



Australian Government

Commonwealth Grants Commission

DISCUSSION PAPER CGC 2003/11

WAGES INPUT COSTS – TECHNICAL UPDATE

SEPTEMBER 2003

INTRODUCTION

1. In this staff technical paper, we report results of further analyses undertaken on the econometric model underlying the wages input cost factors proposed in the *Discussion Paper CGC 2003/04 Wages Input Cost*. The paper also includes staff suggestions on a way of calculating the factors for the six assessment years (1997-98 to 2002-03) and on a mechanism for updating the factors in the future.

2. With the benefit of additional data made available by the ABS, we made some changes to the model to address many of the technical comments made by the States. We have:

- (i) used data on actual levels of wages and hours worked, rather than grouped data (we also introduced break points for very low and very high levels of hours worked to capture diminishing effects of hours on wages);
- (ii) introduced identifiers for small States;
- (iii) used alternative specifications;
- (iv) used weighted regressions with ABS specified weights for the observations;
- (v) introduced occupation at a detailed level to further reduce possible composition effects; and
- (vi) identified and reduced the effects of outliers.

3. We did the analyses using 1997 and 2001 data for the private sector. We report all results based on the revised data and analyses.

4. On the whole, the introduction of these changes made little difference to the broad conclusions or the broad patterns of relativities. The unbanding of hours worked had the most, and downward, impact on Queensland and the Northern Territory and, to a lesser extent, on South Australia. Outlier analysis impacted upwardly on Western Australia in particular.

5. Except as specified below, we retained the conceptual and the empirical premises of the original discussion paper. The location effect variables are defined with respect to the Australian average as was mentioned in the original discussion paper, page 41, paragraph 7.

6. We present the results in the following order to reflect:

- (i) fresh consideration of exclusions of observations and revised specifications;

- (ii) comparison of results based on:
 - a specification using the natural logarithm of earnings; and
 - a specification using the natural logarithm of earnings per hour;
- (iii) the final specification and results;
- (iv) analysis of outliers;
- (v) revised results and a set of possible factors;
- (vi) updateability; and
- (vii) some general information.

7. We ask the States to note that the Commission has indicated that there is in all likelihood little value in pursuing further refinements of the model and that available time of staff would be better spent in ‘plausibility checking’. We therefore ask the States to focus on providing information relating to pattern, growth and ranking of levels of relevant wages by State to facilitate comparison with the model.

8. We are happy to provide further details of results, if readily available, to the States, should they express an interest in having them.

Fresh consideration of exclusions and revised specification;

9. In this work, we have used:
- (i) actual rather than categorised data on levels of wages and hours worked; and
 - (ii) identifiers for small States.

10. We did a preliminary analysis to see if the new data showed any abnormality in wages at the very low end of hours worked (say, less than 15 hours per week) and the very high end of hours worked (say, more than 48 hours per week). It showed that the specification did not work effectively for individuals in the dataset who worked for very low wages (equal to or less than \$2 per week) or very low hours (equal to or less than 5 hours per week). We concluded that these observations were most likely for hobby-type activities and dropped them.

11. The analysis also showed that there was a broad diminishing relationship between earnings and hours worked (that is, wages \propto hours worked ^{β} , $\beta < 1$, other things being equal). Further this relationship was different at the extreme ends of hours worked — that is, β changed when hours worked was very low (say, less than 10 hours per week) or very high (say, more than 54 hours per week).

12. On these bases, we modified¹ our main specification to include terms specified in parenthesis below.

$$\ln(w_t) = X_{it} B_{ijt} + [\ln(\text{hour1}_{it}) * B_{\text{hour1}_t} + \ln(\text{hour}_{it}) * B_{\text{hour}_t} + \ln(\text{hour5}_{it}) * B_{\text{hour5}_t}] + \varepsilon_{jt}$$

Where hour = hours worked;

hour1 = hours worked if hour < 10;

Hour5 = hours worked if hour > 54; and

Bhour1_t, Bhour_t and Bhour5_t represented returns to the three variables.

13. Bhour1_t and Bhour5_t should be interpreted as differences from the normal return to hours worked, Bhour_t, for very low and very high hours worked.

Comparison of results between two alternative specifications

14. We retained the functional form of the main specification because it is conceptually sound, underlies many empirical studies on human capital theory and has some desirable statistical properties. However, in the light of comments made by some States, we considered two alternative specifications based on the natural log of earnings and the natural log of earnings per hour as the dependent variable. Not surprisingly, we obtained identical results.

15. The comparison of the estimated location effects are reproduced in Table 1.

Table 2 LOCATION EFFECTS^(a) (RELATIVE TO AUSTRALIAN AVERAGE) FOR TWO ALTERNATIVE SPECIFICATIONS, 1997 AND 2001, PRIVATE SECTOR

Location Effect	NSW	Vic	Qld	SA	WA	Tas	NT	ACT	Aus
1997									
natural log of earnings	0.017	0.012	-0.033	-0.014	-0.017	-0.060	0.104	0.038	0.000
natural log of earnings per hour	0.017	0.012	-0.033	-0.014	-0.017	-0.060	0.104	0.038	0.000
2001									
natural log of earnings	0.052	-0.004	-0.051	-0.042	-0.028	-0.074	0.053	0.017	0.000
natural log of earnings per hour	0.052	-0.004	-0.051	-0.042	-0.028	-0.074	0.053	0.017	0.000

(a) These estimates were based on unweighted regression and occupation variables at the 1-digit level.

¹ For original specification, see Discussion Paper CGC 2003/4 Wages Input Costs, Page 39.

Final specification and results

16. Because the two specifications gave the same results, we decided to retain the specification used in the discussion paper (that is, natural log of earnings as the dependent variable) for further analysis. In that further analysis, we:

- (i) used a weighted regression using ABS specified weights for individual observations; and
- (ii) introduced occupation variables at a finer 2-digit level.

17. The motivation for (i) was to allow each observation to exert as much weight as was determined by the ABS through the sampling process. This differs from unweighted regression where each observation has implicitly equal weights.

18. The motivation for (ii) was to reduce further plausible composition effects not captured through 1-digit occupation variables.

19. The comparison of the estimated location effects for 1997 and 2001, before the outlier analysis was undertaken, is reproduced in Table 3.

Table 3 LOCATION EFFECTS^(a) (RELATIVE TO AUSTRALIAN AVERAGE) IN 1997 AND 2001, PRIVATE SECTOR, PRE-OUTLIER ANALYSIS

Location Effect	NSW	Vic	Qld	SA	WA	Tas	NT	ACT	Aus
1997									
Location Effects	0.015	0.011	-0.029	-0.016	-0.017	-0.053	0.104	0.043	0.000
2001									
Location Effects	0.051	-0.005	-0.047	-0.043	-0.027	-0.075	0.038	0.016	0.000

(a) These estimates were based on weighted regression using ABS survey weights and occupation variables at 2-digit level.

Analysis of outliers

20. Outlier analysis was done to identify observations that were more influential, relative to other observations, in determining the location effects. The main purpose was to understand the nature of the influential observations and the overall impact they have on the location effects, as one way of checking robustness of results. With that in mind, we removed the outliers following the process below. However, we may not drop all of them for the final results.

21. To proceed, we plotted standardised dfbetas (relating to ‘location effect’ coefficients) for the observations where the absolute value of the standardised dfbetas was more than 0.05 and the corresponding leverage values, by State. We identified observations that were removed from the clusters and closely analysed them to see if there were clear reasons for treating them differently (that is, either reduce their weights or omit them — we decided to omit them from our analysis).

22. We present in Attachment A description of the process followed and characteristics of the influential observations and the outliers omitted.

23. The comparison of the estimated location effects for 1997 and 2001, after the outlier analysis was undertaken, is reproduced in Table 3. We present the detailed regression results in Attachment B and a description of the variables in Attachment C.

Table 4 FINAL ESTIMATED LOCATION EFFECTS^(a) (RELATIVE TO AUSTRALIAN AVERAGE), 1997 AND 2001, PRIVATE SECTOR, POST-OUTLIER ANALYSIS

Location Effect	NSW	Vic	Qld	SA	WA	Tas	NT	ACT	Aus
1997									
Location Effects	0.015	0.010	-0.027	-0.015	-0.016	-0.056	0.094	0.040	0.000
Standard error ^(b)	0.005	0.006	0.007	0.011	0.010	0.022	0.037	0.027	
t-value	3.32	1.75	-4.02	-1.35	-1.62	-2.59	2.57	1.46	
Level of Significance	0.001	0.080	0.000	0.177	0.104	0.010	0.010	0.145	
2001									
Location Effects	0.050	-0.005	-0.048	-0.046	-0.021	-0.079	0.043	0.019	0.000
Standard error	0.005	0.006	0.007	0.011	0.010	0.021	0.042	0.027	
t-value	11.03	-0.90	-7.04	-4.09	-2.18	-3.80	1.02	0.69	
Level of Significance	0.000	0.369	0.000	0.000	0.030	0.000	0.306	0.490	

(a) These estimates were based on weighted regression and occupation variables at 2-digit level, and excluded the outliers.

(b) The coefficient for Tasmania (chosen as reference) is estimated using the constraint that pop weighted location effect is zero. The standard error for Tasmania is calculated using the variance-covariance matrix of location effects. The results are however independent of any specific State chosen as reference.

Final results and possible factors

24. In the discussion paper, we used the location effects from the model for 2001 directly to build the factors, after adjusting them by reference to the wage cost indices to reflect the circumstances of the five assessment years (1998-99 to 2002-03). Since we were proposing to use the wage cost index for annual updating of the factor, it appeared a natural choice, despite concerns we had about whether the wage cost indices were good proxies for location effects.

25. Further reflection on the wage cost indices showed that they were unlikely to be good predictors of location effects. Therefore, we have considered other ways of calculating the factors. We also asked the States at the staff conference to suggest alternative ways of calculating the factors. The Commission may continue to consider a variety of ways to calculate the factors.

26. In general, it would have been nice to draw some general inferences from the model and use other data series to calculate the factors. However, the task is not straightforward or automatic, and can have its own pitfalls.

27. We present below one way — but which is by no means the final accepted version — of calculating the factors using the model results for 1997 and 2001 (based on a simple time-weighted straight line measure, for each State, between the two measures of location effects). The factors on this basis are calculated in Table 5².

Table 5 FACTORS BASED ON LOCATION EFFECTS, 2001 AND 1997, PRIVATE SECTOR

	NSW	Vic	Qld	SA	WA	Tas	NT	ACT	Aus
2001									
Location effect	0.05016	-0.00500	-0.04756	-0.04555	-0.02098	-0.07926	0.04289	0.01855	0.00010
1997									
Location effect	0.01528	0.00968	-0.02724	-0.01538	-0.01552	-0.05604	0.09412	0.03974	0.00004
2001, Adjusted (State – Australia)									
Location effect (A)	0.05006	-0.00509	-0.04766	-0.04565	-0.02108	-0.07936	0.04279	0.01845	0.00000
1997 Adjusted (State – Australia)									
Location effect (B)	0.01525	0.00964	-0.02728	-0.01542	-0.01556	-0.05608	0.09409	0.03970	0.00000
Calculated Location effect For Six Years using Time Weights (TW): (C = A *(1 - TW) + TW*B)									
1997-98 (TW = 0.00)	0.01525	0.00964	-0.02728	-0.01542	-0.01556	-0.05608	0.09409	0.03970	
1998-99 (TW = 0.25)	0.02395	0.00596	-0.03237	-0.02297	-0.01694	-0.06190	0.08126	0.03439	
1999-00 (TW = 0.50)	0.03265	0.00227	-0.03747	-0.03053	-0.01832	-0.06772	0.06844	0.02907	
2000-01 (TW = 0.75)	0.04136	-0.00141	-0.04256	-0.03809	-0.01970	-0.07354	0.05562	0.02376	
2001-02 (TW = 1.00)	0.05006	-0.00509	-0.04766	-0.04565	-0.02108	-0.07936	0.04279	0.01845	
2002-03 (TW = 1.00)	0.05006	-0.00509	-0.04766	-0.04565	-0.02108	-0.07936	0.04279	0.01845	

28. The raw factors are calculated by taking the exponential of the calculated location effects in Table 5. The final factors are calculated by rescaling the raw factors by mean resident population (MRP) to force the Australian average factor to equal one. Table 6 provides the raw and final factors.

² Figures described as adjusted in the table are produced by forcing the Australian average of the location effects derived from the enhanced data to equal zero.

Table 6 ADJUSTED RAW AND FINAL FACTORS

Year	NSW	Vic	Qld	SA	WA	Tas	NT	ACT	Aus
Calculated Raw Factors (D =EXP (C))									
1997-1998	1.01536	1.00969	0.97309	0.98470	0.98456	0.94546	1.09866	1.04050	1.00024
1998-1999	1.02424	1.00597	0.96815	0.97729	0.98320	0.93998	1.08466	1.03498	1.00033
1999-2000	1.03319	1.00227	0.96323	0.96993	0.98185	0.93452	1.07084	1.02950	1.00046
2000-2001	1.04222	0.99859	0.95833	0.96263	0.98050	0.92910	1.05719	1.02405	1.00065
2001-2002	1.05134	0.99492	0.95346	0.95538	0.97914	0.92371	1.04372	1.01862	1.00083
2002-2003	1.05134	0.99492	0.95346	0.95538	0.97914	0.92371	1.04372	1.01862	1.00083
Mean resident population (MRP) proportions									
1997	0.33914	0.24811	0.18375	0.07971	0.09715	0.02538	0.01014	0.01662	1.00000
1998	0.33919	0.24779	0.18446	0.07922	0.09758	0.02503	0.0102	0.01654	1.00000
1999	0.33917	0.24763	0.18524	0.0787	0.09778	0.02471	0.01025	0.01651	1.00000
2000	0.33925	0.24758	0.18612	0.07805	0.09783	0.0244	0.01026	0.0165	1.00000
2001	0.33884	0.24772	0.18728	0.07749	0.09787	0.02414	0.0102	0.01646	1.00000
2002 ^(a)	0.33884	0.24772	0.18728	0.07749	0.09787	0.02414	0.0102	0.01646	1.00000
Final factors									
1997-1998	1.01512	1.00944	0.97286	0.98447	0.98432	0.94524	1.09839	1.04025	1.00000
1998-1999	1.02390	1.00564	0.96783	0.97697	0.98288	0.93967	1.08430	1.03465	1.00000
1999-2000	1.03271	1.00181	0.96278	0.96948	0.98139	0.93409	1.07034	1.02902	1.00000
2000-2001	1.04155	0.99794	0.95771	0.96200	0.97986	0.92850	1.05651	1.02338	1.00000
2001-2002	1.05046	0.99409	0.95267	0.95458	0.97833	0.92294	1.04286	1.01777	1.00000
2002-2003 ^(a)	1.05046	0.99409	0.95267	0.95458	0.97833	0.92294	1.04286	1.01777	1.00000
Average, 6 yrs	1.03982	0.99872	0.95873	0.96352	0.98016	0.92963	1.05937	1.02452	1.00000
Final factors, Discussion Paper									
1998-1999	1.03692	0.98622	0.97993	0.96808	0.97471	0.92976	1.11369	1.01143	1.00000
1999-2000	1.03918	0.98748	0.97753	0.96567	0.97039	0.92672	1.10795	1.01676	1.00000
2000-2001	1.04267	0.98641	0.97393	0.96122	0.97287	0.91841	1.10211	1.02222	1.00000
2001-2002	1.04343	0.98819	0.97161	0.95978	0.97284	0.91488	1.09879	1.02130	1.00000
2002-2003	1.04211	0.98959	0.97000	0.96472	0.97420	0.91449	1.09444	1.01757	1.00000
Average, 5 yrs	1.04086	0.98758	0.97460	0.96389	0.97300	0.92085	1.10340	1.01786	1.00000

(a) The figure is the same as for 2001-02. It will be replaced with the figure for 2002 – 03 when available.

Updateability

29. This way of calculating factors would mean that we would not update the factors for post-2001 years until the 2007 Update. We would continue to use the 2001

factors for future years, subject to rescaling by MRP data, until the 2007 Update. For that update, we would rerun the model based on new data expected to be available from the next ABS survey of education and training in 2006.

Information

30. We reproduce the comparison of location effects between the private and the public sector for 1997 and 2001 below based on the enhanced data we are now using. These correspond to Figure 1 and Figure 2 on page 10 and 11 of the original discussion paper.

Figure 1 LOCATION EFFECTS AS PERCENTAGE OF AUSTRALIAN AVERAGE EARNINGS BY YEAR — PRIVATE SECTOR

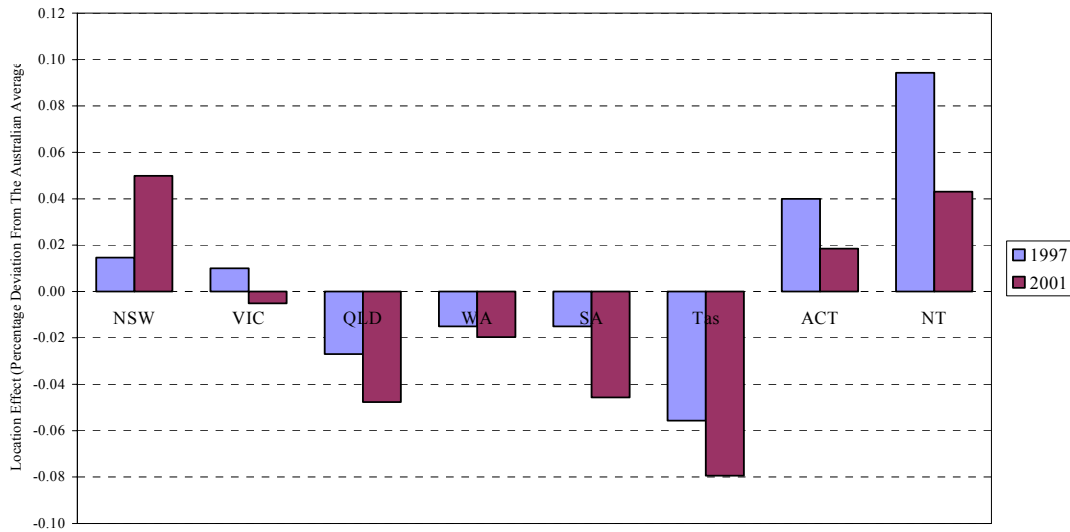
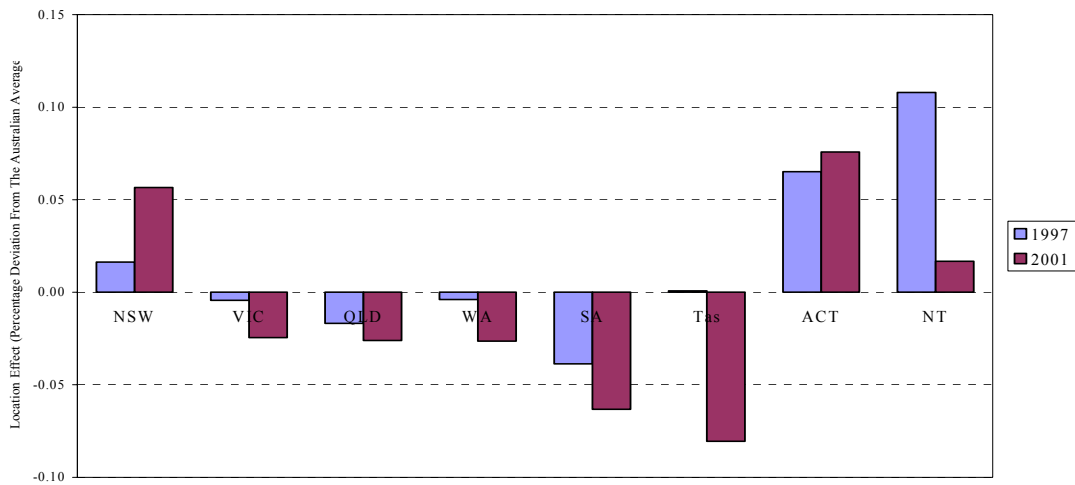


Figure 2 LOCATION EFFECTS AS PERCENTAGE OF AUSTRALIAN AVERAGE EARNINGS BY YEAR — PUBLIC SECTOR



31. We reproduce the comparison of dominant employer effects between for 1997 and 2001 below based on the enhanced data we are now using. These correspond to Figure 1 and Figure 2 on page 10 and 11 of the original discussion paper.

32. No outlier analysis was undertaken for the public sector observations because they were not of direct interest. Some care should be exercised in comparing the results particularly for the ACT because of the inclusion of the Commonwealth sector.

Figure 3 LOCATION EFFECTS AS PERCENTAGE OF AUSTRALIAN AVERAGE EARNINGS FOR DOMINANT AND NON-DOMINANT OCCUPATIONS — PUBLIC SECTOR, 1997

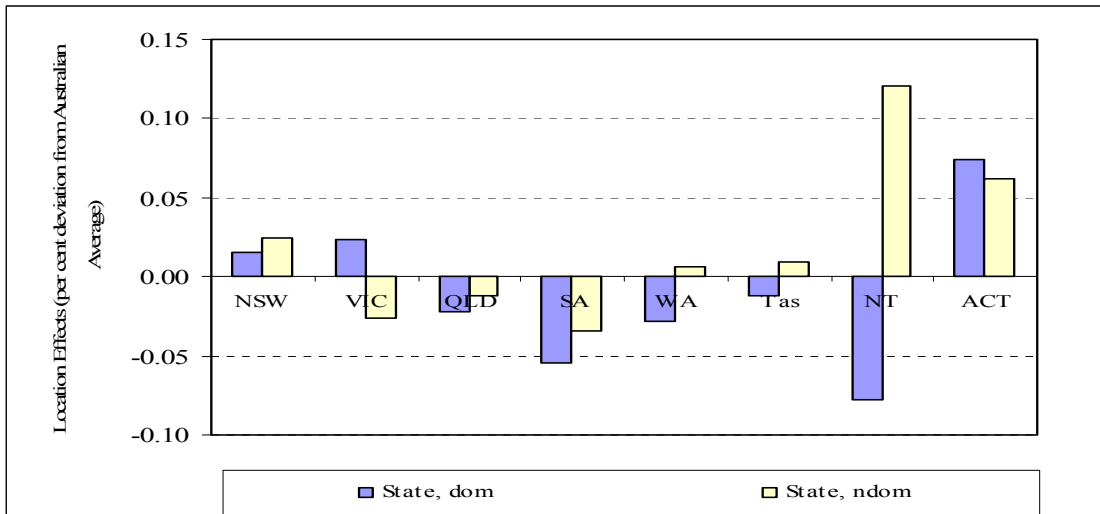
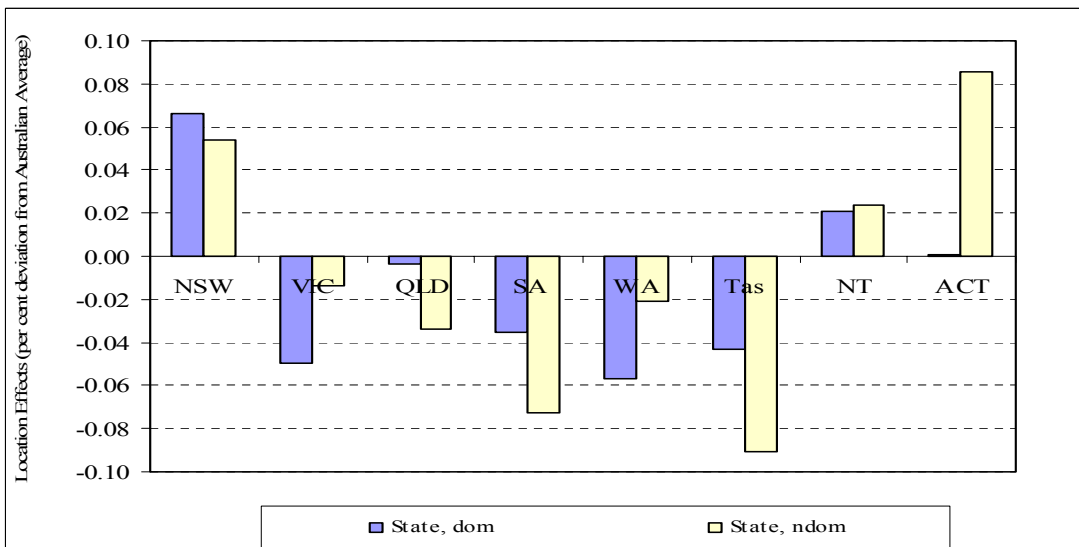


Figure 4 LOCATION EFFECTS AS PERCENTAGE OF AUSTRALIAN AVERAGE EARNINGS FOR DOMINANT AND NON-DOMINANT OCCUPATIONS — PUBLIC SECTOR, 2001

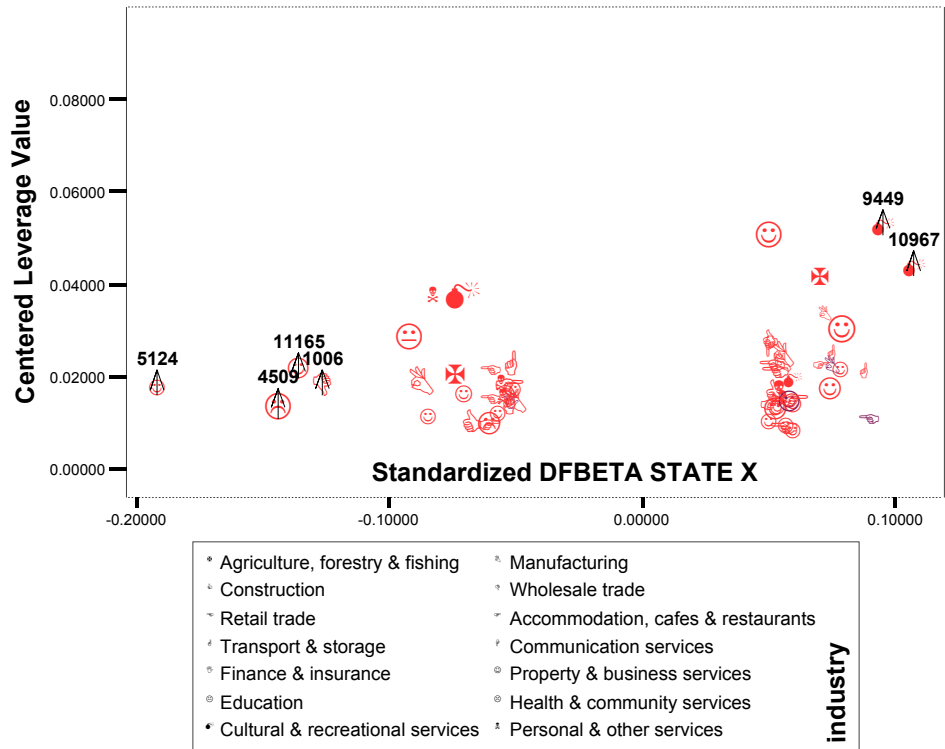


ATTACHMENT A ANALYSIS OF OUTLIERS³

1. Outliers are those observations that exert high degree of influence on the results relative to others because either there are errors in the data or they are ‘abnormal’.
2. The first step in this process was identifying the influential observations by calculating the standardised dfbetas for the location effects. We plotted the standardised dfbetas against the leverage values for each State, where the absolute value of the standardised dfbetas was more than 0.05.
3. We did so using the standard features of SPSS 11.5.
4. As an example, we reproduced a plot for a particular State for 2001.

³ References: *Regression Diagnostics, Identifying Influential Data and Sources of Collinearity* by David A. Belsley, Edwin Kuh and Roy E. Welsh, 1980, John Wiley and Sons.
SPSS Base Applications Guide 10.0.

Figure A-1 PLOT FOR STANDARDISED DFBETAS AGAINST THE LEVERAGE VALUES FOR EACH STATE



5. The numbered observations were isolated for further examination. We repeated this process for each State and year of interest (1997 and 2001).

6. We provide the distribution of the influential observations against a few key indicators — States, industries, occupation and level of earnings. The distribution by State is given in Table A-1.

Table A-1 DISTRIBUTION OF THE INFLUENTIAL OBSERVATIONS BY STATE

Year	NSW and VIC	QLD, SA and WA	Others	Total
1997				
Influential observations	10	9	3	22
Influential observations (per cent)	45	41	14	100
Samples (per cent)	44	45	11	100
2001				
Influential observations	11	16	7	34
Influential observations (per cent)	28	50	22	100
Samples (per cent)	32	47	21	100

7. The number of influential observations did not appear to depend strongly on the size of a State.

8. The distribution of influential observations by industry is given in Table A-2.

Table A-2 DISTRIBUTION OF THE INFLUENTIAL OBSERVATIONS BY STATE

Industry	1997			2001		
	Influential observations (number)	Influential observations (per cent)	Sample (per cent)	Influential observations (number)	Influential observations (per cent)	Sample (per cent)
Agriculture, forestry & fishing	1	5	3.2	1	3	3.2
Mining			2.0			1.6
Manufacturing	1	5	19.4	3	9	17.1
Electricity, gas & water supply			0.3			0.4
Construction	1	5	5.8	2	6	5.8
Wholesale trade	1	5	7.1	1	3	5.7
Retail trade	3	14	18.0	2	6	18.3
Accommodation, cafes & restaurants	2	9	7.0	1	3	7.1
Transport & storage	2	9	4.6	1	3	4.8
Communication services			0.6	3	9	1.2
Finance & insurance			4.8	5	15	4.7
Property & business services	2	9	10.5			12.7
Government administration & defence			0.1			0.1
Education	1	5	2.8			2.9
Health & community services	6	27	8.5	11	32	8.5
Cultural & recreational services	1	5	2.1	4	12	2.4
Personal & other services	1	5	3.3		3	3.3
Total	22	100	100	34	100	100

9. Generally, this suggested that there were potentially more influential observations in the finance and insurance, and health & community services sectors.

10. The distribution of influential observations by occupation is given in Table A-3.

Table A-3 DISTRIBUTION OF THE INFLUENTIAL OBSERVATIONS BY OCCUPATION

Occupation	1997			2001		
	Influential observations (number)	Influential observations (per cent)	Sample (per cent)	Influential observations (number)	Influential observations (per cent)	Sample (per cent)
Managers and administrators	1	5	5	2	4	5
Professionals	4	18	12	4	11	13
Associate professionals	2	9	9	6	15	10
Tradesperson & related	1	5	14	3	7	13
Advance clerical service workers			4	2	7	4
Intermediate clerical & service workers	1	5	20	2	13	19
Intermediate production & transport	2	9	12		2	11
Elementary clerical, sales & service	3	14	12	3	11	13
Labourers & related workers	8	36	13	12	30	12
Total	22	100	100	34	100	100

11. It showed that there were potentially more influential observations in the 'Labourers & related workers' occupation group and possibly in the 'professionals' category in 1997 and the 'Associate professionals' category in 2001.

12. The distribution of influential observations by levels of wages is given in Table A-4.

Table A-4 DISTRIBUTION OF THE INFLUENTIAL OBSERVATIONS BY LEVEL OF EARNINGS

Year	Low Earnings (< \$400)	Medium Earnings (> \$400 , < \$2000)	High Earnings (> \$2000)	Total
1997				
Influential observations (numbers)	14	6	2	22
Influential observations (per cent)	64	27	9	100
Samples (per cent)	39	60	1	100
2001				
Influential observations (numbers)	27	4	3	46
Influential observations (per cent)	79	12	9	100
Samples (per cent)	31	67	2	100

13. It showed that there were potentially more influential observations in the ‘lower earnings category’ (consistent with mote influential observations in the ‘Labourers & related workers’ occupation group) and that every influential observation in the high earning category should be closely examined.

Analysis of Outliers

14. We examined every influential observation in the dataset, particularly those that appear to correspond to the particular features listed above. We ultimately considered as outliers a total of 36 observations (23 in 2001 and 13 in 1997) out of previously identified 56 observations (34 in 2001 and 22 in 1997) as outliers.

15. The overriding reason for dropping most of them was that the wages per hour were either very low below \$6 per hour worked (21 in 2001 and 10 in 1997, mostly in the occupation group ‘labourers & related workers’) or very high above \$70 per hour worked (2 in 2001 and 1 in 1997).

16. The comparison of the influential observations before and after dropping the outliers is given in the following Tables. The distribution of influential observations by State is given in Table A-5.

Table A-5 DISTRIBUTION OF THE OUTLIERS BY STATE

Year	NSW and VIC	QLD, SA and WA	Others	Total
1997				
Influential observations	10	9	3	22
Dropped	4	6	3	13
Retained	6	3	0	9
2001				
Influential observations	11	16	7	34
Dropped	10	11	2	23
Retained	1	5	5	11

17. The distribution of outliers by industry is given in Table A-6.

Table A-6 DISTRIBUTION OF THE OUTLIERS BY INDUSTRY

Industry	1997 (number)			2001 (number)		
	Influential observations	Dropped	Retained	Influential observations (number)	Dropped	Retained
Agriculture, forestry & fishing	1	1	0	1		1
Mining						
Manufacturing	1		1	3	3	0
Electricity, gas & water supply						
Construction	1		1	2		2
Wholesale trade	1		1	1	1	0
Retail trade	3	3	0	2	1	1
Accommodation, cafes & restaurants	2	2	0	1	1	0
Transport & storage	2		2			
Communication services				1	1	0
Finance & insurance				3		3
Property & business services	2	1	1	5	3	2
Government administration & defence						
Education	1		1			
Health & community services	6	5	1	11	10	1
Cultural & recreational services	1		1	4	3	1
Personal & other services	1	1	0			
Total	22	13	9	34	23	11

18. Generally, this suggested that there were potentially more influential observations in the finance and insurance, and health & community services sector.

19. The distribution of influential observations against occupation is given in Table A-7.

Table A-7 DISTRIBUTION OF THE INFLUENTIAL OBSERVATIONS BY OCCUPATION

Occupation	1997			2001		
	Influential observations (number)	Dropped	Retained	Influential observations (number)	Dropped	Retained
Managers and administrators	1		1	2	1	1
Professionals	4	2	2	4	2	2
Associate professionals	2	2	0	6	3	3
Tradesperson & related	1		1	3	2	1
Advance clerical service workers				2	2	0
Intermediate clerical & service workers	1	1	0	2	1	1
Intermediate production & transport	2		2			
Elementary clerical, sales & service	3	1	2	3	2	1
Labourers & related workers	8	7	1	12	10	1
Total	22	13	9	34	23	11

20. It shows that there were potentially more influential observations in the 'Labourers & related workers' occupation group and possibly in the 'professionals' category in 1997 and the 'Associate professionals' category in 2001.

21. The distribution of influential observations by levels of wages is given in Table A-8.

Table A-8 DISTRIBUTION OF THE INFLUENTIAL OBSERVATIONS BY LEVEL OF EARNINGS

Year	Low Earnings (< \$400)	Medium Earnings (> \$400, < \$2000)	High Earnings (> \$2000)	Total
1997				
Influential observations (numbers)	14	6	2	22
Dropped	10	2	1	13
Retained	4	4	1	9
2001				
Influential observations (numbers)	27	4	3	34
Dropped	21		2	23
Retained	6	4	1	11

ATTACHMENT B REGRESSION RESULTS

Table B - 1 PRIVATE SECTOR, 1997

Variables	Coefficients	Standard Error	t-value	Level Of Significance
(Constant)	2.096	0.079	26.637	0.000
NSW	0.015	0.005	3.318	0.001
VIC	0.010	0.006	1.750	0.080
QLD	-0.027	0.007	-4.024	0.000
SA	-0.015	0.011	-1.350	0.177
WA	-0.016	0.010	-1.624	0.104
NT	0.094	0.037	2.575	0.010
ACT	0.040	0.027	1.456	0.145
OCC11	0.451	0.041	10.979	0.000
OCC12	0.410	0.027	14.950	0.000
OCC13	-0.276	0.064	-4.328	0.000
OCC21	0.271	0.035	7.720	0.000
OCC22	0.272	0.029	9.238	0.000
OCC23	0.492	0.084	5.858	0.000
OCC24	0.181	0.063	2.879	0.004
OCC25	0.147	0.037	3.994	0.000
OCC31	0.149	0.033	4.524	0.000
OCC32	0.262	0.028	9.356	0.000
OCC33	0.119	0.028	4.254	0.000
OCC34	0.074	0.053	1.417	0.157
OCC39	0.136	0.060	2.263	0.024
OCC41	0.056	0.028	1.986	0.047
OCC42	0.031	0.033	0.953	0.341
OCC43	0.075	0.032	2.344	0.019
OCC44	0.061	0.034	1.803	0.071
OCC45	-0.020	0.033	-0.620	0.535
OCC46	0.089	0.042	2.130	0.033
OCC49	0.018	0.032	0.576	0.565
OCC51	0.158	0.030	5.272	0.000
OCC59	0.043	0.049	0.876	0.381

Table B-1 PRIVATE SECTOR, 1997 (Continued...)

Variables	Coefficients	Standard Error	t-value	Level Of Significance
OCC62	0.130	0.030	4.342	0.000
OCC63	-0.001	0.027	-0.055	0.956
OCC71	-0.003	0.027	-0.095	0.924
OCC72	0.039	0.032	1.191	0.234
OCC73	-0.043	0.029	-1.480	0.139
OCC79	0.019	0.027	0.727	0.467
OCC81	-0.106	0.049	-2.180	0.029
OCC82	0.030	0.025	1.213	0.225
OCC83	-0.124	0.037	-3.392	0.001
OCC91	-0.064	0.030	-2.134	0.033
OCC92	-0.047	0.027	-1.762	0.078
OCC99	-0.046	0.026	-1.757	0.079
EDU1	0.215	0.039	5.525	0.000
EDU2	0.141	0.039	3.645	0.000
EDU3	0.196	0.028	7.061	0.000
EDU4	0.074	0.029	2.599	0.009
EDU5	0.091	0.029	3.190	0.001
EDU6	0.057	0.022	2.552	0.011
EDU7	0.072	0.011	6.701	0.000
EDU8	0.191	0.035	5.451	0.000
EDU9	0.082	0.013	6.158	0.000
INDU1	0.172	0.029	5.939	0.000
INDU2	0.521	0.032	16.066	0.000
INDU3	0.204	0.023	9.067	0.000
INDU4	0.368	0.058	6.382	0.000
INDU5	0.289	0.025	11.365	0.000
INDU6	0.225	0.023	9.749	0.000
INDU7	0.127	0.021	5.930	0.000
INDU8	0.168	0.023	7.210	0.000
INDU9	0.264	0.025	10.426	0.000
INDU10	0.339	0.052	6.530	0.000
INDU11	0.392	0.032	12.410	0.000
INDU12	0.313	0.025	12.562	0.000
INDU13	0.226	0.146	1.552	0.121

Table B-1 PRIVATE SECTOR, 1997 (Continued...)

Variables	Coefficients	Standard Error	t-value	Level of Significance
INDU14	0.083	0.035	2.411	0.016
INDU15	0.101	0.038	2.642	0.008
INDU16	0.262	0.038	6.844	0.000
FIRM1	-0.089	0.023	-3.918	0.000
FIRM2	0.051	0.022	2.293	0.022
FIRM3	0.131	0.022	5.847	0.000
WEXP	0.031	0.001	26.321	0.000
WEXPSQ	-0.001	0.000	-22.750	0.000
LHRWORK	0.936	0.019	49.259	0.000
LHRWORK1	-0.048	0.018	-2.684	0.007
LHRWORK5	-0.037	0.003	-11.114	0.000
MIG1	0.156	0.017	9.325	0.000
MIG2	0.182	0.025	7.227	0.000
MIG3	0.189	0.026	7.338	0.000
MIG4	0.185	0.023	8.180	0.000
MIG5	0.090	0.020	4.447	0.000
MIG6	0.091	0.023	3.969	0.000
TMPE1	-0.113	0.019	-5.911	0.000
TMPE2	-0.073	0.018	-4.059	0.000
TMPE3	-0.031	0.018	-1.677	0.094
TMPE4	-0.010	0.019	-0.519	0.604
PERM	-0.016	0.013	-1.264	0.206
MARRIED	0.050	0.008	5.923	0.000
CHILD	-0.003	0.011	-0.248	0.804
FEDU1	0.084	0.031	2.673	0.008
FEDU2	0.134	0.067	1.995	0.046
FEDU3	0.104	0.054	1.931	0.054
FEDU4	0.011	0.028	0.383	0.702
FEDU5	0.098	0.033	3.014	0.003
FEDU6	0.076	0.024	3.144	0.002
FEDU7	0.037	0.031	1.190	0.234
FEDU8	-0.005	0.052	-0.090	0.928
YESUNION	0.048	0.011	4.465	0.000

Table B-1 PRIVATE SECTOR, 1997 (Continued...)

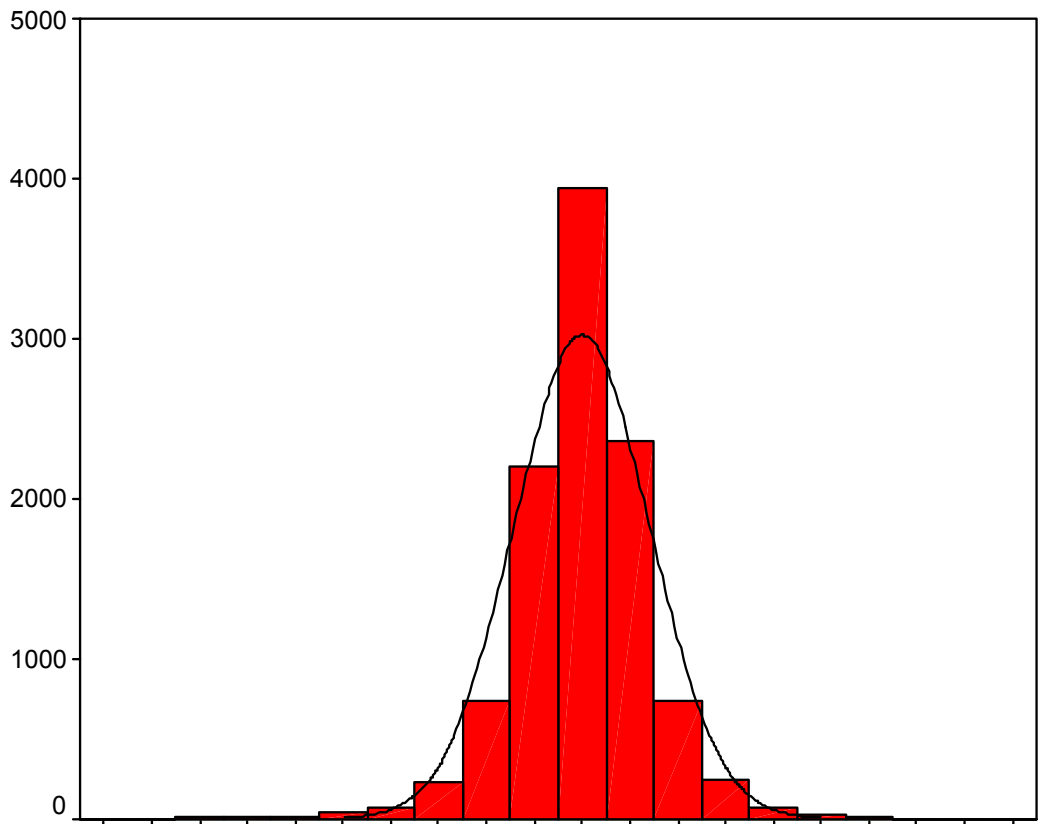
Variables	Coefficients	Standard Error	t-value	Level of Significance
FEMALE	0.101	0.083	1.214	0.225
OCC13F	0.580	0.179	3.237	0.001
OCC22F	0.063	0.035	1.812	0.070
OCC23F	-0.144	0.089	-1.615	0.106
OCC24F	0.209	0.063	3.345	0.001
OCC25F	0.073	0.044	1.658	0.097
OCC44F	-0.706	0.369	-1.911	0.056
OCC49F	0.108	0.045	2.411	0.016
OCC59F	0.196	0.057	3.426	0.001
OCC61F	0.081	0.026	3.149	0.002
EDU8F	-0.139	0.050	-2.794	0.005
EDU9F	-0.037	0.018	-2.032	0.042
INDU3F	-0.025	0.020	-1.266	0.206
INDU10F	-0.225	0.083	-2.731	0.006
INDU11F	-0.191	0.033	-5.790	0.000
INDU12F	-0.101	0.023	-4.337	0.000
INDU15F	0.063	0.037	1.691	0.091
INDU16F	-0.085	0.046	-1.836	0.066
FIRM1F	0.089	0.014	6.400	0.000
WEXPF	-0.004	0.001	-6.907	0.000
LHOURF	-0.068	0.023	-3.014	0.003
LHOURF1	0.045	0.021	2.205	0.028
MIG1F	-0.038	0.017	-2.248	0.025
MIG2F	-0.083	0.033	-2.522	0.012
PERMF	0.070	0.017	3.984	0.000
CHILDF	-0.040	0.014	-2.785	0.005
FEDU1F	-0.081	0.030	-2.676	0.007
FEDU2F	-0.110	0.071	-1.555	0.120
FEDU3F	-0.179	0.056	-3.167	0.002
FEDU8F	-0.288	0.172	-1.673	0.094
UNIONF	0.049	0.016	3.012	0.003
Adj. R ²	78.5%			

Note: Exercise caution in interpreting R² for weighted regression. Excluded variable OCC61.

Code used for regression

```
Regression variables = lnwage  
...List of variables ...  
/regwgt = weight  
/descriptives = mean  
/statistics = defaults cha Bcov  
/dependent = lnwage  
/method = enter nswmtasa to UNIONF  
/RESIDUALS  
/SAVE RESID(RES975).  
execute.  
  
COMPUTE RESID975 = SQRT(weight) * RES975.  
Execute.  
  
COMPUTE RES015SQ = RES015 * RES015.  
Execute.
```

Figure B-1 PLOT OF RESIDUAL (RESID975), 1997



Examination of Residuals (treat this as experimental)

Table B-2 DISTRIBUTION OF SSE BY STATE

State	SSE	Sum Of Weights	ESTIMAED VARIANCE	NUMBER of Samples
NSW	1.774E+09	1.63E+10	0.11	2386
VIC	1.489E+09	1.355E+10	0.11	2337
QLD	1.016E+09	9.508E+09	0.11	2114
SA	347944249	3.685E+09	0.09	1137
WA	555261147	5.379E+09	0.10	1563
Tasmania	148052062	1.068E+09	0.14	620
NT	33027449	381475511	0.09	237
ACT	63409484	677000750	0.09	385

More formally, weighted regression of RES975SQ on the location effects produced the following results. The influence of variance was not significant for any of the location effects.

```

Regression variables = RES975SQ
nswmtasa vicmtasa qldmtasa samtasa wantasa ntmmtasa actmtasa
/regwgt = weight
/statistics = defaults bcov
/dependent = RES975SQ
/method = enter nswmtasa to actmtasa
/RESIDUALS.
execute.

```

Table B-3 RESULT OF REGRESSION OF RESIDUALS ON LOCATION EFFECT, 1997

Location Effect	NSW	Vic	Qld	SA	WA	Tas	NT	ACT	Aus
Coeff of Location Effects	0.002	0.002	0.000	-0.013	-0.004	0.031	-0.021	-0.014	0.002
Standard error ^(b)	0.003	0.004	0.005	0.009	0.007	0.017	0.028	0.021	0.003
t-value	0.43	0.59	-0.08	-1.47	-0.57	1.87	-0.73	-0.65	0.43

Table B - 4 PRIVATE SECTOR, 2001

Variables	Coefficients	Standard Error	t-value	Level Of Significance
(Constant)	2.783	0.084	33.124	0.000
NSW	0.050	0.005	11.028	0.000
VIC	-0.005	0.006	-0.898	0.369
QLD	-0.048	0.007	-7.037	0.000
SA	-0.046	0.011	-4.094	0.000
WA	-0.021	0.010	-2.175	0.030
NT	0.043	0.042	1.023	0.306
ACT	0.019	0.027	0.690	0.490
OCC11	0.549	0.043	12.795	0.000
OCC12	0.430	0.028	15.505	0.000
OCC13	-0.161	0.070	-2.302	0.021
OCC21	0.208	0.036	5.711	0.000
OCC22	0.340	0.026	12.957	0.000
OCC23	0.556	0.083	6.723	0.000
OCC24	0.224	0.044	5.122	0.000
OCC25	0.160	0.031	5.155	0.000
OCC31	0.119	0.033	3.651	0.000
OCC32	0.271	0.033	8.233	0.000
OCC33	0.107	0.027	3.994	0.000
OCC34	0.041	0.059	0.699	0.485
OCC39	0.158	0.055	2.890	0.004
OCC41	0.051	0.029	1.732	0.083
OCC42	-0.028	0.032	-0.891	0.373
OCC43	0.017	0.034	0.513	0.608
OCC44	0.002	0.034	0.052	0.959
OCC45	-0.125	0.034	-3.715	0.000
OCC46	0.014	0.041	0.351	0.726
OCC49	-0.011	0.030	-0.359	0.719
OCC51	0.093	0.032	2.902	0.004
OCC59	0.123	0.033	3.712	0.000

Table B-4 PRIVATE SECTOR, 2001 (Continued...)

Variables	Coefficients	Standard Error	t-value	Level Of Significance
OCC61	0.015	0.024	0.613	0.540
OCC62	0.098	0.030	3.229	0.001
OCC63	-0.032	0.026	-1.209	0.227
OCC71	0.000	0.029	-0.004	0.997
OCC72	-0.081	0.035	-2.288	0.022
OCC73	-0.109	0.029	-3.820	0.000
OCC79	0.014	0.028	0.485	0.627
OCC81	-0.157	0.048	-3.287	0.001
OCC83	-0.131	0.035	-3.775	0.000
OCC91	-0.067	0.040	-1.687	0.092
OCC92	-0.067	0.031	-2.178	0.029
OCC99	-0.073	0.025	-2.846	0.004
EDU1	0.433	0.033	12.960	0.000
EDU2	0.312	0.030	10.456	0.000
EDU3	0.290	0.024	12.250	0.000
EDU4	0.186	0.024	7.761	0.000
EDU5	0.173	0.025	6.942	0.000
EDU6	0.114	0.019	5.839	0.000
EDU7	0.061	0.012	5.089	0.000
EDU8	0.138	0.032	4.340	0.000
EDU9	0.079	0.011	7.525	0.000
INDU1	0.072	0.029	2.518	0.012
INDU2	0.507	0.036	13.913	0.000
INDU3	0.107	0.021	5.045	0.000
INDU4	0.219	0.058	3.773	0.000
INDU5	0.262	0.026	9.952	0.000
INDU6	0.114	0.024	4.769	0.000
INDU7	0.029	0.022	1.362	0.173
INDU8	0.051	0.023	2.201	0.028
INDU9	0.197	0.025	7.943	0.000
INDU10	0.232	0.034	6.880	0.000
INDU11	0.256	0.031	8.292	0.000
INDU12	0.199	0.024	8.369	0.000
INDU13	0.177	0.114	1.560	0.119

Table B-4 PRIVATE SECTOR, 2001 (Continued...)

Variables	Coefficients	Standard Error	t-value	Level of Significance
INDU14	-0.010	0.032	-0.297	0.767
INDU15	-0.090	0.039	-2.306	0.021
INDU16	0.097	0.028	3.458	0.001
FIRM1	-0.016	0.023	-0.702	0.482
FIRM2	0.073	0.024	2.981	0.003
FIRM3	0.201	0.025	8.082	0.000
WEXP	0.028	0.001	24.139	0.000
WEXPSQ	-0.001	0.000	-20.812	0.000
LHRWORK	0.828	0.021	40.005	0.000
LHRWORK1	-0.054	0.013	-4.249	0.000
LHRWORK5	-0.018	0.004	-4.701	0.000
MIG1	0.084	0.017	5.024	0.000
MIG2	0.119	0.022	5.357	0.000
MIG3	0.109	0.024	4.464	0.000
MIG4	0.124	0.025	4.960	0.000
MIG5	0.037	0.022	1.709	0.087
MIG6	-0.008	0.022	-0.351	0.725
TMPE1	-0.127	0.019	-6.734	0.000
TMPE2	-0.076	0.018	-4.223	0.000
TMPE3	-0.051	0.018	-2.798	0.005
TMPE4	-0.009	0.019	-0.491	0.623
PERM	-0.009	0.013	-0.733	0.463
MARRIED	0.081	0.011	7.219	0.000
CHILD	-0.012	0.009	-1.403	0.161
FEDU1	-0.027	0.023	-1.205	0.228
FEDU2	-0.036	0.051	-0.714	0.475
FEDU3	-0.184	0.042	-4.363	0.000
FEDU4	-0.081	0.024	-3.387	0.001
FEDU5	-0.053	0.027	-1.960	0.050
FEDU6	0.023	0.021	1.094	0.274
FEDU7	-0.001	0.028	-0.023	0.981
FEDU8	-0.049	0.040	-1.240	0.215
YESUNION	0.078	0.011	6.775	0.000

Table B-4 PRIVATE SECTOR, 2001 (Continued...)

Variables	Coefficients	Standard Error	t-value	Level of Significance
FEMALE	-0.371	0.098	-3.794	0.000
OCC23F	-0.339	0.087	-3.892	0.000
OCC32F	-0.129	0.035	-3.636	0.000
OCC79F	-0.082	0.042	-1.935	0.053
OCC82F	-0.040	0.025	-1.571	0.116
OCC91F	-0.081	0.045	-1.822	0.068
OCC92F	-0.068	0.036	-1.914	0.056
INDU2F	-0.303	0.073	-4.128	0.000
INDU5F	-0.125	0.046	-2.736	0.006
INDU11F	-0.050	0.031	-1.598	0.110
INDU12F	-0.033	0.021	-1.588	0.112
INDU15F	0.156	0.038	4.130	0.000
FIRM2F	-0.037	0.016	-2.367	0.018
FIRM3F	-0.094	0.017	-5.459	0.000
WEXPF	-0.002	0.001	-3.721	0.000
LHOURF	0.070	0.026	2.637	0.008
LHOURF1	0.052	0.015	3.388	0.001
LHOURF5	-0.022	0.008	-2.655	0.008
PERMF	0.035	0.017	2.047	0.041
FEDU2F	0.035	0.054	0.647	0.518
FEDU3F	0.126	0.045	2.790	0.005
FEDU6F	-0.094	0.047	-1.992	0.046
MARRIEDF	-0.057	0.015	-3.833	0.000
UNIONF	0.072	0.017	4.209	0.000
Adj. R ²	79.1%			

Note: Exercise caution in interpreting R² for weighted regression. Excluded variable OCC82.

Code used for regression

```

Regression variables = lnwage
...List of variables ...
/regwgt = weight
/descriptives = mean
/statistics = defaults cha Bcov
/dependent = lnwage
/method = enter nswmtasa to UNIONF
/RESIDUALS
/SAVE RESID(RES015).

```

execute.

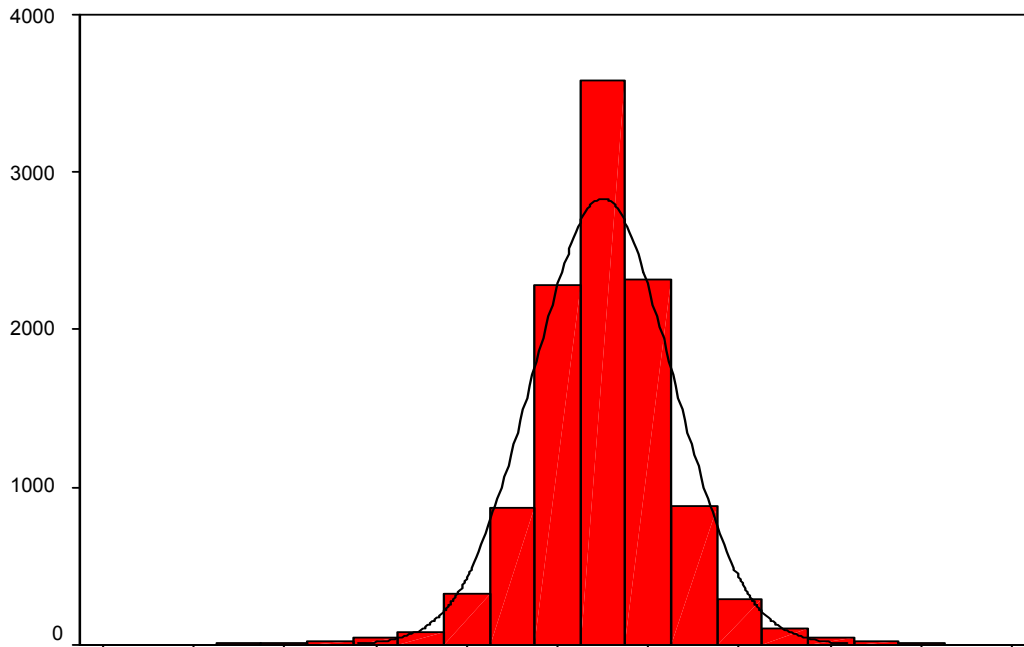
```
COMPUTE RESID015 = SQRT(weight) * RES015.
```

Execute.

```
GRAPH
```

```
/HISTOGRAM(NORMAL)=resid975 .
```

Figure B-2 PLOT OF RESIDUAL (RESID015), 2001



Examination of Residuals (treat this as experimental)

Table B-5 DISTRIBUTION OF SSE BY STATE

State	SSE	Sum Of Weights	Estimated Variance	Number of Samples
NSW	222828	1991440	0.11	2466
VIC	163272	1509804	0.11	2332
QLD	115701	1104492	0.10	2010
SA	45830	452407	0.10	1341
WA	65803	606498	0.11	1555
Tasmania	17733	134935	0.13	663
NT	2179	33582	0.06	134
ACT	8664	81915	0.11	413

More formally, weighted regression of RES015SQ on the location effects produced the following results. The influence of variance was not significant for any of the location effect variables.

```
Regression variables = RES015SQ
nswmtasa vicmtasa qldmtasa samtasa wantasa ntmmtasa actmtasa
/regwgt = fnlwt
/statistics = defaults bcov
/dependent = RES015SQ
/method = enter nswmtasa to actmtasa
/RESIDUALS.
execute.
```

Table B- 6 REGRESSION RESULT OF RESIDUALS ON LOCATION EFFECT, 2001

Location Effect	NSW	Vic	Qld	SA	WA	Tas	NT	ACT	Aus
Coefficient of Location Effects	0.004	0.000	-0.004	-0.007	0.000	0.023	-0.043	-0.003	0.000
Standard error	0.003	0.004	0.005	0.008	0.007	0.016	0.031	0.020	
t-value	1.05	-0.06	-0.73	-0.86	0.02	1.48	-1.39	-0.13	

ATTACHMENT C ABBREVIATIONS FOR VARIABLES USED

Table C-1 DESCRIPTION OF VARIABLES AND VARIABLE ABBREVIATIONS

Variable	Abbreviation
State of residence	
New South Wales	NSW
Victoria	VIC
Queensland	QLD
South Australia	SA
Western Australia	WA
*Tasmania, ACT and the North Territory	
Migrant status	
Born in Australia	MIG1
Born in English-speaking countries, lived in Australia more than 20 years	MIG2
Born in English-speaking countries, lived in Australia more between 10-20 years	MIG3
Born in English-speaking countries, lived in Australia less than 10 years	MIG4
Born in other countries, lived in Australia more than 20 years	MIG5
Born in other countries lived in Australia between 10-20 years	MIG6
*Born in other countries, lived in Australia less than 10 years	
Occupation	
Managers and administrators	OCC1
Professors	OCC2
Associate professors	OCC3
Tradespersons and related workers	OCC4
Advanced clerical and service workers	OCC5
Intermediate clerical, sales and service workers	OCC6
Intermediate production and transport workers	OCC7
Elementary clerical, sales and service workers	OCC8
*Labourers and related workers	

Table C-1 DESCRIPTION OF VARIABLES AND VARIABLE ABBREVIATIONS
(continued)

Variable	Abbreviation
Detailed Occupation	
Generalist Managers	OCC11
Specialist Managers	OCC12
Farmers and Farm Managers	OCC13
Science, Building and Engineering Professionals	OCC21
Business and Information Professionals	OCC22
Health Professionals	OCC23
Education Professionals	OCC24
Social, Arts and Miscellaneous Professionals	OCC25
Science, Engineering and Related Associate Professionals	OCC31
Business and Administration Associate Professionals	OCC32
Managing Supervisors (Sales and Service)	OCC33
Health and Welfare Associate Professionals	OCC34
Other Associate Professionals	OCC39
Mechanical and Fabrication Engineering Tradespersons	OCC41
Automotive Tradespersons	OCC42
Electrical and Electronics Tradespersons	OCC43
Construction Tradespersons	OCC44
Food Tradespersons	OCC45
Skilled Agricultural and Horticultural Workers	OCC46
Other Tradespersons and Related Workers	OCC49
Secretaries and Personal Assistants	OCC51
Other Advanced Clerical and Service Workers	OCC59
Intermediate Clerical Workers	OCC61
Intermediate Sales and Related Workers	OCC62
Intermediate Service Workers	OCC63
Intermediate Plant Operators	OCC71
Intermediate Machine Operators	OCC72
Road and Rail Transport Drivers	OCC73
Other Intermediate Production and Transport Workers	OCC79
Elementary Clerks	OCC81
Elementary Sales Workers	OCC82
Elementary Service Workers	OCC83
Cleaners	OCC91

Table C-1 DESCRIPTION OF VARIABLES AND VARIABLE ABBREVIATIONS
(continued)

Variable	Abbreviation
Factory Labourers	OCC92
Other Labourers and Related Workers.	OCC99
Marital status	
Married	MAR
*Not married	
Whether had any young children	
With children under 15 years old	CHILD
*Without children under 15 year old	
Whether permanent or casual	
Permanent with main period employer	PERM
*Casual with main period employer	
Level of highest education attainment	
Higher degree	EDU1
Postgraduate diploma	EDU2
Bachelor degree	EDU3
Undergraduate diploma	EDU4
Associate diploma	EDU5
Skilled vocational qualifications	EDU6
Basic vocational qualifications	EDU7
With post-school qualifications but level not stated	EDU8
Completed year 12 only	EDU9
*Did not complete year 12/never attended	
Main field of highest educational attainment	
Business and administration	FEDU1
Health	FEDU2
Education	FEDU3
Society and culture	FEDU4
Natural and physical sciences	FEDU5
Engineering	FEDU6
Architecture and building	FEDU7
Agriculture and related fields	FEDU8
*Field not stated/not applicable	

Table C-1 DESCRIPTION OF VARIABLES AND VARIABLE ABBREVIATIONS
(continued)

Variable	Abbreviation
Size of firm (number of employees)	
Less than 20	FIRM1
20-99	FIRM2
100 and over	FIRM3
*Number unknown	
Hours usually worked per week	
1-10 hours	HRWORK1
54+ hours	HRWORK5
Industry	
Agriculture, forestry, fishing and hunting	INDU1
Mining	INDU2
Manufacturing	INDU3
Electricity, gas and water supply	INDU4
Construction	INDU5
Wholesale trade	INDU6
Retail trade	INDU7
Accommodation, cafes and restaurants	INDU8
Transport and storage	INDU9
Communication services	INDU10
Finance and insurance	INDU11
Property and business services	INDU12
Government administration and defence	INDU13
Education	INDU14
Health and community services	INDU15
Cultural and recreational services	INDU16
*Personal and other services	

Table C-1 DESCRIPTION OF VARIABLES AND VARIABLE ABBREVIATIONS
(continued)

Variable	Abbreviation
Cumulative duration of employment	
Under 1 year	TMPE1
1-4 years	TMPE2
5-9 years	TMPE3
10-19 years	TMPE4
*20 years and over	
Trade union membership	
Had trade union membership	UNION
*Did not have trade union membership	
Estimated work experience (years)	
Experience	WEX
Experience square	WEXSQ

(a) For females, a suffix 'F' was used. Variables starting with 'L' means natural logarithm of that variable.

Note: * Denote reference category (omitted dummies) in the model. For the 1997 regression, omitted occupation variable was occ42 and for the 2001 regression, omitted occupation variable was occ81.