HILDA PROJECT DISCUSSION PAPER SERIES
No. 3/09, March 2009

Quality of the Occupation and Industry Coding in the HILDA Survey

Nicole Watson and Michelle Summerfield

The HILDA Project was initiated, and is funded, by the Australian Government Department of Families, Housing, Community Services and Indigenous Affairs

Introduction

Each year, respondents to the HILDA Survey are asked to describe the occupation and industry of their current job and / or their previous job. These verbatim responses are then coded to standard occupation and industry codeframes.

In 2006, the Australian Bureau of Statistics (ABS) updated the standard occupation and industry codeframes (the old and new codeframes are listed in Table 1). With input from the HILDA External Reference Group, the HILDA Survey team decided to have the wave 1 to 6 responses recoded to these new codeframes (this is in addition to the original coding to the old codeframes undertaken each wave). It was also decided that from wave 7 onwards, only the new codeframes would be used.

Analysis of the recoding work has lead to an investigation of the quality of the coding work undertaken. This paper describes the coding process, provides several indicators of the quality of the coding and offers recommendations for future work. Comments are very welcome and should be directed to Nicole Watson (n.watson@unimelb.edu.au).

Table 1: Old and new codeframes for occupation and industry questions

	Old codeframe	*New codeframe*	*Waves coded*
Occupation	Australian Standard Classification of Occupation, Second Edition,1997 (ASCO) (See ABS 1997)	Australian and New Zealand Standard Classification of Occupations, First Edition, 2006 (ANZSCO) (See ABS, 2006a)	Waves 1-6: ASCO & ANZSCO Wave 7 onwards: ANZSCO only
Industry	Australian and New Zealand Standard Industry Classification, First Edition, 1993 (ANZSIC FE) (See ABS 1993)	Australian and New Zealand Standard Industry Classification, Second Edition, 2006 (ANZSIC SE) (See ABS 2006b)	Waves 1-6: ANZSIC FE & ANZSIC SE Wave 7 onwards: ANZSIC SE only

Recoding Process

Both the original coding and the recoding work were undertaken by Nielsen, the fieldwork provider for Waves 1 to 8 of the HILDA Survey.

Assuming that the original coding work was of reasonably high quality, it was decided that the recoding work would only need to occur where there was a one-to-many match in the concordance between the old and new codeframes at the 4 digit level.[1] The ABS provided training to Nielsen staff in the use of the new codeframes. The verbatim responses were manually inspected and coded to the new codeframe. A list of possible codes (according to the concordance of the two codeframes) was provided to the coder, but they were permitted to code outside of these possible codes if they deemed it necessary. All coding was verified by an experienced coder. This verification was not blind – the verifier could see the code specified by the original coder.[2]

[1] Where there was a one-to-one match in the concordance between the old and new codes, the original code was *automatically* coded to the new codeframe.

[2] This process was designed to reflect the procedure adopted for the original coding which occurred directly onto the paper questionnaire prior to data entry (the occupation and industry verbatim text was not captured at

The person coding the responses could view the respondent's occupation title, tasks and industry when coding both occupation and industry. The *ANZSCO Coder* and the *ANZSIC Coder*, which are Windows-based computer assisted coders provided by the ABS, were used to identify a suitable code. The occupation and industry responses were coded to the 4-digit level.[3]

Nielsen began coding the wave 6 data and worked backwards to wave 1. This was because the verbatim text was available for waves 5 and 6 when they started but the verbatim text for waves 1 to 4 had to be data entered before the coding could begin for those waves.

Quality of Nieslen's Coding

On completion of the recoding process, the HILDA Survey team undertook an analysis of the quality of Nielsen's coding. The quality was assessed in three ways:

i) by considering how often the verifier selects a code that is different from the original coder;

ii) by calculated how often Nielsen selected a code outside the bounds of the concordance between the old and new codeframes; and

iii) through an independent evaluation of the coding.

Concordance between original coder and verifier

The top part of Table 2 shows the proportion of codes modified by the Nielsen verifier when recoding cases that matched to multiple codes in the new codeframe. These estimates are averaged across all occupation questions for each wave and across all industry questions for each wave. The high rate of corrections in wave 6 suggest the coders were still learning the codeframe when coding this wave.[4]

Nielsen also recoded a random sample of 500 cases from wave 3 that had a one-to-one match between the new and old codeframes so that an error rate could be established for this part of the codeframe. The rate of agreement between the original coder and the verifier is much higher for this sample (as shown in the last row of Table 2).

As all cases are verified, the verified code is used in the final dataset. The next two sections consider the accuracy of these verified codes.

---

the time for waves 1 to 4). The practice of coding onto the paper questionnaire continued in waves 5 and 6, even though the verbatim text was captured.

[3] 2-digit codes are provided in the General Release, but 4-digit are provided in the In-Confidence Release.

[4] An alternative (and less likely) explanation for the drop in the override rate between wave 6 and wave 5 is that the verifier is choosing not to override codes for ambiguous responses as often.

Table 2: Per cent of cases overridden by verifier at the 4-digit level

Wave	Occupation	Industry
	Manually coded (1-many match)	
1	10.8%	7.9%
2	11.0%	5.7%
3	8.8%	7.5%
4	11.8%	11.3%
5	12.4%	9.2%
6	16.7%	21.0%
	Recoded sample of automatically coded (1-1 match)	
3	5.0%	4.2%

Concordance between old and new codeframes

This section considers how often the new codeframe codes were outside the expected codes given the concordance tables between the old and new codeframes. Our focus here is on the quality of the coding for occupation and industry of the *main* job.[5]

Approximately 44 per cent of the occupation verbatim responses were manually inspected by Nielsen coders (shown in Table 3). Of these, 10 per cent were coded outside the allowable ANZSCO codes given to the ASCO code at the 4-digit level. For industry, there was also approximately 44 per cent of the codes were manually coded. Of these, around 9 per cent were coded outside the allowable ANZSIC SE code for the given ANZSIC FE code at the 4-digit level (see the bottom half of Table 3). There is a greater variability in the error rate across the waves for industry than for occupation.

The error rate at the 1-digit level show a similar pattern, but the overall level of error is less and is around 4 per cent for occupation and 3 to 4 per cent for industry.

From the sample of 500 cases with a one-to-one match between the old and new codeframes, it was found that:

- 61 cases were given a new 4-digit occupation code (12.2%); and
- 69 cases were given a new 4-digit industry code (13.8%).

It is not clear why there is more error in cases with a straight concordance between the old and new codeframes. It was expected that the higher error rate would have occurred in the occupations that were perhaps not well defined and thus were subject greater revision in the construction of the new codeframes.

Note that this analysis does not indicate that the recoded data is more or less accurate than the original data: it only indicates that it is different.

[5] There were up to 6 occupation questions and 4 industry questions each wave.

Table 3: Amount and quality of main job occupation and industry recoding, by wave

Wave	Automatically coded (1 to 1 match)	Manually coded (1-to-many or many-to-many match)				
		Coded within multiples allowed	Coded outside of multiples allowed	% manually coded	% error at 4 digit	% error at 1 digit
			Occupation			
1	4768	3208	538	44.0%	11.6%	4.6%
2	4570	3132	386	43.5%	10.6%	4.2%
3	4499	3104	388	43.7%	10.9%	4.3%
4	4345	3115	361	44.4%	10.0%	4.5%
5	4643	3397	208	43.7%	5.3%	2.0%
6	4697	3253	407	43.8%	10.7%	4.1%
			Industry			
1	4822	3173	499	43.2%	13.6%	4.7%
2	4640	3188	260	42.6%	7.5%	2.6%
3	4574	3080	336	42.8%	9.8%	3.6%
4	4399	3062	360	43.8%	10.5%	4.6%
5	4558	3288	354	44.4%	8.3%	3.1%
6	4607	3408	262	44.3%	6.1%	3.0%

Independent evaluation

To gain an understanding of the difficulties associated with coding and to obtain alternative estimates of the quality of the coding, several people in the HILDA Survey team reviewed the coding of a sample of cases from wave 6. Focusing on the occupation of the main job, we reviewed the coding for the following three groups:

- a random sample of 250 cases that had a one-to-one match between the ASCO and ANZSCO codeframe;
- a random sample of 222 cases that had a one-to-many match between the ASCO and ANZSCO codeframes which Nielsen coded *inside* the possible ANZSCO codes given the ASCO code; and
- all 407 cases that had a one-to-many match between ASCO and ANZSCO codeframes which Nielsen coded *outside* the possible ANZSCO codes given the ASCO code.

The ABS *ANZSCO Coder* was used to help identify the appropriate code. We found that there was, however, still a degree of subjectivity in the coding, particularly when the occupation title or task descriptions are not clear or not specific enough. The codes were reviewed several times to ensure the coding principles were applied correctly and that unjustified deviations from the Nielsen codes were not made.

Similar levels of error were identified in the sample where there was a one-to-one concordance between the codeframes and where Nielsen had coded within the possible codes (see Table 4). Nevertheless, the error rate was high at 14 per cent at the 4 digit level.

There is a much higher error rate amongst the codes where Nielsen decided that a code outside the possible multiple codes was preferable. Interestingly, the Melbourne Institute HILDA team coded 7 per cent of this sample into one of the allowable codes according to the concordance between the old and new codeframes.

Table 4: Accuracy of occupation coding, wave 6

	One-to-one concordance	One-to-many concordance and *within* possible codes	One-to-many concordance but *outside* of possible codes	Total
Number of (randomly selected) cases recoded	250	222	407	879
Number of all cases	4697	3253	407	8357
Proportion of sample coded	5.3%	6.8%	100%	10.5%
Error rate				
At 4 digit	14.4%	13.5%	31.0%	18.2%
At 2 digit	11.6%	9.9%	21.9%	11.4%
At 1 digit	7.6%	5.9%	14.7%	7.3%

We also considered whether the errors were systematic. Table 5 compares the distribution of the occupation codes at the 1 digit level between the codes assigned by Nielsen and by the Melbourne Institute HILDA Survey team. The estimates are constructed by apportioning the outcomes of the three parts of the sample checked to the total sample. The MI coding tends to identify fewer managers, clerical and administrative workers, and labourers than Nielsen's coding, though none of these differences are statistically significant.

An independent coding of a sample of cases to ASCO has not as yet been undertaken as this is a very time consuming task but it would permit an assessment of whether the ANZSCO or ASCO coding is of higher quality. As an ABS Coder is not available for ASCO, it is likely that the ASCO codes are more unreliable than the ANZSCO codes.

Table 5: Estimated differences in occupation codes at ANZSCO 1 digit level, wave 6

	Nielsen codes	MI codes	Difference: (MI – Nielsen)
1: Managers	15.0%	14.0%	-1.0%
2: Professionals	21.8%	22.5%	0.7%
3: Technicians and trade workers	10.2%	10.6%	0.4%
4: Community and personal service workers	9.4%	9.9%	0.5%
5: Clerical and administrative workers	15.1%	14.6%	-0.5%
6: Sales workers	9.6%	9.7%	0.1%
7: Machinery operators and drivers	5.4%	5.9%	0.5%
8: Labourers	13.5%	12.8%	-0.7%
Total	100.0%	100.0%	
Not codeable			
Number	0	89	
Per cent	0.0%	1.1%	

Quality of ABS's Coding for the Labour Force Survey

The ABS was contacted to establish an 'acceptable' level of error in occupation and industry coding. For the Labour Force Survey, the ABS use a combination of automatic coding and manual coding (the automatic coding system is not available outside of the ABS). Approximately 65 per cent of responses are coded via the ABS autocoding system. These cases are very straightforward and testing has shown the accuracy is over 99 per cent. The remaining 35 per cent are manually coded.

A sample of the cases that are manually coded are verified for quality assurance purposes. While the ABS tolerates an error rate of up to 15 per cent at the 1-digit level, it usually varies between 5 and 10 per cent.[6] Most of the error in the ABS coding occurs at the first digit (their testing shows that only 2 per cent of cases have an error at the 2 digit level once the first digit is correct).

This means that the overall error rate in the Labour Force Survey data is generally between 2.4 to 4.1 per cent but would not be higher than 5.9 per cent. Similar error rates occur for occupation and industry coding. There is, therefore, a reasonable level of subjectivity in a portion of the occupation and industry codes, even in the ABS data.

Impact on the Weights

The impact of the quality of the 1-digit codes is also of concern as they have been used in construction of the responding person weights from Release 4 to 6 (as a benchmark to which the weights are adjusted to meet).

The main reason for incorporating occupation into the production of the weights was because there appeared to be an over representation of managers, administrators and professionals in the HILDA sample (in the order of 3 percentage points more under the ASCO framework).

[6] Personal communication with FaHCSIA's ABS Outposted Officer, 27 October 2008.

These ASCO groups broadly relate to the managers and professionals under the ANZSCO framework and, in aggregate, the relative size of this group is largely unchanged by the accuracy of the coding (as shown in Table 5 above). Nevertheless, we have estimated that 7.3 per cent of cases may have been assigned an incorrect 1-digit code (see Table 4). This will influence the weights these respondents are assigned.

Given the level of uncertainty in the coding of occupation that this analysis has demonstrated (which is more than we expected), it is not appropriate to use such a variable in the benchmarking of the weights.[7] Benchmarks should be restricted only to variables that are known with a high degree of certainty, such as age, sex, geographic location, labour force status, marital status and household composition[8].

The responding person weights have been re-estimated without the occupation benchmark and an alternative benchmark based on household composition has been investigated. Tables 6 to 9 show the effect of modifying the benchmarks on the estimates for financial year income, occupation, and household composition.

Removing the occupation benchmark and adding the household composition benchmark will, on average, increase the total financial year income by 1.1 per cent, reduce the proportion of lone person households by 0.6 per cent and increase the proportion of households with three or more adults by 1.0 per cent. The proportion of managers, administrators and professionals will be higher in the HILDA Survey than in the Labour Force Survey by 2.7 percentage points. The addition of the household composition benchmark has a minimal effect on income and occupation estimates but does correct a bias in the composition of the household. This will improve the consistency in estimates between the enumerated person file and the responding person file.

[7] Especially if there are systematic differences in coding practices (which we are unable to ascertain from this analysis). If there is only variability in the coding, including it in the weighting benchmarks would reduce the efficiency of the estimates but would not introduce bias.

[8] For example, the number of adults and children in the household.

Table 6: Effect of changing the benchmarks on estimates of total financial year income

	W1	W2	W3	W4	W5	W6	W7
1. With occupation benchmarks (as per Release 6)							
Wages and salaries	20,592	21,023	21,769	22,748	24,314	26,220	28,550
Benefits	2,437	2,686	2,730	2,834	2,782	2,865	2,839
Business income	1,505	1,689	1,569	1,800	1,917	1,938	2,057
Investments	1,582	1,594	1,673	1,990	2,425	2,868	2,747
Other income	1,131	1,283	1,332	1,314	1,444	1,564	1,695
Total FY income	28,040	29,128	29,931	31,719	33,810	36,535	38,997
2. Without occupation benchmarks							
Wages and salaries	20,965	21,551	22,229	23,201	24,713	26,668	28,879
Benefits	2,428	2,664	2,718	2,823	2,766	2,851	2,834
Business income	1,527	1,707	1,581	1,800	1,908	1,953	2,034
Investments	1,596	1,636	1,711	2,025	2,400	2,883	2,767
Other income	1,131	1,279	1,332	1,309	1,444	1,566	1,706
Total FY income	28,434	29,683	30,419	32,179	34,150	36,988	39,316
3. Without occupation benchmarks but with household composition benchmarks (as per Release 7)							
Wages and salaries	20,955	21,487	22,147	23,119	24,654	26,610	28,834
Benefits	2,428	2,660	2,714	2,819	2,763	2,850	2,833
Business income	1,525	1,692	1,575	1,795	1,902	1,951	2,058
Investments	1,594	1,623	1,707	2,007	2,371	2,856	2,751
Other income	1,130	1,271	1,320	1,304	1,435	1,559	1,702
Total FY income	28,420	29,576	30,301	32,060	34,045	36,894	39,270
Comparison of total financial year income (amount and percent difference)							
2 – 1 ($)	394	555	488	460	340	453	319
2 – 1 (%)	1.4%	1.9%	1.6%	1.5%	1.0%	1.2%	0.8%
3 – 1 ($)	379	448	370	340	235	359	274
3 – 1 (%)	1.4%	1.5%	1.2%	1.1%	0.7%	1.0%	0.7%

Table 7: Effect of changing the benchmarks on occupation estimates, waves 1 – 6

	W1	W2	W3	W4	W5	W6
1. With occupation benchmarks (as per Release 6)						
Managers / Administrators	8.0	7.4	7.4	8.3	8.3	8.3
Professionals	18.7	19.0	19.0	18.9	19.3	19.4
Associate professionals	11.3	11.3	11.6	11.8	11.9	12.3
Trades and related	12.6	12.6	12.8	12.7	12.7	12.8
Adv. clerical/service	4.4	4.3	4.1	3.7	4.1	3.8
Int. clerical/sale/service	17.2	17.3	17.3	16.8	16.6	16.6
Int. production/transport	8.7	8.5	8.3	8.6	8.2	8.6
Elem. Clerical/sales/service	10.2	10.6	10.5	10.2	10.4	9.8
Labourers and related	8.9	9.3	9.1	9.0	8.6	8.5
2. Without occupation benchmarks						
Managers / Administrators	8.7	8.2	8.4	8.4	7.6	8.1
Professionals	21.5	21.3	21.3	21.3	22.3	22.4
Associate professionals	11.6	13.1	12.7	13.4	12.7	13.1
Trades and related	11.9	11.9	11.4	11.4	11.9	12.2
Adv. clerical/service	3.4	3.2	3.3	3.6	3.0	2.9
Int. clerical/sale/service	16.1	16.4	16.3	16.2	17.3	16.9
Int. production/transport	8.2	8.5	7.6	7.7	8.1	7.7
Elem. Clerical/sales/service	9.9	9.4	10.2	9.7	9.5	8.8
Labourers and related	8.5	8.1	8.9	8.2	7.5	8.0
3. Without occupation benchmarks but with household composition benchmarks (as per Release 7)						
Managers / Administrators	8.7	8.2	8.4	8.4	7.6	8.0
Professionals	21.5	21.1	21.2	21.2	22.2	22.3
Associate professionals	11.6	13.1	12.7	13.4	12.6	13.1
Trades and related	11.9	11.9	11.4	11.4	12.0	12.2
Adv. clerical/service	3.4	3.2	3.3	3.6	3.0	2.9
Int. clerical/sale/service	16.2	16.5	16.3	16.3	17.3	16.9
Int. production/transport	8.3	8.5	7.6	7.7	8.2	7.7
Elem. Clerical/sales/service	9.9	9.5	10.2	9.7	9.5	8.8
Labourers and related	8.5	8.1	9.0	8.3	7.6	8.0
Comparison of proportion of managers/administrators/professionals (percent differences)						
2 – 1 (%)	3.5	3.1	3.3	2.5	2.3	2.7
3 – 1 (%)	3.5	2.9	3.2	2.4	2.2	2.7

Note: Occupations provided for W1-6 are the ASCO 1-digit codes, whereas for W7 they are the ANZSCO 1- digit codes.

Table 8: Effect of changing the benchmarks on occupation estimates, wave 7

	1. With occupation benchmarks (as per Release 6)	*2. Without occupation benchmarks*	*3. Without occupation benchmarks but with household composition benchmarks (as per Release 7)*
Managers	12.9	12.2	12.2
Professionals	20.5	23.3	23.2
Technicians/trades	15.4	14.7	14.7
Community/personal service	8.5	10.1	10.0
Clerical/administrative	15.5	15.0	15.1
Sales	9.5	9.3	9.3
Machinery operators/drivers	7.0	6.1	6.1
Labourers	10.8	9.4	9.4

Note: The proportion of manages and professionals changes by 2.0% with the change in benchmarks from using the occupation benchmarks to either of the other two options.

Table 9: Effect of changing the benchmarks on household composition estimates

	W1	W2	W3	W4	W5	W6	W7
1. With occupation benchmarks (as per Release 6)							
Lone person	12.5	13.3	13.4	13.5	13.3	13.3	13.4
Lone parent	2.1	2.2	2.2	2.2	2.1	1.9	1.9
Couple no kids	32.0	32.5	31.8	31.6	31.9	32.3	32.0
Couple w kids	18.1	18.3	18.6	17.9	18.0	17.6	17.6
3+ adult no kids	23.8	22.7	22.9	23.3	23.6	23.0	23.7
3+adult w kids	11.6	11.0	11.1	11.5	11.2	11.9	11.4
2. Without occupation benchmarks							
Lone person	12.5	13.4	13.4	13.6	13.4	13.3	13.5
Lone parent	2.1	2.2	2.2	2.2	2.1	1.9	1.9
Couple no kids	32.0	32.5	31.9	31.7	32.0	32.4	32.1
Couple w kids	18.1	18.3	18.6	17.9	17.9	17.6	17.6
3+ adult no kids	23.7	22.6	22.8	23.2	23.5	23.0	23.6
3+adult w kids	11.6	11.0	11.1	11.4	11.1	11.8	11.4
3. Without occupation benchmarks but with household composition benchmarks (as per Release 7)							
Lone person	12.4	12.6	12.6	12.7	12.7	12.8	12.8
Lone parent	2.1	2.1	2.1	2.1	2.0	1.9	1.8
Couple no kids	31.9	31.8	31.4	31.2	31.4	32.1	31.9
Couple w kids	18.0	17.9	17.9	17.4	17.6	17.3	17.3
3+ adult no kids	23.9	24.0	24.3	24.6	24.7	23.8	24.5
3+adult w kids	11.7	11.6	11.7	12.0	11.6	12.2	11.7
Comparison of proportion (percent differences)							
2-1							
Lone person	0.1	0.1	0.1	0.1	0.1	0.1	0.1
Lone parent	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Couple no kids	0.1	0.1	0.1	0.1	0.1	0.1	0.1
Couple w kids	0.0	0.0	0.0	0.0	-0.1	0.0	0.0
3+ adult w or wo kids	-0.1	-0.2	-0.1	-0.1	-0.1	-0.1	-0.1
3-1							
Lone person	-0.1	-0.7	-0.8	-0.8	-0.6	-0.5	-0.6
Lone parent	0.0	-0.1	-0.1	-0.1	-0.1	0.0	-0.1
Couple no kids	-0.1	-0.7	-0.4	-0.4	-0.5	-0.2	-0.1
Couple w kids	0.0	-0.4	-0.7	-0.4	-0.4	-0.4	-0.3
3+ adult w or wo kids	0.1	1.3	1.4	1.2	1.2	0.8	0.8

Conclusion

It is not clear how to reconcile the different estimates of the error in the occupation and industry codes identified by the three approaches considered above. The errors identified by the verifiers have been corrected and are no longer in the dataset, so they tell us little about the error in the final datasets (other than there is a reasonable level of subjectivity in the coding). When Nielsen coders ventured outside the possible codes based on the concordance between the old and new codeframes, it is not known whether the original code or the revised code is correct. We expect the ANZSCO and ANZSIC SE codes are more likely to be correct as Nielsen have used the ABS Coders to help code the verbatim responses. The HILDA Survey team estimates an error rate of 15.4 per cent for main job occupation in wave 6.

Based on our evaluation, the error rate in the HILDA Survey is approximately double that of the ABS Labour Force Survey.

Given this uncertainty in both the ABS and HILDA estimates of occupation, it is not appropriate to include 1-digit occupation as a benchmark in the HILDA weights. While the inclusion of the household composition benchmark in the responding person weights has only a modest improvement on the income estimates, it is worth including as a benchmark from Release 7 for the greater consistency it will bring with the enumerated person weights.

Note that this paper has not dealt with the longitudinal consistency of occupation and industry – it has only considered the cross-sectional consistency. The coding is undertaken for each wave independently of other waves. We plan to release a future discussion paper on this topic.

Implications for Release 7

1. The ASCO and ANZSIC FE codes are retained in Release 7 together with ANZSCO and ANZSIC SE codes. Users are advised to use ANZSCO and ANZSIC SE in preference to ASCO and ANZSIC FE (as it is available for all waves and is likely to have a lower error rate).
2. From wave 7 onwards, only ANZSCO and ANZSIC SE codes will be available on the datasets.
3. 1-digit occupation has been removed from the benchmarks used in the weighting calculations for Release 7. A household structure benchmark has been incorporated.

Recommendations for further work

1. The HILDA Survey team will routinely assess the quality of the occupation and industry coding by independently coding a sample of cases.
2. The HILDA Survey team will recode a modest sample of cases to the ASCO codeframe to determine whether the ASCO or ANZSCO coding is less prone to error.
3. The HILDA Survey team will assess the wave-to-wave consistency of occupation and industry coding and circulate a discussion paper to users.

References

Australian Bureau of Statistics (ABS) (1997). Australian Standard Classification of Occupations (ASCO). Second Edition, 1997. (ABS Cat.No. 1220.0)

Australian Bureau of Statistics (ABS) (2006a). Australian and New Zealand Standard Classification of Occupations (ANZSCO), First Edition, 2006 (ABS Cat.No. 1220.0)

Australian Bureau of Statistics (ABS) (1993). Australian and New Zealand Standard Industrial Classification (ANZSIC), First Edition. 1993 (ABS Cat.No. 1292.0)

Australian Bureau of Statistics (ABS) (2006b).Australian and New Zealand Standard Industrial Classification (ANZSIC), Second Edition, 2006 (ABS Cat.No.1292.0)