Wage Functions for Demographic Groups in Australia*

John Creedy
Department of Economics,
The University of Melbourne

Alan S. Duncan
University of Nottingham,
Institute for Fiscal Studies
Melbourne Institute of Applied Economic and Social Research,
The University of Melbourne

Mark N. Harris
Melbourne Institute of Applied Economic and Social Research,
The University of Melbourne
Central European University of Hungary

Rosanna Scutella
Melbourne Institute of Applied Economic and Social Research,
The University of Melbourne

Abstract

This paper reports estimates of wage equations for groups of Australian workers, using pooled data from the Income Distribution Surveys for 1995 and 1996, the first two years for which continuous hours information is available for each individual. The problem of using the wage functions to assign a wage rate to non-workers, which is needed in the context of labour supply analysis, is also examined. Special attention is given to the case where the wage equation includes variables that are not available for the unemployed. The use of extraneous information about the occupation and industry characteristics of the unemployed is recommended.

1. Introduction

A major aim of this paper is to report estimates of wage functions for a number of demographic groups in Australia, using pooled information from the 1995 and 1996 Income Distribution Surveys (IDS). The estimation procedure allows for the sample selection bias that would otherwise arise from the fact that only the wage rates of those currently working are observed. Wage functions provide useful descriptive information about the characteristics of individuals, which are associated with relatively high or low wage rates. Earlier Australian wage functions were discussed by Miller and Rummery (1991). They emphasised the need to consider the complete wage distribution (including the unemployed and employed) in attempting to assess the existence of gender discrimination.

Address for correspondence: John Creedy, Department of Economics, The University of Melbourne, Parkville, Vic, 3052, Australia. Tel: +61 3 8344 5336. Email: jcreedy@clyde.its.unimelb.edu.au

* This research was supported by a grant from the Department of Family and Community Services, and a Visiting Scholar Research Grant from the University of Melbourne Faculty of Economics and Commerce. We should like to thank Paul Flatau and two referees for helpful comments on an earlier version of this paper.

A second main aim of the paper is to consider the question of how wage functions can be used to impute wage rates to those who are unemployed. An important context in which this is required is the analysis of labour supply and, in particular, examination of the effects of taxes and transfers.[1] Many tax policies are specially designed in an attempt to stimulate an increase in labour supply. There would therefore be little value in restricting analyses to those currently working, thereby excluding the unemployed whose participation decision may be influenced by taxes and transfers. Labour supply analyses require an individual-specific budget constraint, so a wage rate must be assigned to non-workers. This issue is complicated by the fact that the wage equations contain variables, such as industry and occupation, which are not recorded for non-workers (for the same reason that wage rates are not available). It is nevertheless deemed to be desirable to include as many relevant individual characteristics as possible in the wage functions.

The standard selection model is described briefly in section 2. The data are described in section 3. Estimates of selection and wage equations are reported in section 4. The problem of assigning wage rates to non-workers is examined in section 5. Brief conclusions are in section 6.

2. The Statistical Model

The estimation of wage equations involves two equations, the first of which determines selection (employment) using a probit equation, while the second determines wage rates, conditional on employment. Each individual's observed employment outcome is regarded as being the result of an unobservable index of employability, E_i^*, which varies with observed personal characteristics, z_i. The variables included in z may include both supply and demand side variables. Hence:

$$E_i^* = z_i'\gamma + u_i \tag{1}$$

where u_i is assumed to be independently distributed as $N(0,1)$.[2] The realisation of E_i^* determines whether the individual is employed ($E_i = 1$) or unemployed ($E_i = 0$), such that:

$$E_i = \begin{cases} 1 & \text{if } E_i^* > 0 \text{ with prob } \Phi(z_i'\gamma) \\ 0 & \text{if } E_i^* \le 0 \text{ with prob } 1 - \Phi(z_i'\gamma) \end{cases} \tag{2}$$

where $\Phi(z_i'\gamma)$ is the standard normal distribution function evaluated at $z_i'\gamma$. The associated density function is denoted $\phi(z_i'\gamma)$. The parameters of (2) can be consistently estimated by a standard probit model; see Maddala (1983). Having estimated (2), an estimate, $\hat{\lambda}$, of the inverse Mills ratio for individual i is obtained using:

$$\hat{\lambda}_i = \frac{\phi(z_i'\hat{\gamma})}{\Phi(z_i'\hat{\gamma})} \tag{3}$$

Let w_i denote the logarithm of the wage rate and x_i a vector of characteristics of individual i. The regression model is written as:

$$w_i\big|_{E_i=1} = x_i'\beta + \varepsilon_i \tag{4}$$

[1] The present paper is in fact part of a larger project involving the construction of a behavioural microsimulation model for Australia, the Melbourne Institute Tax and Transfer Simulator (MITTS).

[2] As there is no information about the scale of E^*, the variance of u cannot be identified and is, therefore, set equal to unity; see Maddala (1983).

The u_i from equation (1) and ε_i are assumed to be jointly normally distributed as $N(0,0,1,\sigma_\varepsilon^2,\rho)$.[3] To avoid selectivity bias, this is estimated using:

$$w_i\big|_{E_i=1} = x_i'\beta + \rho\sigma_\varepsilon\hat{\lambda}_i + v_i \tag{5}$$

Equation (5) takes into account the correlation between u_i and ε_i Allowance is made for the fact that the variance, σ_i^2, of v_i is heteroscedastic, since:

$$\sigma_i^2 = \sigma_\varepsilon^2(1-\rho^2\delta_i) \tag{6}$$

where:

$$\delta_i = \lambda_i(\lambda_i + z_i'\gamma) \tag{7}$$

Efficient estimation of this model is carried out using the procedure described in, for example, Greene (1981).

3. The Data

The data used in this analysis are taken from the 1995 and 1996 Income Distribution Surveys, available from the ABS (1995) and (1996) in the form of confidential unit record files (CURFs). The survey collects information on the sources and amounts of income received by persons resident in private dwellings throughout Australia, along with data on a range of characteristics of income units and individuals. The survey is continuous with around 650 households interviewed every month during the financial year. In the 1995 and 1996 surveys, information is available respectively for 14,017 and 14,595 individuals over the age of 15.

Earlier Income Distribution Surveys were carried out, but the 1995 survey is the first to provide published data on the precise hours worked (up to 50 hours per week) by each individual worker in the sample; earlier surveys contain only grouped information, with very broad groups. The details of hours worked are required for the calculation of a wage rate, obtained for each individual as the ratio of total earnings to hours worked. Hence the following analysis ignores the possibility that individuals may obtain overtime premia, or may work in more than one job. Where individuals worked more than 50 hours, the number was set at 50; this obviously produces an upward bias in the wage rate for those people.

In view of the emphasis of the analysis on obtaining results that are useful in labour supply analysis, those individuals permanently not in the labour force, such as the retired, disabled and those in full time education, were omitted from the sample. Hence, only persons employed or unemployed remained.[4] The majority of the data comprising the explanatory variables were recoded as zero-one dummy variables. To keep the variables to scale all of the non-wage income variables were divided by 1,000. Any inconsistencies between wages and hours, that is positive wages for zero hours or zero wages for positive hours, were treated as a zero wage per hour. To avoid the problem of taking the logarithm of zero in the wage equation and to ensure consistency between the employment and wage equations, employment status was determined with reference to the wage rate. Thus, any person with a zero wage rate per hour was treated as being unemployed, regardless of actual employment status as defined in the original data set.

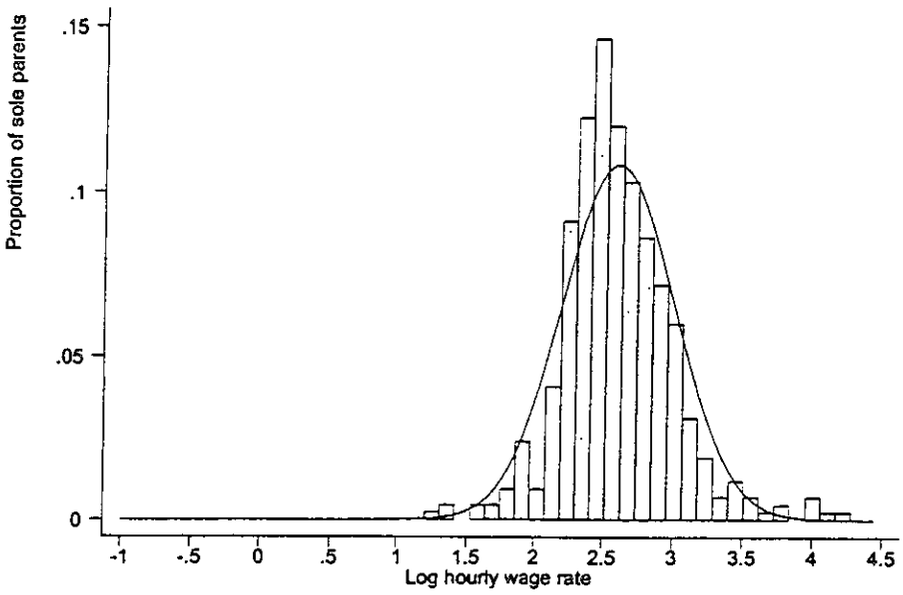
---

[3] The covariance between u_i and ε_i is thus $\rho\sigma_\varepsilon$.

[4] In the two surveys used, there were 1208 people either at school or studying full-time. There were 69 unpaid voluntary workers and 188 individuals permanently unavailable for work. There were 1825 self-employed persons.

The two surveys were pooled (with 1995 values converted to 1996 values using a wage index) and the sample was divided into five broad demographic groups. These are: sole parents; single females without dependents; single males without dependents; married females and married males. Summary tables of sample characteristics are provided for each demographic group in the Appendix. It was not possible to divide the sole parents group into males and females, in view of the small number of male sole parents in the sample.[5] Examples of distributions of the logarithms of observed hourly wage rates for the five demographic groups are shown in figures 1 to 5. These are based on February 1995 wages. The histograms suggest that these distributions are approximately lognormal, although they are slightly more peaked than the corresponding normal distributions with the same mean and variance.

Figure 1 Sole parents: The distribution of hourly wages (February 1995)



[5] There were 104 male sole parents, compared with 857 females.

Figure 2 Single females: The distribution of hourly wages (February 1995)

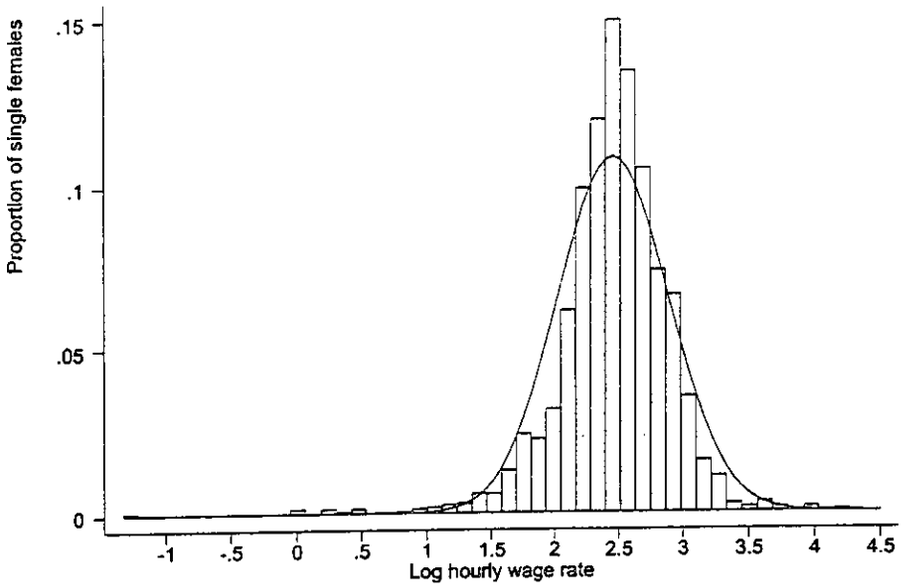


Figure 3 Single males: The distribution of hourly wages (February 1995)

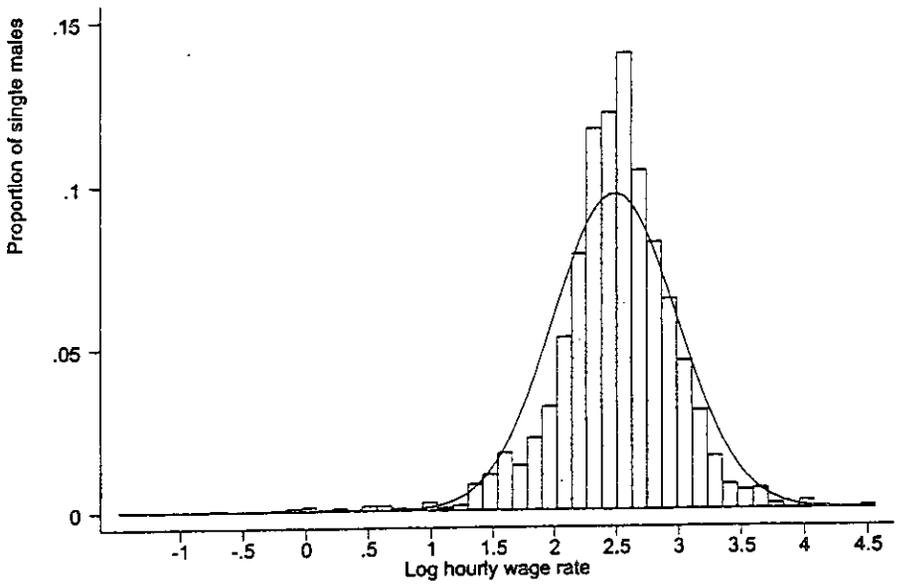


Figure 4 Married females: The distribution of hourly wages (February 1995)

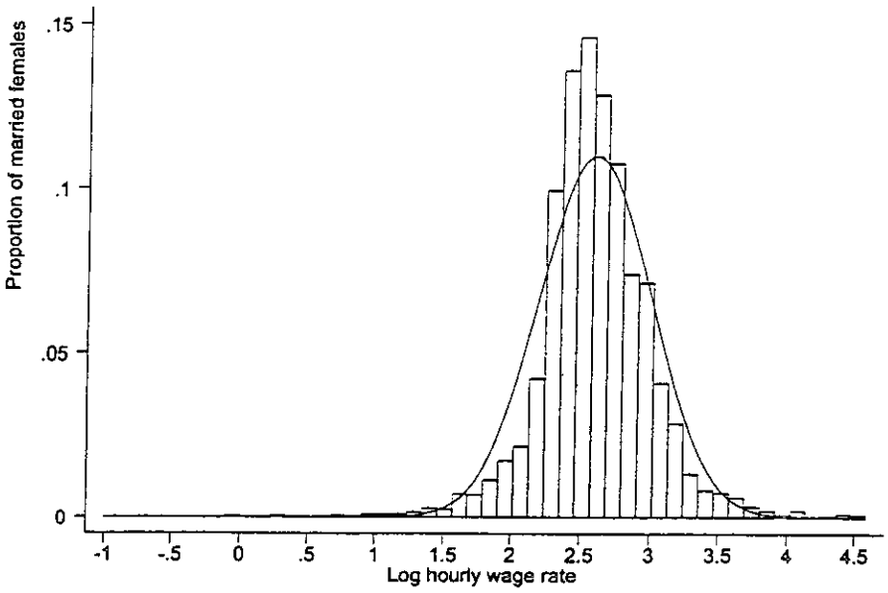
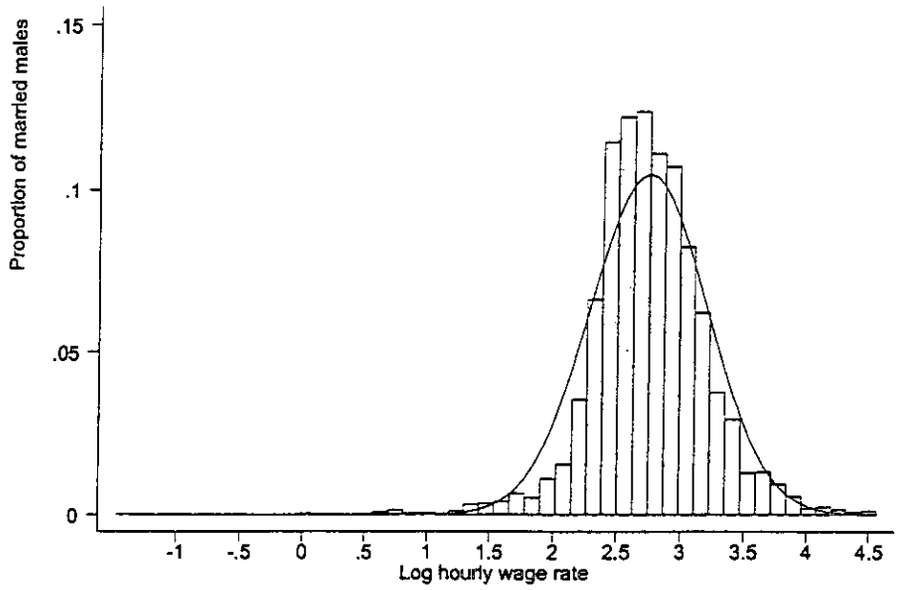


Figure 5 Married males: The distribution of hourly wages (February 1995)



4. Empirical Results

This section presents the main empirical results. The selection equations, along with specification tests and 'hit and miss' tables, are reported in subsection 4.1. The wage equations and associated specification tests are reported in subsection 4.2.

Selection Equations

The results of estimating the selection equations for each demographic group are shown in tables 2, 3 and 4. These are based on sample sizes, for married men, married women, single men, single women and sole parents respectively, of 5689, 4333, 2857, 2264 and 509. The column headed df/dx shows the 'marginal effects' on the probability of being employed, evaluated at sample means of variables (in most cases these are the effects of a discrete change from 0 to 1 in the dummy variable). The majority of coefficients are significantly different from zero and the coefficient signs appear to accord with expectations.[6]

A 'hit and miss' table can be constructed to evaluate how well the selection model predicts. However, if standard techniques are used to do this (based on the maximum probability rule), the models generally tend to overpredict the empirically most frequently chosen outcome. Indeed, this is true of the present models, with the employed being somewhat overpredicted for each demographic group. Such a result stems from the fact that the random elements of the model are explicitly ignored in its evaluation.

An alternative method is to utilise the underlying economic model of equation (1) and to simulate it by repeated draws of the unobserved random variates (that is, from the standard normal distribution). For each random draw, the probabilistic expressions are evaluated and the outcome which yields the maximum probability is the one predicted for that repetition.[7] Each separate hit and miss table is collected, and the simulated hit and miss table is the average of all of these independent ones. The results are reported in table 1, which summarises actual versus predicted values, using 1000 random draws. It can be seen that in each case the probit model provides close predictions of the numbers employed.

Considering the results, state of residence seems to be of more importance to employment patterns of women than for men, whereas living in a capital city is associated with a higher probability of employment for all demographic groups. In the case of married women the probability of employment is higher for those with a tertiary qualification than for those without any qualification, whereas this is not true of married men. The probability of employment is higher for sole parents with a diploma or vocational qualification than for those without any qualification. Couples where the partner is employed are significantly more likely to be in employment than couples where the partner is out of work. Also women whose partner has a postgraduate qualification have a lower probability of employment.

[6] Direct comparisons with the results in Miller and Rummery (1991) are not possible because the latter have just two demographic groups (males and females) and include a much smaller set of variables than are used here.

[7] A predicted probability of ≥ 0.5 gives rise to a predicted value of unity (employed); zero (unemployed) otherwise.

Table 1 Simulated Hit/Miss Results

		Predicted	
Actual	*0*	*1*	*Total*
		Sole Parents	
0	26.5	66.5	93
1	66.3	349.7	416
Total	92.9	416.1	509
		Single Females	
0	57.2	240.8	298
1	241.3	1724.7	1966
Total	298.5	1965.5	2264
		Single Males	
0	122.0	409.0	531
1	409.3	1917.0	2326
Total	531.3	2326.0	2857
		Married Females	
0	128.6	417.4	546
1	418.1	3369.0	3787
Total	546.7	3786.0	4333
		Married Males	
0	146.6	619.4	766
1	620.4	4303.0	4923
Total	767.0	4922.0	5689

As the number of dependent children increases, the probability of employment falls for either member of a couple. However, for single parents, employment status is unaffected by the number of children. The age of the youngest child has opposite effects for married men and women. Coefficients on the age of the youngest child indicator variables are all positive and significant (at least at the ten per cent level) for men. This counteracts the negative coefficient on the number of children. In the case of married women, negative coefficients on the age of youngest child variables reinforces the negative effect of the number of dependent children on the probability of employment. Thus, employment is relatively less likely for women with children under the age of 9. However, for men with children under the age of 15, the probability of employment is higher than for those without children or with older children. All groups, apart from sole parents, are more likely to be employed if their homes are either owned outright or if they have a mortgage.

In addition to variables relating to age, location, qualifications and number and ages of children (in the case of married and sole parent groups), interaction terms for qualification and age were also found to be significant. These suggest, as expected, that age-earnings profiles differ between education groups.

Maximum likelihood estimation of discrete choice models, such as the selection model used here, is based upon a specified distribution of the unobserved elements of the underlying economic model. However, unlike the traditional linear regression model, if the distributional assumptions in such non-linear models are invalid, parameter estimates are both biased and inconsistent. For this reason, diagnostic testing procedures in a discrete choice framework have been suggested; see Olsen (1982), Blundell and Meghir (1987), Chesher and Irish (1987) Gourieroux *et al* (1987) and Pagan and Vella (1989).

Thus the first stage in checking model specification is to test the specification of these selection equations. Following Pagan and Vella (1989), tests for both normality and heteroscedasticity were undertaken. The results are reported in table 5. These results show that the assumption of homoscedasticity appears valid for all of the demographic groups. Also, the assumption of normality cannot be rejected (at the 5 per cent level) for all demographic groups except for sole parents.

Table 2 Selection Terms: Married Men and Women

	Men			*Women*		
Variable	*Coefficient*	*Standard error*	*df/dx*	*Coefficient*	*Standard error*	*df/dx*
Constant	0.8037	0.1175		0.8268	0.1302	
Aged 15 to 19	-0.8862	0.5082	-0.2623			
Aged 20 to 24	0.0596	0.1891	0.0112	-0.1008	0.1721	-0.0180
Aged 25 to 29 (reference)						
Aged 30 to 34	0.0490	0.1279	0.0094	0.0692	0.1328	0.0113
Aged 35 to 39	-0.1290	0.1273	-0.0265	0.2088	0.1414	0.0321
Aged 40 to 44	-0.0154	0.1338	-0.0030	0.0380	0.1381	0.0063
Aged 45 to 49	-0.1611	0.1250	-0.0335	-0.0681	0.1351	-0.0118
Aged 50 to 54	-0.1515	0.1368	-0.0317	-0.1686	0.1520	-0.0311
Aged 55 to 59	-0.1754	0.1441	-0.0373	-0.2893	0.1660	-0.0576
Aged 60 to 64	-0.2780	0.1835	-0.0628	-0.3709	0.2480	-0.0783
Australia (reference)						
Europe/Middle East	-0.0630	0.0571	-0.0126	-0.2125	0.0715	-0.0394
Asia	-0.1630	0.1012	-0.0346	-0.5582	0.1059	-0.1267
America/Africa	-0.0888	0.1602	-0.0182	0.4843	0.2508	0.0597
Postgraduate	0.0712	0.1587	0.0133	0.7528	0.2302	0.0806
Undergraduate	-0.1544	0.1474	-0.0232	0.7657	0.2193	0.0875
Diploma	-0.3613	0.2572	-0.0824	-0.2897	0.4337	-0.0567
Vocational	-0.5214	0.2501	-0.1153	-0.4443	0.4260	-0.0892
No formal qualification (reference)						
Non-labour income unit income	0.0565	0.1056	0.0110	0.0570	0.1055	0.0096
Child support income				0.0245	2.5755	0.0041
New South Wales (reference)						
Victoria	-0.0462	0.0672	-0.0092	-0.0747	0.0827	-0.0130
Queensland	-0.0030	0.0706	-0.0006	-0.1183	0.0854	-0.0210
South Australia	-0.1304	0.0810	-0.0270	-0.1891	0.0966	-0.0350
Western Australia	0.0072	0.0765	0.0014	-0.1952	0.0921	-0.0362
Tasmania	-0.1133	0.0933	-0.0234	-0.1436	0.1158	-0.0263
ACT/Northern Territory	0.2156	0.1019	0.0376	0.2021	0.1205	0.0305
Capital city	0.2759	0.0496	0.0557	0.2458	0.0597	0.0429
Number of children	-0.0569	0.2901	-0.0110	-0.0939	0.0394	-0.0159
Youngest child 0 to 2	0.1857	0.0943	0.0337	-0.3061	0.1173	-0.0600
Youngest child 3 to 4	0.2306	0.1133	0.0399	-0.3287	0.1371	-0.0665
Youngest child 5 to 9	0.1540	0.0937	0.0281	-0.3163	0.1166	-0.0618
Youngest child 10 to 15	0.1662	0.0904	0.0300	-0.1124	0.1096	-0.0200
Owned/mortgaged (reference)						
Rented	-0.2794	0.0557	-0.0600	-0.2376	0.0696	-0.0442
Other tenure	-0.2781	0.1341	-0.0630	-0.3127	0.1566	-0.0636
Partner employed	0.5288	0.0629	0.1085	0.8078	0.0879	0.1728
Partner has postgraduate qualification	-0.0540	0.1279	-0.0108	-0.2443	0.1232	-0.0473
Partner has undergraduate qualification	0.1253	0.0883	0.0230	-0.0600	0.0962	-0.0105
"Older" partner	-0.1177	0.0680	-0.0242	-0.3789	0.2420	-0.0804
"Younger" partner	-0.0137	0.1923	-0.0027	-0.0679	0.0786	-0.0119
University qualification x (20 to 24)	-0.0182	0.4884	-0.0036	-0.6433	0.4625	-0.1567
University qualification x (25 to 29)	1.0848	0.4418	0.1067	-0.5428	0.2971	-0.1247
University qualification x (30 to 34)	0.2481	0.2321	0.0418	-0.4859	0.2920	-0.1084
University qualification x (35 to 39)	0.0673	0.2000	0.0126	-0.4229	0.2787	-0.0909
University qualification x (40 to 44)	0.1639	0.2074	0.0291	-0.4806	0.2375	-0.1065
University qualification x (45 to 49)	0.1399	0.2295				
University qualification x (50 to 54)	0.1322	0.2670	0.0251	-0.1817	0.3688	-0.0344
University qualification x (55 to 59)	-0.0411	0.3392	0.0238			
Vocational qualification x (20 to 24)	0.3396	0.3559	0.0538	0.8797	0.5179	0.0827
Vocational qualification x (25 to 29)	0.7096	0.2648	0.0912	0.4560	0.4404	0.0577
Vocational qualification x (30 to 34)	0.6532	0.2828	0.0882	0.4499	0.4490	0.0578
Vocational qualification x (35 to 39)	0.5559	0.2762	0.0801	0.3149	0.4489	0.0440
Vocational qualification x (40 to 44)	0.4912	0.2805	0.0732	0.6943	0.4552	0.0764
Vocational qualification x (45 to 49)	0.6938	0.2759	0.0926	0.5903	0.4531	0.0692
Vocational qualification x (50 to 54)	0.6425	0.2840	0.0868	0.7229	0.4698	0.0768
Vocational qualification x (55 to 59)	0.5978	0.2596	0.0816	0.1379	0.4811	0.0213
Vocational qualification x (60 to 64)	0.4493	0.3246	0.0666	0.371	0.6330	0.0061

Table 3 Selection Terms: Single Men and Women

	Men			*Women*		
Variable	*Coefficient*	*Standard error*	*df/dx*	*Coefficient*	*Standard error*	*df/dx*
Constant	*1.2671*	*0.1424*		*2.1158*	*0.2072*	
Aged 15 to 19	-0.1691	0.1174	-0.0456	-0.5443	0.1579	-0.1238
Aged 20 to 24	0.0300	0.1133	0.0076	-0.2223	0.1615	-0.0435
Aged 25 to 29 (reference)						
Aged 30 to 34	-0.1042	0.1328	-0.0276	0.0246	0.2060	0.0045
Aged 35 to 39	0.0222	0.1808	0.0056	-0.5705	0.2498	-0.1385
Aged 40 to 44	-0.1903	0.1782	-0.0523	-0.4315	0.2548	-0.0994
Aged 45 to 49	-0.0771	0.2090	-0.0203	-0.3867	0.2382	-.0863
Aged 50 to 54	-0.5788	0.2267	-0.1823	-0.2631	0.2642	-0.0557
Aged 55 to 59	-0.2219	0.2316	-0.0621	-0.3237	0.2687	-0.0709
Aged 60 to 64	0.0832	0.3340	0.0204	1.0595	0.6457	0.0987
Aged 65 to 69	-1.8938	0.6660	-0.6563	0.7190	0.9215	0.0823
Separated/widowed	-0.0200	0.0983	-0.0051	-0.2742	0.1361	-0.0555
Australia (reference)						
Europe/Middle East	-0.1631	0.0979	-0.0441	-0.2881	0.1284	-0.0611
Asia	-0.4797	0.1438	-0.1466	-0.3667	0.1890	-0.0825
Postgraduate	0.7401	0.2634	0.1298	0.1443	0.2483	0.0245
Undergraduate	0.4046	0.1561	0.0874	0.1575	0.2043	0.2070
Diploma	-0.2220	0.4762	-0.0615	-1.5634	0.6089	-0.4877
Vocational	-0.1596	0.4687	-0.0423	-1.8176	0.6031	-0.5388
No formal qualification (reference)						
Non-labour income unit income	-0.4504	0.5276	-0.1149	-1.8168	0.7151	-0.3344
New South Wales (reference)						
Victoria	-0.2895	0.0871	-0.0797	-0.1521	0.1120	-0.0294
Queensland	-0.2396	0.0922	-0.0656	-0.2690	0.1147	-0.0551
South Australia	-0.3588	0.1033	-0.1033	-0.4293	0.1300	-0.0962
Western Australia	-0.2432	0.0968	-0.0670	-0.2084	0.1241	-0.0421
Tasmania	-0.3607	0.1285	-0.1054	-0.1944	0.1572	-0.0396
ACT/Northern Territory	0.1088	0.1337	0.0265	-0.0305	0.1765	-0.0057
Capital city	0.2606	0.0639	0.0684	0.2841	0.0820	0.0557
Owned/mortgaged (reference)						
Rented	-0.3070	0.0911	-0.0763	-0.7058	0.1257	-0.1263
Other tenure	-0.4989	0.1038	-0.1427	-0.9068	0.1411	-0.2200
University qualification x (20 to 24)	0.0600	0.2599	0.0149	0.2287	0.2835	0.0369
University qualification x (35 to 39)	0.0343	0.4247	0.0086	0.8703	0.5081	0.0924
University qualification x (40 to 44)	0.2458	0.4303	0.0556	0.1754	0.4924	0.0289
University qualification x (45 to 49)	-0.4403	0.4265	-0.1341	0.6003	0.5446	0.0750
University qualification x (50 to 54)	-0.0413	0.4932	-0.0107	0.0272	0.6506	0.0049
Vocational qualification x (15 to 19)	-0.0157	0.5294	-0.0040	1.7386	0.6331	0.1138
Vocational qualification x (20 to 24)	0.2899	0.4821	0.0657	1.7282	0.6192	0.1342
Vocational qualification x (25 to 29)	0.3280	0.4874	0.0726	1.5641	0.6161	0.1192
Vocational qualification x (30 to 34)	0.4430	0.5005	0.0915	1.4555	0.6808	0.1082
Vocational qualification x (35 to 39)	0.2882	0.5208	0.0641	2.0729	0.6870	0.1135
Vocational qualification x (40 to 44)	0.5819	0.5314	0.1106	1.9094	0.6975	0.1111
Vocational qualification x (45 to 49)	0.1799	0.5462	0.0421	1.7544	0.6740	0.1121
Vocational qualification x (50 to 54)	0.9002	0.5960	0.1419	1.9444	0.7006	0.1121
Vocational qualification x (55 and over)	-0.2718	0.5580	-0.0779	1.4334	0.6977	0.1059

Table 4 Selection Terms: Sole Parents

Variable	*Coefficient*	*Standard error*	*df/dx*
Constant	1.4189	0.4327	
Aged 15 to 19	-0.8461	0.8229	-0.2744
Aged 20 to 24	-0.9708	0.3994	-0.3171
Aged 25 to 29 (reference)			
Aged 30 to 34	-0.4026	0.3058	-0.1070
Aged 35 to 39	-0.0056	0.3206	-0.0013
Aged 40 to 44	-0.0082	0.3518	-0.0019
Aged 45 to 49	-0.2783	0.4121	-0.0720
Aged 50 to 54	0.3342	0.6466	0.6623
Aged 55 to 59	-0.6849	0.6073	-0.2113
Separated/widowed	-0.1917	0.2033	-0.0424
Australia (reference)			
Europe/Middle East	-0.4270	0.1906	-0.1150
Asia	-0.3466	0.3509	-0.0943
Undergraduate qualification	-0.1359	0.2599	-0.0336
Diploma	1.0094	0.4156	0.1466
Vocational qualification	0.6592	0.3215	0.1269
No formal qualification (reference)			
Non-labour income unit income	3.7022	1.7325	0.8670
Child support income	-0.8672	1.3786	-0.2031
New South Wales (reference)			
Victoria	-0.0777	0.2263	-0.0186
Queensland	0.0876	0.2452	0.0199
South Australia	-0.3420	0.2574	-0.0907
Western Australia	0.1172	0.2578	0.0262
Tasmania	0.4141	0.2969	0.0807
ACT/Northern Territory	0.4408	0.3049	0.0853
Capital city	0.2952	0.1642	0.0707
Number of children	-0.0626	0.1032	-0.0147
Youngest child 0 to 2	-0.2858	0.3696	-0.0748
Youngest child 3 to 4	-0.4499	0.3300	-0.1239
Youngest child 5 to 9	-0.2156	0.2715	-0.0528
Youngest child 10 to 15	0.1991	0.2431	0.0446
Owned/mortgaged (reference)			
Rented	-0.1086	0.1664	-0.0253
Other tenure	0.5529	0.5800	0.0963
Vocational qualification x (25 to 29)	-1.6139	0.5058	-0.5612
Vocational qualification x (35 to 39)	-0.7766	0.4800	-0.2415
Vocational qualification x (40 to 44)	-1.4658	0.4403	-0.4988
Vocational qualification x (45 to 49)	-0.7649	0.5431	-0.2390
Vocational qualification x (50 to 54)	-1.3088	0.8753	-0.4552

Table 5 Specification Tests for Selection Equations (P-Values)

Demographic Group	*Normality*	*Heterosecdasticity*
Sole parents	0.026	0.428*
Single females	0.940*	0.324*
Single males	0.705*	0.492*
Married females	0.157*	0.550*
Married Males	0.549*	0.272*

*Cannot reject at the 5 per cent level.

Wage Equations

The estimated wage equations conditional on being in employment are reported for each demographic group in tables 6, 7 and 8. These are based on sample sizes, for married men, married women, single men, single women and sole parents respectively, of 4923, 3787, 2326, 1966 and 416. The inverse Mills ratio was retained in each case although the estimated t statistics are low (except for single women). For prediction purposes, it is useful to include this variable, along with other non-significant variables. With the exception of the sole parents, the coefficient on the Mills ratio is positive. The interpretation of negative inverse Mills ratios in this context was discussed by Ermisch and Wright (1994).[8] In each table, the 'estimated standard error' and 'estimated correlation coefficient' refer to σ_ε and ρ respectively.

The coefficients display the expected variation with age, industry and educational qualifications. Wage rates of professionals, para-professionals, and clerical/salespersons are significantly higher than for trades persons or labourers across all groups. Wage rates also tend to increase with the level of educational qualification across all groups.

Couples living in New South Wales experience higher wage rates than those living in the other States, while those residing in the Territories receive even higher wages; residents of the ACT dominate this category. Married men and women in capital cities are paid higher wage rates than their counterparts living in other areas of the country. Wage rates of married men and women are generally higher in all industries compared with the agriculture/forestry industry (the reference industry). Single males in the mining industry earn higher wages whereas there is no significant difference in the wage rates across industries for single females without dependents and sole parents.

The specification of the present model is also based on the joint normality of both the selection and regression equations. If the selection equation is mis-specified, the same is true of the correction term in the regression equation, resulting in biased and inconsistent estimates of the determinants of wages; see Olsen (1982). Following Pagan and Vella (1989) it is possible to test the assumption of joint normality by including the product of the linear prediction terms of the selection equation (to powers 1, 2 and 3) and the inverse Mills ratio for each individual. The null hypothesis of joint normality is rejected if these three additional variables are jointly significant. The likelihood ratio tests for these restrictions are reported in table 9.

[8] Miller and Rummery (1991) found a positive value for women and a negative value for men. They also review results found in previous Australian studies.

Table 6 Wage Equations: Married Men and Women

	Men		*Women*	
Variable	*Coefficient*	*Standard error*	*Coefficient*	*Standard error*
Constant	2.3829	0.0551	2.3170	0.0973
Aged 15 to 19	-0.4420	0.2722		
Aged 20 to 24	-0.1231	0.0555	-0.0724	0.0427
Aged 25 to 29 (reference)				
Aged 30 to 34	0.0441	0.0367	0.0585	0.0334
Aged 35 to 39	0.1092	0.0385	0.0464	0.0327
Aged 40 to 44	0.1070	0.0381	-0.0190	0.0312
Aged 45 to 49	0.1023	0.0388	0.0247	0.0313
Aged 50 to 54	0.0662	0.0406	-0.0323	0.0358
Aged 55 to 59	-0.0345	0.0442	-0.0388	0.0464
Aged 60 & over	0.0434	0.0644	0.1426	0.0775
Professional	0.1866	0.0189	0.2947	0.0251
Para-professional	0.1472	0.0222	0.2390	0.0268
Clerical/sales	0.0626	0.0202	0.1064	0.0174
Tradesperson/labourer (reference)				
Agriculture/forestry (reference)				
Mining	0.6171	0.0599	0.2953	0.1278
Manufacturing	0.1824	0.0481	0.1190	0.0970
Construction	0.1538	0.0516	0.1433	0.1072
Sales	0.0452	0.0483	0.0786	0.0949
Transport	0.2003	0.0514	0.2018	0.1028
Communications	0.2499	0.0576	0.2889	0.1067
Financial/business sector	0.1699	0.0491	0.1541	0.0947
Service industries	0.1015	0.0472	0.0969	0.0940
Australia (reference)				
Europe/Middle East	-0.0245	0.0167	-0.0270	0.0184
Asia	-0.1776	0.0312	-0.0762	0.0331
America/Africa	-0.1148	0.0447	-0.0423	0.0378
Postgraduate	0.3163	0.0503	0.2840	0.0478
Undergraduate	0.1885	0.0467	0.2283	0.0425
Diploma	-0.0511	0.1100	0.1132	0.1095
Vocational	-0.1034	0.1089	0.0357	0.1083
No formal qualification (reference)				
New South Wales (reference)				
Victoria	-0.0641	0.0192	-0.0379	0.0189
Queensland	-0.0466	0.0205	-0.0470	0.0204
South Australia	-0.0714	0.0238	0.0092	0.0231
Western Australia	-0.0158	0.0216	-0.0726	0.0222
Tasmania	-0.0218	0.0291	-0.0170	0.0287
ACT/Northern Territory	0.0830	0.0290	0.0784	0.0272
Capital city	0.0860	0.0170	0.0524	0.0157
University qualification x (20 to 24)	-0.0716	0.1494	-0.1765	0.0944
University qualification x (25 to 29)	-0.0532	0.0736	-0.0986	0.0604
University qualification x (30 to 34)	0.0326	0.0633	-0.0550	0.0589
University qualification x (35 to 39)	0.0116	0.0597		
University qualification x (40 to 44)	0.0314	0.0625	0.0177	0.0553
University qualification x (45 to 49)	0.1959	0.0683	-0.0055	0.0560
University qualification x (50 to 54)	0.2006	0.0805	-0.0352	0.0671
University qualification x (55 to 59)	0.2420	0.1202	-0.0124	0.0977
University qualification x (60 to 64)	0.2387	0.1330	-0.0572	0.1411
Vocational qualification x (20 to 24)	0.2146	0.1101	-0.0391	0.1230
Vocational qualification x (25 to 29)	0.2015	0.1146	-0.0246	0.1126
Vocational qualification x (30 to 34)	0.1296	0.1152	-0.0370	0.1138
Vocational qualification x (35 to 39)	0.1791	0.1147	-0.0114	0.1139
Vocational qualification x (40 to 44)	0.1582	0.1152	-0.0382	0.1135
Vocational qualification x (45 to 49)	0.1967	0.1164	-0.0358	0.1136
Vocational qualification x (50 to 54)	0.3070	0.1200	0.0018	0.1164
Vocational qualification x (55 to 59)	0.1271	0.1338	0.0291	0.1331
Inverse Mills ratio	0.0069	0.0645	-0.0190	0.0514
Estimated standard error	0.4083		0.3578	
Estimated correlation coefficient	0.0169		-0.0532	

Table 7 Wage Equations: Single Men and Women

	Men		*Women*	
Variable	*Coefficient*	*Standard error*	*Coefficient*	*Standard error*
Constant	2.5723	0.1370	2.5191	0.2543
Aged 15 to 19	-0.4444	0.0638	-0.3319	0.0478
Aged 20 to 24	-0.1741	0.0492	-0.0268	0.0410
Aged 25 to 29 (reference)				
Aged 30 to 34	0.1289	0.0647	0.0301	0.0533
Aged 35 to 39	0.0895	0.0737	0.1668	0.0697
Aged 40 to 44	0.0865	0.0831	-0.0036	0.0669
Aged 45 to 49	0.1585	0.0883	0.0909	0.0559
Aged 50 to 54	0.2831	0.1371	0.0629	0.0597
Aged 55 to 59	0.0370	0.1075	-0.0295	0.0704
Aged 60 to 64	0.1831	0.1390	-0.0621	0.1009
Aged 65 to 69			-0.1479	0.1380
Professional	0.2518	0.0333	0.2724	0.0335
Para-professional	0.1945	0.0376	0.2537	0.0373
Clerical/sales	0.1504	0.0259	0.1225	0.0229
Tradesperson/labourer (reference)				
Agriculture/forestry (reference)				
Mining	0.5072	0.1220	0.3802	0.2854
Manufacturing	0.0371	0.1016	-0.1490	0.2524
Construction	0.0752	0.1043	-0.1460	02646
Sales	-0.0481	0.1011	-0.1439	0.2512
Transport	0.1664	0.1067	-0.0024	0.2551
Communications	0.1351	0.1132	-0.0207	0.2604
Financial/business sector	0.0179	0.1037	-0.1177	0.2492
Service industries	-0.0756	0.1016	-0.1523	0.2495
Australia (reference)				
Europe/Middle East	0.0500	0.0443	0.0186	0.0355
Asia	0.0657	0.0863	0.0516	0.0519
America/Africa	0.0450	0.0856	-0.0691	0.0711
Postgraduate	0.0955	0.1177	0.3163	0.0752
Undergraduate	0.0237	0.0925	0.2654	0.0694
Diploma	0.2077	0.1539	0.0835	0.0730
Vocational	0.1824	0.1425	0.0903	0.0672
No formal qualification (reference)				
New South Wales (reference)				
Victoria	0.0542	0.0455	-0.0173	0.0264
Queensland	0.0573	0.0447	-0.0202	0.0308
South Australia	0.0113	0.0563	0.0360	0.0370
Western Australia	0.0875	0.0455	-0.0353	0.0321
Tasmania	0.1252	0.0685	-0.0518	0.0421
Act/Northern Territory	0.0615	0.0548	0.0949	0.0471
Capital city	-0.0300	0.0418	-0.0083	0.0247
University qualification x (20 to 24)	0.0168	0.1029	-0.2261	0.0794
University qualification x (25 to 29)	-0.0163	0.0940	-0.1644	0.0808
University qualification x (30 to 34)				
University qualification x (35 to 39)	-0.0549	0.1395	-0.0427	0.1047
University qualification x (35 and over)				
University qualification x (40 to 44)	0.1163	0.1442	-0.0034	0.1267
University qualification x (45 to 49)	0.2528	0.1711	-0.0732	0.1101
University qualification x (50 to 54)	0.0442	0.2224	0.0297	0.1325
University qualification x (55 to 59)	0.1447	0.2239	-0.0199	0.1252
University qualification x (60 to 64)	-0.9790	0.2283	-0.3011	0.1737
Vocational qualification x (24 or less)				
Vocational qualification x (20 to 24)	-0.0529	0.1557	-0.0499	0.0773
Vocational qualification x (25 to 29)	-0.0814	0.1578	-0.0242	0.0833
Vocational qualification x (30 to 34)	-0.2267	0.1716	0.0106	0.1018
Vocational qualification x (35 to 39)	-0.0776	0.1734	-0.0671	0.1124
Vocational qualification x (40 to 44)	-0.0860	0.1900	-0.0074	0.1159
Vocational qualification x (45 to 49)	-0.1857	0.1884	-0.0621	0.1032
Vocational qualification x (50 to 54)	-0.6505	0.2397	0.0500	0.1082
Vocational qualification x (55 to 59)	0.0775	0.2257	0.0344	0.1300
Vocational qualification x (60 to 64)	-0.1567	0.2477	0.1035	0.1845
Inverse Mills ratio	-0.5242	0.2985	-0.2284	0.1150
Estimated standard error	-05165		0.3725	
Estimated correlation coefficient	-1.0151		-0.6131	

Table 8 Wage Equation: Sole Parents

Variable	*Coefficient*	*Standard error*
Constant	2.7077	0.2416
Aged 15 to 19	-0.4127	0.3590
Aged 20 to 24	-0.0409	0.1630
Aged 25 to 29 (reference)		
Aged 30 to 34	-0.0992	0.0962
Aged 35 to 39	-0.0935	0.0790
Aged 40 to 44	-0.1726	0.0816
Aged 45 to 49	-0.0737	0.0943
Aged 50 to 54	-0.1396	0.1218
Aged 55 to 59	-0.1343	0.1731
Aged 60 & over	0.2654	0.0694
Professional	0.1614	0.0806
Para-professional	0.0819	0.0480
Clerical/sales		
Tradesperson/labourer (reference)		
Agriculture/forestry (reference)		
Mining	-0.1950	0.2345
Manufacturing	-0.2390	0.2637
Construction	-0.2876	0.2326
Sales	-0.1453	0.2553
Transport	-0.1280	0.2473
Communications	-0.1588	0.2294
Financial/business sector	-0.2020	0.2256
Service industries		
Australia (reference)	-0.0190	0.0693
Europe/middle east	-0.1326	0.1113
Asia	0.1371	0.1126
America/Africa	0.3932	0.1804
Postgraduate	0.2682	0.1919
Undergraduate	0.0110	0.1652
Diploma	-0.0458	0.1535
Vocational		
No formal qualification (reference)		
New South Wales (reference)	-0.1065	0.0661
Victoria	-0.0948	0.0701
Queensland	-0.1255	0.0821
South Australia	-0.0186	0.0714
Western Australia	-0.0117	0.0784
Tasmania	0.1230	0.0856
ACT/Northern Territory	0.0411	0.0532
Capital city	0.0067	0.1823
University qualification x (35 and over)	0.0652	0.3066
Vocational qualification x (24 or less)	-0.2620	0.2245
Vocational qualification x (25 to 29)	0.1056	0.1890
Vocational qualification x (30 to 34)	0.0031	0.1758
Vocational qualification x (35 to 39)	0.2766	0.1790
Vocational qualification x (40 to 44)	0.1431	0.1878
Vocational qualification x (45 to 49)	0.1738	0.2300
Inverse Mills ratio	0.3547	
Estimated standard error	0.5028	
Estimated correlation coefficient		

Table 9 Joint Normality Tests for the Heckman-Section Model (P-Values)

Demographic Group	*Joint Normality*
Sole parents	0.593*
Single females	0.002
Single males#	-
Married females	0.295*
Married Males	0.873*

Unable to be computed
* Cannot reject joint normality at 5per cent level.

Table 9 shows that the null hypothesis of joint normality cannot be rejected for three out of the five regressions, and for one of the others the test could not be computed.[9] Although the sole parent group fails the normality assumption in the selection equation, the assumption of joint normality cannot be rejected, suggesting confidence in the validity of these results. On the other hand, one might be wary of placing too much emphasis on the results of the determinants of the wages of single females.

5. Wage Predictions

This section considers the question of how a wage rate may be assigned to unemployed individuals. In the simple case where the selection and wage equations contain a common set of variables, consider first the conditional mean log-wage rate, for an individual with given characteristics. For those who are employed, this is given by:

$$E\left(w_i\big|_{E_i=1}\right)= x_i'\hat{\beta} + \hat{\rho}\hat{\sigma}_\varepsilon\hat{\lambda}_i, \tag{8}$$

Imputed wage rates for those who are unemployed can be obtained using the expression:

$$E\left(w_i\big|_{E_i=0}\right)= x_i'\hat{\beta} + \hat{\rho}\hat{\sigma}_\varepsilon\left\{\frac{-\phi(z_i'\hat{\gamma})}{1-\Phi(z_i'\hat{\gamma})}\right\} \tag{9}$$

The use of the conditional mean log-wage is perhaps the most obvious choice for the predicted wage. It is also possible, for example, to take a random draw, for each individual, from the relevant conditional distribution. Indeed, in labour supply analyses there is no necessity to be restricted to using observed wage rates for those employed in the sample period: it would also be possible to take random draws from the relevant conditional distributions.

Missing Variables for Non-workers

In the present context, the expression in (9) cannot be used without modification because some variables used in the estimation of the wage functions are not available for non-workers. In addition to the wage rate, neither the occupation nor the industry of non-workers is known. Although these variables could not be included in the selection equations, they were included in the wage equations because of their demonstrated importance in wage determination.

The treatment of this issue can be illustrated using a simplified example of two occupations. Suppose there are two occupations, denoted 0 and 1. Individual wages, in each occupation are given by:

$$w_i = a_0 + u_i \qquad \text{if in occupation 0} \tag{10}$$

$$= a_1 + u_i \qquad \text{if in occupation 1} \tag{11}$$

where u_i is distributed as $N(0,\sigma_u^2)$. Suppose that n_0 and n_1 are the numbers of individuals in each occupation, and $n= n_0 + n_1$. The average wage, $\bar{w}$, is therefore:

$$\bar{w} = a_0\frac{n_0}{n} + a_1\frac{n_1}{n} \tag{12}$$

*This was caused by the multicolinearity arising from the additional terms in this case.

One approach is to give the non-working individuals the sample average, as in equation (12). This can be achieved as follows. Carry out a dummy variable regression of the form:

$$w_i = a_0 + a_1' d_i + u_i \tag{13}$$

where $a_1' = (a_1 - a_0)$ measures the differential effect of being in occupation 1 compared with occupation 0. Consider the following predicted wage, w, for those individuals whose occupation is not known:

$$\hat{w} = a_0 + a_1' \frac{n_1}{n} \tag{14}$$

A little rearrangement shows that $w = \bar{w}$.

However, it is possible, indeed likely, that the distribution across occupations differs between the employed and the unemployed workers. Therefore the above method would impart a bias in the predictions. However, extraneous information may be used to obtain the proportion of non-workers in occupation 1, say. An alternative predictor is, therefore, simply:

$$\hat{w}^* = a_0 + a_1' \frac{n_1^*}{n} \tag{15}$$

Instead of the above approach, a regression may be carried out of the form:

$$w_i = a + a_0 d_{0i} + a_1 d_{1i} + u_i \tag{16}$$

where $d_{0i} = 1$ if the individual is in occupation 0, and $d_{0i} = 0$ otherwise; and $d_{1i} = 1$ if the individual is in occupation 1, and $d_{1i} = 0$ otherwise. Hence, The regression equation (16) can be estimated subject to the constraint that Thus, using (16) can be arranged to give:

$$w_i = a + a_0 (d_{0i} - d_{1i}) + u_i \tag{17}$$

The predicted wage can easily be obtained from (17) for those whose occupation is known. For non-workers, whose occupation is unknown, consider setting both dummies in (17) equal to zero. This gives the predictor, where:

$$\tilde{w} = a \tag{18}$$

However, from (17) the average wage, is given by:

$$\bar{w} = a + a_0 \left(\frac{n_0}{n} - \frac{n_1}{n} \right) \tag{19}$$

A comparison with (18) shows immediately that $\tilde{w}$ is not equal to $\bar{w}$. Hence, the preferred prediction method is to use (15).

Selected Examples

This subsection provides selected examples of predicted wages obtained when unemployed individuals are assigned the sample occupation and industry characteristics. These are compared with the above procedure involving the use of extraneous information on unemployment rates by industry and occupation.

Consider first a female unemployed lone parent with the following characteristics: aged 30 to 34 years; separated/widowed from a previous relationship; European born; residing in the ACT/NT in a non-capital city; with no other income unit income; with two dependent children aged 5-9 years and 10-15 years; in 'other tenure'. The predicted or

imputed wage obtained using (employed) sample averages for industry and occupation groups is found to be $16.91 per hour. Proportions relating to the unemployed, based on extraneous information from the ABS *Labour Force Survey*, are listed in the appendix. These give an imputed wage of $14.98 per hour.

Second, consider a single female with: no children; never married; aged 20 to 24 years; Australian born; residing outside the Sydney metropolitan region in New South Wales; with a vocational qualification; living in 'other tenure' with no other income. The initial imputed hourly wage is found to be $14.18, whereas it is reduced to $13.98 when extraneous information is used.

Third, consider an unemployed single male with: no children; never married; aged 20 to 24 years; Australian born residing outside the Brisbane metropolitan region in Queensland; in rented accommodation. The imputed wage based on sample averages is $15.32, which in this case is increased to $16.54 per hour when the extraneous information is used.

Fourth, 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, and this is only slightly reduced to $13.11 an hour.

Finally, consider an unemployed married male: aged 45 to 49 years with five dependent children (three of which are aged 5 to 9 years, two are aged 10 to 15 years); European born; residing in Melbourne; with a diploma; partner has no formal qualifications and is currently not employed; no other income; own home outright. The basic hourly rate is $20.54 per hour, and this is reduced to $18.24 by the use of extraneous information.

6. Conclusion

This paper has reported estimates of wage equations for Australian workers, using pooled data from the Income Distribution Surveys for 1995 and 1996, the first two years for which continuous hours information is available for each individual. The problem of assigning a wage rate to non-workers, for example in the context of labour supply analysis, was also examined with special attention given to the case where the wage equation includes variables that are not available for the unemployed. The use of extraneous information, in this context regarding the occupation and industry characteristics of the unemployed, was recommended.

Appendix: Summary Statistics

Summary statistics for the various demographic groups are shown in tables A1 and A2. These show the sample averages. In many cases they are dummy variables taking (0,1) values, in which case the tables show the proportions in each category. Two tables are required in view of the fact that the samples used in the selection equation and the wage equations are different.

Information about the last full time job of those unemployed in June 1995, taken from the *Labour Force Survey* (ABS Catalogue, number 6203, table 28), were used to construct the proportions given in table A3.

Table A1 Sample Averages: Selection Equations

Variable	*Single Parents*	*Single Men*	*Single Women*	*Married Women*	*Men Married*
Aged 15 to 19	0.0059	0.1400	0.1643	0.0012	0.0060
Aged 20 to 24	0.0452	0.2891	0.2765	0.0311	0.0542
Aged 25 to 29 (reference)	0.1198	0.1883	0.1661	0.0976	0.1219
Aged 30 to 34	0.1709	0.1173	0.0830	0.1448	0.1507
Aged 35 to 39	0.2220	0.0739	0.0658	0.1596	0.1629
Aged 40 to 44	0.2338	0.0655	0.0490	0.1580	0.1703
Aged 45 to 49	0.1316	0.0459	0.0742	0.1571	0.1650
Aged 50 to 54	0.0452	0.0340	0.0561	0.1185	0.1018
Aged 55 to 59	0.0196	0.0287	0.0415	0.0796	0.0487
Aged 60 to 64	0.0059	0.0137	0.0150	0.0380	0.0132
Separated/widowed		0.0021	0.0044		
Australia (reference)	0.7760	0.1579	0.2164	0.7311	0.7579
Europe/middle East	0.7780	0.8523	0.8631	0.1192	0.1636
Asia	0.1552	0.1001	0.0848	0.0496	0.0542
America/Africa	0.0413	0.0036	0.0362	0.0202	0.0242
Postgraduate qualification	0.0255	0.0112	0.0159	0.0670	0.0544
Undergraduate qualification	0.0628	0.0308	0.0490	0.1114	0.1165
Diploma	0.0943	0.0970	0.1378	0.1197	0.0997
Vocational qualification	0.0845	0.0809	0.0906	0.2918	0.1941
No formal qualifications (reference)	0.2240	0.2324	0.1811	0.4101	0.5350
Non-labour income unit income	0.5324	0.5586	0.5380	0.3020	0.5605
Child support income	0.0190	0.0089	0.0734	0.0000	0.0008
New South Wales (reference)	0.0242	0.0000	0.0535	0.2238	0.2315
Victoria	0.1847	0.2359	0.2266	0.2144	0.2075
Queensland	0.2043	0.2048	0.2403	0.1733	0.1710
South Australia	0.1513	0.1733	0.1731	0.1085	0.1140
Western Australia	0.1198	0.1085	0.1056	0.1378	0.1290
Tasmania	0.1375	0.1435	0.1334	0.0673	0.0653
ACT/Northern Territory	0.0963	0.0571	0.0658	0.0749	0.0817
Capital city	0.1061	0.0770	0.0552	0.6027	0.5982
Number of children	0.5737	0.6220	0.6789	1.2153	1.0425
Youngest child aged 0 to 2	1.5992			0.1833	0.1163
Youngest child aged 3 to 4	0.0982			0.0745	0.0639
Youngest child aged 5 to 9	0.1139			0.1484	0.1410
Youngest child aged 10 to 15	0.2927			0.1188	0.1336
Owned/mortgaged (reference)	0.2868			0.7715	0.7891
Rented	0.4460	0.1827	0.6718	0.2048	0.1874
Other tenure	0.5265	0.5870	0.5455	0.0237	0.0235
Partner employed	0.0275	0.2303	0.2275	0.5753	0.7334
Partner has postgraduate qualification				0.0410	0.0621
Partner has undergraduate qualification				0.1007	0.1108
"Older" than partner				0.1100	0.0097
"Younger" than partner				0.0123	0.1216
University qualification x (20 to 24)		0.0000	0.0000	0.0229	0.0057
University qualification x (25 to 29)		0.0301	0.0508	0.0118	0.0226
University qualification x (30 to 34)		0.0357	0.0743	0.0239	0.0254
University qualification x (35 to 39)	0.0196	0.0228	0.0261	0.0323	0.0355
University qualification x (40 to 44)	0.0373	0.0109	0.0230	0.0380	0.0321
University qualification x (45 to 49)	0.0413	0.0116	0.0088	0.0341	0.0277
University qualification x (50 to 54)	0.0393	0.0077	0.0119	0.0206	0.0145
University qualification x (55 to 59)	0.0118	0.0042	0.0066	0.0114	0.0046
Vocational qualification x (20 to 24)	0.0019	0.0105	0.0265	0.0116	0.0162
Vocational qualification x (25 to 29)	0.0177	0.0914	0.0857	0.0401	0.0335
Vocational qualification x (30 to 34)	0.0393	0.0714	0.0517	0.0559	0.0501
Vocational qualification x (35 to 39)	0.0550	0.0410	0.0190	0.0679	0.0508
Vocational qualification x (40 to 44)	0.0550	0.0294	0.0186	0.0631	0.0478
Vocational qualification x (45 to 49)	0.0805	0.0231	0.0146	0.0682	0.0455
Vocational qualification x (50 to 54)	0.0394	0.0161	0.0208	0.0503	0.0319
Vocational qualification x (55 to 59)	0.0138	0.0137	0.0168	0.0308	0.0129
Vocational qualification x (60 to 64)	0.0059	0.0112	0.0115	0.0178	0.0025

Table A2 Sample Averages: Wage Equations

Variable	*Single Parents*	*Single Men*	*Single Women*	*MarriedWomen*	*Men Married*
Aged 15 to 19	0.0048	0.1264	0.1455	0.0008	0.0045
Aged 20 to 24	0.0361	0.2945	0.2972	0.0307	0.0549
Aged 25 to 29 (reference)	0.1082	0.1956	0.1740	0.1034	0.1249
Aged 30 to 34	0.1659	0.1199	0.0885	0.1499	0.1484
Aged 35 to 39	0.2356	0.0778	0.0677	0.1568	0.1621
Aged 40 to 44	0.2356	0.0672	0.0488	0.1617	0.1740
Aged 45 to 49	0.1370	0.0460	0.0748	0.1572	0.1682
Aged 50 to 54	0.0505	0.0318	0.0570	0.1180	0.1038
Aged 55 to 59	0.0192	0.0254	0.0407	0.0770	0.0435
Aged 60 to 64	0.0072	0.0142	0.0163	0.0339	0.0188
Aged 65 to 69		0.0004	0.0458		
Professional	0.2332	0.1819	0.2096	0.3120	0.2200
Para-professional	0.1034	0.0744	0.1200	0.1075	0.1397
Clerical/sales	0.4159	0.1986	0.5158	0.1495	0.4658
Traditional labour (reference)	0.2476	0.5451	0.1546	0.4310	0.1745
Agriculture/forestry (reference)	0.0216	0.0357	0.0076	0.0282	0.0151
Mining	0.0048	0.0129	0.0031	0.0223	0.0045
Manufacturing	0.1106	0.1986	0.0814	0.2115	0.0979
Construction	0.0144	0.0881	0.0076	0.0804	0.0187
Sales	0.1418	0.2158	0.2106	0.1607	0.1666
Transport	0.0216	0.0563	0.0204	0.0705	0.0214
Communications	0.0240	0.0305	0.0132	0.0291	0.0137
Financial/business sector	0.0937	0.1135	0.1679	0.1320	0.1524
Service industries	0.5577	0.2395	0.4837	0.2456	0.5049
Australia (reference)	0.7885	0.8573	0.8657	0.7339	0.7668
Europe/middle east	0.1442	0.0980	0.0829	0.1974	0.1590
Asia	0.0384	0.0327	0.0356	0.0481	0.0475
America/Africa	0.0228	0.0120	0.1580	0.0205	0.0267
Postgraduate qualification	0.0769	0.0357	0.0534	0.0703	0.0578
Undergraduate qualification	0.0961	0.1066	0.1490	0.1131	0.1236
Diploma	0.0889	0.0825	0.0941	0.1251	0.1030
Vocational qualification	0.2043	0.2382	0.1750	0.2927	0.1933
No formal qualification (reference)	0.5337	0.5365	0.5254	0.3987	0.5223
New South Wales (reference)	0.1827	0.2498	0.2335	0.2247	0.2337
Victoria	0.2019	0.1991	0.2462	0.2155	0.2094
Queensland	0.1490	0.1694	0.1679	0.1714	0.1685
South Australia	0.1082	0.1045	0.1007	0.1070	0.1122
Western Australia	0.1394	0.1423	0.1328	0.1395	0.1275
Tasmania	0.1034	0.0533	0.0641	0.0646	0.0639
ACT/Northern Territory	0.1154	0.0817	0.0549	0.0772	0.0847
Capital city	0.5793	0.6363	0.6948	0.6185	0.6121
University qualification x (20 to 24)		0.0339	0.0544	0.0223	0.0061
University qualification x (25 to 29)		0.0387	0.0504	0.0134	0.0236
University qualification x (30 to 34)	0.0168	0.0258	0.0285	0.0254	0.0267
University qualification x (35 to 39)	0.0457	0.0125	0.0259	0.0323	0.0375
University qualification x (40 to 44)	0.0433			0.0400	0.0335
University qualification x (35 and over)	0.0457	0.0133	0.0092		
University qualification x (45 to 49)	0.0144	0.0082	0.0132	0.0313	0.0301
University qualification x (50 to 54)	0.0072	0.0043	0.0071	0.0211	0.0158
University qualification x (55 to 59)	0.0021	0.0081	0.0116	0.0053	
University qualification x (60 to 64)		0.0258	0.0041	0.0047	0.0026
Vocational qualification x (24 or less)	0.0120				
Vocational qualification x (20 to 24)		0.0941	0.0860	0.0114	0.0174
Vocational qualification x (25 to 29)	0.0264	0.0744	0.0524	0.0427	0.0341
Vocational qualification x (30 to 34)	0.0649	0.0434	0.0285	0.0589	0.0499
Vocational qualification x (35 to 39)	0.0577	0.0309	0.0259	0.0681	0.0494
Vocational qualification x (40 to 44)	0.0697	0.0249	0.0092	0.0648	0.0502
Vocational qualification x (45 to 49)	0.0408	0.0159	0.0132	0.7049	0.0467
Vocational qualification x (50 to 54)	0.0144	0.0146	0.0071	0.0508	0.0335
Vocational qualification x (55 to 59)	0.0072	0.0086	0.0081	0.0303	0.0106
Vocational qualification x (60 to 64)	0	0.0047	0.0041	030158	0.0021

Table A3 Occupation and Industry Proportions: Unemployed June 1995

Category	*Males*	*Females*
Industry Division		
Agriculture, Forestry and Fishing	0.06568	0.03792
Manufacturing	0.24968	0.17465
Construction	0.17768	0.01896
Wholesale Trade	0.03958	0.02595
Retail Trade	0.13684	0.19661
Accommodation, Cafes and Restaurants	0.04968	0.09980
Transport and Storage	0.02894	0.04797
Property and Business Services	0.05684	0.08483
Government Administration and Defence	0.04547	0.04291
Education	0.01979	0.05389
Health and Community Services	0.01389	0.11177
Cultural and Recreational Services	0.01853	0.02894
Personal and Other Services	0.02021	0.03992
Other Industries	0.05010	0.05489
Occupational Group		
Managers and Administrators	0.04755	0.02095
Professionals	0.05597	0.06686
Para-professionals	0.03072	0.04690
Tradespersons	0.22601	0.04291
Clerks	0.04545	0.25149
Sales and Personal Service Workers	0.09932	0.29441
Plant and Machine Operators and Drivers	0.14351	0.04391
Labourers and Related Workers	0.35143	0.23253

References

Australian Bureau of Statistics (1995), *Income and Housing Costs Survey*: Confidentialised Unit Record File. AGPS, Canberra.

Australian Bureau of Statistics (1996), *Income and Housing Costs Survey*: Confidentialised Unit Record File. AGPS, Canberra.

Blundell, R. and Meghir, C. (1987), 'Bivariate alternatives to the Tobit model', *Journal of Econometrics*, 34, 179-200.

Chesher, A. and Irish, M. (1987), 'Residual Analysis in the Grouped and Censored Normal Linear Model', *Journal of Econometrics*, 34, 33-61.

Ermisch, J.F. and Wright, R.E. (1994), 'Interpretation of Negative Sample Selection Effects in Wage Offer Equations', *Applied Economic Letters*, 1, 187-189.

Gourieroux, C., Monfort, A., Renault, E. and Trognon, A. (1987), 'Generalised Residuals', *Journal of Econometrics*, 34, 5-32.

Gourieroux, C., Monfort, A., Renault, E. and Trognon, A. (1987a), 'Simulated Residuals', *Journal of Econometrics*, 34, 201-252.

Greene, W. (1981), 'Sample Selection Bias as a Specification Error: Comment', *Econometrica*, 49, 795-798.

Heckman, J. (1979), 'Sample Selection Bias as a Specification Error', *Econometrica*, 47, 153-161.

Maddala, G.S. (1983), *Limited Dependent and Qualitative Variables in Econometrics*, Cambridge: Cambridge University Press.

McFadden, D. (1987), 'Regression-based Specification Tests for the Multinomial Logit Model, *Journal of Econometrics*, 34, 63-82.

Miller, P. and Rummery, S. (1991), 'Male-female Wage Differentials in Australia: a Reassessment', *Australian Economic Papers*, 30, 50-69.

Olsen, R. (1982), 'Distributional Tests for Selectivity Bias and a More Robust Estimator, *International Economic Review*, 23, 223-240.

Pagan, A. and Vella, F. (1989), 'Diagnostic Tests for Models Based on Individual Data: a Survey, *Journal of Applied Econometrics*, 4, s29-s59.