C Results

This appendix presents, in section C.1, the results of estimation of the model described in appendix A. In section C.2 the possible approaches to estimating the marginal effects of education and health are described and a preferred approach is selected.

C.1 Regression results

This section presents estimates of the coefficients for the participation and wage equations, estimated separately for men and women. Table C.1 sets out the estimated coefficients for the Heckman selection equation.

Table C.2 sets out the estimated coefficients for the wage equation, including λ, the coefficient that accounts for sample selection bias (appendix A). The estimated coefficients show that there is no sample selection bias present for women in the sample (because the estimated value of the sample selection coefficient λ is not significantly different from zero).

Table C.1 **Probit selection equation coefficient estimates**[a]

	Male			Female		
Variable			*Standard error*			*Standard error*
Age 15–24	0.243	***	0.083	0.674	***	0.062
Age 45–64	-0.992	***	0.078	-1.026	***	0.056
Vic	-0.068		0.060	0.032		0.050
Qld	-0.027		0.065	0.006		0.051
SA	0.018		0.079	0.030		0.072
WA	-0.008		0.081	-0.081		0.064
Tas	-0.240	**	0.115	0.257	**	0.105
NT	0.089		0.232	0.305		0.310
ACT	0.032		0.166	0.112		0.121
Region	-0.108	**	0.049	-0.080	**	0.041
Indigenous	-0.353	**	0.138	-0.207	*	0.117
Married	0.247	***	0.056	-0.148	***	0.042
Unemployment history	-2.194	***	0.191	-0.880	***	0.158
Experience	0.062	***	0.008	0.102	***	0.007
Experience^2	-0.001	***	0.000	-0.001	***	0.000
Degree or higher	0.350	***	0.071	0.708	***	0.054
Diploma or certificate	0.159	***	0.055	0.342	***	0.048
Year 12	0.257	***	0.073	0.421	***	0.052
PCS	0.045	***	0.002	0.034	***	0.002
MCS	0.022	***	0.002	0.012	***	0.002
NESB	-0.259	***	0.067	-0.270	***	0.056
Studying	-0.864	***	0.076	-0.570	***	0.066
Children 0-4	0.007		0.056	-0.769	***	0.033
Children 5-14	0.007		0.035	-0.199	***	0.023
Children 15-24	0.259	***	0.043	0.146	***	0.033
Wave 2	0.077	**	0.034	-0.002		0.027
Wave 3	0.167	***	0.039	-0.009		0.030
Wave 4	0.212	***	0.040	0.024		0.032
Wave 5	0.190	***	0.042	0.035		0.033
Constant	-2.960	***		-2.546	***	

*** significant at 1 per cent; ** significant at 5 per cent; * significant at 10 per cent.

[a] The dependent variable is a binary indicator of employment.

Source: Productivity Commission estimates based on HILDA release 5.1, waves 1–5.

Table C.2 Wage equation coefficient estimates[a]

	Male			Female		
Variable			*Standard error*			*Standard error*
Age 15–24	-15.567	***	2.130	-11.127	***	1.776
Age 25–44	-1.682		2.318	-0.619		1.542
Vic	-4.379	***	1.581	-6.012	***	1.405
Qld	-6.611	***	1.730	-8.975	***	1.432
SA	-10.926	***	2.224	-9.027	***	2.018
WA	-4.094	*	2.283	-9.099	***	1.869
Tas	-7.294	**	3.449	-4.776	*	2.626
NT	5.279		8.598	2.032		5.180
ACT	7.024	**	3.556	1.230		3.967
Region	-9.439	***	1.359	-6.393	***	1.170
Indigenous	2.005		5.021	8.060	**	3.573
Married	10.790	***	1.386	4.530	***	1.092
Experience	1.556	***	0.224	2.027	***	0.213
Experience^2	-0.019	***	0.005	-0.038	***	0.005
Degree or higher	38.373	***	1.900	38.180	***	1.573
Diploma or certificate	13.668	***	1.497	11.992	***	1.436
Year 12	12.402	***	2.111	10.743	***	1.636
PCS	0.329	***	0.094	0.328	***	0.065
MCS	0.158	***	0.060	0.162	***	0.047
NESB	-6.087	***	1.928	-6.130	***	1.940
Studying	0.465		2.901	2.575		2.497
Part time	-2.945		2.038	-0.162		1.048
Wave 2	2.873	***	0.958	3.265	***	1.065
Wave 3	7.791	***	1.018	7.854	***	0.932
Wave 4	12.878	***	1.015	10.971	***	0.981
Wave 5	17.242	***	1.050	15.726	***	1.016
λ	-7.637	**	2.666	2.00		1.301
ρ	-0.194	**	0.067	0.055		0.036
Constant	234.396	***		224.402	***	

*** significant at 1 per cent; ** significant at 5 per cent; * significant at 10 per cent.

[a] The dependent variable is the natural logarithm of hourly wage multiplied by 100.

Source: Productivity Commission estimates based on HILDA release 5.1, waves 1–5.

C.2 Estimating marginal effects

Because the model is not a simple linear regression, estimating the marginal effects of changes in education and health status on wages is not simply a matter of reporting the estimated coefficients.

Cameron and Trivedi (2009) describe three common methods for the evaluation of the marginal effects of independent variables in nonlinear models:

1. The average of the marginal effects at each observation (AME)
2. The marginal effect at the sample mean (MEM)
3. The marginal effect at a representative value of the independent variables (MER).

Cameron and Trivedi state that the marginal effects that are calculated using the different approaches 'can differ appreciably' (p. 340). Bartus (2005) prefers the AME approach to evaluating marginal effects. He states:

> The main argument in favour of AME is based on a demand for realism: the sample means used during the calculation of MEM might refer to either nonexistent or inherently nonsensical observations, a problem typically encountered when there are dummies among the regressors. (Bartus 2005, pp. 309–310)

Cameron and Trivedi argue that for nonlinear models, using the MEM approach is:

> … better than doing nothing, because it does provide a rough gauge of the magnitude of the [marginal effect]. However, for policy analysis, one should use either the MER for targeted values of the regressors … or the AME … (Cameron and Trivedi 2009, p. 340)

Greene (2003) states that:

> … in large samples [the MEM and AME approaches] will give the same answer. But that is not so in small or moderate-sized samples. Current practice favours averaging the individual marginal effects when it is possible to do so. (Greene 2003, p. 668)

In the empirical literature in this area there are examples of the MEM and MER approaches.

Breusch and Gray (2004) used the MEM approach to estimate the marginal effects of education on male and female wages in Australia. Their model (a Heckman model similar to that used in this study) used HILDA data, and included educational attainment through four binary variables ('incomplete high school', 'year 12', 'trade' and 'degree') that are analogous to the variables used in this study.

To estimate the marginal effects of continuous variables, the variable of interest was increased from just below the sample mean to just above the sample mean, while all other variables were held at their sample means. For binary variables (including level of education attained), the marginal effect was measured by changing the value of the variable from 0 to 1, and comparing the change with the base case of incomplete high school education.

Creedy et al. (2000) used the MER approach to evaluate marginal effects. They used data from the Income Distribution Surveys for 1995 and 1996 and a model that was similar to the model used for this project to estimate the wages of different demographic groups in Australia. Creedy et al. presented a sample of 'case studies' that demonstrated how their model could be used to estimate the wages of various demographic groups. For example:

> … consider an unemployed married female: aged 40 to 44 years; with one dependent child aged over 15 years; European born; residing in Perth; with no formal educational qualifications; partner has vocational qualification but is currently not employed; other income is $25 per week; owns home outright. The basic imputed wage is $13.49 per hour. (Creedy et al. 2000, p. 313)

Following the arguments put forward by Greene (2003), Bartus (2005) and Cameron and Trivedi (2009), the AME approach was chosen as the most suitable for evaluating the marginal effects of education and health status on wages.[1] The results are set out in chapter 5.

---

[1] Marginal effects were estimated using the Stata program 'margeff' version 8. Bartus (2005) describes the program.