Name...Last Name ....

Telephone number *: (include area code)*   **Mobile number :** *(if applicable)*

....

Facsimile number : *(if applicable - include area code)*   **Email Address :** *(if applicable)*

....

Organisation Name :

....

Branch : *(if applicable, e.g. Regional Affairs Unit)*

....

Alternate Contact Name within organisation :

....

Postal Address : *(Invoice will be posted to this address)*

....

....

....

....

....

City/Suburb... State ....Postcode ....

Country *(if not Australia)* ....

Delivery Address :*(if different from Postal Address. If not write 'as above'. Note: The product/s will be delivered to this address)*

....

....

....

....

....

City/Suburb... State ....Postcode ....

Country *(if not Australia)* ....

SECTION C - Order Details. *(Client Manager/Salesperson to complete)*

Office of Sale:

ABS Central Office ❑	ABS NSW ❑	ABS VIC ❑	ABS QLD ❑
ABS SA ❑	ABS WA ❑	ABS TAS ❑	ABS NT ❑
ABS ACT ❑	MapInfo ❑		

Enter the number of product/s and the basic geographic level for each product/s.

SEIFA 2001 – Basic Geographic Level (A)	$ Price (includes GST) (B)	Number of Stand-alone products. i.e. 1 Licence (C)	Networked product (D)	Number of Multi User Defined Accesses (Network Licences) (E)	Total Number of Licences = (D x E) +C (F)	Total Cost $ = (F) X (B) (G)
AUSTRALIA	$2500					
NSW	$1750					
VICTORIA	$1750					
QUEENSLAND	$1750					
SOUTH AUST	$1250					
WA	$1250					
TASMANIA	$1000					
NT	$1000					
ACT	$1000					
					TOTAL	$
Less Discount for bulk purchases and networks. * (See table below)					%	$
						-
					(H) Cost	$

* SEIFA 2001 Discounts for bulk purchases and networks

Number of Copies or Licences	Percentage Discount %
1	0.00 %
2 to 5	20%
6 and over	30%

Total Cost of SEIFA 2001 Order		
SEIFA 2001 Order	enter (H) cost	$
TOTAL *(Includes GST)*	(H)	$
GST Component	(Total Cost of Sale x .09090)	$

Signature　　　　　　　　　　　　　　　**Name**

Date　　　　　/　　　/

SECTION D – PAYMENT DETAILS *(Client to complete)*   (ABS ABN Number 26 331 428 522)

Please raise an invoice on the basis of my Official Order   ❑

or

Please debit my Credit Card. *(tick the appropriate box)*   American Express   ❑   Bankcard   ❑

                                                                  Mastercard   ❑   Visa   ❑

To the value of: $ _____

Card Number: _____ Expiry Date: _____ / _____

Name on Card: _____

Signature: _____ Date: _____/_____/_____

or

My crossed cheque/ purchase order for $_____

Payable to the Australian Bureau of Statistics, Collector of Public Monies is attached to this order form.

Accounts address if different to the contact / postal address you specified earlier in Section B:

Billing Contact: ....

Title..

Telephone ....Fax..

Address

....

....

....

....

....

....

City/Suburb .... State ....Postcode ....

Country *(if not Australia)* ....

SECTION E

ABS CONDITIONS OF SALE

Australian Bureau of Statistics (ABS) products are sold by the Commonwealth of Australia ('The Commonwealth') through the ABS. Prices are subject to change without notice.

COPYRIGHT

Copyright in ABS products is owned by the Commonwealth. The Client agrees not to reproduce, distribute or commercialise the product, or any product or service derived from or incorporating it or part of it (whether or not amounting to a copyright reproduction) other than as allowed by these Conditions of Sale or with the prior written consent of the Commonwealth.

Where such consent is sought the Commonwealth reserves the right to set an appropriate charge or to require a revenue sharing arrangement.

The Client is permitted to quote an insubstantial part of the statistical data contained in the products (whether or not copyright subsists in the products), providing that:

- an 'insubstantial part' means a 'fair dealing' as defined in Sections 40, 41, 42 and 43 of the *Copyright Act 1968*;
- the ABS is cited as the source of the data used;
- the terminology used is that used by the ABS for describing data; and
- any analysis or transformation of the data is not attributed to the ABS.

Where a product consists of data in computer readable form, or software, the Commonwealth authorises the Client to use on a non-transferable and non-exclusive basis, this data or software personally or for internal purposes of its organisation (as the case may be). This authorisation does not permit the client to reproduce this data or software other than for the purposes of this personal or internal use.

In these Conditions, 'commercialise' in respect of a product or a derivative of that product, means to manufacture, sell, distribute, hire or otherwise exploit a product or process, or to provide a service incorporating the product or any other product or service derived from that product, or to license a third party to do any of those things.

GOODS AND SERVICES TAX

The price includes Goods and Services Tax (GST) for that part of the supply made by the ABS to the Client under this Agreement which is a taxable supply as determined under *A New Tax System (Goods and Services Tax) Act 1999* (GST Act).

In relation to taxable supplies made by ABS to the Client under this Agreement, the ABS agrees to issue the Client with either:

(a) a tax invoice in accordance with the GST Act; or

(b) a document satisfying the minimum information requirements to entitle the recipient of a taxable supply to claim an input tax credit without holding a tax invoice.

GENERAL

The Commonwealth gives no warranty, other than a warranty that may be implied by law, that the products are free from errors, are complete, have any particular quality, are suitable for any purpose or otherwise.

Subject to any warranty which may be implied by law, the Commonwealth's liability to the Client for any loss, damage or injury howsoever caused by the Commonwealth, whether due to negligence or otherwise, in relation to a product shall be limited to providing a replacement copy of that product.

The Client agrees to indemnify the Commonwealth (and its servants and agents) in respect of all liability for loss (including all legal costs) or liability from any claim, suit, demand, action or proceeding brought by any third person in connection with these Conditions of Sale or from a Client's use of the products.

These limitation of liability provisions shall survive the expiration or earlier termination of these Conditions of Sale. This Agreement shall be construed in accordance with the law of the Australian Capital Territory and the parties submit to the jurisdiction of the courts in that Territory (including the federal courts).

FURTHER INFORMATION

If the Client wishes to commercialise the product, or reproduce the product other than as permitted by these Conditions, or has any further questions relating to these Conditions, they should contact:

The Manager
Secondary Distribution
Australian Bureau of Statistics (2N 202)
Locked Bag 10
BELCONNEN ACT 2616

FOR MORE INFORMATION . . .

INTERNET	**www.abs.gov.au** the ABS web site is the best place to start for access to summary data from our latest publications, information about the ABS, advice about upcoming releases, our catalogue, and Australia Now—a statistical profile.
LIBRARY	A range of ABS publications is 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 web site for a list of libraries.
CPI INFOLINE	For current and historical Consumer Price Index data, call 1902 981 074 (call cost 77c per minute).
DIAL-A-STATISTIC	For the latest figures for National Accounts, Balance of Payments, Labour Force, Average Weekly Earnings, Estimated Resident Population and the Consumer Price Index call 1900 986 400 (call cost 77c per minute).

INFORMATION SERVICE

Data already published that can be provided within five minutes will be free of charge. Our information consultants can also help you to access the full range of ABS information—ABS user pays services can be tailored to your needs, time frame and budget. Publications may be purchased. Specialists are on hand to help you with analytical or methodological advice.

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

WHY NOT SUBSCRIBE?

ABS subscription services provide regular, convenient and prompt deliveries of ABS publications and products as they are released. Email delivery of monthly and quarterly publications is available.

PHONE	1300 366 323
EMAIL	subscriptions@abs.gov.au
FAX	03 9615 7848
POST	Subscription Services, ABS, GPO Box 2796Y, Melbourne Vic 3001

2203900001018

RRP $10.00