Australian Bureau of Statistics

2039.0.55.001

Technical Paper

Census of Population and Housing: Socio-Economic Indexes for Area's (SEIFA)

Australia

2001

Technical Paper

Census of Population and Housing: Socio-Economic Indexes for Area's (SEIFA)

Australia

2001

Dennis Trewin
Australian Statistician

AUSTRALIAN BUREAU OF STATISTICS

EMBARGO: 11.30AM (CANBERRA TIME) TUES 20 JAN 2004

ABS Catalogue No. 2039.0.55.001

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 Robert Tanton on Canberra 02 6252 5506.

CONTENTS

ABBREVIATIONS

ABS	Australian Bureau of Statistics
ACT	Australian Capital Territory
ASCED	Australian Standard Classification of Education
ASCO	Australian Standard Classification of Occupations
Aust.	Australia
CD	Collection District
HES	Household Expenditure Survey
NSW	New South Wales
NT	Northern Territory
OMG	Occupational Major Group
OT	Other Territories
PCA	Principal Components Analysis
POA	Postal Area
Qld	Queensland
SA	South Australia
SEIFA	Socio-Economic Indexes for Areas
SES	socioeconomic status
SLA	statistical local area
Tas.	Tasmania
TAFE	Technical and Further Education
Vic.	Victoria
WA	Western Australia

INTRODUCTION

INTRODUCTION

This is the first Technical Reference Manual produced for the SEIFA indexes. The manual gives more detailed information on the SEIFA calculation, and the nature of the data and indexes, so users have a better understanding of the indexes. The accompanying Information Paper shows how SEIFA can be used, but has little technical detail about the indexes.

This manual includes a description of the four indexes (Chapter 1); a full description of the data and limitations of the data (Chapter 2); a full description of the method used to calculate the indexes (Chapter 3); a description of how each index was constructed (Chapter 4); the validation process (Chapter 5); and a description of the limitations of the indexes (Chapter 6).

CHAPTER 1 THE 2001 SEIFA

THE NATURE OF THE 2001 INDEXES

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. The Index of Economic Resources and the Index of Education and Occupation were focused on specific aspects of advantage/disadvantage.

In 2001, there are four indexes. Again, the most general index is the Index of Disadvantage. This index includes all the available variables that either reflect or measure disadvantage. The inclusion of variables that reflect disadvantage (e.g. Indigenous, Separated/Divorced) means that while the index may reflect an area's disadvantage, it is not possible to say what aspects of disadvantage are 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 simultaneously (so it is the 'nett' effect which is measured). For any CD, information on advantaged persons in the CD will offset information on disadvantaged persons in the CD.

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

CHAPTER 2 THE DATA

THE DATA

This chapter describes the data used for SEIFA 2001, how we defined disadvantage, and how we selected variables to include in the indexes. We also explore the variables included in the index, including looking at the distribution of each variable, and looking at how it has changed from 1996 to 2001. Limitations of the data are also described.

Variable Selection — The Decision Tree

Reviewing the variables included in the SEIFA was a major part of the 2001 review of SEIFA. The original selection of variables was done for the 1986 indexes, and published in 1990.

In the review of SEIFA 2001, we tried to define the term 'disadvantage', to give a stronger conceptual basis to the variable selection. The concept of disadvantage was based on Ainley et al. (1995), in which disadvantage was viewed as an extension to socioeconomic status, which in turn was measured through education, occupation and income. Aspects of socioeconomic disadvantage are those factors that put someone at a disadvantage compared to someone else. Examples include wealth, residential conditions, health, access to services, and language. From this base, we developed a decision tree for selecting variables. This section describes the decision tree and the rationale behind it.

THE DECISION TREE AND ITS RATIONALE

Chart 2.1 shows the decision tree used for variable selection.

Following the literature on socioeconomic status (SES) and socioeconomic disadvantage, potential variables can be grouped into levels.

1) Level 1 are core variables — variables in this category are agreed in the literature (see Marks et al. (2000), and Ainley, et al. (1995)) to measure SES or one of the key aspects of SES. The key aspects of socioeconomic status are:

- income
- education and/or qualification
- occupation.

2) Level 2 are direct measures of an aspect of socioeconomic disadvantage. These variables include:

- number of motor vehicles, number of rooms in house (wealth or assets)
- unemployment (employment status)
- type of residence; number of bedrooms (residential conditions)
- low fluency in English (language disadvantage).

3) Level 3 are variables that reflect measures of disadvantage, but are not direct measures. Examples include Indigenous status (which may be associated with poor health or living conditions) and divorced/separated (which may be associated with low income). These level 3 variables attempt to capture aspects of socioeconomic disadvantage which might not (due to inadequate data) be fully captured in level 1 and

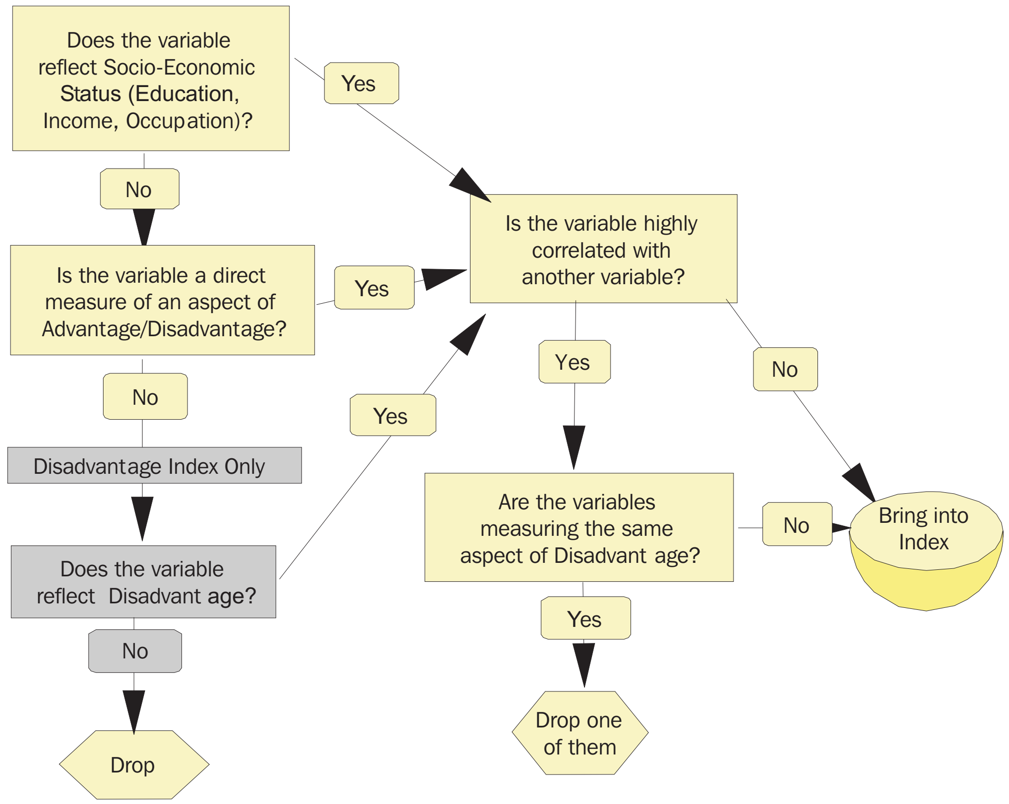
Variable Selection — The Decision Tree continued

level 2 variables. Level 3 variables can be thought of as indicators which signal that an area has some disadvantage.

In 2001, the indexes of Advantage/Disadvantage; Economic Resources; and Education and Occupation excluded such level 3 variables. But the Index of Disadvantage did include level 3 variables, making it directly comparable to the 1996 Index of Disadvantage.

Once a set of variables are selected as a theoretical base for the principal components analysis, one needs to ensure that variables are not highly correlated, since high correlations between two variables can lead to instability in the weights. However, if the two variables were highly correlated but measuring different aspects of disadvantage, we left them both in the model, to get a better picture of all aspects of disadvantage.

2.1 DECISION TREE FOR 2001 SEIFA VARIABLE SELECTION



2001 SEIFA Variables and their specifications

We then looked for variables available from the 2001 census that would fit into this theoretical base.

The Appendix gives detailed specifications of the variables used in the SEIFA calculation. The census codes for each variable extracted are given. The mnemonics given are the same as the ones used in the 2001 Census Dictionary. The short form names used in the tables throughout this technical paper are defined fully in this appendix.

There were some changes made to some variables between 1996 and 2001. The main changes were:

- Any variables with dollar value cut-offs (e.g. income and rent variables) had the dollar value cut-offs updated.

2001 SEIFA Variables and their specifications continued

- Income was changed to income by family type, and cut-offs were calculated for each family type.
- The education variable 'Age left school' was replaced with 'Year left school', due to a census question change.

INCOME

The income variables in SEIFA are calculated as the proportion of families in the CD with a high income and the proportion of families in the CD with a low income. The cut-offs for high and low income are calculated as the top and bottom quintiles of family income for each family type. Since census income is grouped, we chose our cut-offs as the income groups containing the top and bottom quintiles. Table 2.2 shows the percentage in each income group.

For single persons in a house, we needed to use individual income, rather than family income, and the income ranges were slightly different. The frequency distribution is shown in table 2.3. Note that the rule used was to take the group that the top or bottom quintile (20%) was in. In the case of single person households with low income, this is a group going from 18.61% to 43.3%. This is a huge range in the income group, but it is the only family type that had this problem.

2.2 DISTRIBUTION OF WEEKLY FAMILY INCOME BY FAMILY TYPE

	Couple family with dependents only	*Couple family with dependents and non-dependents*	*Couple family with non-dependents only*	*Couple only family*	*Lone parent family with dependents only*	*Lone parent family with dependents and non-dependents*	*Lone parent family with non-dependents only*
Income	Cumulative %	Cumulative %	Cumulative %	Cumulative %	Cumulative %	Cumulative %	Cumulative %
Negative income	0.17	0.06	0.13	0.30	0.20	0.08	0.12
Nil income	0.39	0.16	0.24	0.83	1.15	0.33	0.33
$1 to $39	0.45	0.18	0.26	0.93	1.34	0.38	0.38
$40 to $79	0.51	0.21	0.29	0.98	1.65	0.48	0.44
$80 to $119	0.59	0.24	0.32	1.05	2.21	0.58	0.53
$120 to $159	0.72	0.29	0.38	1.32	3.59	0.76	0.77
$160 to $199	0.87	0.36	0.47	1.55	7.49	1.36	1.38
$200 to $299	1.52	0.58	0.70	2.65	20.40	3.80	3.58
$300 to $399	5.21	1.33	1.70	21.07	44.36	10.78	11.76
$400 to $499	9.82	2.94	3.90	31.93	61.38	21.65	26.47
$500 to $599	14.82	5.39	7.66	37.00	72.26	33.11	34.36
$600 to $699	21.72	8.68	12.09	45.41	79.54	43.81	43.81
$700 to $799	27.86	12.70	16.76	49.82	84.76	53.44	52.62
$800 to $999	42.77	21.79	27.55	59.43	91.43	68.30	66.88
$1,000 to $1,199	54.61	32.52	38.93	67.73	96.23	79.09	77.05
$1,200 to $1,499	70.25	49.53	54.68	77.66	97.57	89.05	87.79
$1,500 to $1,999	87.00	72.71	75.55	89.78	99.67	96.39	95.35
$2,000 or more	100.00	100.00	100.00	100.00	100.00	100.00	100.00

2001 SEIFA Variables and their specifications continued

2.3 DISTRIBUTION OF WEEKLY INDIVIDUAL INCOME FOR SINGLE PERSONS

Income	Cumulative %
Negative income	0.39
Nil income	2.38
$1 to $39	2.70
$40 to $79	3.21
$80 to $119	4.47
$120 to $159	8.44
$160 to $199	18.61
$200 to $299	43.35
$300 to $399	52.23
$400 to $499	61.25
$500 to $599	69.78
$600 to $699	76.45
$700 to $799	82.16
$800 to $999	89.58
$1,000 to $1,499	96.61
$1,500 or more	100.00

RENT AND MORTGAGE PAYMENTS

Rent and mortgage payments were updated in the same way, except the census has exact dollar values for these. The quintiles were calculated for households, and table 2.4 shows the frequency distribution. Note that rent and mortgage payments may be higher in capital cities than rural towns, but we have not tested this, and did not adjust for it.

2.4 RENT AND MORTGAGE DECILES FOR HOUSEHOLDS

Decile	Rent $ / week	Mortgage $ / month
1	0 to 49	1 to 432
2	50 to 87	433 to 554
3	88 to 114	555 to 649
4	115 to 134	650 to 779
5	135 to 149	780 to 866
6	150 to 169	867 to 999
7	170 to 189	1000 to 1147
8	190 to 224	1148 to 1359
9	225 to 299	1360 to 1799
10	greater than 300	greater than 1800

EDUCATION

There were also changes made to the census question on 'Year left School' in the 2001 census. In 1996, the variable was 'Age left school' (15, 16 or 17 years, etc.) Because of different education systems in each state, the age left school means different levels of education across the country. Standardised schooling in Australia had been in place for a number of years, and so in 2001, the census question was changed to 'Highest year of schooling completed' (Year 10, Year 11, etc.) The question also included a category 'Still at school', and this was used as a separate variable in the Index of Advantage/Disadvantage; but not in the Index of Disadvantage. The 'Still at School' category was not included in the 'Highest year of schooling completed' variable.

2001 SEIFA Variables and their specifications continued

Year 11 was chosen as the cut-off for the Advantage/Disadvantage and Education and Occupation indexes, as those without a Year 12 certificate experience greater disadvantage than those with a Year 12 certificate. In 1996, the corresponding variable was 'left school at age 15 years and under' (which corresponds to Year 10 in most states). For maximum comparability with the 1996 index, the Index of Disadvantage used Year 10 as the cut-off. The increase to Year 11 in the Advantage/Disadvantage and Education and Occupation indexes reflects the increasing educational requirements in most occupations.

Changes to census questions and classification practices for non-school qualifications also meant we changed some of the non-school qualification variables. In 1996, we used the following variables in SEIFA:

- % Persons aged 15 years or over with degree or higher (which included associate diplomas).
- % Persons aged 15 years or over with a trade or other qualification.

In 2001, we used:

- % Persons aged 15 years and over with degree or higher (which excludes diplomas).
- % Persons aged 15 years and over having an advanced diploma or diploma qualification.
- % Persons aged 15 years and over with certificate qualification.

These new variables were based on the analysis of income and education shown in the section 'Relationship between income, qualifications and occupation' below. We compared the split with the Australian Standard Classification of Education (ASCED), and there is comparability in the ranking. Persons with a bachelor degree or higher are most advantaged (ASCED classifications 1 to 3). Advanced diplomas and diplomas (ASCED 4) were the second most advantaged in terms of income. Certificates (ASCED 5) were the least advantaged of the post school qualifications.

NOT STATED

Generally we excluded the 'Not Stated' values from the numerator and denominator of the ratios used in SEIFA. There were some exceptions, where we had some additional information due to the sequencing of the census questions. Appendix 1 shows the variables where we have included the not stateds.

Data Extraction, Validation and Clearance

This section describes the procedure involved in data derivation and validation.

DATA EXTRACTION

All variables (except average number of bedrooms) in SEIFA are expressed as proportions. For each variable (including average number of bedrooms), a numerator and a denominator were required. We specified the numerator and denominator for each variable based on previous SEIFA variable specifications, the 2001 Census Dictionary and consultations with the Australian Bureau of Statistics (ABS) census area and other ABS subject matter areas. The Census Output Section at the ABS produced the data for each numerator and denominator for each variable for each CD using 2001 census data.

Data Extraction, Validation and Clearance continued

VALIDATION OF THE NUMERATORS AND DENOMINATORS

The numerators and denominators were first validated. For each numerator and denominator we calculated a state total which was compared with the published census summary statistics. If a numerator (or denominator) had no published figure available to compare with, we calculated the figure from the census unit record file.

EXCLUSION OF SOME CDS FROM THE ANALYSIS

To ensure that the indexes were meaningful, CDs with a very low population or a low proportion of people responding to some census questions were excluded from the analysis. Those CDs with any one of the following characteristics were excluded from the analysis:

- CDs with population smaller than or equal to 10.
- CDs with five people or fewer employed.
- CDs where 70% or more of people did not respond to at least one of the census questions on family income (FINF), occupation (OCCP), labour force status (LFSP), type educational institution (TYPP), and qualifications (QALLP).
- CDs where more than 20% of dwellings are non-private.
- Off-shore and migratory CDs. These CDs contain people who are enumerated on off-shore oil rigs, drilling platforms and the like, aboard ship in Australian waters, or on an overnight journey by train or bus.

The number of CDs excluded in 1996 and 2001 are shown in table 2.5.

2.5 NUMBER OF EXCLUDED CDs, 1996 AND 2001

Requirement	*1996*	*2001*
Population = 0	219	909
Population > 0 and < or = 10	258	57
Employed persons < or = 5	288	1 040
Family Income not stated > or = 70%	26	14
Occupation not stated > or = 70%	1	—
Labour Force Status not stated > or = 70%	18	12
Type of educational institution attending not stated > or = 70%	277	363
Level of education (Non-school qualification) not stated > or = 70%	208	214
> 20% dwellings Non-private	109	95
Off-shore & Migratory CDs	25	42
Total number of CD's excluded	417	1 514

— nil or rounded to zero (including null cells)

In this table, one CD can fulfill a number of criteria. For instance, a CD might have fewer than six employed people and also have a population less than 11. Therefore, summing the CDs satifying each criteria will not give the total number of CDs excluded.

The number of CDs excluded in the 2001 SEIFA was much larger than in 1996. The difference is mainly due to a significant increase in CDs with population less than 11. For the 2001 census, a number of new 'point' CDs were created, especially in remote areas. Such CDs were often used to enumerate Indigenous communities. After enumeration, if the population in the point CD was less than 30, the population was moved to the surrounding CDs in the output stage. This was done to reduce the number of CDs with

Data Extraction, Validation and Clearance continued

small populations. But the 'point' CDs were left on the census dataset and given a zero population. In 2001, there were 909 CDs with nil population, in 1996 there were only 219.

CALCULATING THE VARIABLE VALUES AND TREATING MISSING VALUES

Each variable used in SEIFA was a ratio. If the denominator was zero, the ratio could not be calculated. For all but one of the SEIFA variables (AVEBED), the numerator was a subset of the denominator, so if the denominator was zero, then the numerator must also be zero. Therefore, missing ratios were given a value of zero. This is the same procedure we have used for previous indexes.

SUMMARY OF 2001 SEIFA VARIABLES

After those CDs described above were dropped and missing values were set to zero, we looked at summary statistics for each variable to give us an idea of the distribution in the data. Table 2.6 presents summary statistics for 2001 variables. There were 35,695 CDs included in the 2001 SEIFA calculation. Only the variable, AVEBED (average bedrooms per person), has a permitted maximum greater than one.

Where the variable definition has changed since 1996 — because of updated cut-offs or a change in the question — we include information on the variable as defined in 1996 along with information on the 2001 definition. The 2001 variables are shown with the same variable name and a _01 postscript. For the 1996 variable left school at age 15 years (SHRTSCH) we used Left school at Year 10 or below from the 2001 census.

For most variables the mean is close to the median, indicating the variable is not skewed. RENTHI_01 (rented occupied dwellings with rent>$225/week) and RENTLO_01 (rented occupied dwellings with rent < $88/week) appear to be two exceptions where the median is substantially lower than the mean, suggesting these two variables are skewed to low values. The distribution for 'Caravan' is also highly skewed; the upper quartile shows that 75% of CDs have no permanent residents in caravan parks; yet there is at least one CD with 100% of permanent residents in a caravan park. We did not correct this skewness, as it only affected a small number of variables.

Data Extraction, Validation and Clearance continued

2.6 SUMMARY STATISTICS FOR 2001 SEIFA VARIABLES

Variable	Mean	Standard Deviation	Minimum	Lower Quartile	Median	Upper Quartile	Maximum
ATSCH	0.04	0.02	—	0.02	0.03	0.05	0.65
AVEBED	1.15	0.13	0.25	1.08	1.16	1.23	2.08
CAEUNI	0.05	0.05	—	0.02	0.04	0.06	0.98
CARAVAN	—	0.03	—	—	—	—	1.00
CERT	0.17	0.05	—	0.14	0.18	0.21	0.57
CFD_NDINCHI	0.23	0.21	—	0.05	0.19	0.33	1.00
CFD_NDINCLO	0.29	0.22	—	0.13	0.25	0.40	1.00
CFDOINCHI	0.28	0.21	—	0.12	0.22	0.40	1.00
CFDOINCLO	0.25	0.18	—	0.12	0.22	0.35	1.00
CFNKINCHI	0.22	0.17	—	0.09	0.18	0.31	1.00
CFNKINCLO	0.22	0.13	—	0.12	0.20	0.30	1.00
COMP	0.41	0.12	—	0.33	0.41	0.50	0.89
DEGREE	0.22	0.12	—	0.12	0.18	0.29	0.73
DEGREE_01	0.14	0.11	—	0.06	0.11	0.19	0.64
DIP	0.07	0.03	—	0.05	0.06	0.08	0.24
F_UNEMP	0.07	0.06	—	0.04	0.06	0.09	1.00
FADVCLRK	0.07	0.04	—	0.05	0.07	0.09	0.50
FELECLSW	0.14	0.06	—	0.10	0.14	0.18	1.00
FEWBED	0.06	0.09	—	0.01	0.03	0.07	1.00
FINCLO	0.04	0.04	—	0.02	0.03	0.06	0.67
FINTCLRK	0.26	0.08	—	0.21	0.26	0.31	0.75
FINTPRTW	0.03	0.03	—	0.01	0.02	0.04	0.50
FLABOUR	0.08	0.07	—	0.03	0.07	0.11	1.00
FMGR_ADM	0.07	0.10	—	0.02	0.04	0.08	1.00
FPARA_PF	0.12	0.05	—	0.09	0.11	0.14	0.75
FPROF	0.21	0.10	—	0.14	0.19	0.27	1.00
FTRADE	0.03	0.03	—	0.02	0.03	0.04	0.67
GRENT	0.05	0.11	—	—	0.01	0.04	1.00
GROUP	0.04	0.04	—	0.01	0.03	0.05	0.55
HIGHBED	0.26	0.18	—	0.13	0.22	0.35	1.00
HIGHCAR	0.14	0.09	—	0.08	0.12	0.19	0.86
IMPDWEL	0.01	0.04	—	—	—	0.01	1.00
INDIG	0.02	0.08	—	—	0.01	0.02	1.00
INTNET	0.18	0.06	—	0.14	0.18	0.22	0.48
M_UNEMP	0.09	0.06	—	0.04	0.07	0.11	0.75
MADVCLRK	0.01	0.01	—	—	—	0.01	0.17
MELECLSW	0.06	0.03	—	0.04	0.06	0.08	0.50
MINTCLRK	0.08	0.04	—	0.06	0.08	0.11	0.50
MINTPRTW	0.13	0.08	—	0.07	0.12	0.18	1.00
MLABOUR	0.11	0.08	—	0.06	0.10	0.14	1.00
MMGR_ADM	0.14	0.13	—	0.06	0.10	0.16	1.00
MORTHI	0.25	0.22	—	0.08	0.19	0.38	1.00
MORTHI_01	0.20	0.20	—	0.05	0.14	0.31	1.00
MPARA_PF	0.12	0.05	—	0.08	0.12	0.15	0.67
MPROF	0.15	0.10	—	0.08	0.13	0.21	0.78
MTRADE	0.20	0.08	—	0.14	0.21	0.26	1.00
MULTFAM	0.01	0.02	—	—	0.01	0.01	0.71
NGRENT	0.22	0.14	—	0.12	0.18	0.28	1.00
NOCAR	0.11	0.10	—	0.03	0.08	0.15	1.00
NOQUAL	0.57	0.12	0.10	0.49	0.58	0.66	1.00
NOSCH	0.01	0.02	—	—	—	0.01	0.57
ONEPARDP	0.09	0.06	—	0.05	0.08	0.11	0.50
OWNING	0.41	0.15	—	0.31	0.42	0.52	0.93
PINCLO	0.08	0.06	—	0.04	0.07	0.10	1.00
PRFLUEN	0.03	0.05	—	—	0.01	0.03	0.83

— nil or rounded to zero (including null cells)

Data Extraction, Validation and Clearance continued

2.6 SUMMARY STATISTICS FOR 2001 SEIFA VARIABLES *(continued)*

Variable	Mean	Standard Deviation	Minimum	Lower Quartile	Median	Upper Quartile	Maximum
PURCHAS	0.26	0.13	—	0.16	0.24	0.33	0.89
RECMIG	0.02	0.04	—	—	0.01	0.02	0.69
RENT	0.26	0.17	—	0.13	0.21	0.35	1.00
RENTHI	0.15	0.23	—	—	0.04	0.17	1.00
RENTHI_01	0.18	0.25	—	0.01	0.06	0.25	1.00
RENTLO	0.21	0.22	—	0.06	0.13	0.29	1.00
RENTLO_01	0.25	0.24	—	0.07	0.16	0.35	1.00
SEPDIV	0.11	0.04	—	0.08	0.11	0.13	0.50
SHRTSCH	0.46	0.16	—	0.35	0.47	0.58	1.00
SPD_NDINCHI	0.22	0.24	—	—	0.17	0.33	1.00
SPD_NDINCLO	0.25	0.24	—	—	0.20	0.38	1.00
SPDOINCHI	0.23	0.22	—	0.07	0.19	0.33	1.00
SPDOINCLO	0.19	0.18	—	0.05	0.17	0.28	1.00
SPINCHI	0.23	0.15	—	0.13	0.21	0.32	1.00
SPINCLO	0.43	0.17	—	0.31	0.42	0.54	1.00
TAFE	0.03	0.02	—	0.02	0.03	0.04	0.30
TRADEQL	0.14	0.05	—	0.11	0.14	0.17	0.53
YR11LO	0.57	0.16	0.01	0.47	0.59	0.68	1.00

— nil or rounded to zero (including null cells)

COMPARISON BETWEEN 1996 AND 2001 DATA

Before doing any analysis, the 2001 data were compared with 1996 data to see whether there were unusual changes in the value of each variable.

Table 2.7 presents average changes of variable values between 1996 and 2001. For variables common to both the 1996 and 2001 census, the change in variable value was calculated by subtracting the 1996 value from 2001 value for each CD. The average was then calculated by averaging all CDs. This is shown in table 2.7. Several points need to be made about this table:

- Only changes for CDs common to both the 2001 and 1996 census were calculated. This meant only 33,352 of the 35,695 2001 CDs were included.
- Only variables common to both sets of census data were compared.
- Changes in some CDs might be caused by a CD boundary change and CD renumbering rather than actual variable value changes. The results in table 2.7 did not take into account the impact of these factors. However, because we are looking at the average change across CDs, the problems arising from CD boundary change and CD renumbering should be mitigated to some degree.
- The definitions for DEGREE (persons aged 15 years and over with a degree or higher) and SHRTSCH (persons aged 15 years and over who left school at Year 11 or lower) for 2001 data are different from 1996. This may explain the big changes in these two variables.

For all variables, the change was not large over five years (a maximum change of 8%), with most variables experiencing a change so small that at two decimal places, it is rounded to zero.

Data Extraction, Validation and Clearance continued

2.7 AVERAGE CHANGE OF VARIABLE VALUES ACROSS CDs

Variable	Mean Change (% points)
ATSCH	–0.01
AVEBED	0.05
CAEUNI	0.01
DEGREE	0.07
FADVCLRK	–0.01
FELECLSW	0.01
FEWBED	—
FINTCLRK	—
FINTPRTW	–0.01
FLABOUR	–0.01
FMGR_ADM	—
FPARA_PF	0.01
FPROF	0.01
FTRADE	—
F_UNEMP	–0.02
GRENT	—
GROUP	—
HIGHBED	0.03
HIGHCAR	0.02
IMPDWEL	—
INDIG	—
MADVCLRK	—
MELECLSW	—
MINTCLRK	—
MINTPRTW	—
MLABOUR	0.01
MMGR_ADM	—
MORTHI	0.08
MPARA_PF	—
MPROF	0.01
MTRADE	–0.01
MULTFAM	—
M_UNEMP	–0.02
NGRENT	—
NOCAR	–0.01
NOQUAL	–0.03
NOSCH	—
ONEPARDP	0.04
OWNING	–0.01
PRFLUEN	—
PURCHAS	—
RECMIG	–0.06
RENTHI	0.07
RENTLO	–0.05
SEPDIV	0.01
SHRTSCH	0.10
TAFE	—
TRADEQL	0.02

— nil or rounded to zero (including null cells)

Exploration of some selected variables

Before calculating the SEIFA indices, we undertook an investigation of the interaction between the key variables of socioeconomic status, namely education, occupation and income. This analysis is shown below.

Exploration of some selected variables continued

We also examine the relationship between computer (and Internet) use at home and age. It was recommended in some of the responses to the SEIFA review that computer (or Internet) use should be included in SEIFA because families with a computer (or Internet access) are more advantaged than families without. One major concern with including these two variables was whether using computers or Internet use was largely determined by a person's age, in which case including computer/Internet use in SEIFA would reflect the differences in age rather than advantage/disadvantage across areas.

RELATIONSHIP BETWEEN INCOME, QUALIFICATION AND OCCUPATION

This section examines the relationship between Income, Qualifications and Occupation. For all the graphs, the income groups used are shown in table 2.8.

2.8 GROUPS FOR WEEKLY FAMILY INCOME CUT-OFFS

Group	$ per week
0	<= 0
1	> 0
2	> 39
3	> 79
4	> 119
5	> 159
6	> 199
7	> 299
8	> 399
9	> 499
10	> 599
11	> 699
12	> 799
13	> 999
14	> 1499

Occupation Major Groups (OMGs) and income

OMGs are the broadest level of occupational classification in the Australian Standard Classification of Occupation (ASCO 2nd edition). The nine OMGs in ASCO (2nd edition) are distinguished from each other on the basis of skill level. The nine OMGs and their skill levels are: Manager and Administrators (skill level 1), Professionals (skill level 1), Associate Professionals (skill level 2), Tradespersons and Related Workers (skill level 3), Advanced Clerical, Sales and Service Workers (skill level 3), Intermediate Clerical, Sales and Service Workers (skill level 4), Intermediate Production and Transport Workers (skill level 4), Elementary Clerical, Sales and Service Workers (skill level 5), and Labourers and Related Workers (skill level 5).

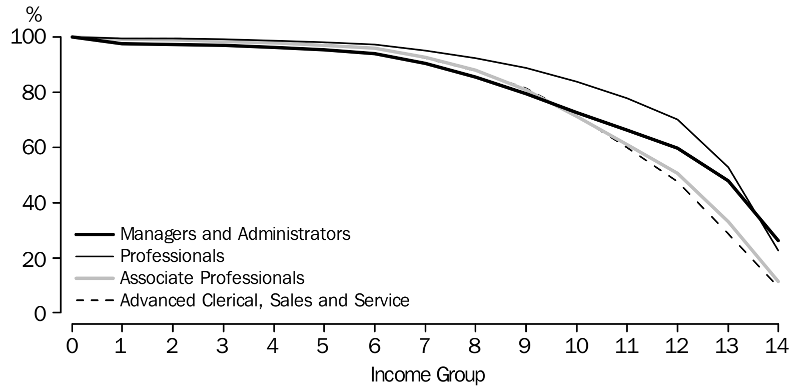
Graphs 2.9 to 2.14 present the relationship between OMGs and income. We look at the relationship separately for men and women because of their different employment and occupational patterns. We have also split the graphs into Advantaged; Middle (i.e. no advantage or disadvantage); and disadvantaged occupations. These were natural splits when all the data were shown on one graph; but due to the amount of information, we have split the graph into the three groups.

Exploration of some selected variables continued

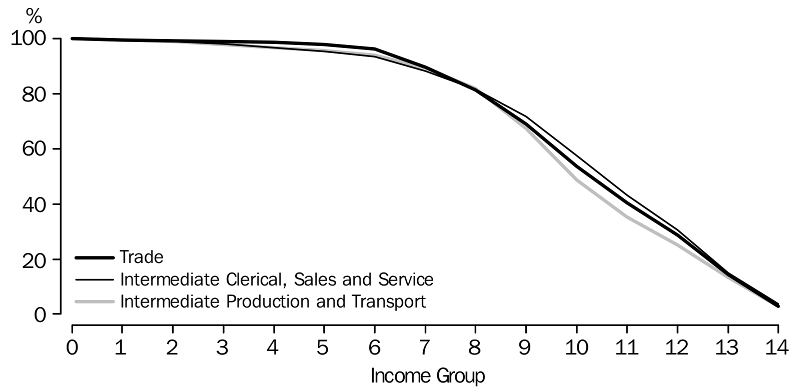
For each graph, the point on a curve indicates the proportion of persons (read from the Y axis) who have a weekly income greater than a certain level (read from the X axis). As such, the higher an OMGs curve, the more advantaged it is in terms of income.

Graph 2.9 for men shows that four OMGs, the Professionals, Managers & Administrators, Associate Professionals and Advanced Clerical, Sales & Service Workers, are in a natural group. Graph 2.11 shows that two OMGs, the Labourers & Related Workers and Elementary Clerical, Sales & Service Workers can be grouped as relatively disadvantaged. Graph 2.10 shows the Intermediate Clerical, Sales & Service Workers, Tradespersons and Intermediate Production & Transport Workers groups lie between.

2.9 CUMULATIVE PER CENT MALES ABOVE INCOME GROUP, Advantaged Occupations

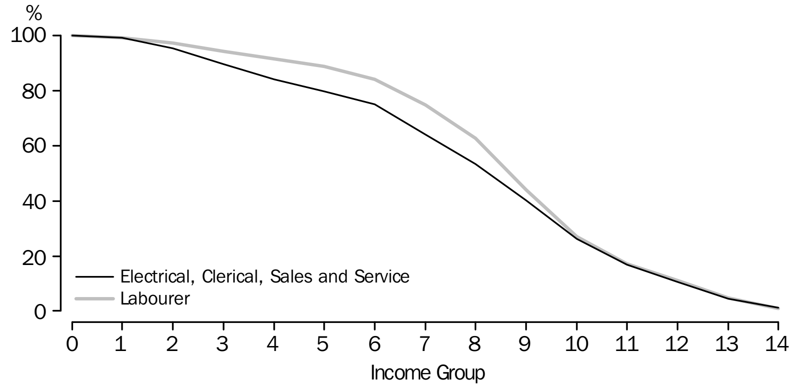


2.10 CUMULATIVE PER CENT MALES ABOVE INCOME GROUP, Middle Occupations



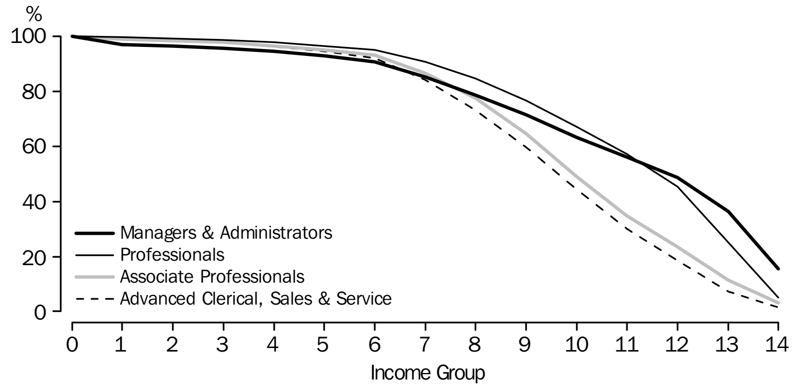
Exploration of some selected variables continued

2.11 CUMULATIVE PER CENT MALES ABOVE INCOME GROUP, Disadvantaged Occupations

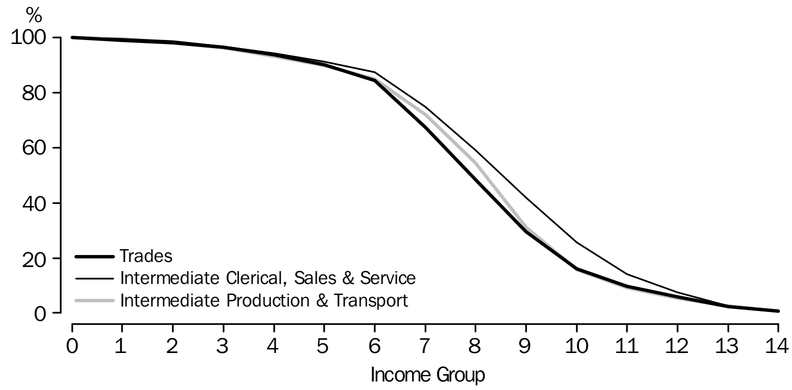


Graphs 2.12 to 2.14 shows the relationship between OMG and income for women. The interpretation is the same as male charts; and the natural splits occur in about the same places (although the split into three groups is not as clear for women).

2.12 CUMULATIVE PER CENT FEMALES ABOVE INCOME GROUP, Advantaged Occupations

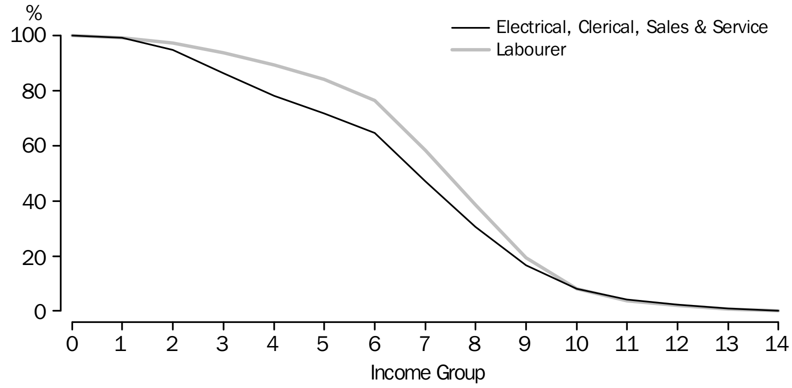


2.13 CUMULATIVE PER CENT FEMALES ABOVE INCOME GROUP, Middle Occupations



Exploration of some selected variables continued

2.14 CUMULATIVE PER CENT FEMALES ABOVE INCOME GROUP, Disadvantaged Occupations



This analysis gives us some understanding of the relationship between occupation and income; but income is only one element of advantage/disadvantage.

We have included all OMGs in the initial Principal Component Analysis (PCA) for the 2001 SEIFA and calculated the correlation of each OMG with the first principal component. If an OMG had a correlation with the first component greater than 0.3 or less than –0.3, it was retained in the final calculation (0.2 and –0.2 for the Index of Disadvantage). The reason for using this criteria is that if an OMG contributes significantly to the index, it should not be dropped.

OMGs and qualifications

We used the same approach as above to examine the relationship between OMGs and qualifications. The qualification groups used are shown in table 2.15.

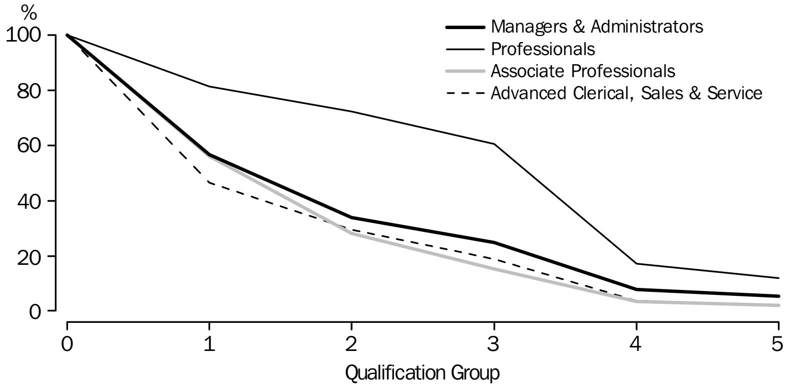
2.15 QUALIFICATION GROUPS

Group	*Qualification*
0	All Persons
1	Certificate or greater
2	Advanced Diploma or greater
3	Bachelor Degree or greater
4	Graduate Diploma/Certificate or greater
5	Postgraduate or greater

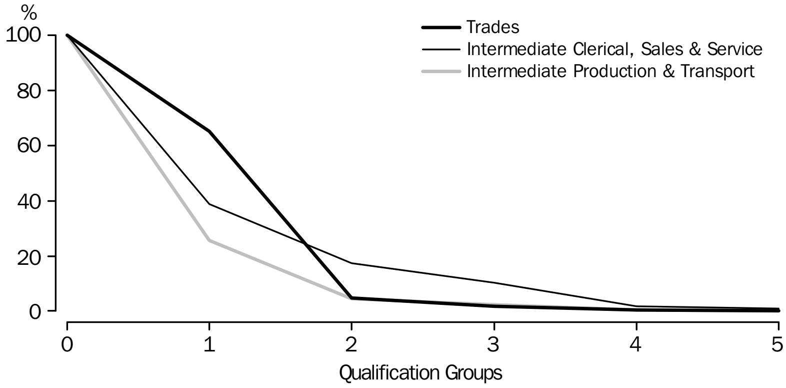
Graphs 2.16 to 2.21 show the relationship between OMGs and qualifications for males and females, respectively. It can be seen that there isn't the same grouping in these graphs as there was in the OMG and Income graphs. However, we have used the same occupations as advantaged and disadvantaged; but when looked at together, the only occupation that stands out is Professionals, who have a much higher education. This supports the assignment of this OMG into the advantage index. However, we cannot say much about the other occupations, because there is not a great difference between them.

Exploration of some selected variables continued

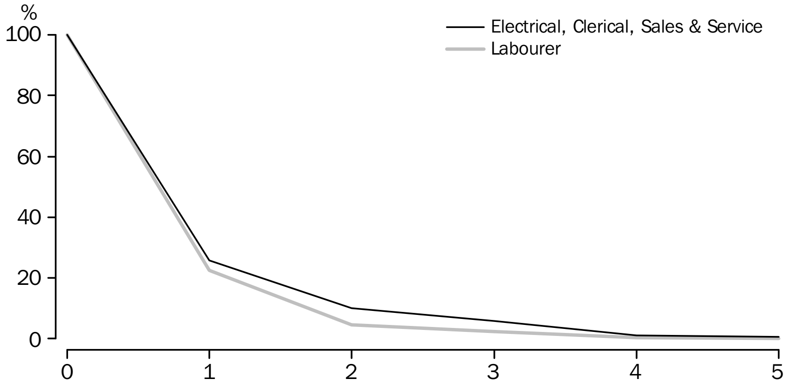
2.16 CUMULATIVE PER CENT MALES HIGHER QUALIFICATIONS, Advantaged Occupations



2.17 CUMULATIVE PER CENT MALES HIGHER QUALIFICATIONS, Middle Occupations

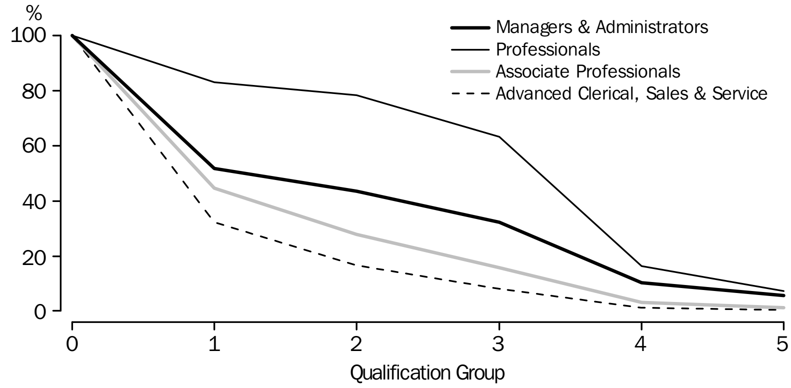


2.18 CUMULATIVE PER CENT MALES HIGHER QUALIFICATIONS, Disadvantaged Occupations

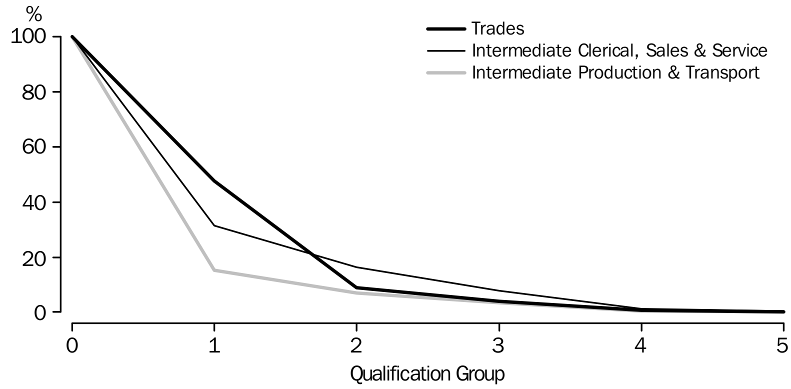


Exploration of some selected variables continued

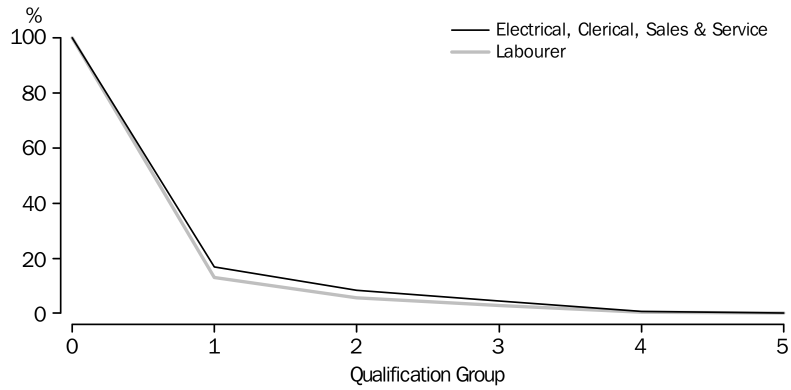
2.19 CUMULATIVE PER CENT FEMALES HIGHER QUALIFICATIONS, Advantaged Occupations



2.20 CUMULATIVE PER CENT FEMALES HIGHER QUALIFICATIONS, Middle Occupations



2.21 CUMULATIVE PER CENT FEMALES HIGHER QUALIFICATIONS, Disadvantaged Occupations



Income and qualification

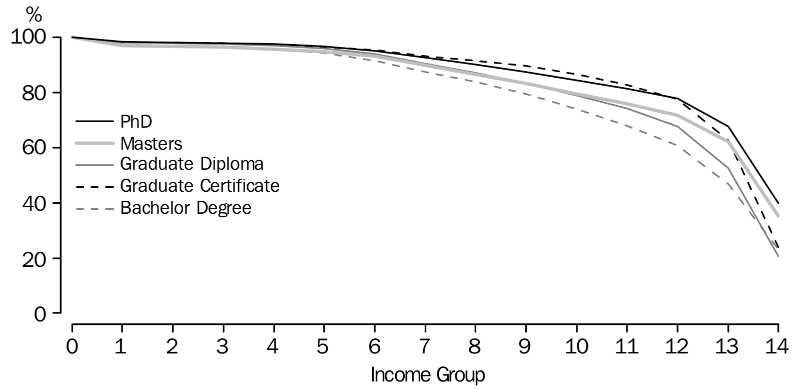
We looked at the relationship between income and qualifications. This gives us some idea of which qualifications are more advantaged due to the additional income earned by people with this qualification. The income groups used are shown in table 2.8.

Exploration of some selected variables continued

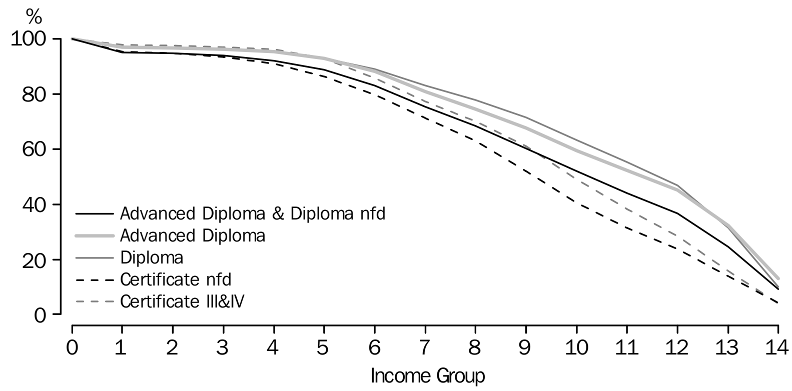
Graphs 2.22 to 2.25 present the relationship between income and qualifications for males and females, respectively. Each curve shows the proportion of population (with that qualification) having weekly income greater than a specified level. A graph with all educations plotted showed a natural break at the Bachelor degree; and the graphs shown here use this break to split this one graph into two graphs, showing bachelor degree and above; and below bachelor degree.

For both males and females, it was clear that individuals with a bachelor degree or higher earn more than those without such degrees.

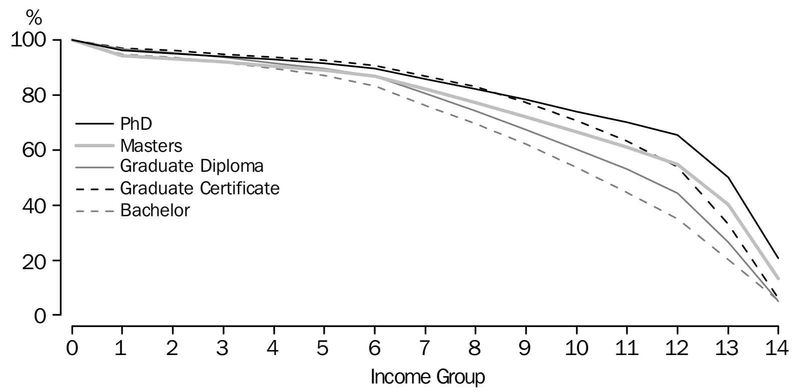
2.22 CUMULATIVE PER CENT MALES ABOVE INCOME GROUP, Advantaged Qualifications



2.23 CUMULATIVE PER CENT MALES ABOVE INCOME GROUP, Disadvantaged Qualifications

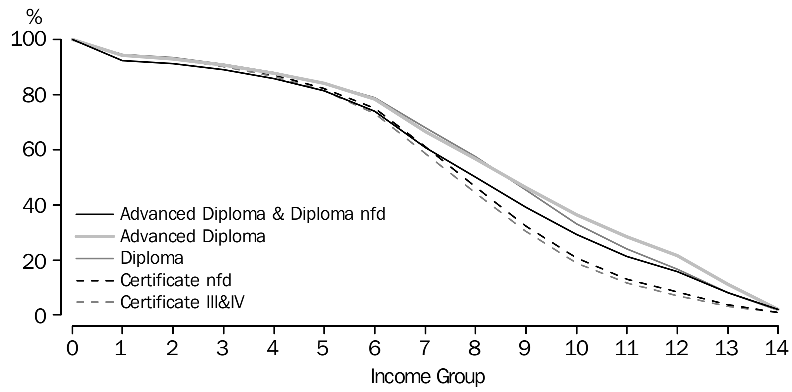


2.24 CUMULATIVE PER CENT FEMALES ABOVE INCOME GROUP, Advantaged Qualifications



Exploration of some selected variables continued

2.25 CUMULATIVE PER CENT FEMALES ABOVE INCOME GROUP, Disadvantaged Qualifications



The 1996 SEIFA defined three variables based on qualifications: (a) persons aged 15 years and over with a degree or higher (defined as having an associate diploma or higher qualification); (b) persons aged 15 years and over with a certificate (defined as having a trade or other non-diploma qualifications); and (c) persons aged 15 years and over having no qualification. Graphs 2.22 to 2.25 suggest that these groups are no longer sensible reflections of income related to advantage/disadvantage.The figures suggest that we can group the 'bachelor' and 'higher' degree holders into one advantaged group (% Persons aged 15 years and over with degree or higher), and 'Certificate' can be classified as less advantaged (% Persons aged 15 years and over with certificate qualification). We have also included a Diploma variable (% Persons aged 15 years and over having an advanced diploma or diploma qualification), as less advantaged; and % Persons aged 15 years and over with no qualifications as disadvantaged.

RELATIONSHIP BETWEEN COMPUTER (AND INTERNET) USE AND AGE

The 2001 census asked questions about whether a computer was used at home and how an individual got access to the Internet. In user consultations, it was suggested these may be related to advantage. However, some users also suggested that these would be more closely related to age. The following two figures (graphs 2.26 and 2.27) show the proportion of the population using computers and Internet at home by age. Due to the wide access to the Internet in workplaces and public libraries nowadays, most people can get access to the Internet. So access to the Internet itself may not be a good measure of advantage due to its wide public availability. This is the reason we looked at Internet use at home only.

Understandably, the shape of the age profiles for computer and Internet use at home is very similar (to use the Internet at home, you generally need a computer at home), but Internet use is lower than computer use. It is clear from the figures that computer and Internet use at home are closely related to age. Computers and the Internet are most likely to be used by people aged 15–19 years, perhaps because they are most likely to be students who use computers/Internet as study tools. People aged 35–44 years are the second most likely population to use computers and internet at home. Older people,

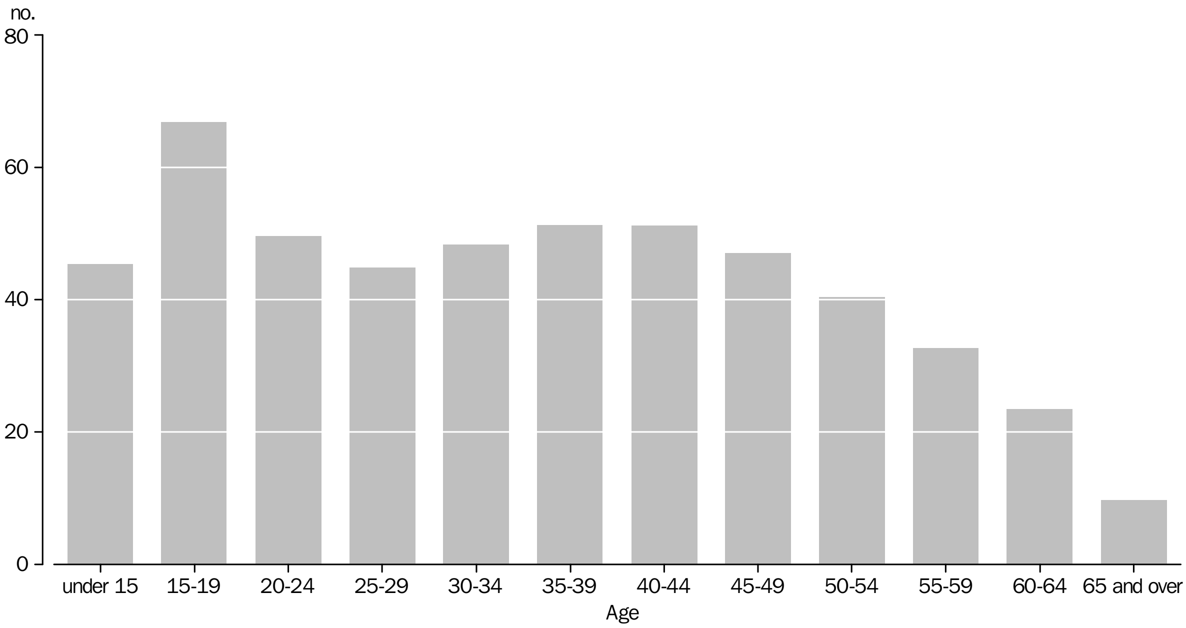
Exploration of some selected variables continued

especially those aged 65 years and over, are the least likely to use computers and Internet at home.

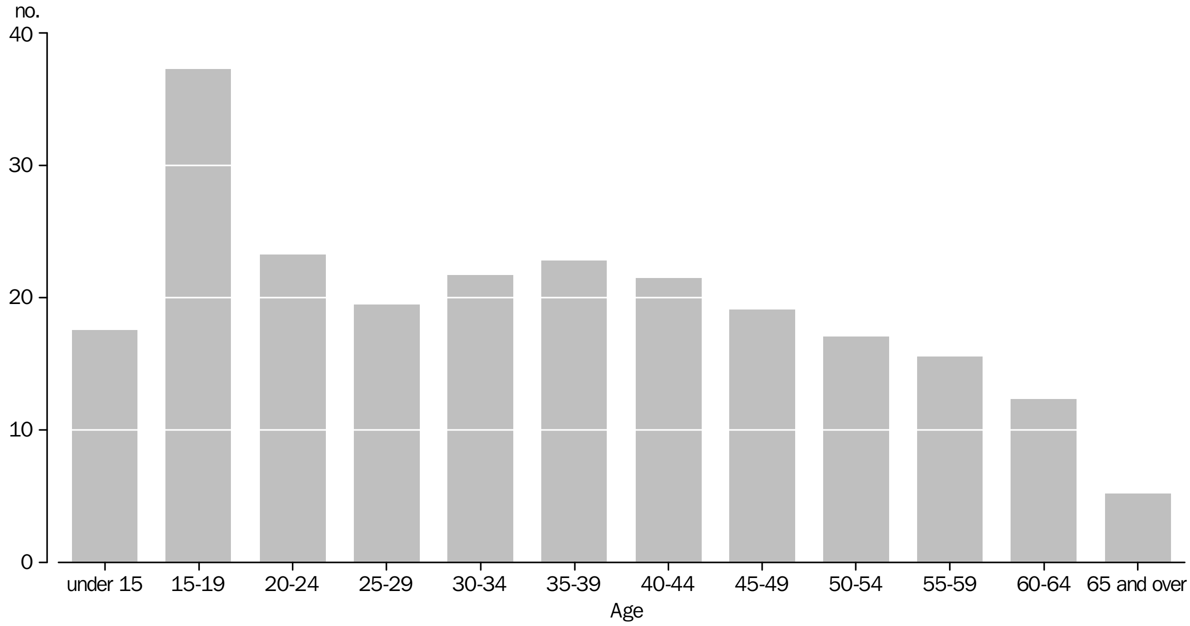
The close relationship between computer/Internet use at home and age might have invalidated the inclusion of these variables in SEIFA because they would reflect differences in age rather than advantage across areas. We decided to include them in the initial set of variables before taking a decision about including them in the final index, to see if they contributed anything extra to the index of advantage/disadvantage.

After including both variables in the initial variable list for the advantage/disadvantage index, we found that they were highly correlated. We decided to include the Internet use variable only. This variable has a final correlation of 0.671 with the advantage/disadvantage index. Because this was fairly high, we decided it was contributing to the index as an indicator of access to services, and left it in the model.

2.26 PER CENT OF POPULATION USING COMPUTERS AT HOME BY AGE GROUP



2.27 PER CENT OF POPULATION USING INTERNET AT HOME BY AGE GROUP



Limitations with the data

The data used for the 2001 SEIFA Indexes is from the 2001 census. There are some limitations to these data, which are outlined in this section.

First, there is limited information on an important component of disadvantage — wealth. We included some wealth variables, being the number of cars at a dwelling, number of bedrooms in a dwelling, and whether the occupant is owning or purchasing the dwelling. Ideally, though, we would have used variables like the value of owned dwellings, value of cars, and value of other investments.

Second, the census has no information on remoteness. The ABS remoteness indicators provide an indication of where remote areas are, but no indicators of remoteness (e.g. distance to major town) are available directly from the census.

Third, the census has no information on access to infrastructure such as schools, shops, community services, transport and hospitals. We have used Internet use at home as a broad measure of such access.

Fourth, the census does not collect data on Health status.

Fifth, the income variable on the census is measured in intervals, not exact values. This makes calculating quintiles difficult. An extreme example of this was low income for single persons — about 24% of single people were in one income group. This group covered the 19th to 43rd percentile of the income distribution for single people. While the first quintile is in this range, so is the second.

The groups also mean we cannot update the income range each census; and when it is updated, the change is quite large, as low income moves into the next grouped income range. It also makes it difficult to equivalise family income to reflect family size.

While it is important to understand the limitations of the data being used, we consider that the census is the best dataset for calculating SEIFA.

CHAPTER 3 METHOD FOR CALCULATING THE 2001 INDEXES ..

THE REVIEW OF SEIFA 2001

In August 2002, the ABS started reviewing the methods and concepts applied in the SEIFA indexes in preparation for publishing the indexes using 2001 census data. A number of stakeholders were consulted in the review, including researchers, Commonwealth and state government departments, users of the indexes, and information groups (groups of technical experts and users formed to discuss ABS data and surveys). In total, approximately 200 stakeholders were contacted about the indexes. Meetings were held with four Commonwealth departments, and we contacted several other departments who provided verbal or written comment. Meetings were held in all states, with representatives from state government departments present at each meeting. Comments from many of the information groups were also received. The main comments we had were that:

- some users felt that SEIFA should not contain the '% Indigenous Persons' as a variable because while the proportion of Indigenous people in an area might reflect disadvantage, it was not a direct measure of disadvantage
- including some of the computer/Internet use variables from the census would add value to the indexes
- there was a demand for postal area indexes
- component rotation should be investigated
- making available more technical information would help users understand the indexes.

The review of SEIFA included an extensive literature search, looking at best practice overseas. We then wrote a discussion paper, which included a variable selection strategy, and a way to prioritise the variables associated with disadvantage. This discussion paper was the basis of SEIFA review meetings held in December 2002.

We would like to thank everyone who contributed to the SEIFA Review. A special thanks is due to our Technical Reference Group, who contributed their time and knowledge to enable the ABS to deliver a better SEIFA index.

INDEX CALCULATION — THE PROCEDURE

This section describes the main steps used for producing all SEIFA indexes. The details of the calculation for each index are given in the next chapter.

The Socio-Economic Indexes for Areas use a method of summarising a number of variables called Principal Components Analysis (PCA). This procedure creates a number of new uncorrelated components, each of which is a linear combination of the original variables. This smaller set of uncorrelated components is easier to understand and use in further analysis than a large number of correlated variables.

INDEX CALCULATION — THE PROCEDURE *continued*

The procedure gives an eigenvalue for each component, which indicates how much of the variance in the original data is explained by the component. The eigenvalue can be converted to the per cent of variance explained by dividing it by the number of variables in the analysis.

Each variable in the analysis will be correlated with each component. This correlation is called the loading. Loadings help to interpret what aspects of disadvantage a component may represent. For instance, if mainly education variables have high loadings on a component, then this component mainly reflects education disadvantage. Alternatively, a component may be a very general measure of disadvantage, so no one aspect of disadvantage has high loadings on a component; but all aspects have moderate loadings.

The loading can be converted to a weight by dividing it by the square root of the eigenvalue. The loadings are more useful in analysing the results from using different sets of original variables, and these are what we have shown in most tables. The product of the weight and standardised variable values are summed to produce a raw score. In the results, we have only shown the final set of weights used to get the SEIFA values. This raw score is standardised to a mean of 1,000 and standard deviation of 100 to give the SEIFA value.

Step 1. Making an initial list of variables

The initial list of variables was based on the variable selection strategy outlined in Chapter 2. Our user consultation also informed our variable selection. Chapter 2 outlined the process in detail.

Step 2. Removing highly correlated variables

Highly correlated variables were removed to ensure the stability of the resultant variable weights. Whenever two variables had a correlation coefficient greater than 0.8 they were considered to be highly correlated. Generally one of the highly correlated variables was removed. However, if two variables had a high correlation, but were measuring different aspects of advantage/disadvantage, both were retained. This is similar to the criteria used in previous SEIFA indexes.

For example, the variable '% Employed Females classified as Professionals' (FPROF) is highly correlated with the variable '% Persons aged 15 years and over with degree or higher' (DEGREE). Because these two variables are measuring different concepts (education and occupation), both were retained.

Step 3. Initial PCA and rotation

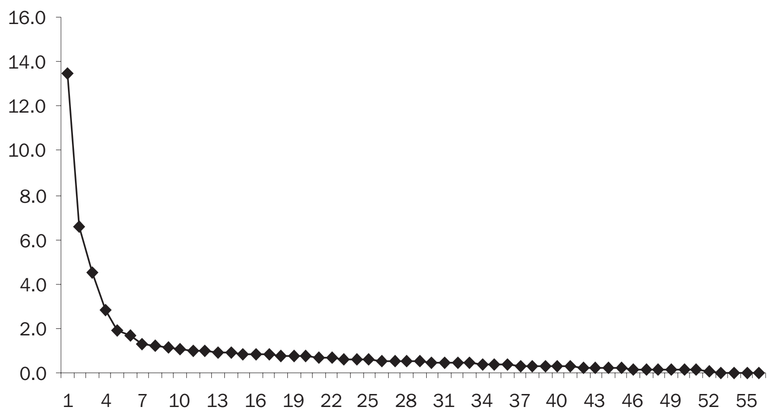
After removing some of the highly correlated variables in step 2, an initial PCA was conducted to produce the unrotated loadings. This analysis was done using the correlation matrix. In the past, variables with a positive loading were used for the Advantage indexes; and variables with a negative loading were used for the Disadvantage index. This was not done in 2001. The Disadvantage index in 2001 used the same variables as the 1996 index, to maintain consistency; and the Advantage index was replaced with the Advantage/Disadvantage index, which used all variables.

Step 3. Initial PCA and rotation continued

Rotation simplifies the pattern of component loadings to obtain a more readily interpretable solution. It will give variables higher loadings on one component, making the components more readily interpretable. Running a principal components analysis with rotation will give a different set of loadings for each variable compared to an unrotated principal components analysis. The loadings for a rotated analysis will also depend on how many components are retained.

For all indexes except the Disadvantage index, we looked at whether rotating the components gave a better first component; or whether the second, third and further components after rotation were meaningful. For rotation, we need a criteria to decide how many components to keep, as the results after rotation will be different depending on how many components are kept. One criteria is to keep any of the unrotated components with an eigenvalue greater than 1. Another criteria is to use Cattell's scree test. Cattell's Scree Test plots the eigenvalues for all unrotated components. At some point, the graph levels out; and the number of components just before the plot levels out is the number of components to keep. A scree plot is shown in figure 3.1. From this scree plot, we can see the levelling out at component number 5; so we keep 4 components. We used both these criteria to decide how many components to retain; but have used the number of components identified by the Cattells Scree Test for the examples shown in this paper.

3.1 SAMPLE SCREE PLOT



The orthogonal VARIMAX rotation method was used for all three indexes. We also looked at an oblique rotation, but found the results harder to interpret. After comparing the rotated and unrotated weights, we did not find rotated weights easier to interpret. Second and higher components were not showing any obvious interpretation; and the first unrotated component appeared to give a better indication of general advantage/disadvantage, since no one variable was loading strongly on this component.

We decided that the first unrotated component gave a better general index of advantage/disadvantage. This was confirmed by the expert group used to advise the project.

Step 4. Remove variable if low correlation with the first component

The correlation between a variable and a principal component is known as a loading. If a variable is poorly correlated with a principal component, it makes a small contribution to the component because its weight will normally be small. Excluding the variable will not make much difference to the final index score. In the 1996 SEIFA, if the absolute value of the correlation coefficient of a variable and the first component was less than 0.2, the variable was removed. Based on comments from experts, and a literature search, this cut-off was increased to 0.3 in 2001 SEIFA (see Jolliffe, (1986), p. 108, 111). However, for the Disadvantage index, where we tried to follow the 1996 procedure closely, the cut-off was left at 0.2.

Step 5. Calculating CD level raw and standardised index scores

Once variables poorly correlated with the first unrotated component were removed, we conducted a new PCA on the remaining variables. The new weights were applied to the standardised raw variables to calculate the raw index score for each CD. For the Disadvantage index, the weights were multiplied by –1, so lower numbers were more disadvantaged. For the Index of Education and Occupation, the procedure assigned negative weights to variables associated with advantage; so we multiplied the raw score by –1 so higher numbers represented advantage. For presentation convenience, the raw scores are standardised to a mean of 1,000 and standard deviation of 100.

Step 6. Calculating indexes for higher area levels

Index scores for higher area levels than CD are calculated by weighting together constituent CDs scores by their populations, so the formula is:

$SEIFA_{SLA}=Sum_{SLA}(SEIFA_{CD}*Pop_{CD})/Pop_{SLA}$

where $Sum_{SLA}(SEIFA_{CD}*Pop_{CD})$ is the summed population weighted SEIFA value for every CD in the SLA.

This aggregation method is discussed further in Chapter 6.

CHAPTER 4 THE CONSTRUCTION OF EACH INDEX

OVERVIEW

This chapter shows how each index has been constructed, showing eigenvalues, loadings and weights. Summary results are also given for Collection District (CD), Statistical Local Area (SLA) and Postal Area (POA) indexes. Where the indexes are not new to 2001, we have also plotted the CD values against the 1996 values to give an idea of how much the index has changed.

Only one index is conceptually comparable to the 1996 index, being the index of disadvantage. The Index of Economic Resources is similar, but uses different income variables and new cut-offs for income, rent and mortgage payments. The Index of Education and Occupation uses new Education variables. We have shown how the 2001 CD rankings compare to the 1996 CD rankings for comparison, but for the Index of Education and Occupation and the Index of Economic Resources, the change can be attributed to both data and new variables. For the Index of Disadvantage, the only change has been in the data (although some variables changed due to different census questions).

Index of Advantage/Disadvantage

This index included variables measuring both advantage and disadvantage. No variables reflecting advantage or disadvantage (the level 3 variables in the variable selection strategy) were included.

INITIAL LIST OF VARIABLES

The initial list of variables was selected using the decision tree. Both advantage and disadvantage variables were included. A full range of education and economic resource variables were not included, because they were included in the Index of Economic Resources and Index of Education and Occupation.

The list of variables included is shown in table 4.1. The Appendix has a detailed definition of all the variables.

4.1 INITIAL LIST OF VARIABLES USED

VARIABLE	DESCRIPTION
ATSCH	% Persons aged 15 years and over who are still at school
AVEBED	Average number of bedrooms per person
CAEUNI	% Persons aged 15 years and over at University or other tertiary institution
CARAVAN	% Persons living in caravan park
CERT	% Persons aged 15 years and over with a certificate qualification
CFD_NDINCHI	% Couple families with dependents and non-dependents or with non-dependents only with annual income greater than $103,999
CFD_NDINCLO	% Couple families with dependents and non-dependents or with non-dependents only with annual income less than $52,000
CFDOINCHI	% Couple families with dependent child(ren) only with annual income greater than $77,999
CFDOINCLO	% Couple families with dependent child(ren) only with annual income less than $36,400
CFNKINCHI	% Couple families with no children with annual income greater than $77,999
CFNKINCLO	% Couple families with no children with annual income less than $20,800
COMP	% Persons using computer at home
DEGREE_01	% Persons aged 15 years and over with degree or higher
DIP	% Persons aged 15 years and over having an advanced diploma or diploma qualification
FADVCLRK	% Employed Females classified as 'Advanced Clerical & service Workers'
FELECLSW	% Employed Females classified as 'Elementary Clerical, Sales & Service Workers'
FEWBED	% Dwellings with one or no bedrooms
FINTCLRK	% Employed Females classified as 'Intermediate Clerical, Sales & Service Workers'
FINTPRTW	% Employed Females classified as 'Intermediate Production & Transport Workers'
FLABOUR	% Employed Females classified as 'Labourers & Related Workers'
FMGR_ADM	% Employed Females classified as 'Managers or Administrators'
FPARA_PF	% Employed Females classified as 'Associate Professionals'
FPROF	% Employed Females classified as 'Professionals'
FTRADE	% Employed Females classified as 'Tradespersons'
FUNEMP	% Females (in labour force) unemployed
GROUP	% Households who are group households
HIGHBED	% Dwellings with four or more bedrooms
HIGHCAR	% Dwellings with three or more motor cars
IMPDWEL	% Households living in improvised dwellings
INTNT	% Persons using Internet at home
MADVCLRK	% Employed Males classified as 'Advanced Clerical & Service Workers'
MELECLSW	% Employed Males classified as 'Elementary Clerical, Sales & Service Workers'
MINTCLRK	% Employed Males classified as 'Intermediate Clerical, Sales & Service Workers'
MINTPRTW	% Employed Males classified as 'Intermediate Production & Transport Workers'
MLABOUR	% Employed Males classified as 'Labourers & Related Workers'
MMGR_ADM	% Employed Males classified as 'Managers or Administrators'
MPARA_PF	% Employed Males classified as 'Associate Professionals'
MPROF	% Employed Males classified as 'Professionals'
MTRADE	% Employed Males classified as 'Tradespersons'
MULTFAM	% Occupied private dwellings with two or more families
MUNEMP	% Males (in Labour Force) unemployed
NOCAR	% Dwellings with no motor car at dwelling
NOQUAL	% Persons aged 15 years and over with no qualifications
NOSCH	% Persons aged 15 years and over who did not go to school
ONEPARDP	% One-parent families with dependent offspring only
OWNING	% Households owning dwelling
PRFLUEN	% Lacking fluency in English
PURCHAS	% Households purchasing dwelling
RENT	% Rental dwellings
SPD_NDINCHI	% Single parent families with dependents and non-dependents or with non-dependents with annual income greater than $62,399
SPD_NDINCLO	% Single parent families with dependents and non-dependents or with non-dependents with annual income less than $26,000
SPDOINCHI	% Single parent families with dependent child(ren) only with annual income greater than $36,399
SPDOINCLO	% Single parent families with dependent child(ren) only with annual income less than $15,600
SPINCHI	% Single person households with income greater than $36,399
SPINCLO	% Single person household with income less than $15,600
TAFE	% Persons aged 15 years and over at TAFE
YR11LO	% Persons aged 15 years and over who left school at Year 11 or lower

Index of Advantage/Disadvantage continued

REMOVAL OF HIGHLY CORRELATED VARIABLES

Highly correlated variables lead to instability of the resulting weights and the index so they were removed.

Index of Advantage/Disadvantage continued

COMP and INTNET were highly correlated, so we removed COMP from the analysis, because while access to computers at home is becoming more prevalent, lack of access to the Internet may imply a lack of access to certain infrastructure.

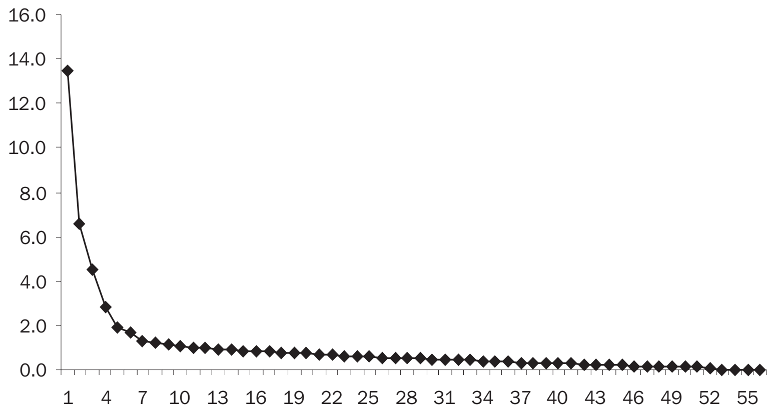
There was a high degree of correlation between many of the education and occupation variables (FPROF-DEGREE, MPROF-DEGREE, MMGR_ADM - FMGR_ADM, NOQUAL - YR11LO), but none of them were removed because they were measuring different aspects of advantage/disadvantage (education and occupation). NOQUAL and YR11LO, while measures of education, measure different aspects (school and post-school). Both aspects were considered important to disadvantage.

ROTATION

Table 4.3 shows the eigenvalues for each component from a PCA run using all variables.

The scree plot is shown in graph 4.2. The number of components before the levelling out of the curve at the 5th component suggests retaining four components.

4.2 SCREE PLOT FOR INDEX OF ADVANTAGE/DISADVANTAGE



Index of Advantage/Disadvantage continued

4.3 EIGENVALUES AND VARIANCE EXPLAINED FROM INITIAL PCA

Component	Eigenvalue	Variance explained	Cumulative
1	13.4796	0.2407	0.2407
2	6.6216	0.1182	0.3590
3	4.5132	0.0806	0.4395
4	2.8206	0.0504	0.4899
5	1.9456	0.0347	0.5247
6	1.6720	0.0299	0.5545
7	1.2855	0.0230	0.5775
8	1.2039	0.0215	0.5990
9	1.1471	0.0205	0.6194
10	1.0648	0.0190	0.6385
11	0.9903	0.0177	0.6561
12	0.9659	0.0172	0.6734
13	0.9267	0.0165	0.6899
14	0.8877	0.0159	0.7058
15	0.8606	0.0154	0.7212
16	0.8518	0.0152	0.7364
17	0.8067	0.0144	0.7508
18	0.7887	0.0141	0.7649
19	0.7340	0.0131	0.7780
20	0.7279	0.0130	0.7910
21	0.7070	0.0126	0.8036
22	0.6538	0.0117	0.8153
23	0.6305	0.0113	0.8265
24	0.5976	0.0107	0.8372
25	0.5851	0.0104	0.8476
26	0.5677	0.0101	0.8578
27	0.5611	0.0100	0.8678
28	0.5318	0.0095	0.8773
29	0.5001	0.0089	0.8862
30	0.4826	0.0086	0.8948
31	0.4639	0.0083	0.9031
32	0.4393	0.0078	0.9110
33	0.4290	0.0077	0.9186
34	0.3882	0.0069	0.9256
35	0.3690	0.0066	0.9322
36	0.3548	0.0063	0.9385
37	0.3370	0.0060	0.9445
38	0.3067	0.0055	0.9500
39	0.2972	0.0053	0.9553
40	0.2873	0.0051	0.9604
41	0.2701	0.0048	0.9652
42	0.2413	0.0043	0.9696
43	0.2373	0.0042	0.9738
44	0.2224	0.0040	0.9778
45	0.2064	0.0037	0.9814
46	0.1844	0.0033	0.9847
47	0.1791	0.0032	0.9879
48	0.1715	0.0031	0.9910
49	0.1609	0.0029	0.9939
50	0.1359	0.0024	0.9963
51	0.1152	0.0021	0.9984
52	0.0511	0.0009	0.9993
53	0.0363	0.0006	0.9999
54	0.0032	0.0001	1.0000
55	0.0014	—	1.0000
56	—	—	1.0000

— nil or rounded to zero (including null cells)

The results from rotating the matrix using an orthogonal (varimax) rotation and retaining four components is shown in table 4.4.

Index of Advantage/Disadvantage continued

4.4 RESULTS AFTER ROTATION WITH FOUR COMPONENTS RETAINED

	Component 1	*Component 2*	*Component 3*	*Component 4*
Eigenvalues	12.8458	5.7749	5.2608	3.5534
Variance explained	0.2294	0.1031	0.0939	0.0635

The aim of the rotation was to to check whether rotation makes interpretation of the index easier, or gives any interpretable components after the first component. To check this, we need to look at the loadings of the variables on each component. This is shown in table 4.5 for four components retained.

Index of Advantage/Disadvantage continued

4.5 LOADINGS WITH ROTATION WITH FOUR COMPONENTS RETAINED

Variable	Component 1	Component 2	Component 3	Component 4
ATSCH	0.0165	–0.3863	–0.0977	0.2073
AVEBED	–0.0052	–0.1021	0.1689	–0.6680
CAEUNI	0.5984	0.3360	–0.1320	0.0560
CARAVAN	–0.1197	0.1210	0.0526	–0.1128
CERT	–0.2472	–0.5007	–0.2405	–0.4095
CFDOINCHI	0.8368	–0.0172	0.1028	–0.0868
CFDOINCLO	–0.6767	0.3006	0.1488	0.2412
CFD_NDINCHI	0.6207	–0.1758	0.0730	0.0303
CFD_NDINCLO	–0.5168	0.2607	0.0756	0.0332
CFNKINCHI	0.8397	0.0667	0.0990	0.0017
CFNKINCLO	–0.6849	0.1791	–0.0171	0.1975
DEGREE_01	0.8887	0.2990	0.1172	–0.0540
DIP	0.7233	0.0189	0.1138	–0.1868
FADVCLRK	0.3005	–0.2811	0.0036	–0.1401
FELECLSW	–0.2948	–0.1379	–0.5621	0.0070
FEWBED	0.0618	0.6275	0.0557	–0.0122
FINTCLRK	–0.1297	–0.1947	–0.6608	–0.0182
FINTPRTW	–0.3705	–0.1155	–0.2297	0.3714
FLABOUR	–0.6322	0.0608	0.0080	0.3894
FMGR_ADM	–0.0243	0.0008	0.8636	0.0031
FPARA_PF	0.1620	0.1136	–0.1342	–0.3768
FPROF	0.7433	0.2819	0.1174	–0.1287
FTRADE	–0.2852	–0.0266	–0.0156	–0.0521
F_UNEMP	–0.4796	0.3505	–0.3415	0.1600
GROUP	0.3514	0.5969	–0.1868	–0.0336
HIGHBED	0.2263	–0.7006	0.2462	–0.0096
HIGHCAR	–0.0228	–0.6176	0.5693	0.0372
IMPDWEL	–0.1238	0.1930	0.2313	0.0481
INTNET	0.6314	–0.4683	0.0188	–0.1380
MADVCLRK	0.3253	0.0593	–0.1275	–0.0567
MELECLSW	0.0281	0.1375	–0.5185	0.0286
MINTCLRK	0.3363	0.1436	–0.5178	–0.0306
MINTPRTW	–0.6125	–0.1564	–0.3547	0.2097
MLABOUR	–0.6345	0.1340	0.0882	0.3891
MMGR_ADM	0.0499	–0.1036	0.8919	–0.0553
MPARA_PF	0.4750	0.0767	–0.2699	–0.3105
MPROF	0.8325	0.3112	–0.0115	–0.1020
MTRADE	–0.4608	–0.3852	–0.5092	–0.1446
MULTFAM	–0.0484	–0.0374	–0.0049	0.6446
M_UNEMP	–0.5073	0.4257	–0.3239	0.0436
NOCAR	–0.0565	0.7706	–0.2358	0.2890
NOQUAL	–0.8796	0.0119	–0.0147	0.2503
NOSCH	–0.1275	0.1379	–0.0832	0.7391
ONEPARDP	–0.3780	0.3518	–0.4721	0.0471
OWNING	–0.0515	–0.3436	0.5175	–0.2276
PRFLUEN	0.0265	0.1977	–0.1652	0.7440
PURCHAS	0.1334	–0.6942	–0.3447	–0.0819
RENT	–0.0247	0.7376	–0.3526	0.2167
SPDOINCHI	0.6080	–0.0652	–0.0204	–0.0159
SPDOINCLO	–0.2235	0.1655	0.0307	0.2383
SPD_NDINCHI	0.4758	–0.1574	–0.0196	0.1295
SPD_NDINCLO	–0.3584	0.1931	–0.0261	–0.0942
SPINCHI	0.7232	–0.1511	–0.0145	0.0878
SPINCLO	–0.6452	0.2672	–0.0373	–0.0684
TAFE	0.0357	0.0304	–0.3748	0.1312
YR11LO	–0.9144	–0.1870	0.0098	–0.0339

Index of Advantage/Disadvantage continued

The first component had some income variables loading strongly; but not all (e.g. SPDOINCLO). Also, some of the education variables were loading (NOQUAL, YR11LO); but not all (e.g. TAFE). The second component had some wealth variables loading (HIGHCAR, HIGHBED, PURCHAS, NOCAR), but not all wealth variables were loading strongly (e.g. OWNING); and some non-wealth variables were loading (e.g. GROUP). The third component had some occupations loading strongly (e.g. Managers and Administrators), but there was no general trend. The fourth component had a range of variables loading strongly (AVEBED, MULTFAM, NOSCH, PRFLUEN), with no obvious relationship between them.

Overall, the rotated weights did not show any coherent or distinct patterns. In addition, because we are looking for a general measure for advantage/disadvantage, the weights associated with the first unrotated component appear to load better on a variety of variables. We therefore decided to use the first unrotated component to produce this index.

REMOVAL OF VARIABLES POORLY CORRELATED WITH THE FIRST UNROTATED COMPONENT

Variables with a loading on the first unrotated component in the range –0.3 to 0.3 were dropped from the final index calculation because their contribution to the index was negligible. This was an iterative stage — removing some of the low loading variables meant weights for other variables changed, which gave them a lower loading. The step continued until there were no more variables with a loading in the range –0.3 to 0.3.

The loadings and weights for the initial run and the final run are shown in table 4.6. These weights and loadings were calculated after specifying that we only wanted to retain one component. There were three iterations.

Index of Advantage/Disadvantage continued

4.6 LOADINGS AND WEIGHTS, ADVANTAGE/DISADVANTAGE

Variable	Initial Loading	Initial Weight	Final Loading	Final Weight
ATSCH	0.0040	0.0010	na	na
AVEBED	–0.1690	–0.0460	na	na
CAEUNI	–0.5100	–0.1390	–0.5285	–0.1482
CARAVAN	0.0960	0.0260	na	na
CERT	0.1410	0.0380	na	na
CFDOINCHI	–0.8400	–0.2290	–0.8490	–0.2381
CFDOINCLO	0.7070	0.1920	0.7048	0.1977
CFD_NDINCHI	–0.6190	–0.1690	–0.6268	–0.1758
CFD_NDINCLO	0.5180	0.1410	0.5220	0.1464
CFNKINCHI	–0.8150	–0.2220	–0.8291	–0.2325
CFNKINCLO	0.7190	0.1960	0.7143	0.2003
DEGREE_01	–0.8540	–0.2330	–0.8699	–0.2440
DIP	–0.7490	–0.2040	–0.7529	–0.2111
FADVCLRK	–0.3460	–0.0940	–0.3416	–0.0958
FELECLSW	0.3580	0.0980	0.3522	0.0988
FEWBED	–0.0090	–0.0020	na	na
FINTCLRK	0.2040	0.0560	na	na
FINTPRTW	0.4570	0.1240	0.4450	0.1248
FLABOUR	0.6930	0.1890	0.6839	0.1918
FMGR_ADM	–0.1100	–0.0300	na	na
FPARA_PF	–0.2010	–0.0550	na	na
FPROF	–0.7310	–0.1990	–0.7437	–0.2086
FTRADE	0.2630	0.0720	na	na
F_UNEMP	0.5810	0.1580	0.5630	0.1579
GROUP	–0.2570	–0.0700	na	na
HIGHBED	–0.3270	–0.0890	–0.3011	–0.0845
HIGHCAR	–0.1190	–0.0330	na	na
IMPDWEL	0.1120	0.0310	na	na
INTNET	–0.6840	–0.1860	–0.6712	–0.1882
MADVCLRK	–0.2990	–0.0810	na	na
MELECLSW	0.0730	0.0200	na	na
MINTCLRK	–0.2350	–0.0640	na	na
MINTPRTW	0.6710	0.1830	0.6636	0.1861
MLABOUR	0.6900	0.1880	0.6786	0.1903
MMGR_ADM	–0.2080	–0.0570	na	na
MPARA_PF	–0.4710	–0.1280	–0.4829	–0.1354
MPROF	–0.7890	–0.2150	–0.8090	–0.2269
MTRADE	0.4540	0.1240	0.4542	0.1274
MULTFAM	0.1770	0.0480	na	na
M_UNEMP	0.5890	0.1600	0.5744	0.1611
NOCAR	0.2260	0.0620	na	na
NOQUAL	0.9010	0.2450	0.9070	0.2544
NOSCH	0.3010	0.0820	na	na
ONEPARDP	0.4810	0.1310	0.4608	0.1292
OWNING	–0.1110	–0.0300	na	na
PRFLUEN	0.1730	0.0470	na	na
PURCHAS	–0.1600	–0.0440	na	na
RENT	0.1960	0.0530	na	na
SPDOINCHI	–0.5910	–0.1610	–0.5969	–0.1674
SPDOINCLO	0.2760	0.0750	na	na
SPD_NDINCHI	–0.4430	–0.1210	–0.4531	–0.1271
SPD_NDINCLO	0.3480	0.0950	0.3515	0.0986
SPINCHI	–0.6900	–0.1880	–0.7039	–0.1974
SPINCLO	0.6380	0.1740	0.6428	0.1803
TAFE	0.0540	0.0150	na	na
YR11LO	0.8520	0.2320	0.8717	0.2445

The PCA assigned positive weights to variables associated with disadvantage, and negative weights to variables associated with advantage. In calculating the final indexes, these signs were reversed, so lower numbers represented disadvantage; and higher numbers advantage.

Index of Advantage/Disadvantage continued

Variables associated with disadvantage were CFDOINCLO, CFD_NDINCLO, CFNKINCLO, FELECLSW, FINTPRTW, FLABOUR, FUNEMP, MINTPRTW, MLABOUR, MTRADE, MUNEMP, NOQUAL, ONEPARDP, SPD_NDINCLO, SPINCLO and YR11LO. These related to low income, unemployment, low status occupations and low education.

Variables associated with advantage were CAEUNI, CFDOINCHI, CFD_NDINCHI, CFNKINCHI, DEGREE_01, DIP, FADVCLRK, FPROF, HIGHBED, INTNET, MPARA_PF, MPROF, SPDOINCHI, SPD_NDINCHI and SPINCHI. These related to high income, well paid occupations, high education and high wealth.

The variables with the highest loadings for both advantage and disadvantage were the education variables.

The first component had an eigenvalue of 12.71 and explained 41 per cent of the total variance.

RESULTS OF INDEX SCORES

Tables 4.7 to 4.9 present summary results of the advantage/disadvantage index scores.

The summary indexes at higher geographies (POA and SLA) were calculated using a population weighted average; and the summary to state used a simple average. Care should be taken when interpreting the aggregated indexes, and Chapter 6 outlines some of the issues associated with population weighted averages.

The rankings based on the simple average of CD indexes are also shown in table 4.7. In some cases, a state will be ranked over another state because of a difference at the first or second decimal place. NSW and Vic. show an example of this.

Table 4.7 shows that Tas. is the most disadvantaged state; and the ACT is the most advantaged.

4.7 SUMMARY OF THE ADVANTAGE/DISADVANTAGE CD INDEX

State	Average	Rank	10th percentile	25th percentile	50th percentile	75th percentile	90th percentile
NSW	1 011	7	882	932	999	1 085	1 165
Vic.	1 011	6	889	942	1 003	1 080	1 147
Qld	980	3	878	919	971	1 035	1 104
SA	973	2	859	907	966	1 036	1 108
WA	1 001	5	886	936	994	1 067	1 127
Tas.	943	1	831	881	935	1 007	1 073
NT	987	4	812	914	1 010	1 074	1 114
ACT	1 120	8	1 047	1 082	1 117	1 167	1 208
OT	968	—	826	858	989	1 093	1 148
Aust.	1 000	—	880	928	990	1 068	1 141

— nil or rounded to zero (including null cells)

Index of Advantage/Disadvantage continued

4.8 SUMMARY OF THE ADVANTAGE/DISADVANTAGE POA INDEX

State	Average	10th percentile	25th percentile	50th percentile	75th percentile	90th percentile
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
OT	980	947	947	978	1 015	1 015
Aust.	978	892	923	957	1 023	1 105

4.9 SUMMARY OF THE ADVANTAGE/DISADVANTAGE SLA INDEX

State	Average	10th percentile	25th percentile	50th percentile	75th percentile	90th percentile
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
OT	980	947	947	978	1 015	1 015
Aust.	994	906	930	969	1 053	1 120

These tables show that as the geographic level that the index is calculated at increases in size, the NTs rank in advantage increases. If Advantage/Disadvantage is calculated at the SLA level, the NT is the second most advantaged state, but it is the fourth most advantaged based on CD level indexes. This is because in the NT, more populous CDs are more advantaged. These CDs have more of an impact when the population is used as a weight. This highlights the problems to be aware of when population weighting the CD level indexes to larger geographies.

Index of Disadvantage

This index is a general measure of disadvantage only. Variables reflecting, as well as measuring, socioeconomic status and disadvantage are included.

INITIAL LIST OF VARIABLES

The first step for this index was to take the advantage and disadvantage variables used in 1996, and run a principal components analysis over them. Any variables associated with disadvantage were used for the initial Disadvantage index. Some variables had very small disadvantage weights in the initial run (TRADEQL, ATSCH and TAFE). The loadings for these variables were very low (0.00 – 0.01), and they were excluded because of this low weight.

The initial list of variables used for the Disadvantage index is shown in table 4.10.

4.10 INITIAL LIST OF VARIABLES USED

VARIABLE	DESCRIPTION
FELECLSW	% Employed Females classified as 'Elementary Clerical, Sales & Service Workers'
FEWBED	% Dwellings with one or no bedrooms
FINCLO	% Families with income less than $15,600
FINTCLRK	% Employed Females classified as 'Intermediate Clerical, Sales & Service Workers'
FINTPRTW	% Employed Females classified as 'Intermediate Production & Transport Workers'
FLABOUR	% Employed Females classified as 'Labourers & Related Workers'
FTRADE	% Employed Females classified as 'Tradespersons'
F_UNEMP	% Females (in Labour Force) unemployed
GRENT	% Households renting (government authority)
IMPDWEL	% Households living in improvised dwellings
INDIG	% Indigenous
MELECLSW	% Employed Males classified as 'Elementary Clerical, Sales & Service Workers'
MINTPRTW	% Employed Males classified as 'Intermediate Production & Transport Workers'
MLABOUR	% Employed Males classified as 'Labourers & Related Workers'
MTRADE	% Employed Males classified as 'Tradespersons'
MULTFAM	% Occupied private dwellings with two or more families
M_UNEMP	% Males (in Labour Force) unemployed
NGRENT	% Households Renting (non-government authority)
NOCAR	% Dwellings with no motor car at dwelling
NOQUAL	% Persons aged 15 years and over with no qualifications
NOSCH	% Persons aged 15 years and over who did not go to school
ONEPARDP	% One-parent families with dependent offspring only
PINCLO	% Families with offspring and parental income less than $15,600
PRFLUEN	% Lacking fluency in English
RECMIG	% Recent migrants from non-English speaking countries
SEPDIV	% Persons aged 15 years and over separated or divorced
SHRTSCH	% Persons aged 15 years and over whose highest level of schooling completed was year 10 or lower

Index of Disadvantage continued

See Appendix 1 for a detailed variable definition.

REMOVAL OF HIGHLY CORRELATED VARIABLES

Highly correlated variables lead to instability of the resulting weights and the index. After examination of the correlation matrix of the initial variables, we found no highly correlated variables.

ROTATION

For the other indexes, we examined the use of rotation. But for the disadvantage index, since our aim was to maintain consistency with the 1996 method, rotation was not considered. We have included the eigenvalues and variance explained from the initial analysis in table 4.11.

Index of Disadvantage continued

4.11 EIGENVALUES AND VARIANCE EXPLAINED FROM INITIAL PCA

Component	*Eigenvalue*	*Variance explained*	*Cumulative*
1.0000	13.2804	0.2290	0.2290
2.0000	7.9048	0.1363	0.3653
3.0000	6.5707	0.1133	0.4785
4.0000	3.0499	0.0526	0.5311
5.0000	2.3477	0.0405	0.5716
6.0000	1.8209	0.0314	0.6030
7.0000	1.7096	0.0295	0.6325
8.0000	1.6192	0.0279	0.6604
9.0000	1.4921	0.0257	0.6861
10.0000	1.3518	0.0233	0.7094
11.0000	1.3065	0.0225	0.7320
12.0000	1.1636	0.0201	0.7520
13.0000	1.0843	0.0187	0.7707
14.0000	0.9372	0.0162	0.7869
15.0000	0.8907	0.0154	0.8022
16.0000	0.8224	0.0142	0.8164
17.0000	0.7848	0.0135	0.8299
18.0000	0.7361	0.0127	0.8426
19.0000	0.6940	0.0120	0.8546
20.0000	0.6730	0.0116	0.8662
21.0000	0.6476	0.0112	0.8774
22.0000	0.6035	0.0104	0.8878
23.0000	0.5941	0.0102	0.8980
24.0000	0.5608	0.0097	0.9077
25.0000	0.5509	0.0095	0.9172
26.0000	0.4597	0.0079	0.9251
27.0000	0.4200	0.0072	0.9323
28.0000	0.4066	0.0070	0.9394
29.0000	0.3741	0.0064	0.9458
30.0000	0.3518	0.0061	0.9519
31.0000	0.3410	0.0059	0.9578
32.0000	0.2856	0.0049	0.9627
33.0000	0.2718	0.0047	0.9674
34.0000	0.2648	0.0046	0.9719
35.0000	0.2538	0.0044	0.9763
36.0000	0.2281	0.0039	0.9802
37.0000	0.2050	0.0035	0.9838
38.0000	0.1842	0.0032	0.9869
39.0000	0.1629	0.0028	0.9898
40.0000	0.1419	0.0024	0.9922
41.0000	0.1373	0.0024	0.9946
42.0000	0.1028	0.0018	0.9963
43.0000	0.0921	0.0016	0.9979
44.0000	0.0373	0.0006	0.9986
45.0000	0.0331	0.0006	0.9991
46.0000	0.0110	0.0002	0.9993
47.0000	0.0079	0.0001	0.9995
48.0000	0.0069	0.0001	0.9996
49.0000	0.0069	0.0001	0.9997
50.0000	0.0048	0.0001	0.9998
51.0000	0.0045	0.0001	0.9999
52.0000	0.0034	0.0001	0.9999
53.0000	0.0020	—	1.0000
54.0000	0.0013	—	1.0000
55.0000	0.0010	—	1.0000
56.0000	—	—	1.0000
57.0000	—	—	1.0000
58.0000	—	—	1.0000

— nil or rounded to zero (including null cells)

Index of Disadvantage continued

REMOVING VARIABLES POORLY CORRELATED WITH THE FIRST COMPONENT

The loadings and weights from the initial and final runs are shown in table 4.12. These weights and loadings were calculated after specifying that we only wanted one component retained. Variables with a loading in the range –0.2 to 0.2 were dropped from the inital index calculation because their contribution to the index is negligible. The cut-off was the same as used in the 1996 Index of Disadvantage. This was an iterative step, and resulted in three iterations. Only the first and last are shown in table 4.12.

4.12 LOADINGS AND WEIGHTS, INDEX OF DISADVANTAGE

Variable	Initial Loading	Initial Weight	Final Loading	Final Weight
FELECLSW	0.3423	0.1327	0.3409	0.1337
FEWBED	0.1408	0.0546	na	na
FINCLO	0.5842	0.2264	0.5859	0.2298
FINTCLRK	0.1968	0.0763	na	na
FINTPRTW	0.4622	0.1791	0.4718	0.1850
FLABOUR	0.6642	0.2574	0.6850	0.2687
FTRADE	0.1814	0.0703	na	na
F_UNEMP	0.7080	0.2744	0.7001	0.2746
GRENT	0.5566	0.2157	0.5592	0.2193
IMPDWEL	0.1135	0.0440	na	na
INDIG	0.4547	0.1762	0.4601	0.1804
MELECLSW	0.1890	0.0733	na	na
MINTPRTW	0.5920	0.2294	0.6039	0.2369
MLABOUR	0.6685	0.2591	0.6856	0.2689
MTRADE	0.2797	0.1084	0.2877	0.1128
MULTFAM	0.3222	0.1249	0.3283	0.1288
M_UNEMP	0.6970	0.2701	0.6876	0.2697
NGRENT	0.2010	0.0779	na	na
NOCAR	0.5299	0.2054	0.4882	0.1915
NOQUAL	0.7571	0.2934	0.7788	0.3054
NOSCH	0.4770	0.1848	0.4717	0.1850
ONEPARDP	0.6563	0.2543	0.6461	0.2534
PINCLO	0.7471	0.2895	0.7465	0.2928
PRFLUEN	0.3950	0.1531	0.3750	0.1471
RECMIG	0.1314	0.0509	na	na
SEPDIV	0.5156	0.1998	0.4957	0.1944
SHRTSCH	0.6103	0.2365	0.6393	0.2507

The eigenvalue for the final run was was 6.50, and the variance explained was 32.5 per cent of total variance.

Variables with the highest loadings related to education, some occupations (labourers) and income . All these variables measure socioeconomic status, so in our variable selection strategy, they are level 1 variables. Weights for some occupation variables were low (MTRADE and FELECLSW). Variables reflecting disadvantage (i.e. our level 3 variables — INDIG and SEPDIV), had moderate weights.

The final list of variables, and their weights in 1996 and 2001, is shown in table 4.13.

Index of Disadvantage
continued

4.13 1996 AND 2001 WEIGHTS

Variable	1996	2001
% Persons aged 15 years and over with no qualifications	0.30	0.30
% Families with offspring having parental income less than $15,600	0.30	0.29
% Females (in Labour Force) unemployed	0.29	0.27
% Males (in Labour Force) unemployed	0.28	0.27
% Employed Males classified as 'Labourers & Related Workers'	0.26	0.27
% Employed Females classified as 'Labourers & Related Workers'	0.26	0.27
% One-parent families with dependent offspring only	0.23	0.25
% Persons aged 15 years and over who left school at Year 10 or lower	0.24	0.25
% Employed Males classified as 'Intermediate Production & Transport Workers'	0.24	0.24
% Families with income less than $15,600	0.30	0.23
% Households Renting (Government Authority)	0.22	0.22
% Persons aged 15 years and over separated or divorced	0.19	0.19
% Dwellings with no motor car at dwelling	0.19	0.19
% Employed Females classified as 'Intermediate Production & Transport Workers'	0.19	0.19
% Persons aged 15 years and over who did not go to school	0.16	0.18
% Indigenous	0.16	0.18
% Lacking fluency in English	0.12	0.15
% Employed Females classified as 'Elementary Clerical, Sales & Service Workers'	0.12	0.13
% Occupied private dwellings with two or more families	0.12	0.13
% Employed Males classified as 'Tradespersons'	0.13	0.11

The weights have not changed much between 1996 and 2001. The weight for families with incomes less than $15,600 has dropped, and this was investigated. This decrease was valid; the distribution of this variable had changed considerably between 1996 and 2001.

RESULTS OF INDEX SCORES

Table 4.14 presents summary results of the CD disadvantage index scores. These are the simple average of the CD level scores up to a state.

4.14 SUMMARY OF THE CD INDEX

State	Average	10th percentile	25th percentile	50th percentile	75th percentile	90th percentile
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
OT	834	499	583	952	1 032	1 038
Aust.	1 000	882	951	1 013	1 068	1 110

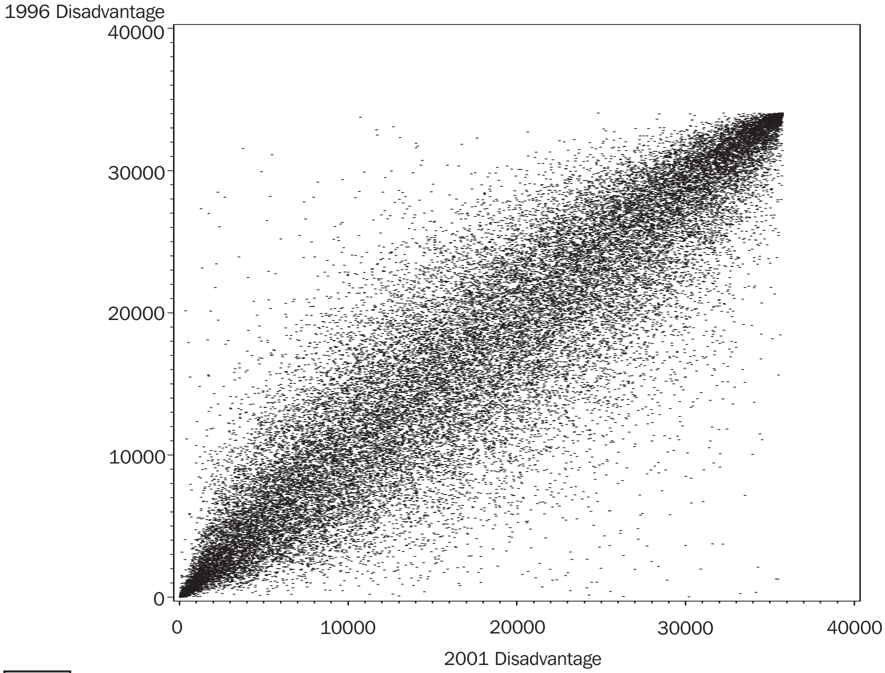
We looked at changes in CD rankings as a result of the revised data (there were no method changes in the Index of Disadvantage). Graphs 4.15 and 4.16 show the new 2001 index compared to the 1996 index. These graphs are used throughout this chapter. They plot the rankings of CDs before and after a change. On the vertical axis is the old (reference) rank; and the horizontal axis shows the new rank. If there are only small changes for most CDs, the first graph will appear as a diagonal line. The larger the change, the more disperse the points in the graph.

Index of Disadvantage continued

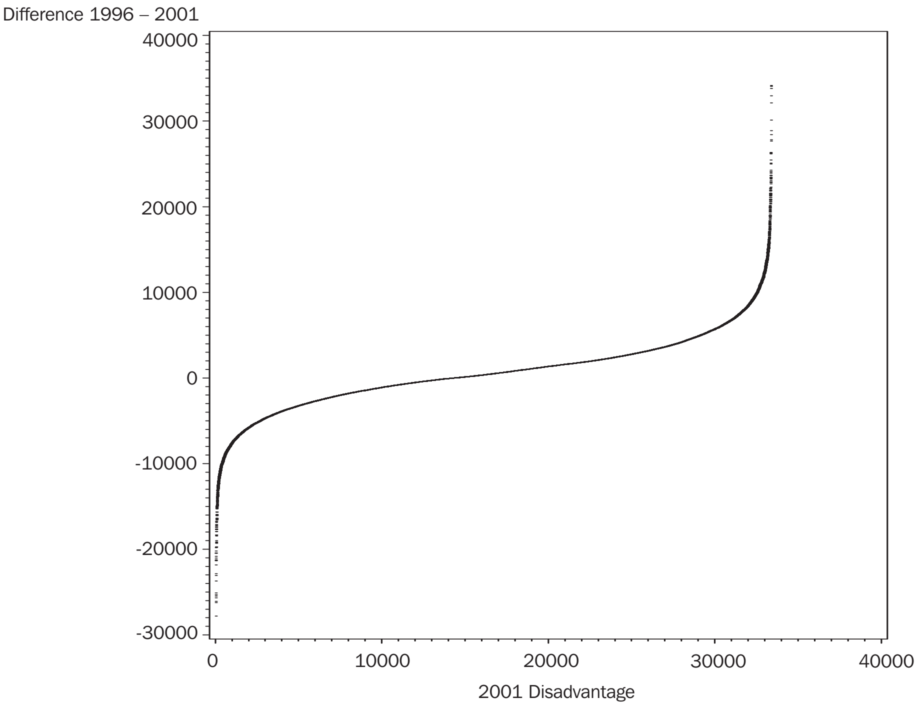
The second graph shows how far the CD rankings have changed. The vertical axis on this graph is the change from 1996 to 2001; and the horizontal axis is a count for the CD. For a small change, we would expect a horizontal line around zero on the vertical axis. A large change in SEIFA will appear as many CDs in the 'tails' of the graph.

It can be seen that there are some CDs with large changes, which we expected. The majority of CDs experienced little change.

4.15 RANKINGS FOR SEIFA 1996 TO SEIFA 2001, DISADVANTAGE



4.16 DIFFERENCE IN RANKING FROM SEIFA 1996 TO SEIFA 2001, DISADVANTAGE



Index of Disadvantage continued

Average values for all CDs in each state in 1996 and 2001, along with the state rankings, are shown in table 4.17. The state values are a simple average of the CD level values in each state. Rankings can be misleading because a state can be ranked over another state because of a difference at the first or second decimal place.

There is not much change in rank from 1996 to 2001. The only states that have changed rank are Qld and SA. The NT is the most disadvantaged state according to this index.

4.17 CHANGE IN STATE RANKINGS, DISADVANTAGE CD INDEXES

State	SEIFA 1996 Index	SEIFA 1996 Rank	SEIFA 2001 Index	SEIFA 2001 Rank
NSW	1 006	6	1 000	6
Vic.	1 015	7	1 016	7
Qld	983	4	989	3
SA	982	3	994	4
WA	1 000	5	996	5
Tas.	974	2	969	2
NT	922	1	903	1
ACT	1 086	8	1 076	8

The next step was to look at summary indexes by SLA and POA. The summary indexes were calculated using a population weighted average; and the summary to state was done using a simple average, as was done with the CD level indexes.

The indexes are shown in tables 4.18 and 4.19.

The summary by postal area and SLA still show the NT as the most disadvantaged state.

4.18 SUMMARY OF THE DISADVANTAGE POA INDEX

State	Average	10th percentile	25th percentile	50th percentile	75th percentile	90th percentile
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
OT	832	672	672	896	928	928
Aust.	998	923	962	1 000	1 044	1 085

Index of Disadvantage continued

4.19 SUMMARY OF THE DISADVANTAGE SLA INDEX

State	Average	10th percentile	25th percentile	50th percentile	75th percentile	90th percentile
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
OT	832	672	672	896	928	928
Aust.	999	926	963	1 000	1 046	1 087

Index of economic resources

This index only includes variables measuring economic disadvantage. This index used the 2001 variable selection strategy to identify relevant variables to include. There are no level 3 variables in this index.

INITIAL LIST OF VARIABLES

The list of variables used for this index are all associated with economic disadvantage. We included all income, expenditure and wealth variables. The full list is shown in table 4.20.

4.20 INITIAL LIST OF VARIABLES USED

Variable	Specification
AVEBED	Average number of bedrooms per person
CFDOINCHI	% Couple families with dependent child(ren) only with annual income greater than $77,999
CFDOINCLO	% Couple families with dependent child(ren) only with annual income less than $36,400
CFD_ NDINCHI	% Couple families with dependents and non-dependents or with non-dependents only with annual income greater than $103,999
CFD_NDINCLO	% Couple families with dependents and non-dependents or with non-dependents only with annual income less than $52,000
CFNKINCHI	% Couple families with no children with annual income greater than $77,999
CFNKINCLO	% Couple families with no children with annual income less than $20,800
FEWBED	% Dwellings with one or no bedrooms
GROUP	% Households who are group households
HIGHBED	% Dwellings with four or more bedrooms
HIGHCAR	% Dwellings with three or more motor cars
IMPDWEL	% Households living in improvised dwellings
MORTHI_01	% Households paying mortgage greater than $1,360 per month
NOCAR	% Dwellings with no motor car at dwelling
OWNING	% Households owning dwelling
PURCHAS	% Households purchasing dwelling
RENT	% Rental dwellings
RENTHI_01	% Households paying rent greater than $225 per week
RENTLO_01	% Households paying rent less than $88 per week
SPDOINCHI	% Single parent families with dependent child(ren) only with annual income greater than $36,399
SPDOINCLO	% Single parent families with dependent child(ren) only with annual income less than $15,600
SPD_NDINCHI	% Single parent families with dependents and non-dependents or with non-dependents with annual income greater than $62,399
SPD_ NDINCLO	% Single parent families with dependents and non-dependents or with non-dependents with annual income less than $26,000
SPINCHI	% Single person households with income greater than $36,399
SPINCLO	% Single person household with income less than $15,600

REMOVAL OF HIGHLY CORRELATED VARIABLES

There were no highly correlated variables among the set of initial variables.

Index of economic resources continued

ROTATION

Table 4.21 shows the eigenvalues for each component from a PCA run using all variables.

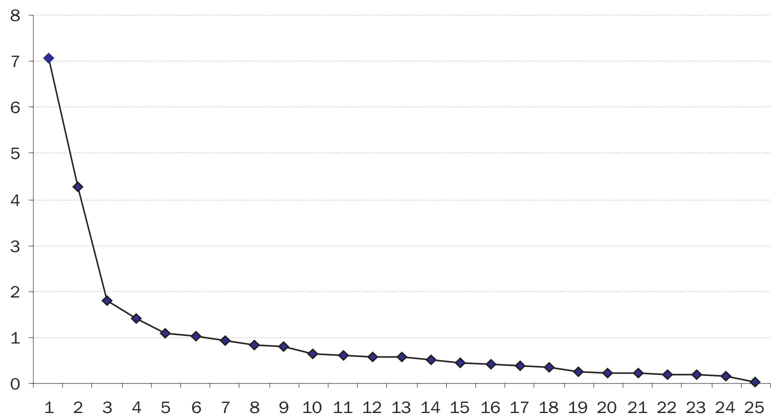
It can be seen that the first component explains 28.3 per cent of the variance between the variables.

The scree plot shown in graph 4.22 suggests retaining two components.

4.21 EIGENVALUES AND VARIANCE EXPLAINED FROM INITIAL PCA

Component	*Eigenvalue*	*Variance explained*	*Cumulative*
1	7.0755	0.2830	0.2830
2	4.2575	0.1703	0.4533
3	1.8014	0.0721	0.5254
4	1.4233	0.0569	0.5823
5	1.0822	0.0433	0.6256
6	1.0174	0.0407	0.6663
7	0.9298	0.0372	0.7035
8	0.8434	0.0337	0.7372
9	0.8059	0.0322	0.7695
10	0.6422	0.0257	0.7951
11	0.6132	0.0245	0.8197
12	0.5774	0.0231	0.8428
13	0.5645	0.0226	0.8653
14	0.5162	0.0206	0.8860
15	0.4424	0.0177	0.9037
16	0.4053	0.0162	0.9199
17	0.3797	0.0152	0.9351
18	0.3514	0.0141	0.9491
19	0.2460	0.0098	0.9590
20	0.2365	0.0095	0.9684
21	0.2119	0.0085	0.9769
22	0.1941	0.0078	0.9847
23	0.1806	0.0072	0.9919
24	0.1640	0.0066	0.9985
25	0.0384	0.0015	1.0000

4.22 SCREE PLOT FOR INDEX OF ECONOMIC RESOURCES



The results from rotating the matrix using an orthogonal (varimax) rotation and retaining two components is shown in table 4.23.

Index of economic resources continued

4.23 RESULTS AFTER ROTATION WITH TWO COMPONENTS RETAINED

	Component 1	*Component 2*
Eigenvalues	6.9412	4.3917
Variance explained	0.2776	0.1757

The aim of the rotation was to check whether rotation makes interpretation of the index easier, or gives any interpretable components after the first component. To check this, we need to look at the loadings of the variables on each component. This is shown in table 4.24.

4.24 LOADINGS WITH ROTATION WITH TWO COMPONENTS RETAINED

Variable	*Component 1*	*Component 2*
AVEBED	–0.0053	–0.3766
CFDOINCHI	0.8399	–0.0105
CFDOINCLO	–0.7314	0.1955
CFD_NDINCHI	0.6683	–0.1197
CFD_NDINCLO	–0.5746	0.1432
CFNKINCHI	0.8513	0.1362
CFNKINCLO	–0.7241	0.0974
FEWBED	0.0110	0.6389
GROUP	0.2807	0.6374
HIGHBED	0.2858	–0.7239
HIGHCAR	0.0131	–0.7347
IMPDWEL	–0.1351	0.1503
MORTHI_01	0.7367	0.1000
NOCAR	–0.1423	0.8557
OWNING	–0.0659	–0.6249
PURCHAS	0.2251	–0.5221
RENT	–0.0752	0.8793
RENTHI_01	0.7668	0.1084
RENTLO_01	–0.5247	–0.0586
SPDOINCHI	0.6337	–0.0136
SPDOINCLO	–0.2800	0.1603
SPD_NDINCHI	0.5169	–0.0604
SPD_NDINCLO	–0.3939	0.1062
SPINCHI	0.7790	0.0029
SPINCLO	–0.7070	0.1155

This table shows that Income loads highly on the first component; but so does mortgage and high rent payments. Some income variables were not loading on this component (e.g. SPD_NDINCLO); and some wealth measures were (MORTHI, RENTHI and RENTLO). The second component had mainly wealth variables loading on it.

The rotated weights do not appear to produce any coherent or distinct patterns. There is more of a pattern in this index than in the index of disadvantage and advantage/disadvantage; but the pattern is still unclear because of the loadings of some variables.

Index of economic resources continued

In addition, we were looking for a general measure of economic advantage/disadvantage; a rotated matrix will load more stongly onto particular aspects of advantage/disadvantage. We therefore decided to use the first unrotated component to produce the final index.

FINDING VARIABLES POORLY CORRELATED WITH THE FIRST UNROTATED COMPONENT

Variables with a correlation coefficient (loading) with the first unrotated component in the range –0.3 to 0.3 were dropped from the final index calculation because their contribution to the index was negligible. This was an iterative stage — removing some of the low loading variables meant weights and loadings for other variables changed. This step continued until there were no variables with a loading in the range –0.3 to 0.3.

The initial and final loadings and weights are shown in table 4.25. These weights and loadings were calculated after specifying that we only wanted one component retained.

4.25 LOADINGS AND WEIGHTS, INDEX OF ECONOMIC RESOURCES

Variable	Initial Loading	Initial Weight	Final Loading	Final Weight
AVEBED	0.0770	0.0290	na	na
CFDOINCHI	0.8220	0.3090	0.8480	0.3264
CFDOINCLO	–0.7560	–0.2840	–0.7379	–0.2840
CFD_NDINCHI	0.6780	0.2550	0.6875	0.2646
CFD_NDINCLO	–0.5920	–0.2230	–0.5855	–0.2254
CFNKINCHI	0.8010	0.3010	0.8406	0.3235
CFNKINCLO	–0.7280	–0.2740	–0.7208	–0.2774
FEWBED	–0.1290	–0.0480	na	na
GROUP	0.1350	0.0510	na	na
HIGHBED	0.4370	0.1640	0.3483	0.1341
HIGHCAR	0.1730	0.0650	na	na
IMPDWEL	–0.1650	–0.0620	na	na
MORTHI_01	0.6970	0.2620	0.7429	0.2859
NOCAR	–0.3260	–0.1220	na	na
OWNING	0.0720	0.0270	na	na
PURCHAS	0.3340	0.1250	na	na
RENT	–0.2650	–0.1000	na	na
RENTHI_01	0.7250	0.2720	0.7668	0.2951
RENTLO_01	–0.4990	–0.1880	–0.4950	–0.1905
SPDOINCHI	0.6210	0.2340	0.6285	0.2419
SPDOINCLO	–0.3080	–0.1160	na	na
SPD_NDINCHI	0.5180	0.1950	0.5262	0.2025
SPD_NDINCLO	–0.4080	–0.1530	–0.4036	–0.1554
SPINCHI	0.7600	0.2860	0.7765	0.2988
SPINCLO	–0.7150	–0.2690	–0.7080	–0.2725

The final set of variables included variables relating to incomes, rent, mortgage payments and wealth. All weights were in the expected direction.

The first component had an eigenvalue of 6.75 and explained 45.0 per cent of the total variance.

The final list of variables, with definitions and their weights in 1996 and 2001, is shown in table 4.26. An 'na' indicates that the variables either aren't comparable, or weren't used, in that year.

4.26 1996 AND 2001 WEIGHTS

Variable	*1996*	*2001*
% Couple families with dependent child(ren) only with annual income greater than $77,999	na	0.33
% Couple families with no children with annual income greater than $77,999	na	0.32
% Households paying rent greater than $225 per week	0.21	0.30
% Single person households with income greater than $36,399	na	0.30
% Households paying mortgage greater than $1,360 per month	0.21	0.29
% Couple families with dependents and non-dependents or with non-dependents only with annual income greater than $103,999	na	0.27
% Single parent families with dependent child(ren) only with annual income greater than $36,399	na	0.24
% Single parent families with dependents and non-dependents or with non-dependents with annual income greater than $62,399	na	0.20
% Dwellings with four or more bedrooms	0.27	0.13
% Single parent families with dependents and non-dependents or with non-dependents with annual income less than $26,000	na	–0.16
% Households paying rent less than $88 per week	–0.17	–0.19
% Couple families with dependents and non-dependents or with non-dependents only with annual income less than $52,000	na	–0.23
% Single person household with income less than $15,600	na	–0.27
% Couple families with no children with annual income less than $20,800	na	–0.28
% Couple families with dependent child(ren) only with annual income less than $36,400	na	–0.28
% Households owning or purchasing dwelling	0.28	na
% Families with family structure other than two parent or single parent with dependent offspring or consisting of a couple only, and income greater than $77,999	0.27	na
% Families consisting of a two parent family with dependent offspring and income greater than $77,999	0.26	na
% Families consisting of a couple only with income greater than $62,399	0.24	na
% Families consisting of a single parent with dependent offspring with income greater than $31,199	0.22	na
% Households purchasing dwelling	0.20	na
% Dwellings with three or more motor cars	0.19	na
% Households owning dwelling	0.16	na
Average number of Bedrooms per person	0.11	na
% Families with family structure other than two parent or single parent with dependent offspring or consisting of a couple only, and income less than $26,000	–0.23	na
% Households in improvised dwellings	–0.09	na
% Households renting (non-government authority)	–0.15	na
% Dwellings with one or no bedrooms	–0.16	na
% Families consisting of a single parent with dependent offspring with income less than $26,000	–0.18	na
% Households renting (government authority)	–0.20	na
% Families consisting of a couple only with income less than $15,600	–0.23	na
% Dwellings with no motor cars	–0.26	na
% Families consisting of a two parent family with dependent offspring and income less than $26,000	–0.27	na

Index of economic resources continued

The variables in the 2001 index are quite different to the variables in the 1996 index. Many of the variables were dropped because they had a low weight in 2001. This is partly due to the increased loading cut-off in 2001 — four of the ten dropped variables were excluded because of the increased cut-off (RENT, PURCHAS, NOCAR and SPDOINCLO). The income variables are also different in 2001. In fact, there are only four variables common to 1996 and 2001. For all these variables, the weights changed (and for some, they more than halved).

RESULTS OF INDEX SCORE

Table 4.27 presents summary results of the CD Index of Economic Resources index scores. These are simple averages of the CD level scores for each state.

Index of economic resources continued

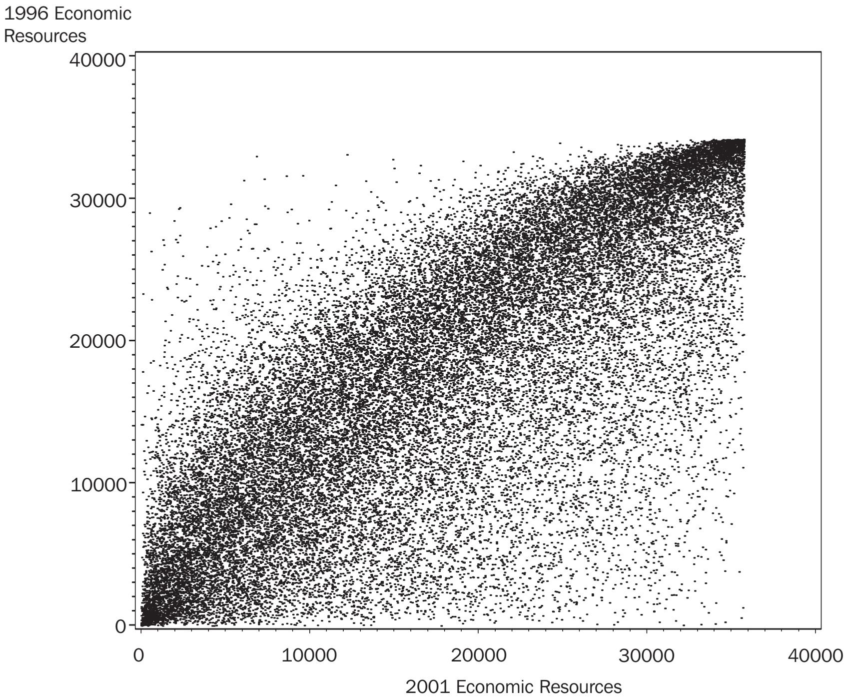
4.27 SUMMARY OF THE ECONOMIC RESOURCE CD INDEX

State	Average	10th percentile	25th percentile	50th percentile	75th percentile	90th percentile
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
OT	979	867	898	983	1 064	1 100
Aust.	1 000	883	925	986	1 065	1 142

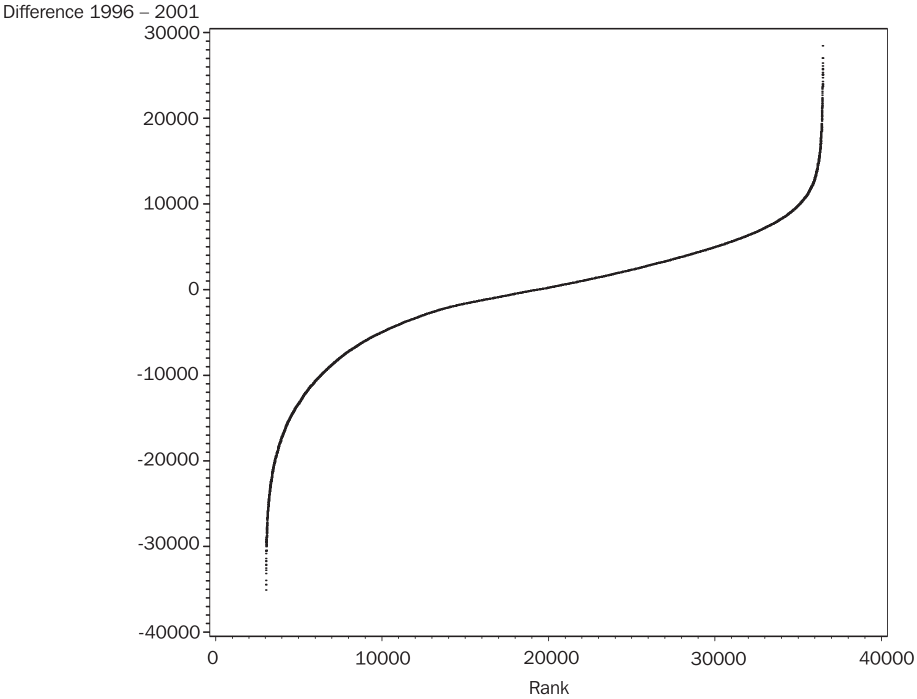
We looked at how CDs had changed ranking as a result of the 2001 method and data. Graphs 4.28 and 4.29 show the change from the 1996 index to the 2001 index. These graphs plot the rankings of CDs before and after a change. On the vertical axis of the graph is the old (reference) rank; and the horizontal axis shows the new rank.If there are only small changes for most CDs, the first graph will appear as a diagonal line.

The second graph shows how far the CD rankings have changed. The vertical axis on this graph is the change from 1996 to 2001; and the horizontal axis is a count for the CD. A small change would give a horizontal line around zero on the vertical axis. A large change will show as many CDs in the 'tails' of the graph.

4.28 RANKINGS FOR SEIFA 1996 TO SEIFA 2001, INDEX OF ECONOMIC RESOURCES



4.29 DIFFERENCE IN RANKING FROM SEIFA 1996 TO SEIFA 2001, INDEX OF ECONOMIC RESOURCES



Index of economic resources continued

These graphs show that the index changed considerably from 1996 to 2001.

State rankings for 1996 and 2001, along with the indexes, are shown in table 4.30. Rankings based on state averages suffer not only from the aggregation, but also from problems associated with ranking. In some cases, a state will be ranked over another state because of a difference at the first or second decimal place.

There are some large changes in rank from 1996 to 2001. The NT, in particular, has gone from being the most disadvantaged state to the fifth most disadvantaged. All states apart from the ACT changed rank.

In this index, Tasmania is the most economically disadvantaged state.

4.30 CHANGE IN STATE RANKINGS, ECONOMIC RESOURCE CD INDEXES

State	SEIFA 1996 Index	SEIFA 1996 Rank	SEIFA 2001 Index	SEIFA 2001 Rank
NSW	1 009	6	1 021	7
Vic.	1 008	5	1 006	6
Qld	984	4	980	3
SA	977	3	963	2
WA	1 011	7	997	4
Tas.	966	2	928	1
NT	914	1	1 002	5
ACT	1 075	8	1 107	8

The next step was to look at summary indexes by SLA and POA. The summary indexes were calculated using a population weighted average; and the summary to state was done using a simple average.

Index of economic resources continued

The summary by Postal Area is shown in table 4.31, and by SLA in table 4.32.

4.31 SUMMARY OF THE ECONOMIC RESOURCE POA INDEX

State	Average	10th percentile	25th percentile	50th percentile	75th percentile	90th percentile
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
OT	990	956	956	1 006	1 010	1 010
Aust.	973	880	910	954	1 023	1 101

4.32 SUMMARY OF THE ECONOMIC RESOURCE SLA INDEX

State	Average	10th percentile	25th percentile	50th percentile	75th percentile	90th percentile
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	1 140	1 195
OT	990	956	956	1 006	1 010	1 010
Aust.	990	893	924	971	1 047	1 111

When comparing the state average of indexes above the CD level, the NTs ranking increases significantly. When aggregated to PA and SLA, the NT is the second most advantaged state. This is due to the population weighted average and the correlation between a CDs population and advantage in the NT (more populous CDs tend to be more advantaged in the NT).

Index of Education and Occupation

This index only includes variables measuring the educational and occupational aspects of advantage/disadvantage. There are no level 3 variables in this index.

INITIAL LIST OF VARIABLES

The variables used for this index are associated with advantage and disadvantage in Education and Occupation. We have included all occupation variables; and a richer set of education variables than we used in the Index of Advantage/Disadvantage.

The list of variables used is shown in table 4.33.

Index of Education and Occupation continued

4.33 INITIAL LIST OF VARIABLES USED

Variable	*Specification*
ATSCH	% Persons aged 15 years and over who are still at school
CAEUNI	% persons aged 15 years and over at University or other tertiary institution
CERT	% Persons aged 15 years and over with a certificate qualification
DEGREE_01	% Persons aged 15 years and over with degree or higher
DIP	% Persons aged 15 years and over having an advanced diploma or diploma qualification
FADVCLRK	% Employed Females classified as 'Advanced Clerical & Service Workers'
FELECLSW	% Employed Females classified as 'Elementary Clerical, Sales & Service Workers'
FINTCLRK	% Employed Females classified as 'Intermediate Clerical, Sales & Service Workers'
FINTPRTW	% Employed Females classified as 'Intermediate Production & Transport Workers'
FLABOUR	% Employed Females classified as 'Labourers & Related Workers'
FMGR_ADM	% Employed Females classified as 'Managers or Administrators'
FPARA_PF	% Employed Females classified as 'Associate Professionals'
FPROF	% Employed Females classified as 'Professionals'
FTRADE	% Employed Females classified as 'Tradespersons'
F_UNEMP	% Females (in Labour Force) unemployed
MADVCLRK	% Employed Males classified as 'Advanced Clerical & Service Workers'
MELECLSW	% Employed Males classified as 'Elementary Clerical, Sales & Service Workers'
MINTCLRK	% Employed Males classified as 'Intermediate Clerical, Sales & Service Workers'
MINTPRTW	% Employed Males classified as 'Intermediate Production & Transport Workers'
MLABOUR	% Employed Males classified as 'Labourers & Related Workers'
MMGR_ADM	% Employed Males classified as 'Managers or Administrators'
MPARA_PF	% Employed Males classified as 'Associate Professionals'
MPROF	% Employed Males classified as 'Professionals'
MTRADE	% Employed Males classified as 'Tradespersons'
M_UNEMP	% Males (in Labour Force) unemployed
NOQUAL	% Persons aged 15 years and over with no qualifications
NOSCH	% Persons aged 15 years and over who did not go to school
TAFE	% Persons aged 15 years and over at TAFE
YR11LO	% Persons aged 15 years and over who left school at Year 11 or lower

FIND HIGHLY CORRELATED VARIABLES AMONG THE INITIAL LIST OF VARIABLES

After examination of the correlation matrix of the initial variables, we found three sets of highly correlated variables; NOQUAL and YR11LO (r=0.83), FPROF and DEGREE_01 (r=0.82), and MPROF and DEGREE_01 (r=0.82). All the variables were left in the index because they appeared to be measuring different aspects of educational attainment; and measured socioeconomic status, which we wanted left in the model.

ROTATION

Table 4.34 shows the eigenvalues for each component from a PCA run using all variables.

The scree plot shown in graph 4.35 suggests retaining three components.

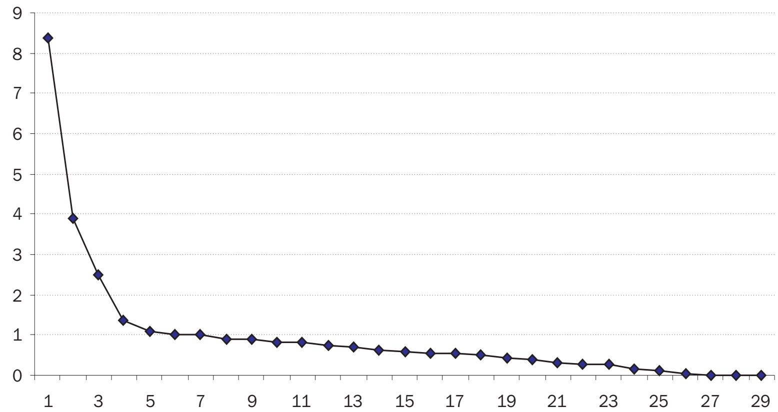
Index of Education and Occupation continued

4.34 EIGENVALUES AND VARIANCE EXPLAINED FROM INITIAL PCA

Components	*Eigenvalue*	*Variance explained*	*Cumulative*
1	8.3939	0.2894	0.2894
2	3.9122	0.1349	0.4243
3	2.4804	0.0855	0.5099
4	1.3532	0.0467	0.5565
5	1.0963	0.0378	0.5943
6	1.0260	0.0354	0.6297
7	1.0073	0.0347	0.6645
8	0.8939	0.0308	0.6953
9	0.8825	0.0304	0.7257
10	0.8343	0.0288	0.7545
11	0.8091	0.0279	0.7824
12	0.7549	0.0260	0.8084
13	0.7016	0.0242	0.8326
14	0.6314	0.0218	0.8544
15	0.5935	0.0205	0.8748
16	0.5507	0.0190	0.8938
17	0.5399	0.0186	0.9124
18	0.4941	0.0170	0.9295
19	0.4195	0.0145	0.9440
20	0.3911	0.0135	0.9574
21	0.3228	0.0111	0.9686
22	0.2760	0.0095	0.9781
23	0.2733	0.0094	0.9875
24	0.1708	0.0059	0.9934
25	0.1284	0.0044	0.9978
26	0.0583	0.0020	0.9998
27	0.0033	0.0001	1.0000
28	0.0014	—	1.0000
29	—	—	1.0000

— nil or rounded to zero (including null cells)

4.35 SCREE PLOT FOR INDEX OF EDUCATION AND OCCUPATION



The results from rotating the matrix using three components are shown in table 4.36. The method of rotation used was an orthoganal (Varimax) rotation.

Index of Education and Occupation continued

4.36 RESULTS AFTER ROTATION WITH THREE COMPONENTS RETAINED

	Component 1	Component 2	Component 3
Eigenvalues	6.8731	4.0733	3.8400
Variance explained	0.2370	0.1405	0.1324

The aim of the rotation was to check whether rotation makes interpretation of the index easier, or gives any interpretable components after the first component. To check this, we need to look at the loadings of the variables on each component. This is shown in table 4.37.

4.37 LOADINGS WITH ROTATION WITH THREE COMPONENTS RETAINED

Variable	Component 1	Component 2	Component 3
ATSCH	–0.0678	–0.0963	–0.0096
CAEUNI	0.7363	–0.1524	0.0415
CERT	–0.5949	–0.2103	–0.5762
DEGREE_01	0.9161	0.1134	–0.2320
DIP	0.6195	0.1060	–0.4686
FADVCLRK	0.0542	–0.0084	–0.4975
FELECLSW	–0.2833	–0.5824	0.1166
FINTCLRK	–0.2179	–0.6589	–0.0860
FINTPRTW	–0.3274	–0.2221	0.3897
FLABOUR	–0.4769	0.0319	0.6051
FMGR_ADM	–0.0189	0.8865	0.0291
FPARA_PF	0.1179	–0.1580	–0.3226
FPROF	0.7977	0.1028	–0.2387
FTRADE	–0.3080	–0.0171	0.0306
F_UNEMP	–0.1626	–0.3727	0.6109
MADVCLRK	0.3102	–0.1450	–0.1698
MELECLSW	0.1574	–0.5715	0.1381
MINTCLRK	0.3621	–0.5328	–0.1067
MINTPRTW	–0.6065	–0.3346	0.3435
MLABOUR	–0.4443	0.1067	0.6192
MMGR_ADM	0.0067	0.9044	–0.0892
MPARA_PF	0.3961	–0.3028	–0.4417
MPROF	0.8856	–0.0308	–0.2293
MTRADE	–0.6569	–0.4904	–0.1849
M_UNEMP	–0.1525	–0.3755	0.5770
NOQUAL	–0.7087	–0.0225	0.5909
NOSCH	0.0369	–0.1135	0.5572
TAFE	0.0481	–0.3572	0.0529
YR11LO	–0.8905	0.0049	0.2537

The first component has education loading strongly; and Professionals (which we found was highly correlated with Education). The second component has some occupations loading strongly (but not all). The third component has FLABOUR, MLABOUR and FUNEMP loading strongly. There does not appear to be any pattern in the loadings.

There does not appear to be any coherent or distinct patterns emerging in the components. In addition, we were looking for a general measure for disadvantage; a rotated matrix will load more stongly onto particular aspects of disadvantage. We therefore decided to use the first unrotated component to produce the final index.

Index of Education and Occupation continued

REMOVAL OF VARIABLES POORLY CORRELATED WITH THE FIRST UNROTATED COMPONENT

Variables with a loading in the range –0.3 to 0.3 were dropped from the final index calculation because their contribution to the index is negligible. This was an iterative procedure, with three iterations conducted. The loadings and weights from the inital and final (first and third) runs are shown in table 4.38. These weights and loadings were calculated after specifying that we only wanted one component retained.

4.38 LOADINGS AND WEIGHTS, INDEX OF EDUCATION AND OCCUPATION

Variable	Initial Loading	Initial Weight	Final Loading	Final Weight
ATSCH	0.0720	0.0250	na	na
CAEUNI	–0.5830	–0.2010	–0.5955	–0.2121
CERT	0.2750	0.0950	na	na
DEGREE_01	–0.9180	–0.3170	–0.9217	–0.3282
DIP	–0.7750	–0.2680	–0.7807	–0.2780
FADVCLRK	–0.2830	–0.0980	na	na
FELECLSW	0.4090	0.1410	0.3864	0.1376
FINTCLRK	0.2710	0.0930	na	na
FINTPRTW	0.5090	0.1760	0.4994	0.1778
FLABOUR	0.6920	0.2390	0.6993	0.2490
FMGR_ADM	–0.1380	–0.0480	na	na
FPARA_PF	–0.2250	–0.0780	na	na
FPROF	–0.8180	–0.2820	–0.8264	–0.2943
FTRADE	0.2820	0.0970	na	na
F_UNEMP	0.5020	0.1730	0.4939	0.1759
MADVCLRK	–0.3200	–0.1100	–0.3293	–0.1173
MELECLSW	0.0390	0.0140	na	na
MINTCLRK	–0.2600	–0.0900	na	na
MINTPRTW	0.7480	0.2580	0.7335	0.2612
MLABOUR	0.6570	0.2270	0.6659	0.2371
MMGR_ADM	–0.2200	–0.0760	na	na
MPARA_PF	–0.4930	–0.1700	–0.5116	–0.1822
MPROF	–0.8630	–0.2980	–0.8757	–0.3118
MTRADE	0.5680	0.1960	0.5352	0.1906
M_UNEMP	0.4780	0.1650	0.4708	0.1677
NOQUAL	0.8940	0.3090	0.9109	0.3244
NOSCH	0.2560	0.0880	na	na
TAFE	0.0520	0.0180	na	na
YR11LO	0.8840	0.3050	0.8954	0.3188

The Principal Components Analysis procedure assigned negative weights to variables representing advantage; and positive weights to variables representing disadvantage. We reversed these when calculating the final index values.

All forms of higher education were measures of advantage, which we expected. Professionals, associate professionals and advanced clerical workers were all advantaged occupations.

Not completing Year 12 and having no qualification was strongly associated with disadvantage; as was unemployment and the occupations 'Elementary clerical, sales and service workers', 'Intermediate Production and Transport Workers', 'Labourers' and Male 'Trade' workers. All these results were in line with expectations.

The eigenvalue for this component was 7.89, and the variance explained was 46.4 per cent of total variance.

Index of Education and Occupation continued

The final list of variables, with definitions and the weights for 1996 and 2001, is shown in table 4.39. An 'na' indicates that the variable was specified differently in 1996 (e.g. the Year left School variable); was dropped in 2001 because of a low weight; or was new to the 2001 census.

4.39 1996 AND 2001 WEIGHTS

Variable	*1996*	*2001*
% Persons aged 15 years and over with degree or higher	na	0.33
% Employed Males classified as 'Professionals'	0.34	0.31
% Employed Females classified as 'Professionals'	0.31	0.29
% Persons aged 15 years and over having an advanced diploma or diploma qualification	na	0.28
% Persons aged 15 years and over at University or other tertiary institution	0.24	0.21
% Employed Males classified as 'Associate Professionals'	0.19	0.18
% Employed Males classified as 'Advanced Clerical & Service Workers'	0.14	0.12
% Employed Females classified as 'Elementary Clerical, Sales & Service Workers'	–0.12	–0.14
% Males (in Labour Force) unemployed	–0.21	–0.17
% Females (in Labour Force) unemployed	–0.22	–0.18
% Employed Females classified as 'Intermediate Production & Transport Workers'	–0.20	–0.18
% Employed Males classified as 'Tradespersons'	–0.19	–0.19
% Employed Males classified as 'Labourers & Related Workers'	–0.26	–0.24
% Employed Females classified as 'Labourers & Related Workers'	–0.29	–0.25
% Employed Males classified as 'Intermediate Production & Transport Workers'	–0.29	–0.26
% Persons aged 15 years and over who left school at Year 11 or lower	na	–0.32
% Persons aged 15 years and over with no qualifications	–0.35	–0.32
% Employed Females classified as 'Advanced Clerical & service Workers'	0.14	na
% Employed Males classified as 'Intermediate Clerical, Sales & Service Workers'	0.11	na
% Employed Females classified as 'Tradespersons'	–0.12	na
% Persons aged 15 years and over who left school at or under 15 years	–0.31	na

The weights in 2001 are similar to those in 1996. There were no variables changing sign; and the largest change was about 0.04.

RESULTS OF INDEX SCORE

Table 4.40 presents summary results of the CD Index of Education and Occupation index scores. These are a simple average of the CD level scores in a state.

4.40 SUMMARY OF THE EDUCATION AND OCCUPAITON CD INDEX

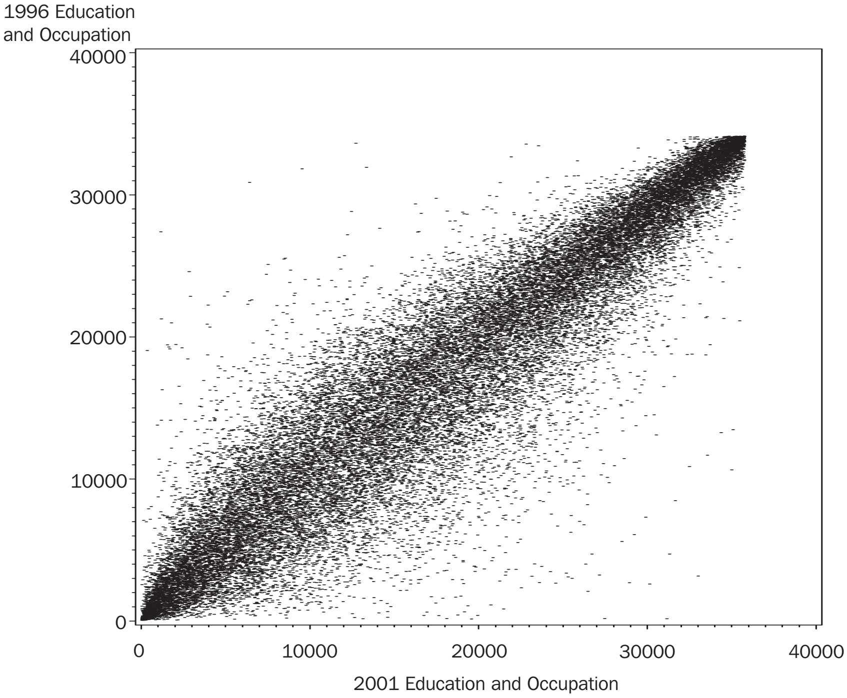
State	*Average*	*10th percentile*	*25th percentile*	*50th percentile*	*75th percentile*	*90th percentile*
NSW	1 009	889	936	996	1 078	1 162
Vic.	1 012	889	940	999	1 084	1 161
Qld	980	881	919	967	1 034	1 104
SA	978	859	910	969	1 037	1 122
WA	998	880	928	986	1 064	1 136
Tas.	959	840	888	945	1 026	1 102
NT	980	838	928	998	1 048	1 094
ACT	1 116	1 031	1 071	1 113	1 161	1 213
OT	959	792	812	993	1 104	1 141
Aust.	1 000	882	928	987	1 067	1 146

Index of Education and Occupation continued

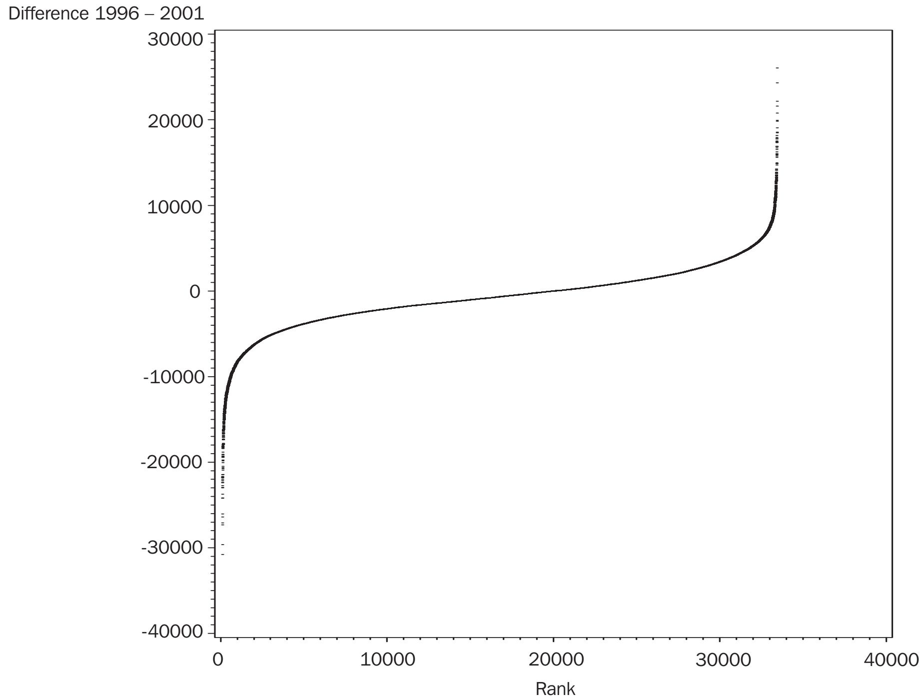
We looked at how CDs had changed ranking as a result of the 2001 method and data. Graphs 4.41 and 4.42 show the new 2001 index compared to the 1996 index. These graphs plot the rankings of CDs before and after a change. In the first graph, the vertical axis is the old (reference) rank; and the horizontal axis shows the new rank. A small change will show as a diagonal line in the first graph. The larger the change, the more disperse the points will appear around the diagonal line.

The second graph shows how far the CD rankings have changed. For a small change, we would expect a horizontal line around zero on the vertical axis. A large change will show as many CDs in the 'tails' of the graph.

4.41 RANKINGS FOR SEIFA 1996 TO SEIFA 2001, INDEX OF EDUCATION AND OCCUPATION



4.42 DIFFERENCE IN RANKING FROM SEIFA 1996 TO SEIFA 2001, INDEX OF EDUCATION AND OCCUPATION



Index of Education and Occupation continued

There are some CDs with large changes, which we expected. However, the majority of CDs in graph 4.42 lie around the zero axis, so experienced little change.

State rankings for 1996 and 2001, along with the indexes, are shown in table 4.43.

There was not a great change in rank from 1996 to 2001.

4.43 CHANGE IN STATE RANKINGS, EDUCATION AND OCCUPATION CD INDEXES

State	SEIFA 1996 Index	SEIFA 1996 Rank	SEIFA 2001 Index	SEIFA 2001 Rank
NSW	1 011	6	1 009	6
Vic.	1 016	7	1 012	7
Qld	971	2	980	3
SA	982	4	978	2
WA	990	5	998	5
Tas.	965	1	959	1
NT	979	3	980	4
ACT	1 123	8	1 116	8

In 1996 and 2001, Tas. was the most disadvantaged state for the Index of Education and Occupation. The NT is slightly better off in 2001, but SA has slipped two places to be second most disadvantaged in 2001. Qld is slightly better off in 2001; and all other states rank the same.

We also looked at summary indexes by SLA and POA. The summary indexes were calculated using a population weighted average; and the summary to state was done using a simple average.

Index of Education and Occupation continued

The indexes are shown in tables 4.44 and 4.45.

4.44 SUMMARY OF THE EDUCATION AND OCCUPATION POA INDEX

State	Average	10th percentile	25th percentile	50th percentile	75th percentile	90th percentile
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	1 115	1 049	1 074	1 122	1 155	1 178
OT	971	943	943	948	1 022	1 022
Aust.	979	896	925	959	1 017	1 104

4.45 SUMMARY OF THE EDUCATION AND OCCUPATION SLA INDEX

State	Average	10th percentile	25th percentile	50th percentile	75th percentile	90th percentile
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
OT	971	943	943	948	1 022	1 022
Aust.	993	906	932	965	1 046	1 129

The aggregation to SLA has put the NT at second most advantaged, compared to 4th most advantaged when looking at the CD level indexes averaged to state; and 4th most advantaged when looking at PA indexes averaged to state. This is due to the population weighted average and the correlation between a CDs population and advantage in the NT (more populous CDs tend to be more advantaged in the NT).

CHAPTER 5 VALIDATION OF THE INDEXES

VALIDATION OF THE INDEXES

This section summarises the validation done on the 2001 SEIFA indexes. Nine methods of validation were used:

- Check the CDs with extreme values, and explain why the CD has that value.
- Look at the relationship between the indexes.
- Confirm whether the indexes are consistent with local knowledge of areas of advantage/disadvantage.
- Use SEIFA 2001 data to replicate earlier analysis done using the 1996 SEIFA.
- Analyse the indexes against other ABS data.
- Compare the 1996 and 2001 indexes and identify the drivers of the change.
- Use an external reference group of experts to peer review the method, data and indexes in SEIFA 2001.
- Look at colour maps of the indexes to check whether geographic patterns of the indexes line up with local knowledge.

Examination of extreme values

We identified the top and bottom ranked five CDs in each index, and found the variables contributing to the extreme value.

We calculated the contribution of each variable to each CDs raw score, to identify the variables that contributed the most to ensure their values seemed sensible. For the Advantage/Disadvantage index, in all cases of the extreme values, the disadvantaged CDs had very high proportions of variables associated with disadvantage; and very low proportions of variables associated with advantage (compared to the Australian mean). Some CDs had 100% of the people in the CD earning low incomes, and 0% experiencing high incomes. This is shown in table 5.1.

Conversely, all the top five most advantaged CDs had a very low proportion of people experiencing disadvantage and a high proportion of people experiencing advantage (compared to the Australian mean). This is shown in table 5.2.

These same tables were done for each index, although only the Index of Advantage/Disadvantage have been reproduced here. The tables showed similar patterns for each index.

Examination of extreme values continued

5.1 VARIABLE VALUES FOR BOTTOM FIVE CDs

	CD NUMBER					
Name	*7013603*	*5010701*	*7020206*	*7020404*	*5020507*	*Aust. mean*
INTNET	—	—	—	—	—	0.182
HIGHBED	—	—	0.174	0.286	—	0.257
NOQUAL	0.974	0.957	0.983	0.972	0.986	0.572
NOSCH	0.079	0.152	0.094	0.119	0.186	0.010
MPROF	—	—	—	—	—	0.154
FPROF	—	—	—	—	—	0.206
MPARA_PF	—	—	—	0.200	—	0.118
FPARA_PF	—	—	0.167	—	—	0.116
MTRADE	—	—	—	—	—	0.199
FTRADE	—	—	—	—	—	0.032
MADVCLRK	—	—	—	—	—	0.007
FADVCLRK	—	—	—	—	—	0.071
MINTCLRK	—	—	—	—	—	0.083
MINTPRTW	—	—	—	—	—	0.131
FINTPRTW	—	—	—	—	—	0.026
FELECLSW	—	—	0.500	—	—	0.139
MLABOUR	1.000	1.000	1.000	0.800	1.000	0.112
FLABOUR	1.000	1.000	0.333	1.000	1.000	0.082
M_UNEMP	0.500	0.545	0.636	0.643	—	0.087
F_UNEMP	0.667	0.400	0.600	0.500	—	0.072
ONEPARDP	0.111	—	0.088	0.094	0.333	0.087
GROUP	—	—	—	0.057	—	0.036
SPINCLO	—	1.000	1.000	1.000	1.000	0.428
SPINCHI	—	—	—	—	—	0.233
SPDOINCLO	0.500	—	0.500	0.300	0.875	0.191
SPDOINCHI	—	—	—	0.100	—	0.231
SPD_NDINCLO	0.500	0.667	—	0.400	0.500	0.245
SPD_NDINCHI	—	—	—	—	—	0.218
CFNKINCLO	1.000	0.667	0.667	0.667	0.600	0.217
CFNKINCHI	—	—	—	—	—	0.220
CFDOINCLO	1.000	1.000	0.667	0.800	1.000	0.252
CFDOINCHI	—	—	—	—	—	0.277
CFD_NDINCLO	1.000	1.000	1.000	1.000	1.000	0.286
CFD_NDINCHI	—	—	—	—	—	0.227
YR11LO	1.000	0.957	0.991	0.972	0.971	0.566
DEGREE_01	—	—	—	—	—	0.141
DIP	0.026	—	—	—	—	0.066
CAEUNI	—	—	—	—	—	0.048

— nil or rounded to zero (including null cells)

Examination of extreme values continued

5.2 VARIABLE VALUES FOR TOP FIVE CDs

Name	CD NUMBER *1390312*	*5090319*	*1412601*	*2291107*	*1391708*	*Aust. mean*
INTNET	0.249	0.333	0.266	0.224	0.344	0.182
HIGHBED	0.134	0.051	0.215	0.167	0.777	0.258
NOQUAL	0.273	0.397	0.239	0.294	0.305	0.573
NOSCH	—	—	—	—	—	0.010
MPROF	0.466	0.370	0.406	0.423	0.352	0.154
FPROF	0.329	0.550	0.474	0.514	0.377	0.207
MPARA_PF	0.121	0.370	0.109	0.171	0.254	0.118
FPARA_PF	0.138	0.050	0.123	0.181	0.131	0.116
MTRADE	0.017	—	0.051	0.072	0.014	0.199
FTRADE	—	—	0.007	0.010	—	0.032
MADVCLRK	0.012	—	0.023	—	0.014	0.007
FADVCLRK	0.132	0.200	0.065	0.076	0.115	0.071
MINTCLRK	0.075	—	0.091	0.099	0.085	0.083
MINTPRTW	0.012	0.037	0.029	0.009	—	0.131
FINTPRTW	0.006	—	—	0.010	—	0.026
FELECLSW	0.030	0.100	0.026	0.038	0.131	0.139
MLABOUR	0.023	—	0.006	0.018	0.014	0.112
FLABOUR	—	—	0.007	0.010	—	0.082
M_UNEMP	0.027	—	0.022	0.017	0.026	0.087
F_UNEMP	0.012	—	0.019	0.009	—	0.072
ONEPARDP	0.051	0.067	0.035	0.021	0.025	0.088
GROUP	0.051	0.051	0.073	0.073	—	0.037
SPINCLO	0.067	0.174	0.101	0.095	—	0.428
SPINCHI	0.756	0.522	0.685	0.536	1.000	0.233
SPDOINCLO	0.125	—	—	—	—	0.191
SPDOINCHI	0.875	1.000	0.750	1.000	1.000	0.231
SPD_NDINCLO	—	—	—	—	—	0.245
SPD_NDINCHI	1.000	—	0.800	1.000	1.000	0.219
CFNKINCLO	0.011	—	—	0.020	—	0.217
CFNKINCHI	0.807	0.625	0.779	0.784	0.733	0.221
CFDOINCLO	—	—	0.029	—	0.057	0.252
CFDOINCHI	0.800	1.000	0.912	1.000	0.886	0.277
CFD_NDINCLO	—	—	—	—	0.100	0.286
CFD_NDINCHI	0.778	1.000	1.000	1.000	0.600	0.227
YR11LO	0.141	0.153	0.144	0.154	0.113	0.566
DEGREE_01	0.500	0.362	0.525	0.526	0.520	0.141
DIP	0.102	0.155	0.106	0.100	0.070	0.066
CAEUNI	0.126	0.119	0.078	0.138	0.140	0.048

— nil or rounded to zero (including null cells)

Relationship between the indexes

We examined how similar the indexes were to each other. We expected a high correlation between some of the indexes, but not too high (otherwise we could drop one of the indexes); and lower correlations between those indexes where we only included certain dimensions of disadvantage (the index of education and occupation, and the index of economic resources).

Table 5.3 shows the correlations. The correlations between the Index of Advantage/Disadvantage and the Index of Economic Resources and Index of Education and Occupation are very high. This is not surprising, as the Index of Advantage/Disadvantage includes all the elements of advantage and disadvantage. The correlation between the indexes of Education and Occupation and Economic Resources is lower at 0.81. This was expected; the two indexes share no common variables.

Relationship between the indexes continued

The correlations between the Index of Disadvantage and the other indexes are lower because the Index of Disadvantage is the only index that includes level 3 variables (variables reflecting aspects of disadvantage), and excludes aspects of advantage.

5.3 CORRELATIONS BETWEEN THE INDEXES

	Index of Disadvantage	Index of Advantage/Disadvantage	Index of Economic Resources	Index of Education and Occupation
Index of Disadvantage	1.00	0.94	0.87	0.88
Index of Advantage/Disadvantage	0.94	1.00	0.93	0.96
Index of Economic Resources	0.87	0.93	1.00	0.81
Index of Education and Occupation	0.88	0.96	0.81	1.00

The next step was to investigate the difference between the indexes further. The following tables present the number of CDs which changed quintile from one index to another. Table 5.4 shows how CDs change quintiles between indexes. For instance, the first table compares the Index of Disadvantage and the Index of Advantage/Disadvantage. The figures in the first row show that among the 7,138 CDs in the bottom quintile of the Index of Disadvantage, 6,158 CDs were also in the bottom quintile of the Index of Advantage/Disadvantage; and 930 were in the second quintile of the Index of Advantage/Disadvantage; and so on.

CDs with extreme differences in quintiles (defined as a change of three or more quintiles) were further examined.

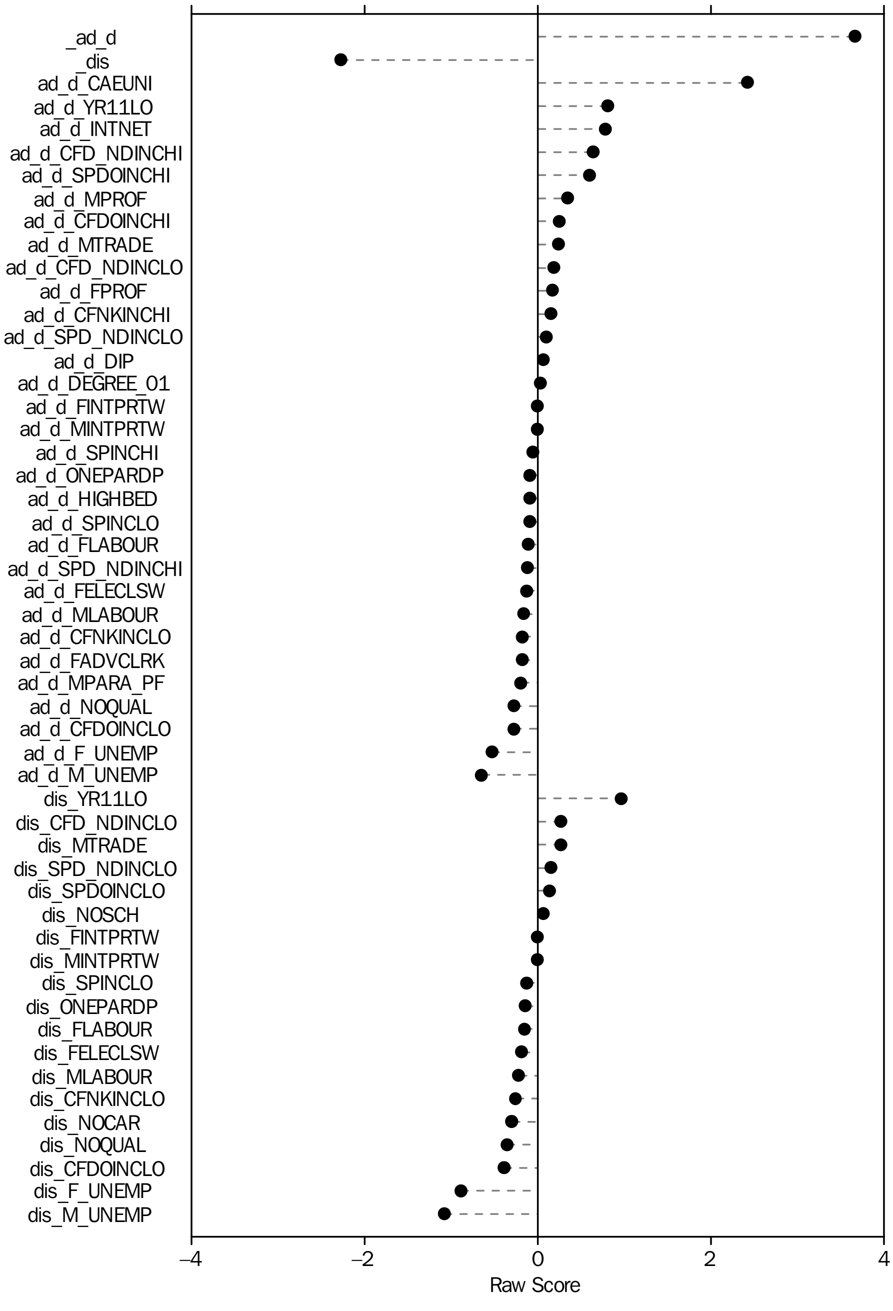
The seven CDs that were identified as being extremely different between indexes were examined further (as shown in graph 5.5). The chart decomposes the raw score for a CD into its contributing variables. The first two rows show the raw scores for each index (_dis for disadvantage and _ad_d for advantage/disadvantage). These raw scores are the weights multiplied by standardised values of the variables (standardised to a mean of zero and standard deviation of one). The other rows show how each variable contributed to the raw score for each index. In this example the main contributions to the higher Index of Advantage/Disadvantage were the education variables; and the main contributors to the lower Index of Disadvantage were the unemployment variables. This suggests a rather heterogeneous CD with both a significant number of educated people and unemployed people.

5.4 QUINTILES FOR DISADVANTAGE AND ADVANTAGE/DISADVANTAGE

Disadvantage Quintile	1	2	3	4	5	Total
1	6 158	930	47	2	1	7 138
2	970	4 832	1 253	80	4	7 139
3	10	1 294	4 528	1 281	26	7 139
4	—	83	1 253	5 018	785	7 139
5	—	—	58	758	6 324	7 140
Total	7 138	7 139	7 139	7 139	7 140	35 695

— nil or rounded to zero (including null cells)

5.5 EFFECT OF VARIABLES ON RAW SCORE, Advantage/Disadvantage and Disadvantage—Sample CD



We found in many cases where there were large differences between indexes, it was because the CD was heterogeneous.

State office validation

We sent the index values to the ABS state offices for validation using their local knowledge. Most ABS state office staff found the indexes were what they expected. Where they did raise concerns about an area, we looked at it in more detail, and found (in all cases) the value could be explained.

Validation using Social Health Atlas

In 1999, the Public Health Information Development Unit (PHIDU) at the University of Adelaide published the *Social Health Atlas*. This atlas brought together a wide variety of health statistics for every state.

Validation using Social Health Atlas continued

We investigated the correlation between the 1996 and 2001 Index of Disadvantage with health and socio-demographic data from the atlas. We wanted to ensure that:

- the strong correlations that were expected based on previous analysis still existed
- there were no major changes in the strength of the correlation between the atlas data and 1996 and 2001 SEIFAs.

This analysis used Pearson correlation coefficients and was done only for capital cities. The main point to make about the results was that the correlations with mortality were lower using the 2001 Index of Disadvantage.

The correlations with the 1996 Index of Disadvantage are available from the 1999 *Social Health Atlas*, which can be downloaded from www.publichealth.gov.au/atlas.htm.

Validation using Household Expenditure Survey

We regressed all the SEIFA indexes at the SLA level with the Household Expenditure Survey (HES) data. The SEIFA was population weighted to SLA level, then matched to HES data by SLA. Note that the Household Expenditure Survey data had high relative standard errors for some SLAs.

This analysis used actual SEIFA values, so we assumed a linear relationship between the index and underlying disadvantage. A better method would be to arrange the index values in quantiles, and use dummy variables. We didn't do this because for validation purposes, we considered that the assumption that there was a linear relationship between the index and the underlying disadvantage was acceptable. We would not make this assumption if we were making policy decisions or interpreting the results.

The results of this analysis are summarised in table 5.6. Coefficients are given, with standard errors in brackets. Overall, disadvantage was associated with lower housing costs; higher expenditure on tobacco; and lower expenditure on income tax. All these results seemed reasonable, given the limitation of using the actual indexes rather than quantiles.

Only those variables that were significantly correlated with at least one of the indexes are shown in the table. Expenditure items that were not associated with disadvantage at all were Food, Fuel and power, Clothing and footwear, Household furnishings, Medical and health expenses, Recreation, Personal care, Miscellaneous goods and services, Other capital housing, and Superannuation and life insurance.

Validation using Household Expenditure Survey continued

5.6 COEFFICIENTS FROM REGRESSION OF SEIFA WITH HES

Selected independent variables	Disadvantage	Advantage-disadvantage	Economic resource	Education & occupation
Housing costs	0.20	0.45	0.52	0.42
Standard Error(a)	(0.05)***	(0.09)***	(0.09)***	(0.10)***
Income tax	0.08	0.14	0.16	0.12
Standard Error(a)	(0.04)*	(0.06)**	(0.07)**	(0.06)**
Tobaco	–1.02	–1.35	- 0.96	–1.61
Standard Error(a)	(0.26)***	(0.34)***	(0.33)***	(0.36)***
Alcohol	0.19	0.33	0.19	0.46
Standard Error(a)	(0.14)	(0.19)*	(0.17)	(0.21)**
Transport	–0.02	–0.08	-0.07	–0.09
Standard Error(a)	(0.02)	(0.04)*	(0.04)*	(0.045)*
Mortgage (principal)	–0.01	–0.12	–0.06	–0.20
Standard Error(a)	(0.06)	(0.08)	(0.09)	(0.08)**
Intercept	948.70	927.80	911.30	947.30
Standard Error(a)	(9.0)***	(11.0)***	(10.1)***	(12.5)***

(a) * = significant at 10%, ** = significant at 5%, *** = significant at 1%.

Comparison between 1996 and 2001 rankings

We compared the 1996 indexes by CD with three of the 2001 indexes (there was no Index of Advantage/Disadvantage in 1996). Only those CDs that were comparable within a 10% dwelling or population limit were used, i.e. CDs where there were no major boundary changes between 1996 and 2001. This limited the analysis to about 29,000 CDs.

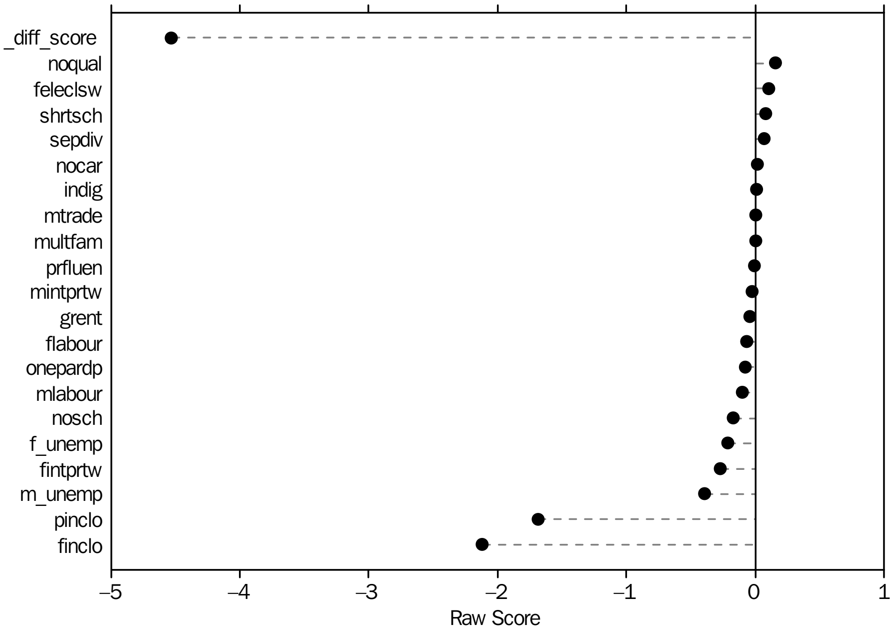
Table 5.7 shows our analysis for the Disadvantage index. We took extreme changes to be those that spanned four quartiles. 15 CDs moved from the top to the bottom quartiles (or vice-versa) between SEIFAs. We investigated why.

Graph 5.8 is an example of a CD which was in the fourth quartile in 1996 and is in the first quartile in 2001. It can be seen that the main change came from a change to the contribution of low family income (FINCLO) and low parental income (PINCLO). The weight for these variables changed considerably in 2001, caused by a change to the data.

5.7 CHANGE IN QUARTILES FROM 1996 TO 2001, DISADVANTAGE

1996 Quartile	1 2001 Quartile	2 2001 Quartile	3 2001 Quartile	4 2001 Quartile	Total 2001 Quartile
1	6 178	1 323	171	12	7 684
2	1 491	4 179	1 519	168	7 357
3	69	1 700	3 943	1 275	6 987
4	3	100	1 354	5 552	7 009
Total	7 741	7 302	6 987	7 007	29 037

5.8 EFFECT OF CHANGES IN VARIABLES, Disadvantage Index



Comparison between 1996 and 2001 rankings continued

The Economic Resource index showed more of a change from 1996 to 2001, as shown in table 5.9. The number of CDs going from the fourth quartile in 1996 to the first quartile in 2001 is 50; and the number going from the first quartile in 1996 to the fourth quartile in 2001 was 311. Changes from the fourth quartile in 1996 to the first quartile in 2001 were mainly due to the new low and high income variables in 2001. Income for 2001 was split into family types, and different cut offs were calculated for each family type.

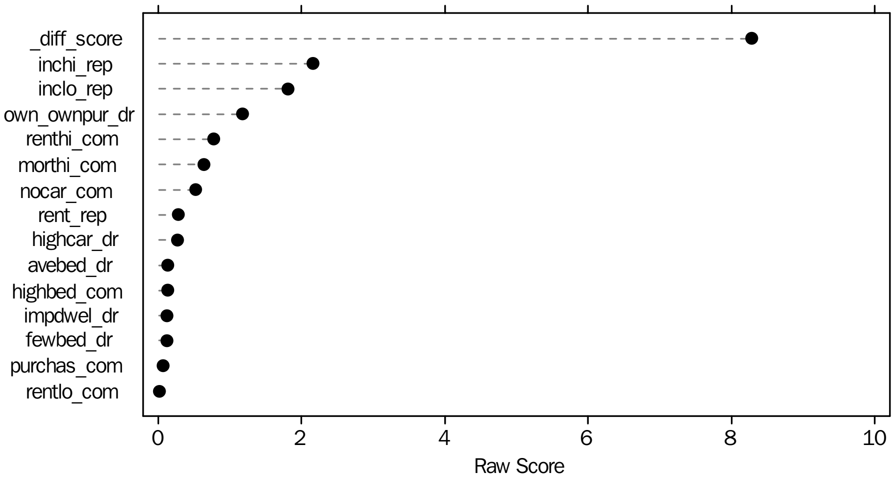
Examples of the graphs used to investigate which variables were affecting a CD index are shown in Graphs 5.10 and 5.11. In these graphs, a '_dr' suffix indicates the variable was dropped from 2001; a '_com' suffix indicated the variable was common to 1996 and 2001; and a '_rep' suffix indicates the variable replaced a 1996 variable (or a number of 1996 variables) in 2001.

The change from the first quartile in 1996 to the fourth quartile in 2001 was again driven by changes to income variables; but there was also some contribution by dropping the NoCar variable as an indicator of disadvantage; dropping the Owning and Purchasing a dwelling variable; and changing the financial cut offs for rent and mortgage payments.

5.9 CHANGES IN QUARTILE FROM 1996 TO 2001, INDEX OF ECONOMIC RESOURCES

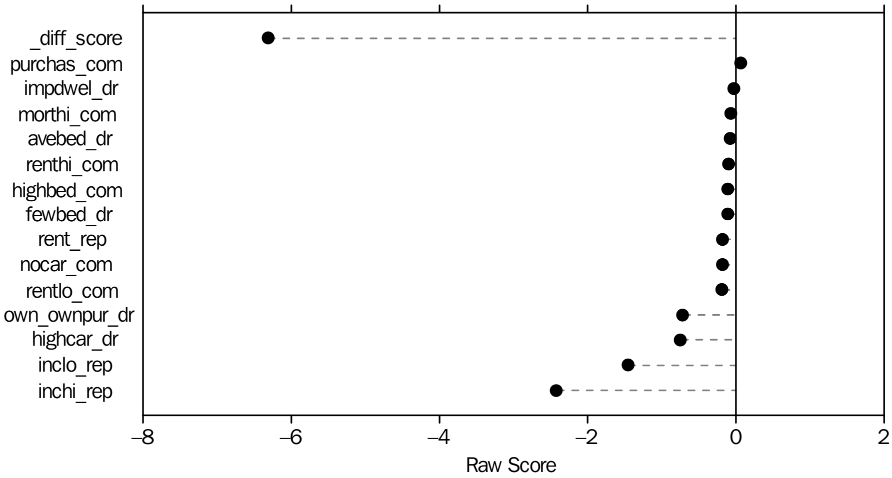
	1	*2*	*3*	*4*	*Total*
1996 Quartile	2001 Quartile	2001 Quartile	2001 Quartile	2001 Quartile	2001 Quartile
1	4 851	1 801	826	311	7 789
2	2 406	3 024	1 383	682	7 495
3	407	2 410	2 846	1 367	7 030
4	50	221	1 949	4 503	6 723
Total	7 714	7 456	7 004	6 863	29 037

5.10 EFFECT OF CHANGE IN VARIABLES, Quartile 1 to Quartile 4, Economic Resources



Comparison between 1996 and 2001 rankings continued

5.11 EFFECT OF CHANGES IN VARIABLES, Quartile 4 to Quartile 1, Economic Resources



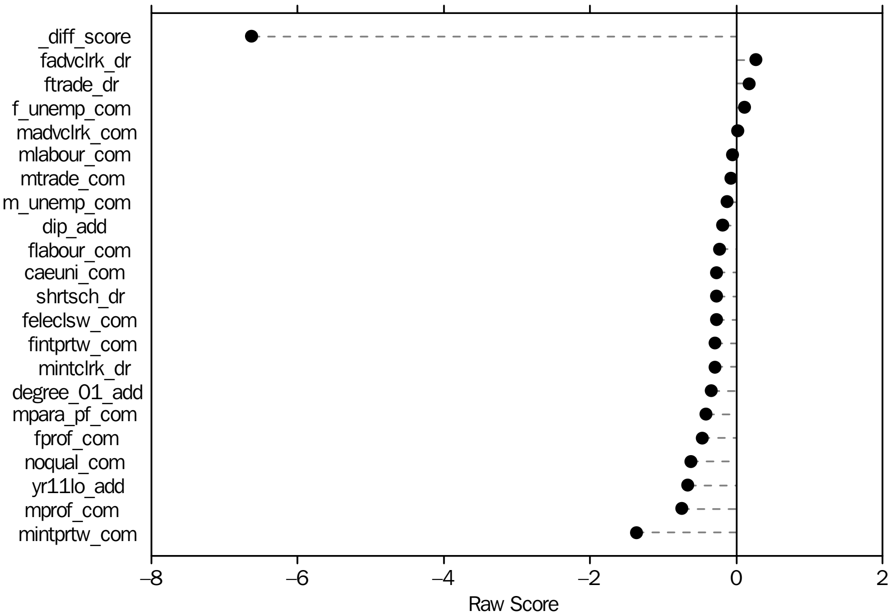
The Index of Education and Occupation showed the least change from 1996 to 2001. Table 5.12 shows that only two CDs went from quartile 4 in 1996 to quartile 1 in 2001; and five CDs went from quartile 1 in 1996 to quartile 4 in 2001. No one variable drove the change; in all CDs, it was a combination of a number of variables operating in the same direction. Graph 5.13 shows an example of this.

Comparison between 1996 and 2001 rankings continued

5.12 CHANGES IN QUARTILE FROM 1996 TO 2001, INDEX OF EDUCATION AND OCCUPATION

	1	*2*	*3*	*4*	*Total*
1996 Quartile	2001 Quartile	2001 Quartile	2001 Quartile	2001 Quartile	2001 Quartile
1	6 181	1 297	113	5	7 596
2	1 429	4 393	1 312	35	7 169
3	41	1 483	4 609	839	6 972
4	2	8	881	6 409	7 300
Total	7 653	7 181	6 915	7 288	29 037

5.13 EFFECT OF CHANGES IN VARIABLES, Quartile 4 to Quartile 1, Education and Occupation



Overall, we were satisfied that we could explain why the CDs had moved so much.

External Expert Group

In 2001, a group of six experts reviewed the method and the results of the index calculation before the indexes were released.

The expert group consisted of people from commonwealth and state public service departments and a number of Universities. The group were selected because they were either experts in Principal Component Analysis (PCA), in Socio-Economic research, or advanced users of SEIFA indexes.

The expert group provided much helpful advice and guidance, and their efforts in validating the indexes were appreciated. However, the ABS remains responsible for the methods used to construct the indexes and is accountable for any errors.

Colour maps of the indexes

We used colour maps to look for patterns in the indexes. Patterns for capital cities show areas of high advantage in the centre of capital cities; then areas of disadvantage on the outskirts. The areas of advantage and disadvantage tend to cluster together. Overall, we found that these patterns were visible in each state.

Colour maps of the indexes continued

Map 5.14 shows the Index of Advantage/Disadvantage for NSW. There are areas of high disadvantage clustered along the north coast; but areas of advantage through the south coast. There is also a band of advantage through the middle of the state.

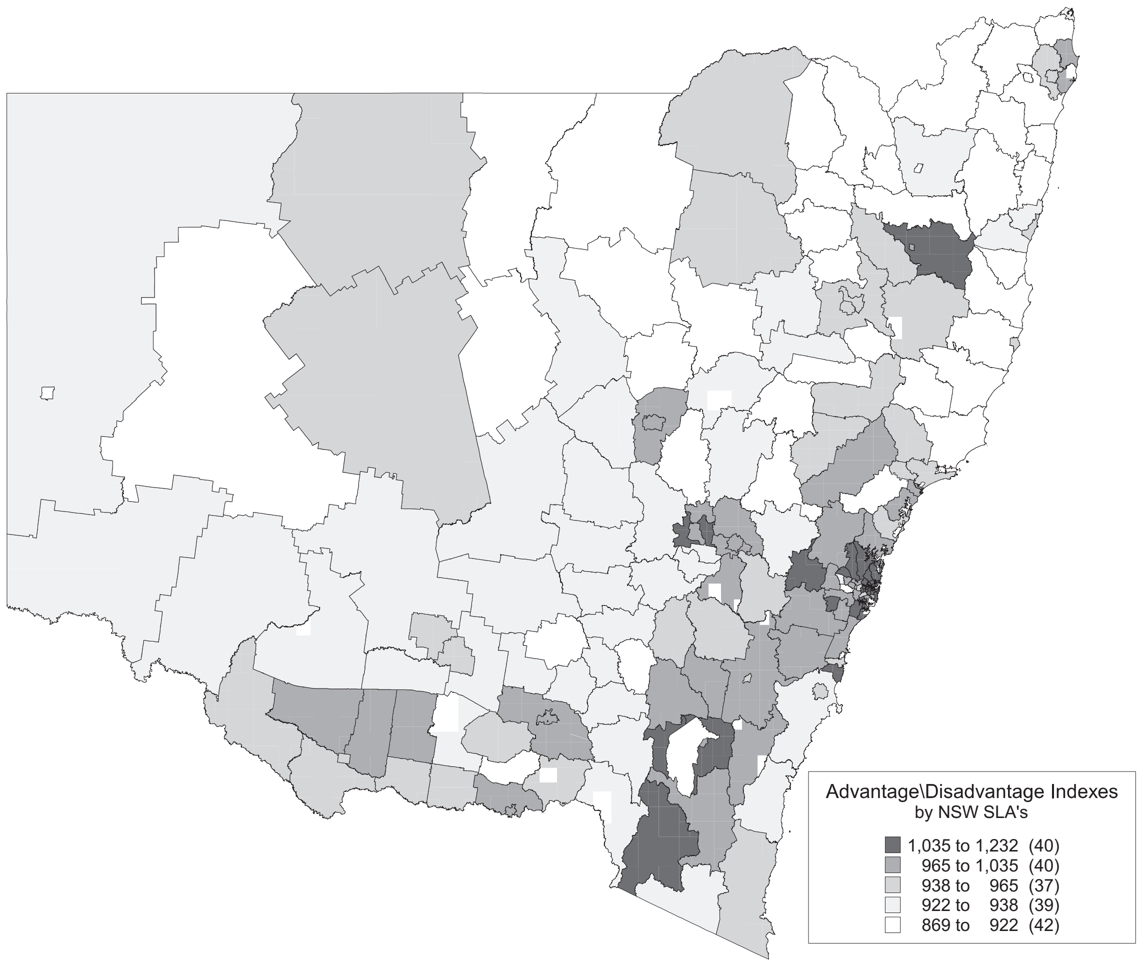
Map 5.15 shows the Index of Advantage/Disadvantage for Sydney. There is an area of advantage through the Northern suburbs; and an area of disadvantage to the west of Sydney. The centre of Sydney also experiences high advantage.

Map 5.16 shows a similar pattern for Hobart. The areas of advantage and disadvantage tend to be clustered (although there are exceptions to this rule, where we find some highly disadvantaged suburbs among highly advantaged ones).

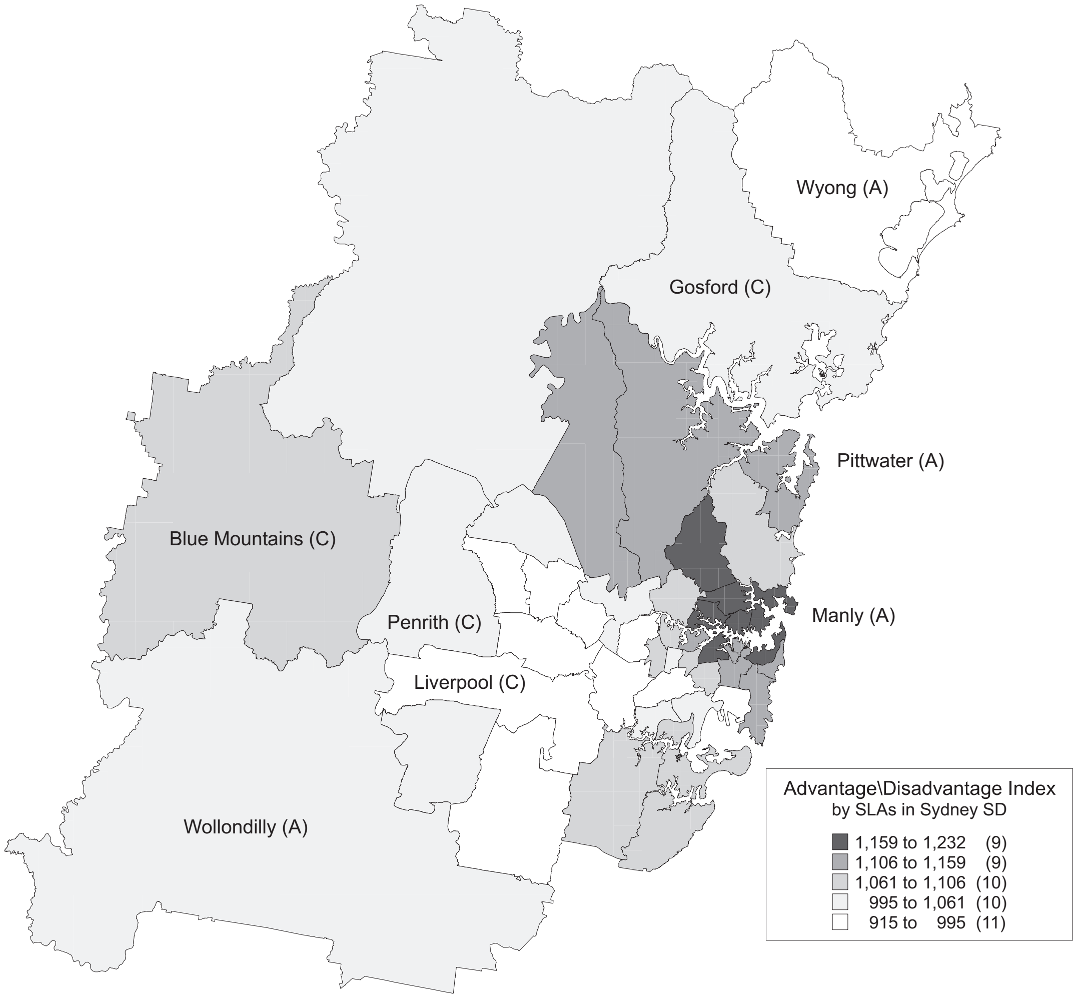
In all these maps, the numbers in brackets are the number of areas in each quintile.

Similar analyses were done with the other indexes.

5.14 INDEX OF ADVANTAGE/DISADVANTAGE FOR NEW SOUTH WALES

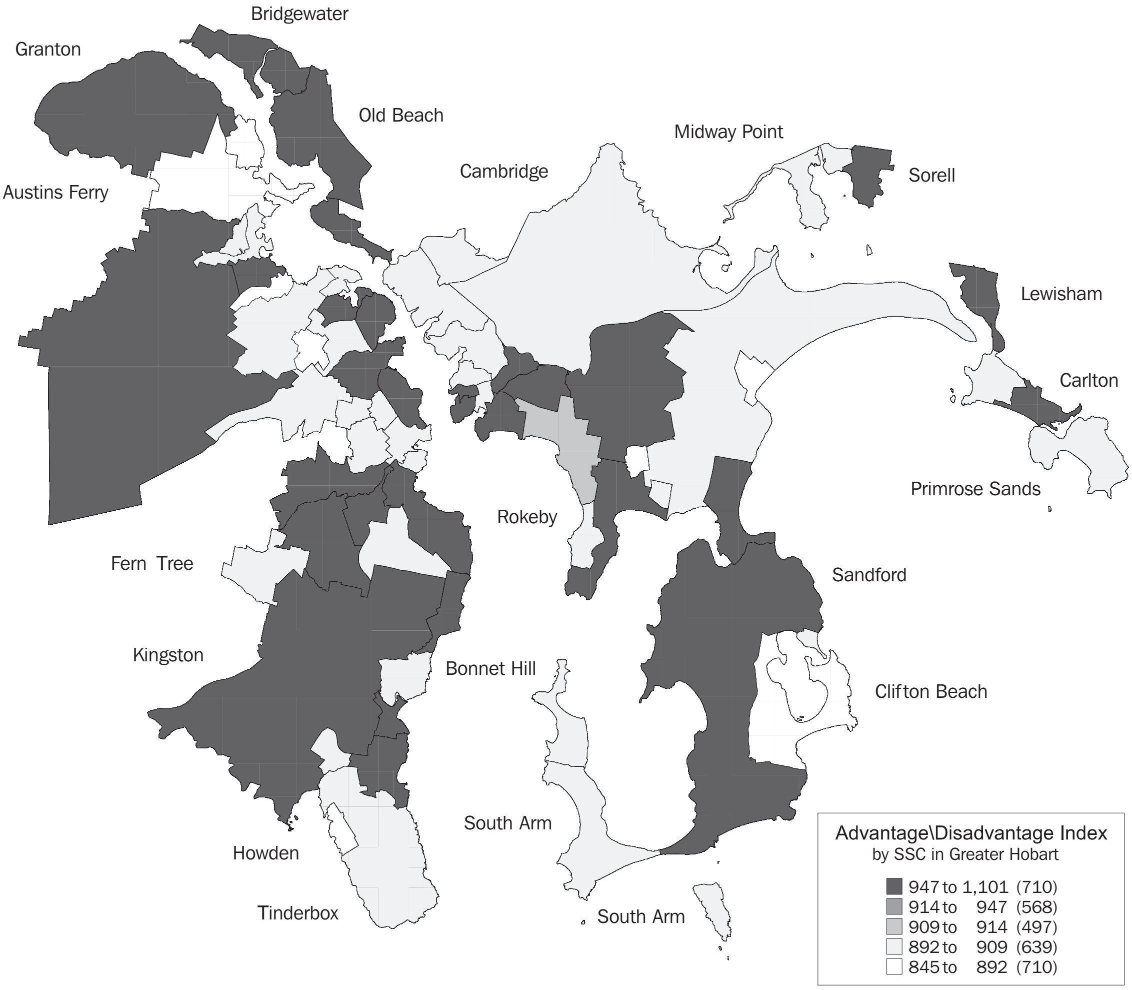


5.15 INDEX OF ADVANTAGE/DISADVANTAGE FOR SYDNEY



Colour maps of the indexes continued

5.16 INDEX OF ADVANTAGE/DISADVANTAGE FOR HOBART



CHAPTER **6**

THE NATURE OF THE INDEXES

THE NATURE OF THE SEIFA SCORES

This chapter outlines some of the issues users should be aware of when using SEIFA. They are:

- as a measure of underlying socioeconomic disadvantage, the indexes are ordinal
- differences between the Index of Disadvantage and the other three indexes
- SEIFA scores are more variable between less populous CDs
- SEIFA scores for remote areas in some states are most variable
- there are problems with aggregating the indexes to larger geographies, but there are methods that can be used to overcome this
- some variables in SEIFA are context specific
- there is a relationship between some of the SEIFA variables and age.

Ordinal indexes

SEIFA indexes are ordinal measures of the underlying socioeconomic disadvantage/advantage of areas. Ordinal measures allow the CDs to be ranked, but the distances between two CD values (with equal differences in SEIFA scores) are not necessarily equivalent. That is the difference in the underlying socioeconomic disadvantage between two CDs with index values of 1,000 and 1,100 is not necessarily the same as the underlying difference between CDs with index values of 900 and 1,000.

The SEIFA indexes themselves are interval scaled, and so analysts could use the indexes directly in techniques such as regression analysis. However, as indicators of the underlying concept that we are trying to measure (socioeconomic disadvantage), they should be used as ordinal scaled variables. We explain why below.

Suppose a CD has a particular level of socioeconomic disadvantage denoted by U, which cannot be measured as it is not directly observed, but manifests itself through the observed census socioeconomic variables $Y_1, Y_2, \ldots, Y_h$. U may be related to each of the Y variables in the sense that areas with a high value of disadvantage would tend to show a high (if the index weight is positive) value of Y. But this relationship might not be linear and one cannot verify it as U is not observable. An areas principal component score Z, on the other hand, is a linear summary of the CDs Y variables.

For areas i, j, and k, and their corresponding values of U and Z, we can say, that

(1) if $Z_i > Z_j$ then $U_i > U_j$

But we cannot say that

(2) if $Z_i - Z_j = Z_j - Z_k$ then $U_i - U_j = U_j - U_k$.

In other words Us and Zs are monotonically related, but are not necessarily linearly related.

Disadvantage vs other indexes

The Index of Relative Socio-Economic Disadvantage is somewhat different to the other three indexes. It includes only variables that are measures of or indicators of disadvantage (rather than advantage). This means that for CDs where there are pockets of advantage and disadvantage, the pockets of advantage will not offset the pockets of disadvantage. In the other indexes (the Index of Advantage/Disadvantage, the Index of Economic Resources and the Index of Education and Occupation), variables associated with both advantage and disadvantage are included. Thus within a CD, the pockets of advantage will offset pockets of disadvantage.

The Index of Advantage/Disadvantage, Index of Economic Resources and Index of Education and Occupation are on a continuum of disadvantage (low numbers) to advantage (high numbers); whereas the Index of Disadvantage is on a continuum of high disadvantage (low numbers) to low disadvantage (high numbers). Low disadvantage does not equate to high advantage.

The Index of Disadvantage is also different, in that it alone includes variables that reflect rather than measure disadvantage (level three variables). It is therefore the best general measure of disadvantage, if the user is not concerned about what aspects of disadvantage are being measured. The other three indexes allow a clearer identification of the aspects of advantage and disadvantage being measured.

The choice of index will depend on the use. If a user wants a general measure of disadvantage, and does not want any aspects of advantage, then the Index of Disadvantage is the best index to use. This is also the index to use if the user wants an index similar to the 1996 Index of Relative Socio-Economic Disadvantage.

If a user is interested in both advantage and disadvantage in an area, and wants a more targetted measure, then the other three indexes should be used, depending on whether the user wants general advantage/disadvantage; economic advantage/disadvantage; or education and occupation advantage/disadvantage.

Variation of SEIFA between less populous CDs

This section looks at whether less populous CDs have higher or lower SEIFA values.

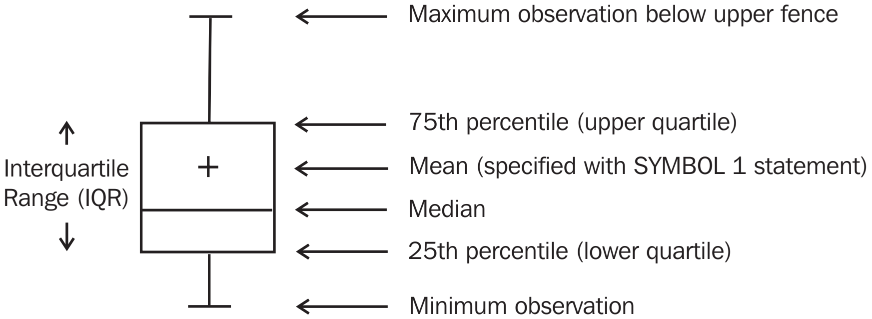
We looked at the relationship between population size of the CD and the indexes by plotting the index on the X axis of a graph and the population on the Y axis. If we see a strong relationship in some states but not in others, then we need to be aware that when using the population to weight the CD index to a larger area, the weighting pattern is not random. There will be a bias towards advantage or disadvantage, depending on which type of CDs are more populous.

The plots can be reproduced using the SEIFA indexes by CD. The plots for Advantage/Disadvantage show that for the NT, the less populous CDs tend to have a lower Advantage/Disadvantage index (i.e. they are more disadvantaged), but for most states, the relationship is not apparent.

The next set of plots we looked at provide a view of the distribution of the CD values in a larger geographic area. They give some idea of whether certain areas experience more variability in the CD values. The larger areas we have chosen for this analysis are ABS Remoteness Areas.

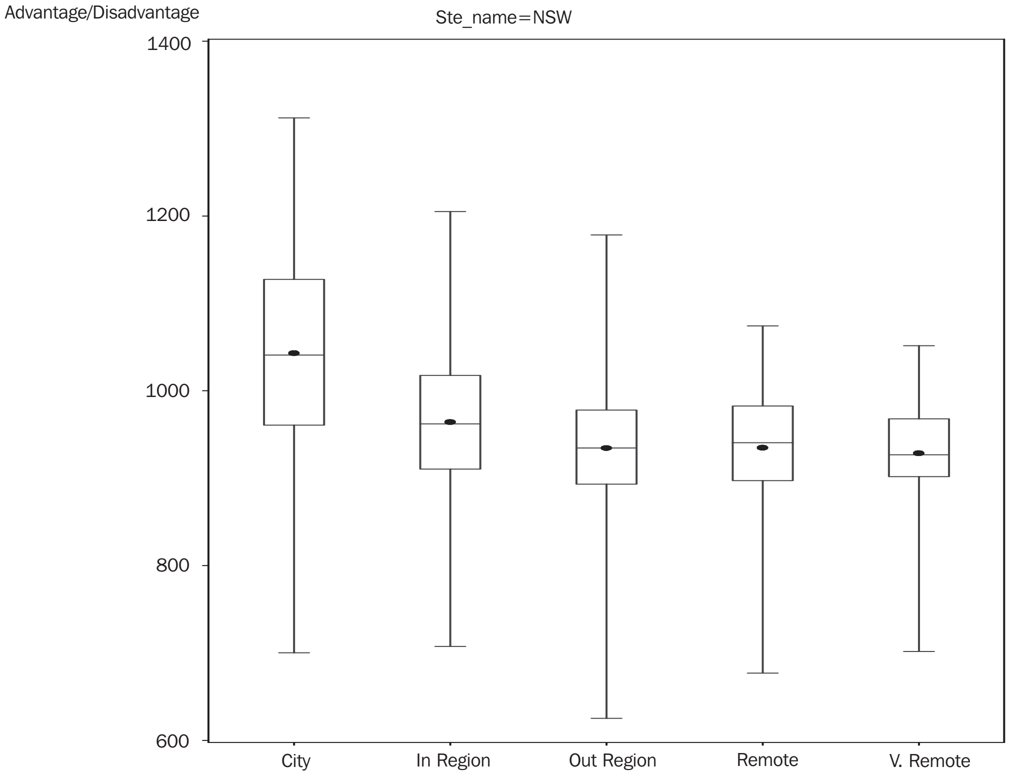
The plots below were of the following format:

6.1 FORMAT OF BOXPLOTS

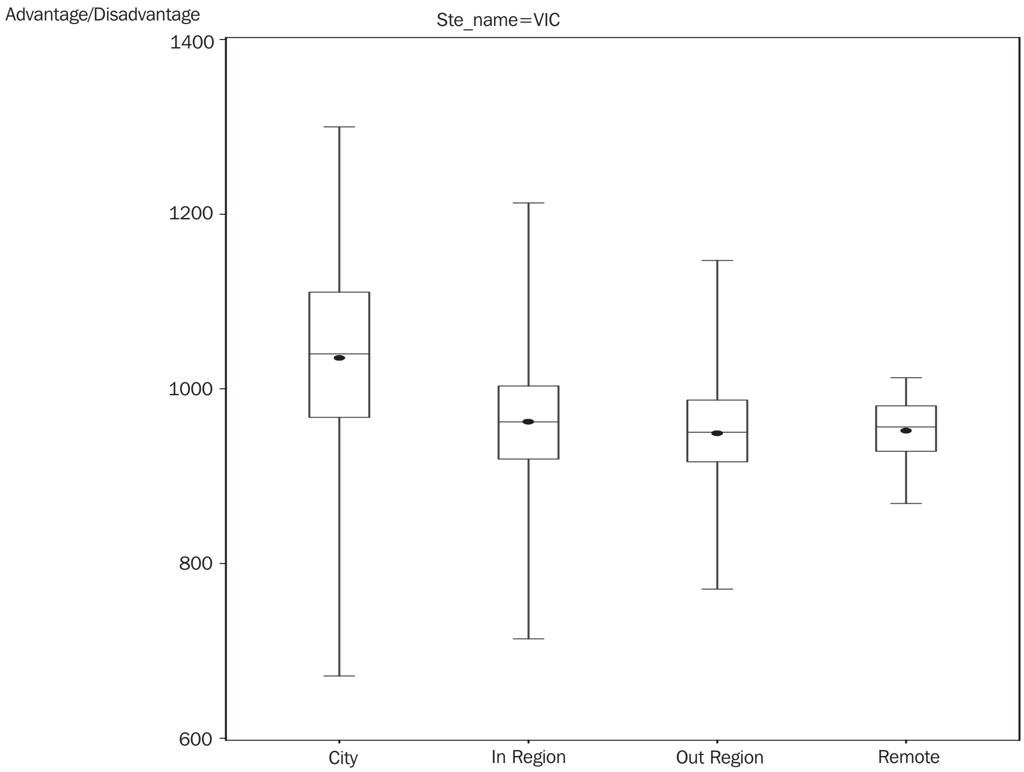


We found there was little relationship between the remoteness of the area and the range of values. Urban, rural and remote areas experience variability in the CD values.

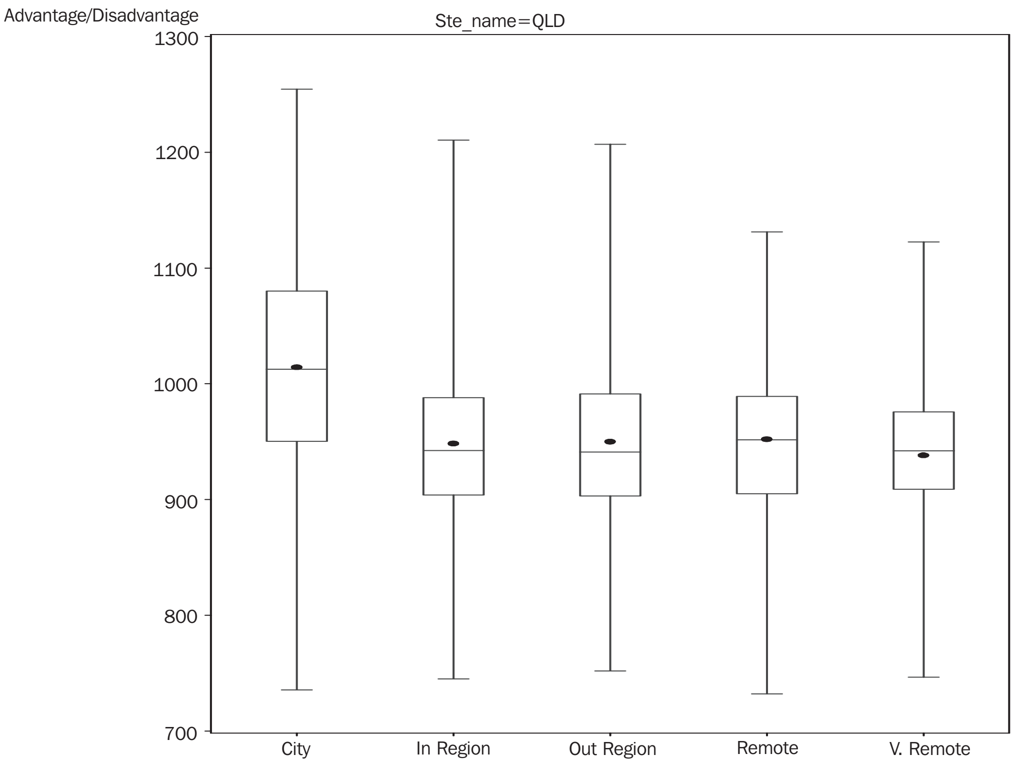
6.2 DISTRIBUTION OF ADVANTAGE/DISADVANTAGE CD VALUES WITHIN REMOTENESS AREA, NSW



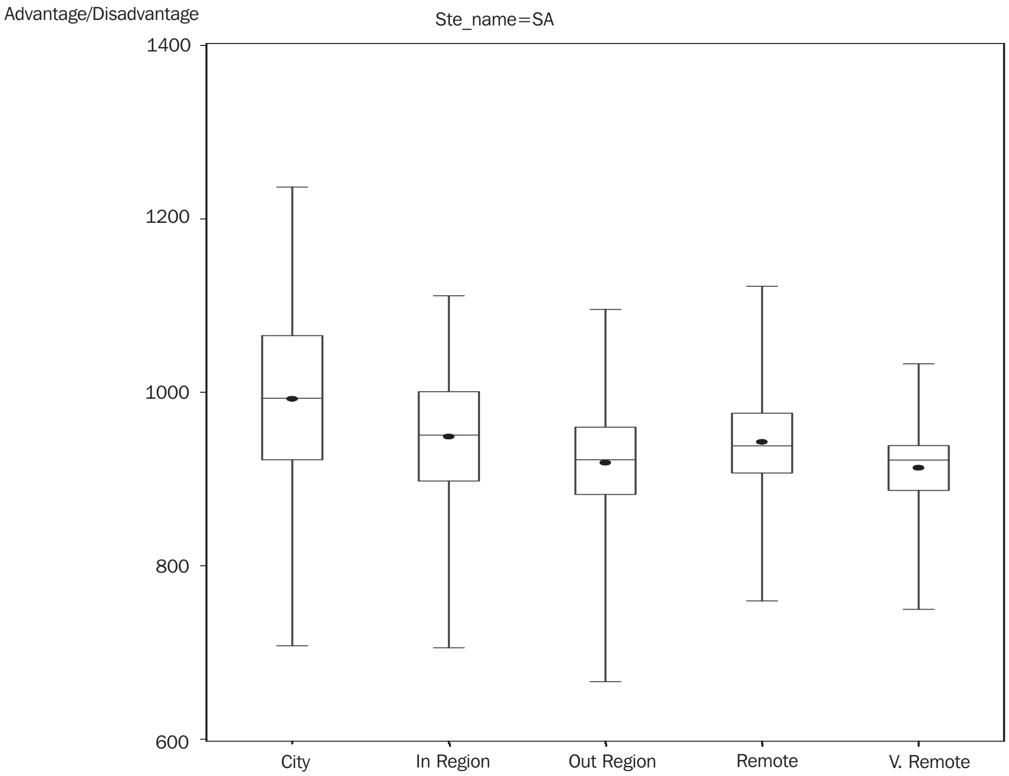
6.3 DISTRIBUTION OF ADVANTAGE/DISADVANTAGE CD VALUES WITHIN REMOTENESS AREA, VIC



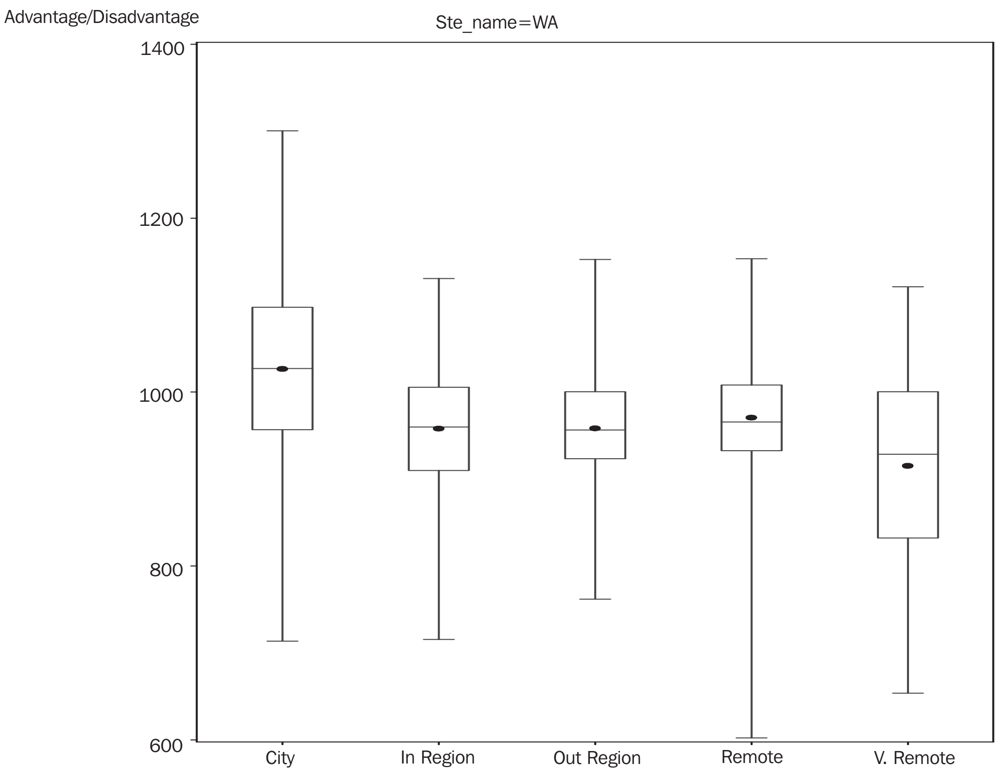
6.4 DISTRIBUTION OF ADVANTAGE/DISADVANTAGE CD VALUES WITHIN REMOTENESS AREA, QLD



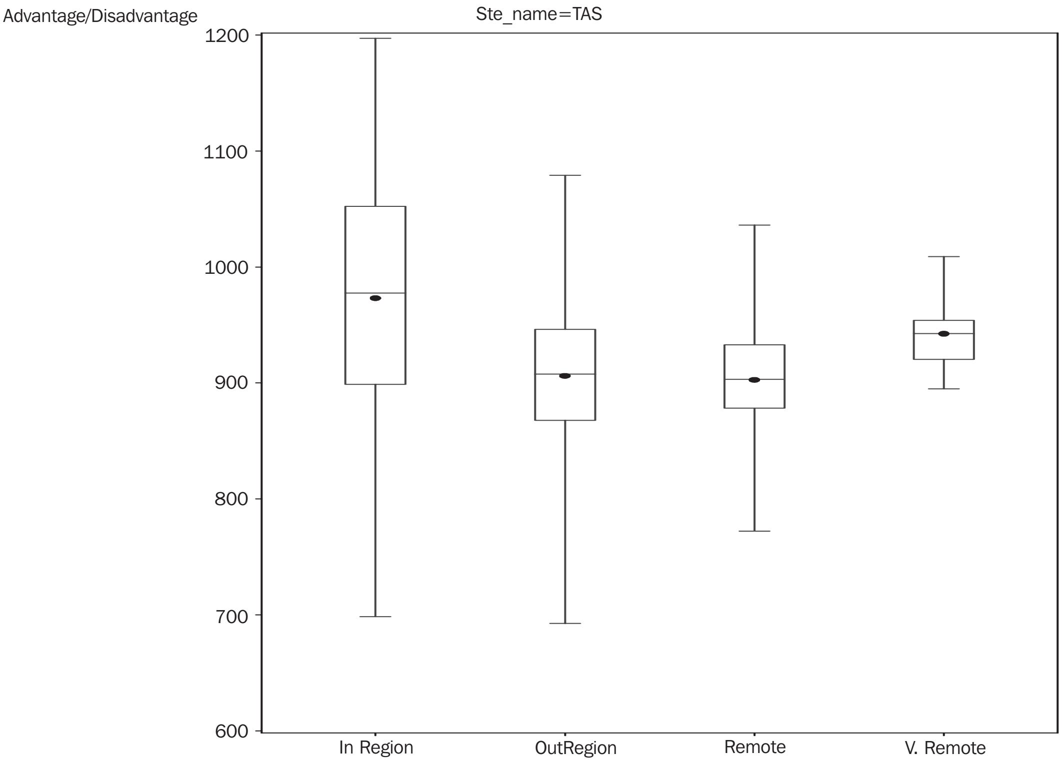
6.5 DISTRIBUTION OF ADVANTAGE/DISADVANTAGE CD VALUES WITHIN REMOTENESS AREA, SA



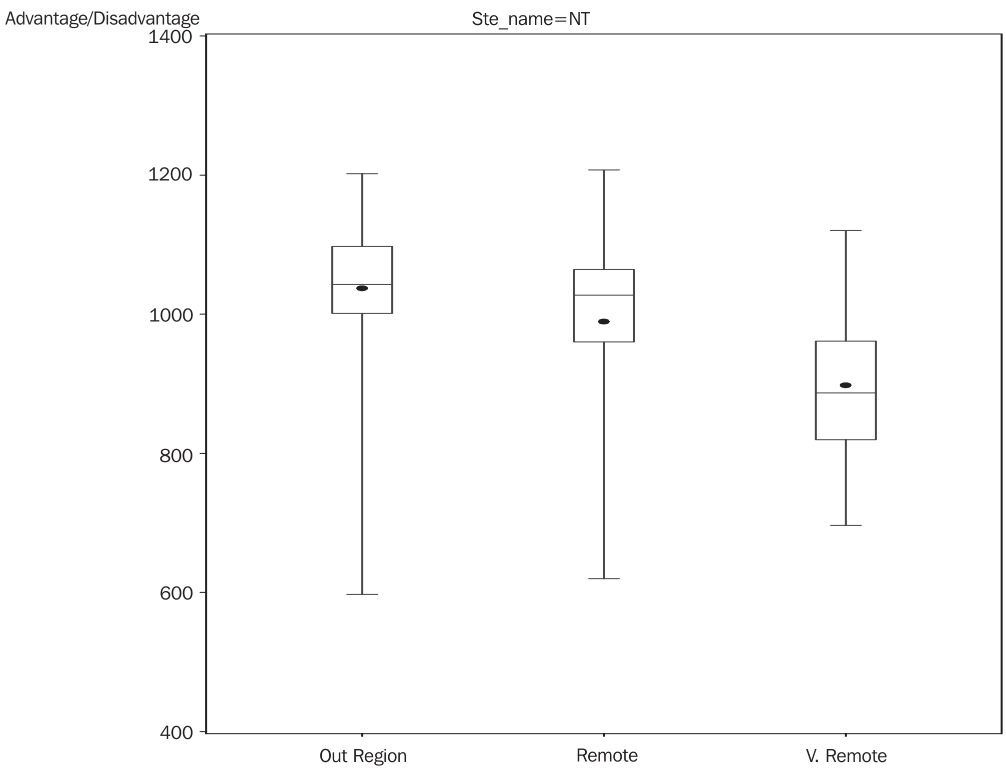
6.6 DISTRIBUTION OF ADVANTAGE/DISADVANTAGE CD VALUES WITHIN REMOTENESS AREA, WA



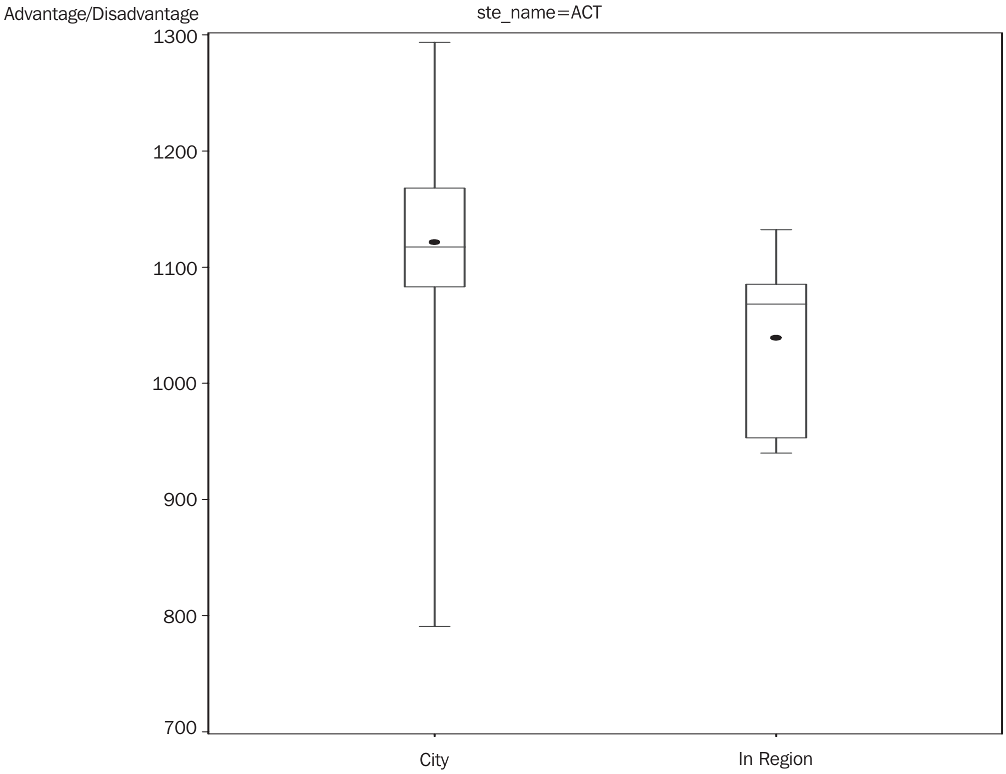
6.7 DISTRIBUTION OF ADVANTAGE/DISADVANTAGE CD VALUES WITHIN REMOTENESS AREA, TAS



6.8 DISTRIBUTION OF ADVANTAGE/DISADVANTAGE CD VALUES WITHIN REMOTENESS AREA, NT



6.9 DISTRIBUTION OF ADVANTAGE/DISADVANTAGE CD VALUES WITHIN REMOTENESS AREA, ACT



Variation of SEIFA between less populous CDs continued

These types of graphs could be produced for any geography, giving the researcher some idea of the variability of the CD values in the areas they were looking at.

SEIFA by remoteness

When validating the SEIFA indexes, we found some interesting relationships with ABS remoteness indicators. This section investigates those relationships further.

Table 6.9 shows the percentage change from 1996 to 2001 in the Disadvantage index by ABS Remoteness Area classification for each state. Very Remote areas in some states experienced more change than other areas. Across Australia (which is a simple average of all the state values) the Very Remote classification had the largest change (–1.03%).

The previous section suggests that this change is not due to greater variability in remote areas, since we found little difference in the variability of the CD level indexes in urban and remote areas. There may have been more variability over time (i.e. from 1996 to 2001) rather than within the areas (which the previous section looked at). Or it may be due to people moving from urban to city areas over this time.

6.10 PERCENTAGE CHANGE FROM 1996 TO 2001 DISADVANTAGE BY REMOTENESS AREA

	NSW	Vic.	Qld	SA	WA	Tas.	NT	ACT	Aust.
Remoteness Area	per cent	per cent	per cent	per cent	per cent	per cent	per cent	per cent	per cent
Major Cities of Australia	–0.88	–0.39	—	0.91	–0.20	—	—	–1.10	–0.29
Inner Regional Australia	–0.30	0.30	0.93	1.42	0.72	–0.61	—	1.90	0.63
Outer Regional Australia	0.31	1.41	–0.10	2.11	–0.61	–0.11	–1.95	—	0.13
Remote Australia	1.07	1.79	1.04	1.83	–0.81	0.95	–1.10	—	0.68
Very Remote Australia	–0.55	—	0.56	–0.87	–3.44	–0.61	–1.28	—	–1.03

— nil or rounded to zero (including null cells)

Problems with summarising the indexes to larger geographies

To get indexes for larger geographies than CD, a population weighted average of the CD level indexes is calculated. Suppose there are three CDs in an SLA, the formula used is:

$I_{SLA}=(Pop_{CD1}*I_{CD1} + Pop_{CD2}*I_{CD2} + Pop_{CD3}*I_{CD3})/Pop_{SLA}$

where I_{SLA} is the SEIFA Index for an SLA, Pop_{CDn} is the population in CD n, I_{CDn} is the SEIFA index for CD n and Pop_{SLA} is the population in an SLA.

Summarising the indexes in this way means information is lost. If smaller populations are more disadvantaged (which in the NT they are), then a population weighted average will mask information about these CDs because the more populous advantaged CDs are given a greater weight.

There may be better ways to calculate an index for an SLA than using a population weighted average. For instance, the analyst could look at the distribution of CD indexes in the SLA, and (if they are looking at disadvantage) might use the SEIFA index for the bottom quintile of CDs in each SLA. Or they might look at the proportion of CDs in an SLA that are in the bottom quintile of Australian CDs.

The other averaging method we use for the indexes is a simple average. The simple average is usually used to summarise indexes at a higher geography, which we have calculated using a population weighted average. We may not want to list the index values for all 1,336 SLAs, so we average them to a state.

Simple averages put the value of a less populous CD on an equal footing as more populous CDs. This can mean that in some areas, there will be a large difference between a population weighted average and a simple average.

In summary, the weighted average is used to create indexes at different geographies; and a simple average can be used for presentation (e.g. SLA level indexes summarised to state).

Context-specific variables

Some of the variables in SEIFA may be context specific. For instance, a low mortgage in Sydney might be a high mortgage in Dubbo. All SEIFA variables were chosen as indicators of advantage/disadvantage at the Australia level.

Relationships between some of the variables and Age.

Some of the variables we use in SEIFA are associated with age. For instance, income increases as employees gain experience. Similarly, younger people are less likely to own a home. This is something to be aware of when using the indexes; but cannot be corrected in SEIFA, since the fact is that in terms of socioeconomic advantage/disadvantage, older people will tend to be more economically advantaged, due to a longer period saving, paying mortgages, etc.

INCOME VARIABLES

CFD_NDINCLO

% COUPLE FAMILIES WITH DEPENDENTS AND NON-DEPENDENTS OR WITH NON-DEPENDENTS ONLY WITH ANNUAL INCOME LESS THAN $52,000

- number of families with dependents and non-dependents or with non-dependents only with income <$ 52,000 [FMTF = 111, 121,131,141 and FINF = 1–14]
- number of families with dependents and non-dependents or with non-dependents only with fully stated income [FMTF = 111, 121,131,141 and FINF = 1–18]

CFD_NDINCHI

% COUPLE FAMILIES WITH DEPENDENTS AND NON-DEPENDENTS OR WITH NON-DEPENDENTS ONLY WITH ANNUAL INCOME GREATER THAN $103,999

- number of families with dependents and non-dependents or with non-dependents only with income > $ 103,999 [FMTF = 111, 121,131,141 and FINF = 18]
- number of families with dependents and non-dependents or with non-dependents only with fully stated income [FMTF = 111, 121,131,141 and FINF = 1–18]

CFDOINCLO

% COUPLE FAMILIES WITH DEPENDENT CHILD(REN) ONLY WITH ANNUAL INCOME LESS THAN $36,400

- number of couple families with dependent child(ren) only with stated annual income < $36,400 [FMTF = 112, 122, 132 and FINF = 1–12]
- number of couple families with dependent child(ren) only with fully stated annual income [FMTF = 112, 122, 132 and FINF=1–18]

CFDOINCHI

% COUPLE FAMILIES WITH DEPENDENT CHILD(REN) ONLY WITH ANNUAL INCOME GREATER THAN $77,999

- number of couple families with dependent child(ren) only with stated annual income > $77,999 [FMTF = 112, 122, 132 and FINF =17,18]
- number of couple families with dependent child(ren) only with fully stated annual income [FMTF = 112, 122, 132 and FINF=1–18]

CFNKINCLO

% COUPLE FAMILIES WITH NO CHILDREN WITH ANNUAL INCOME LESS THAN $20,800

- number of couple only families with stated annual income < $20,800 [FMTF = 2 and FINF = 1–9]
- number of couple families with fully stated annual income [FMTF = 2 and FINF=1–18]

CFNKINCHI

% COUPLE FAMILIES WITH NO CHILDREN WITH ANNUAL INCOME GREATER THAN $77,999

- number of couple only families with stated annual income > $77,999 [FMTF = 2 and FINF =17,18]
- number of couple only families with fully stated annual income [FMTF = 2 and FINF=1–18]

FINCLO

% FAMILIES WITH INCOME LESS THAN $15,600

- number of families with all members present and with stated annual income < $15,600 (FINF = 1–8)
- number of families with all members present and fully stated income (excludes visitors, lone person & group households) (FINF = 1–18)

PINCLO

% FAMILIES WITH OFFSPRING AND PARENTAL INCOME LESS THAN $15,600

- number of families with some offspring and all members present and with stated annual income < $15,600 (FINF = 1–8 and FMTF = 1 or 3)
- number of families with all members present and fully stated income (excludes visitors, lone person & group households) (FINF = 1–18 and FMTF = 1 or 3)

SPD_NDINCLO

% SINGLE PARENT FAMILIES WITH DEPENDENTS AND NON-DEPENDENTS OR WITH NON-DEPENDENTS WITH ANNUAL INCOME LESS THAN $26,000

- number of single parent families with dependents and non-dependents or with non-dependents only with stated annual income < $ 26,000 [FMTF = 311, 321, 331, or 341 and FINF = 1–10]
- number of single parent families with dependents and non-dependents or with non-dependents only with fully stated annual income [FMTF = 311, 321, 331, or 341 and FINF = 1–18]

SPD_NDINCHI

% SINGLE PARENT FAMILIES WITH DEPENDENTS AND NON-DEPENDENTS OR WITH NON-DEPENDENTS WITH ANNUAL INCOME GREATER THAN $62,399

- number of single parent families with dependents and non-dependents or with non-dependents only with stated annual income > $ 62,399 [FMTF = 311, 321, 331, or 341 and FINF = 16–18]
- number of single parent families with dependents and non-dependents or with non-dependents only with fully stated annual income [FMTF = 311, 321, 331, or 341 and FINF = 1–18]

SPDOINCLO

% SINGLE PARENT FAMILIES WITH DEPENDENT CHILD(REN) ONLY WITH ANNUAL INCOME LESS THAN $15,600

- number of single parent families with dependent child(ren) only with stated annual income < $15,600 [FMTF =312, 322, or 332 and FINF =1–8]
- number of single parent families with dependent child(ren) only with fully stated annual income [FMTF = 312, 322, or 332 and FINF = 1–18]

SPDOINCHI

% SINGLE PARENT FAMILIES WITH DEPENDENT CHILD(REN) ONLY WITH ANNUAL INCOME GREATER THAN $36,399

- number of single parent families with dependent child(ren) only with stated annual income > $ 36,399 [FMTF =312, 322, or 332 and FINF =13–18]
- number of single parent families with dependent child(ren) only with fully stated annual income [FMTF = 312, 322, or 332 and FINF = 1–18]

SPINCLO

% SINGLE PERSON HOUSEHOLDS WITH ANNUAL INCOME LESS THAN $15,600

- number of single person households with stated annual income < $15,600 (group household members are treated as a single person household) [RLHP = 72, 73 and INCP = 1–8]
- number of single person households with fully stated income [HHTD = 21, 22 and INCP = 1–16]

SPINCHI

% SINGLE PERSON HOUSEHOLDS WITH ANNUAL INCOME GREATER THAN $36,399

- number of single person households with stated annual income > $36,399 (group household members are treated as a single person household) [RLHP = 72, 73 and INCP = 13–16]
- number of single person households with fully stated income [HHTD = 21, 22 and INCP = 1–16]

EDUCATION VARIABLES

ATSCH

% PERSONS AGED 15 YEARS AND OVER WHO ARE STILL AT SCHOOL

- number of persons aged 15 years and over who are still at school [HSCP = 1]
- number of persons aged 15 years and over (excluding not stated 'Highest level of schooling completed') [HSCP = 1–7]

CAEUNI

% PERSONS AGED 15 YEARS AND OVER AT UNIVERSITY OR OTHER TERTIARY INSTITUTION

- number of persons aged 15 years and over who are attending a university or other tertiary institution [AGEP>=15 and TYPP = 9]
- number of persons aged 15 years and over (excluding not stated) [AGEP >=15 and TYPP<>&&]

CERT

% PERSONS AGED 15 YEARS AND OVER WITH CERTIFICATE QUALIFICATION

- number of persons aged 15 years and over with a certificate qualification [QALLP = 5]
- number of persons aged 15 years and over (excluding not stated level of qualification) [QALLP = 1–5 or ((QALLP=@@ or =01) and AGEP >=15)]

DEGREE

% PERSONS AGED 15 AND OVER WITH DEGREE OR HIGHER

- number of persons aged 15 years and over with a bachelor degree or higher [QALLP = 1–3]
- number of persons aged 15 years and over (excluding not stated level of qualification) [QALLP = 1–5 or ((QALLP=@@ or =01) and AGEP >=15)]

DIP

% PERSONS AGED 15 YEARS AND OVER HAVING AN ADVANCED DIPLOMA OR DIPLOMA QUALIFICATION

- number of persons aged 15 years and over with an advanced diploma or diploma [QALLP = 4]
- number of persons aged 15 years and over (excluding not stated level of qualification) [QALLP = 1–5 or ((QALLP=@@ or =01) and AGEP >=15)]

NOQUAL

% PERSONS AGED 15 YEARS AND OVER WITH NO QUALIFICATIONS

- number of persons aged 15 years and over with no qualifications (excluding still at school) [QALLP=@@ and AGEP>=15 and HSCP <> 1]
- number of persons aged 15 years and over (excluding not stated level of qualification) [QALLP = 1–5 or ((QALLP=@@ or =01) and AGEP >=15)]

NOSCH

% PERSONS AGED 15 YEARS AND OVER WHO DID NOT GO TO SCHOOL

- number of persons aged 15 years and over who did not go to school [HSCP = 2]
- number of persons aged 15 years and over (excluding not stated 'Highest level of schooling completed') [HSCP = 1–7]

SHRTSCH

% PERSONS AGED 15 YEARS AND OVER WHOSE HIGHEST LEVEL OF SCHOOLING COMPLETED WAS YEAR 10 OR LOWER

- number of persons aged 15 years and over whose highest level of schooling completed was year 10 or lower (HSCP = 2–5)
- number of persons aged 15 years and over (excluding not stated 'Highest level of schooling completed') (HSCP = 1–7)

YR11LO

% PERSONS AGED 15 YEARS AND OVER WHO LEFT SCHOOL AT YEAR 11 OR LOWER

- number of persons aged 15 years and over whose highest level of schooling completed was year 11 or lower [HSCP = 2–6]
- number of persons aged 15 years and over (excluding not stated 'Highest level of schooling completed') [HSCP = 1–7]

OCCUPATION

FADVCLRK

% EMPLOYED FEMALES CLASSIFIED AS 'ADVANCED CLERICAL & SERVICE WORKERS'

- number of females aged 15 years and over classified as 'Advanced Clerical and Service Workers' [SEXP = 2 and OCCP = 5]
- number of females aged 15 years and over with stated occupation [SEXP = 2 and OCCP = 1–9]

MADVCLRK

% EMPLOYED MALES CLASSIFIED AS 'ADVANCED CLERICAL & SERVICE WORKERS'

- number of males aged 15 years and over classified as 'Advanced Clerical and Service Workers' [SEXP = 1 and OCCP = 5]
- number of males aged 15 years and over with stated occupation [SEXP = 1 and OCCP = 1–9]

FELECLSW

% EMPLOYED FEMALES CLASSIFIED AS 'ELEMENTARY CLERICAL, SALES & SERVICE WORKERS'

- number of females aged 15 years and over classified as 'Elementary Clerical, Sales and Service Workers' [SEXP = 2 and OCCP = 8]
- number of females aged 15 years and over with stated occupation [SEXP = 2 and OCCP = 1–9]

MELECLSW

% EMPLOYED MALES CLASSIFIED AS 'ELEMENTARY CLERICAL, SALES & SERVICE WORKERS'

- number of males aged 15 years and over classified as 'Elementary Clerical, Sales and Service Workers' [SEXP = 1 and OCCP = 8]
- number of males aged 15 years and over with stated occupation [SEXP = 1 and OCCP = 1–9]

FINTCLRK

% EMPLOYED FEMALES CLASSIFIED AS 'INTERMEDIATE CLERICAL, SALES & SERVICE WORKERS'

- number of females aged 15 years and over classified as 'Intermediate Clerical, Sales and Service Workers' [SEXP = 2 and OCCP = 6]
- number of females aged 15 years and over with stated occupation [SEXP = 2 and OCCP = 1–9]

MINTCLRK

% EMPLOYED MALES CLASSIFIED AS 'INTERMEDIATE CLERICAL, SALES & SERVICE WORKERS'

- number of males aged 15 years and over classified as 'Intermediate Clerical, Sales and Service Workers' [SEXP = 1 and OCCP = 6]
- number of males aged 15 years and over with stated occupation [SEXP = 1 and OCCP = 1–9]

FINTPRTW

% EMPLOYED FEMALES CLASSIFIED AS 'INTERMEDIATE PRODUCTION & TRANSPORT WORKERS'

- number of females aged 15 years and over classified as 'Intermediate Production and Transport Workers' [SEXP = 2 and OCCP = 7]
- number of females aged 15 years and over with stated occupation [SEXP = 2 and OCCP = 1–9]

MINTPRTW

% EMPLOYED MALES CLASSIFIED AS 'INTERMEDIATE PRODUCTION AND TRANSPORT WORKERS'

- number of males aged 15 years and over classified as 'Intermediate Production and Transport Workers' [SEXP = 1 and OCCP = 7]
- number of males aged 15 years and over with stated occupation [SEXP = 1 and OCCP = 1–9]

FLABOUR

% EMPLOYED FEMALES CLASSIFIED AS 'LABOURERS & RELATED WORKERS'

- number of females aged 15 years and over classified as 'Labourers and Related Workers' [SEXP = 2 and OCCP = 9]
- number of females aged 15 years and over with stated occupation [SEXP = 2 and OCCP = 1–9]

MLABOUR

% EMPLOYED MALES CLASSIFIED AS 'LABOURERS & RELATED WORKERS'

- number of males aged 15 years and over classified as 'Labourers and Related Workers' [SEXP = 1 and OCCP = 9]
- number of males aged 15 years and over with stated occupation [SEXP = 1 and OCCP = 1–9]

FMGR_ADM

% EMPLOYED FEMALES CLASSIFIED AS 'MANAGERS OR ADMINISTRATORS'

- number of females aged 15 years and over classified as 'Managers or Administrators' [SEXP = 2 and OCCP = 1]
- number of females aged 15 years and over with stated occupation [SEXP = 2 and OCCP = 1–9]

MMGR_ADM

% EMPLOYED MALES CLASSIFIED AS 'MANAGERS OR ADMINISTRATORS'

- number of males aged 15 years and over classified as 'Managers or Administrators' [SEXP = 1 and OCCP = 1]
- number of males aged 15 years and over with stated occupation [SEXP = 1 and OCCP = 1–9]

FPARA_PF

% EMPLOYED FEMALES CLASSIFIED AS 'ASSOCIATE PROFESSIONALS'

- number of females aged 15 years and over classified as 'associate professionals' [SEXP = 2 and OCCP = 3]
- number of females aged 15 years and over with stated occupation [SEXP = 2 and OCCP = 1–9]

MPARA_PF

% EMPLOYED MALES CLASSIFIED AS 'ASSOCIATE PROFESSIONALS'

- number of males aged 15 years and over classified as 'associate professionals' [SEXP = 1 and OCCP = 3]
- number of males aged 15 years and over with stated occupation [SEXP = 1 and OCCP = 1–9]

FPROF

% EMPLOYED FEMALES CLASSIFIED AS 'PROFESSIONALS'

- number of females aged 15 years and over classified as 'Professionals' [SEXP = 2 and OCCP = 2]
- number of females aged 15 years and over with stated occupation [SEXP = 2 and OCCP = 1–9]

MPROF

% EMPLOYED MALES CLASSIFIED AS 'PROFESSIONALS'

- number of males aged 15 years and over classified as 'Professionals' [SEXP = 1 and OCCP = 2]
- number of males aged 15 years and over with stated occupation [SEXP = 1 and OCCP = 1–9]

FTRADE

% EMPLOYED FEMALES CLASSIFIED AS 'TRADESPERSONS'

- number of females aged 15 years and over classified as 'tradepersons' [SEXP = 2 and OCCP = 4]
- number of females aged 15 years and over with stated occupation [SEXP = 2 and OCCP = 1–9]

MTRADE

% EMPLOYED MALES CLASSIFIED AS 'TRADESPERSONS'

- number of males aged 15 years and over classified as 'tradepersons' [SEXP = 1 and OCCP = 4]
- number of males aged 15 years and over with stated occupation [SEXP = 1 and OCCP = 1–9]

UNEMPLOYMENT

F_UNEMP

% FEMALES (IN LABOUR FORCE) UNEMPLOYED

- number of females aged 15 years and over unemployed and looking for work [SEXP = 2 and LFSP = 5–6]
- number of females aged 15 years and over in labour force [SEXP = 2 and LFSP = 1–6]

M_UNEMP

% MALES (IN LABOUR FORCE) UNEMPLOYED

- number of males aged 15 years and over unemployed and looking for work [SEXP = 1 and LFSP = 5–6]
- number of males aged 15 years and over in labour force [SEXP = 1 and LFSP = 1–6]

DWELLINGS AND LIVING CONDITIONS

AVEBED

AVERAGE NUMBER OF BEDROOMS PER PERSON

- estimated number of bedrooms in occupied private dwellings in CD [number of bedrooms using BEDD = 1–5]
- estimated number of persons in CD living in private dwellings (including no bedrooms) [number of persons with BEDD = 0–5]

CARAVAN

% PERSONS LIVING IN CARAVAN PARKS

- number of persons living in caravan parks and enumerated at home [DLOD = 1 and UAICP = 1]
- number of persons enumerated at home [UAICP = 1]

FEWBED

% DWELLINGS WITH ONE OR NO BEDROOMS

- number of occupied private dwellings (includes caravans in parks) with zero or one bedrooms [BEDD = 0 or 1]
- number of occupied private dwellings (includes caravans in parks) with a stated number of bedrooms (including no bedrooms) [BEDD = 0–5]

GRENT

% HOUSEHOLDS RENTING (GOVERNMENT AUTHORITY)

- number of households who rent from a government authority (LLDD = 3 or 5)
- number of occupied private dwellings (includes those with occupancy not stated and caravans in parks) (TEND = 1–7 or &)

GROUP

% HOUSEHOLDS WHO ARE GROUP HOUSEHOLDS

- number of private dwellings (including caravans in parks) occupied by group households [HHTD = 22]
- number of occupied private dwellings (including caravans in parks) [HHTD = 11–32]

HIGHBED

% DWELLINGS WITH FOUR OR MORE BEDROOMS

- number of occupied private dwellings with four (4) or more bedrooms [BEDD = 4 or 5]
- number of occupied private dwellings with a stated number of bedrooms (including 0) [BEDD = 0–5]

IMPDWEL — % HOUSEHOLDS LIVING IN IMPROVISED DWELLINGS

- number of households in occupied private dwellings of type: a. non-park caravan, houseboat etc., b. improvised home, campers out, c. house/flat attached to shop or office etc. [(STRD = 91 (but excluding those where DLOD = 1, 2, 3 or 4), 93 or 94) (exclude non–occupied private dwellings)]
- number of households in occupied private dwellings (with stated structure type) [STRD = 11, 21, 22, 31–34, 91, 93, or 94 (exclude non–occupied private dwellings)]

MORTHI_01 — % HOUSEHOLDS PAYING MORTGAGE GREATER THAN $1,360 PER MONTH

- number of occupied private dwellings with monthly mortgage repayments > $1,360 per month [HLRD > $1,360]
- number of mortgaged occupied private dwellings (excluding those with mortgage not stated) [HLRD >= 0]

MULTFAM — % OCCUPIED PRIVATE DWELLINGS WITH TWO OR MORE FAMILIES

- number of occupied dwellings (including caravans in parks) with two or more families (excludes group households) [HHTD = 12–13]
- number of occupied dwellings (including caravans in parks, excluding not classifiable households) [HHTD = 1–3]

NGRENT — % HOUSEHOLDS RENTING (NON-GOVERNMENT AUTHORITY)

- number of households who rent from other (non-government) landlords (LLDD = 1, 2, 3, 6 or 7)
- number of occupied private dwellings (includes those with occupancy not stated and caravans in parks) (TEND = 1–7 or &)

OWNING — % HOUSEHOLDS OWNING DWELLING

- number of households owning the dwelling they occupy (includes caravans in parks) [TEND = 1]
- number of occupied private dwellings (includes those with occupancy not stated) [TEND = 1–7 or &]

PURCHAS — % HOUSEHOLDS PURCHASING DWELLING

- number of households purchasing the dwelling they occupy [TEND = 2–3]
- number of occupied private dwellings (includes those with occupancy not stated) [TEND = 1–7 or &]

RENT — % RENTAL DWELLINGS

- number of household renting dwelling [TEND = 4]
- number of occupied private dwellings (includes those with occupancy not stated) [TEND = 1–7 or &]

RENTHI_01 — % HOUSEHOLDS PAYING RENT GREATER THAN $225 PER WEEK

- number of rented occupied private dwellings with rent > $225 per week [RNTD > $225]
- number of rented occupied private dwellings with stated rent [RNTD >= 0]

RENTLO_01 — % HOUSEHOLDS PAYING RENT LESS THAN $88 PER WEEK

- number of rented occupied private dwellings with rent < $88 per week [RNTD>= 0 and RNTD < $88]
- number of rented occupied private dwellings with stated rent [RNTD >= 0]

MOTOR CARS

HIGHCAR

% DWELLINGS WITH THREE OR MORE MOTOR CARS

- number of occupied private dwellings (includes caravans in parks) which had three or more registered motor vehicles at or the near dwelling [VEHD = 3 or 4]
- number of occupied private dwellings (includes caravans in parks) with a stated number of vehicles (including 0) [VEHD = 0–4]

NOCAR

% DWELLINGS WITH NO MOTOR CAR AT DWELLING

- number of occupied private dwellings (includes caravans in parks) which did not have a registered motor vehicle at or the near dwelling [VEHD = 0]
- number of occupied private dwellings (includes caravans in parks) with a stated number of vehicles (including 0) [VEHD = 0–4]

OTHERS

COMP

% PERSONS USING COMPUTER AT HOME

- number of persons using computers at home [COMP = 2]
- number of persons (including not stated) [COMP = 1,2 or &]

INDIG

% INDIGENOUS PERSONS

- number of persons identified as either/both an Aboriginal or Torres Strait Islander (INGP = 2–4)
- number of persons with stated indigenous status (INGP = 1–4)

INTNET

% PERSONS USING INTERNET AT HOME

- number of persons using the internet at home (NETP = 2)
- number of persons (including not stated NETP) (NETP = 1–8 or &)

ONEPARDP

% ONE-PARENT FAMILIES WITH DEPENDENT OFFSPRING ONLY

- number of one-parent families with dependent offspring only [FMTF = 3122, 3222 or 3322]
- number of families [FMTF = 1–3 or 9]

PRFLUEN

% LACKING FLUENCY IN ENGLISH

- number of persons aged five years and over born in non-English speaking countries who speak English either not well or not at all [AGEP >= 5 and ENGP = 3–4]
- number of persons aged five years and over (excluding those who did not state their proficiency or language [AGEP >=5 and (ENGP <> 5 and ENGP <> &)]

RECMIG

% RECENT MIGRANTS FROM NON-ENGLISH SPEAKING COUNTRIES

- number of persons born overseas in non-English speaking countries who have stated arrival in Australia or an external territory in 1996 or later (YARP = 4–9: All values of BPLP except the following: 1101, 1201, 2100, 2102, 2104-2106, 2201, 8102, 8104, 9225, and supplementary codes for BPLP).
- number persons excluding those migrants with year of arrival not stated and/or birth place not stated (AGEP = 0–99 and YARP <> && and BPLP <> supplementary codes for BPLP)

SEPDIV

% PERSONS AGED 15 YEARS AND OVER SEPARATED OR DIVORCED

- number of persons aged 15 years and over are separated or divorced (MSTP = 3 or 4)
- number of persons aged 15 years and over (MSTP = 1–5)

GLOSSARY

Component	In Principal Components Analysis, a linear combination of the original variables specified. There will be a number of components in an analysis, with each component explaining less of the variance in the original dataset.
Correlation Matrix	A matrix where the diagonal elements are 1, and the off-diagonal elements are correlations between pairs of variables.
Covariance Matrix	A matrix where the diagonal elements are the variances of the variables, and the off-diagonal elements are the covariances of pairs of variables.
Eigenvalue	The amount of total variance accounted for. It is measured in units of variance, and is typically converted to a percent of variance by dividing it by the number of variables and multiplying by 100.
Interval Value	Also called Continuous variables, equal sized differences on different parts of the scale are equivalent.
Loadings	A standardised measure of the relationship between a component and a variable. In SEIFA, it is the correlation between a component and a variable.
Monotonically Related	For a monotonic increasing series, each value is greater than or equal to the previous value. For a monotonic decreasing series, each value is less than or equal to the previous value.
Ordinal Value	A measurement that allows ranking of values; but differences at different points of the scale are not necessarily equivalent.
Pearsons Correlation Coefficient	An index that identifies the linear relationship between 2 variables using a value between –1 and 1. The sign indicates the direction of the relationship, and the number the strength. Values of –1 and 1 indicate a perfect correlation. A value of zero indicates a lack of any linear relationship.
Quantiles	Quantiles are divisions of a frequency distribution into equal, ordered subgroups. It is the general term for quartiles, quintiles, deciles, etc.
Raw Score	The score calculated by summing the product of all the variables and their weights.
Remoteness Area	ABS geographies based on the Accessibility/Remoteness Index of Australia (ARIA).
Rotation	Used in Principal Components Analysis to achieve a more interpretable solution. Typically, a number of variables will load moderately onto the first component, making interpretation of the component difficult. Rotation will result in a simpler structure, meaning each variable will have a higher loading on only one factor. Rotation can be Orthogonal or Oblique. Assuming for simplicity only two components are retained, an orthogonal rotation keeps the two components at right angles. An oblique rotation allows the reference axes to be rotated without constraint to achieve an optimal fit.
Scree Plot	A method of deciding how many components to retain. For SEIFA, we always use the first component; however, a scree plot was used to decide how many components to retain for rotation.
Weight	The weight applied to each variable's value to produce a raw score. It is calculated as the variables loading divided by the square root of the component's eigenvalue.

BIBLIOGRAPHY

BIBLIOGRAPHY

Ainley, J, Graetz, B, Long, M & Batten, M (1995). 'Socio-economic status and School education', Australian Centre for Educational Research, June 1995

Glover, J , Harris, K & Tennant, S (1999). *A Social Health Atlas of Australia*, Public Health Information Development Unit, Adelaide

Jolliffe, I (1986). *Prinipal Components Analysis*, Springer-Verlag, New York

Marks, G, McMillan, J, Jones, F & Ainley, J (2000). *The measurement of Socioeconomic Status for the Reporting of National Comparable Outcomes of Schooling*, Draft report of the National Education Performance Monitoring Taskforce, Australian Council for Educational Research & Sociology Program, Research School of Social Sciences, Australian National University

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

2000001451199

RRP $0.00