Unobserved Heterogeneity and Inter-Industry Wage Premiums

Mark N. Harris and Joanne Loundes*
Melbourne Institute of Applied Economic and Social Research
The University of Melbourne

Melbourne Institute Working Paper No. 4/99

ISSN 1328-4991
ISBN 0 7340 1455 4

February 1999

*We are grateful to John Creedy, Peter Dawkins, Alan Duncan, Steven Ellacott, Sebastien Laurent and Marno Verbeek for helpful suggestions. The usual caveats apply.

Melbourne Institute of Applied Economic and Social Research
The University of Melbourne
Parkville, Victoria 3052 Australia
***Telephone* (03) 9344 5288**
***Fax* (03) 9344 5630**
***Email* melb.inst@iaesr.unimelb.edu.au**
***WWW Address* http://www.ecom.unimelb.edu.au/iaesrwww/home.html**

Abstract

In the majority of applied work on the determinants of individual wages, the existence of significant industry wage differentials is typically used as evidence against competitive wage theories. The contention of this paper is that such inter-industry wage premiums are generally a manifestation of unobserved individual heterogeneity. We control for individual heterogeneity by utilising a sample selection panel model, that is, a model that allows for endogenous employment outcomes whilst controlling for unobserved heterogeneity. Estimation of such a selectivity-corrected wage equation using panel data is more computationally demanding than the standard Heckman (1979) cross-section case. As a consequence, there are few empirical examples in the literature.

Contents

1 INTRODUCTION **1**

2 MICRO WAGE EQUATIONS **2**

3 A PANEL SAMPLE SELECTION MODEL **4**

3.1 EMPLOYMENT OUTCOMES 4

3.2 THE WAGE EQUATION 6

3.2.1 Wages with Exogenous Employment 6

3.2.2 Wages with Endogenous Employment and a Simple Test 7

4 THE DATA **10**

5 RESULTS **11**

5.1 CROSS-SECTION ESTIMATES 11

5.1.1 Employment Outcomes 11

5.1.2 The Wage Equation 12

5.2 PANEL ESTIMATES 13

5.2.1 Employment Outcomes 13

5.2.2 The Wage Equation 13

6 CONCLUSION **14**

1 Introduction

In the majority of applied work on the determinants of individual wages, the existence of significant industry wage differentials is typically used as evidence against competitive wage theories. The contention of this paper is that such inter-industry wage premiums are generally a manifestation of unobserved individual heterogeneity. Following Keane *et al*, (1988) and Nijman and Verbeek, (1992) this model controls for individual heterogeneity by utilising a sample selection panel model, i.e. a model that allows for endogenous employment outcomes whilst controlling for unobserved heterogeneity. Estimation of such a selectivity-corrected wage equation using panel data is more computationally demanding than the standard Heckman (1979) cross-section case, and as a consequence, there are few empirical examples in the literature. This is the first application using Australian data.

Use of panel data in general (as opposed to a time-series or cross-section model) can account for a number of difficulties. Firstly, both observed *and* unobserved heterogeneity can be adequately controlled for. Unobserved heterogeneity causes a number of difficulties in empirical work, for example the possibility of making erroneous inferences on the effects of measured variables.[1] Secondly, the chances of a particular cross-section being in some sense atypical are reduced and finally, it enables the incorporation of time varying factors (including important demand-side variables).

The plan of this paper is as follows. Section 2 provides a brief review of the literature on wages and their determinants using micro data. The panel sample selection model is described in Section 3. Section 4 describes the data used and some related estimation issues. The results are discussed in Section 5 and some conclusions are drawn in Section 6.

---

1 There are numerous contexts (in addition to wage rates) in which unobserved heterogeneity plays an important role. Examples include transport policy (Stern, 1991), labour market analysis (Crossley, 1998, Bell and Ritchie, 1996), firm profitability (Haskel and Martin, 1992) and social security provision (Samwick, 1997).

2 Micro Wage Equations

Adequately accounting for unobserved heterogeneity is required to avoid the possibility of misinterpreting results obtained for observable characteristics. The use of panel data is one technique available to control for this problem (see Mátyás and Sevestre, 1996). However, as a consequence of a combination of a paucity of panel data sets and estimation difficulties, cross-sectional studies tend to dominate. A number of studies estimate models that include industry dummies as explanatory variables, and these are invariably found to be significant (see for example Dickens and Katz, 1987, Kao *et al*, 1994, Kim, 1998, Krueger and Summers, 1988, Moghadam, 1990, Murphy and Topel, 1987, Stegwee, 1990).[2] Examples of Australian work on individual wage equations include: union effects on wages (Miller and Mulvey, 1995, Miller, Mulvey and Neo, 1997); gender and union pay gaps (Körösi *et al*, 1993) and; industry premiums (Borland and Suen, 1990, Vella and Woodbridge, 1993). These industry estimates are typically used as evidence against competitive wage theories, that is:

> The essential feature of a perfectly competitive labor market is that workers who accept jobs can expect to receive compensation equal to their opportunity cost. Firms pay a wage that is just sufficient to attract workers of the quality they desire and no higher…If an employee's industry is a significant factor in determining wages after controlling for labor quality and working conditions we must look beyond the standard competitive theory and ask why firms choose to pay workers more than their alternative wage. (Krueger and Summers, 1988, pp. 259, 263)

Some alternative explanations to the competitive model that attempt to explain this phenomenon include efficiency wage theory, the rent-sharing hypothesis, worker sorting and match specific productivity.

Using a sample of both genders from the *1993 Training and Education Survey*, Miller and Mulvey (1995) reported that in Australia, out of the 11 industry variables that were included in their model, only 2 had no apparent impact on the determination of hourly wages. Körösi *et al* (1993) estimate a model in which they examine the hourly

---

[2] Haskel and Martin (1991), through the incorporation of firm characteristics, present evidence of *no* inter-industry wage premiums.

wage gaps of young individuals from the Australian Longitudinal Survey, using the years 1985 through to 1988. They also find significant inter-industry differences, with the (relative) premium for mining workers estimated at around three times as large as that for agricultural workers.

Borland and Suen (1990) and Vella and Woodbridge (1993) look at the existence of industry wage premia in Australia. Utilising a sample of males from the 1986 *Australian Income Distribution Survey*, the former find that of the 11 industry dummies included in their equation, only the wholesale and retail trade sector were an insignificant determinant of relative wages. The fact that these industry dummies are significant even after the inclusion of a range of variables designed to take into account observed worker heterogeneity and working conditions leads them to conclude that significant wage premia exist across industries. The authors suggest that differences in unobserved labour quality across industries are the most likely explanation for these results, as they feel that “non-competitive explanations…appear to be inconsistent with the nature of the wage differentials.” (Borland and Suen, 1993, p 33).

Vella and Woodbridge (1993), using a sample of males from the *Current Population Survey*, suggest that their ‘single factor’ estimate (which represents unobserved attributes including motivation, communication skills and the ability to learn) is able to explain away inter-industry wage differentials. They therefore argue that—while their results are not inconsistent with efficiency wage theory—the observed industry wage premia are the result of individuals of varying skill (that is unobserved worker heterogeneity) sorting themselves into higher paying industries.

This paper adopts a different approach to Vella and Woodbridge (1993) to control for unobserved heterogeneity. Following Keane *et al*, (1988) and Nijman and Verbeek, (1992) individual heterogeneity is controlled for by utilising a sample selection panel model.

3 A Panel Sample Selection Model

3.1 Employment Outcomes

Before any wages can be earned by individual *i*, in period *t*, the individual has to be in some form of paid employment. It is assumed that the individual's observed employment outcome over time is the result of an index of employability, r_{it}^*. This index will vary with personal characteristics (which affect both the supply of, and demand for, labour) z_{it}^* and aggregate variables (affecting the demand for labour) z_i, with unknown weights γ. In addition to z_{it}, where $z_{it} = \left(z_{it}^*, z_i\right)$, there will also be unobserved random individual behaviour over time v_{it}. It is useful to decompose v_{it} into a strictly time and individual varying component η_{it}, and one which is time invariant but which varies across individuals ε_i. In this way we can adequately account for any remaining unobserved heterogeneity across individuals. Assuming a linear relationship yields

$$r_{it}^* = z_{it}'\gamma + v_{it}, \qquad v_{it} = \varepsilon_i + \eta_{it}, \qquad i = 1,\ldots,N \text{ and } t = 1,\ldots,T. \tag{3.1}$$

However r_{it}^* is not directly observed, but the observed realisation of this latent variable r_{it}, which is unity if $r_{it}^* \geq 0$ (the individual is employed) and zero otherwise (the individual is unemployed).[3] As is well known, the usual regression techniques are not appropriate when the dependent variable is dichotomous (Maddala, 1983).

If, as usual, the assumption that the two error terms of (3.1) are independently normally distributed, with respective variances σ_ε^2 and σ_η^2 a *panel probit* model results (Butler and Moffit, 1982), with the associated correlation between composite error terms over time being equal to

$$corr(v_{it}v_{is}) = \rho = \sigma_\varepsilon^2 \big/ \left(\sigma_\varepsilon^2 + \sigma_\eta^2\right) \qquad t \neq s. \tag{3.2}$$

---

3 We do not focus in this paper on those individuals who are either in full time education or not in the labour force.

For identification purposes, a normalisation of the variances has to be made. Choosing $\sigma_{\varepsilon}^{2} = 1$ gives

$$P(r_{it} = 1 | z_{it}, \gamma, \varepsilon_i) = \Phi(z'_{it}\,\gamma/\sigma_\eta + \varepsilon_i/\sigma_\eta), \tag{3.3}$$

where Φ denotes the cumulative distribution function of the normal distribution. Integrating out (or conditioning on) the individual effects yields

$$P(r_{it} = 1 | z_{it}, \gamma) = \int_{-\infty}^{\infty} \Phi(z'_{it}\,\gamma/\sigma_\eta + \varepsilon_i/\sigma_\eta)\phi(\varepsilon_i)d\varepsilon_i, \tag{3.4}$$

where ϕ is the standard normal probability density function. However, from (3.2) the observations corresponding to the same individual are dependent, such that

$$\begin{aligned} P(r_{i1} = 1, \ldots, r_{it} = 1) &= \int_{-\infty}^{\infty} \prod_{t=1}^{T} \Phi(z'_{it}\,\gamma/\sigma_\eta + \varepsilon_i/\sigma_\eta)\phi(\varepsilon_i)d\varepsilon_i \\ &\neq \prod_{t=1}^{T} P(r_{it} = 1)\ . \end{aligned} \tag{3.5}$$

By symmetry for $r_{it} = 0$, yields

$$P(r_i) = \int_{-\infty}^{\infty} \frac{1}{(2\pi)^{1/2}} \exp^{-\varepsilon_i/2} \prod_{t=1}^{T} \Phi\left\{\left[z'_{it}\,\gamma/\sigma_\eta + \varepsilon_i \left(\frac{\rho}{1-\rho}\right)^{1/2}\right][2r_{it} - 1]\right\} d\varepsilon_i, \tag{3.6}$$

where $(\rho/(1-\rho))^{1/2} = 1/\sigma_\eta$ due to the normalisation of σ_{ε}^{2}. Finally, by defining $\tilde{\varepsilon}_i = \varepsilon_i/\sqrt{2}$ and $\theta = (2\rho/1-\rho)^{1/2}$ (3.6) can be written as

$$P(r_i) = \int_{-\infty}^{\infty} \frac{1}{(2\pi)^{1/2}} \exp^{-\tilde{\varepsilon}_i^{2}} \prod_{t=1}^{T} \Phi\{[z'_{it}\,\gamma/\sigma_\eta + \tilde{\varepsilon}_i\theta][2r_{it} - 1]\} d\tilde{\varepsilon}_{\mathrm{i}}. \tag{3.7}$$

The log-likelihood function is then

$$\log L = \sum_i \log P(r_i). \tag{3.8}$$

A major computational problem is that the integral in (7) has no analytical closed form. However, it can be approximated using Gaussian quadrature numerical integration techniques, using the Hermite integration formula

$\int_{-\infty}^{\infty} \exp^{-a^2} g(a)da \approx \sum_{j=1}^{J} w_j g(a_j)$. Under weak regularity conditions, maximisation of (3.8) with respect to θ and γ/σ_η yields consistent and efficient (maximum likelihood) parameter estimates (Butler and Moffit, 1982.).

A fairly standard set of personal demographic variables is expected to enter into z_{it}^* (see for example, Harris, 1996), such as gender, educational achievement and racial origin. An obvious candidate for a macro variable would be the national unemployment rate, although including separate time dummies for each period would simultaneously condition on *all* macroeconomic variables.

Following Chamberlain (1980, 1984), possible correlations of the individual effects and z_{it} are allowed for by including the time average of the time varying variables $\bar{z}_i$ (experience) as an additional explanatory variable.

3.2 The Wage Equation

3.2.1 Wages with Exogenous Employment

Having allowed the employment outcome to be endogenous, attention is now turned to the wage equation. Adopting a standard human-capital approach, wages are postulated to be a function of a similar set of personal characteristics, such as education and age/experience. In addition, both occupation and (potentially) industry are likely to affect wages, as will any important macroeconomic variables (capturing demand-side effects). Thus

$$y_{it} = x'_{it}\beta + e_{it},\ e_{it} = \mu_i + u_{it} \qquad i = 1,\ldots,N \text{ and } t = 1,\ldots,T_i, \qquad (4.1)$$

where y_{it} is a measure of wages; x_{it} a vector of explanatory variables with unknown weights β; μ_i a random individual effect (again to account for unobserved heterogeneity) and; u_{it} the usual error term. The two error terms are assumed to be independently normally distributed with zero means and respective variances σ_μ^2 and σ_u^2. Note that only if all individuals are employed for the full duration of their recorded observations, will $T_i = T\ \forall\ i$ and there is no reason to expect $t_i,\ldots,T_i$ to be contiguous.

3.2.2 Wages with Endogenous Employment and a Simple Test

The problem with equation (4.1) is however, that it ignores the potential endogeneity of employment and the consequent selection bias. In the standard cross-sectional approach this has been well recognised and adequately dealt with and consists of augmenting the regression equation with an additional regressor which takes into account the correlation between the two error terms in the probit and regression equations (see the seminal paper by Heckman, 1979). However, in a panel data setting there are two error terms in both equations and therefore two, not one, additional correction terms required – the conditional expectations of μ_i and u_{it} *given* selection (here, "selection" refers to employment). The covariances between ε_i and μ_i $\left(\sigma_{\varepsilon\mu}\right)$ and η_{it} and u_{it} $\left(\sigma_{\eta u}\right)$ are the parameters for these correction terms. In general, this additional correction term is invariably ignored, resulting in potentially biased parameter estimates and erroneous inference.

Nijman and Verbeek (1992) show that $E\left\{\mu_i \middle| \underline{r}_i\right\} = \sigma_{\varepsilon\mu} A_{1i}$ and $E\left\{u_{it} \middle| \underline{r}_i\right\} = \sigma_{\eta u} A_{2it}$, where

$$A_{1i} = \frac{1}{\sigma_\eta^2 + T\sigma_\varepsilon^2} \sum_{s=1}^{T} E\left\{\varepsilon_i + \eta_{is} \middle| \underline{r}_i\right\} \tag{4.2}$$

and

$$A_{2it} = \frac{1}{\sigma_\eta^2}\left[E\left\{\varepsilon_i + \eta_{is} \middle| \underline{r}_i\right\} - \frac{\sigma_\varepsilon^2}{\sigma_\eta^2 + T\sigma_\varepsilon^2} \sum_{s=1}^{T} E\left\{\varepsilon_i + \eta_{is} \middle| \underline{r}_i\right\}\right]. \tag{4.3}$$

Unlike the usual cross-sectional case, evaluation of these correction terms is not straightforward, as they require evaluation of $E\left\{\varepsilon_i + \eta_{it} \middle| \underline{r}_i\right\}$ which is given by

$$E\left\{\varepsilon_i + \eta_{it} \middle| \underline{r}_i\right\} = \int_{-\infty}^{\infty} \left[\varepsilon_i + E\left\{\eta_{it} \middle| \underline{r}_i, \varepsilon_i\right\}\right] f\left(\varepsilon_i \middle| \underline{r}_i\right) d\varepsilon_i, \tag{4.4}$$

where

$$E\{\eta_{it}|\underline{r}_i,\varepsilon_i\}=(2r_{it}-1)\sigma_\eta \frac{\Phi\left(\dfrac{z_{it}'\gamma+\varepsilon_i}{\sigma_\eta}\right)}{\Phi\left[(2r_{it}-1)\left(\dfrac{z_{it}'\gamma+\varepsilon_i}{\sigma_\eta}\right)\right]}$$

and

$$f(\varepsilon_i|\underline{r}_i)=\frac{\prod_{s=1}^{T}\Phi\left[(2r_{is}-1)\left(\dfrac{z_{it}'\gamma+\varepsilon_i}{\sigma_\eta}\right)\right]\dfrac{1}{\sigma_\varepsilon}\phi(\varepsilon_i/\sigma_\varepsilon)}{\int_{-\infty}^{\infty}\prod_{s=1}^{T}\Phi\left[(2r_{is}-1)\left(\dfrac{z_{it}'\gamma+\varepsilon}{\sigma_\eta}\right)\right]\dfrac{1}{\sigma_\varepsilon}\phi(\varepsilon/\sigma_\varepsilon)d\varepsilon}.$$

Although equations (4.2) to (4.4) are simplified somewhat by the normalising assumption of $\sigma_\varepsilon^2=1$, the evaluation of (4.4) - and hence the two necessary correction terms – requires a significant amount of numerical integration. Firstly, it is required to obtain first round estimates of γ and σ_η, where $\sigma_\eta^2=((1-\rho)/\rho)$. Secondly, even with these parameter estimates in hand, equation (4.4) can only be evaluated numerically, again requiring Gaussian quadrature and Hermite integration. Such computational burden has meant that there are very few such empirical examples of correctly specified sample selection panel models in the literature.

Once A_{1i} and A_{2it} have been estimated they are simply added to equation (4.1) as additional regressors, and consistent parameter estimates are obtained by performing OLS or GLS on the augmented model

$$y_{it}=x_{it}'\beta+A_{it}^*\sigma+e_{it}, \qquad i=1,\ldots,N \text{ and } t=1,\ldots,T_i, \qquad (4.5)$$

where $A_{it}^*=(A_{1i}\otimes\iota_T,A_{2it})$ and $\sigma=(\sigma_{\varepsilon\mu},\sigma_{\eta u})'$.[4] A test for an ignorable selection rule can be based on GLS parameter estimates and standard errors that are valid under the

---

4 Note that for the dimensions to be correct, the appropriate rows of A^*, and indeed x_{it}, have to be deleted when y_{it} is not observed.

null hypothesis of an ignorable selection rule. Thus, standard *F*- or Wald tests for joint significance of σ can be used based on GLS estimates.

Estimating equation (4.5) by feasible GLS requires an estimate of the covariance matrix of $e_i = Var\left(\mu_i \iota_{T_i} + u_i\right) = \Omega_i$ (where ι_{T_i} is a T_i vector of ones). By (assumed) independence, this will be a function of the two respective variances σ_μ^2 and σ_u^2, although its dimensions will vary across *i* such that

$$\Omega_i = \sigma_\mu^2 + \sigma_u^2, \qquad \forall T_i = 1 \tag{4.6}$$

and

$$\Omega_i = \sigma_\mu^2 J_{T_i} + \sigma_u^2 I_{T_i}, \qquad \forall T_i \geq 1, \tag{4.7}$$

where J_{T_i} is a $T_i \times T_i$ matrix of ones and I_{T_i} the identity matrix of order T_i. Consistent estimates of σ_μ^2 and σ_u^2 can be based on quadratic estimates of the *Within* and *Between* residuals (see, for example, Hsiao, 1985) as

$$\hat{\sigma}_u^2 = \frac{1}{H^*} \sum_{i, T_i > 1} \sum_{t = t_i}^{T_i} \left[\left(y_{it} - \bar{y}_{i.}\right) - \left(x_{it} - \bar{x}_{i.}\right)' \hat{\beta}_{Within} \right]^2 \tag{4.8}$$

and

$$\hat{\sigma}_\mu^2 = \frac{1}{N^*} \sum_{i, T_i > 0} \left[\left(\bar{y}_{i.} - \bar{x}_{i.}' \hat{\beta}_{Between}\right)^2 - \frac{1}{T_i} \hat{\sigma}_u^2 \right], \tag{4.9}$$

where $H^* = \sum_i T_i$, $\forall T_i > 1$ and N^* is the number of individuals who were observed *at least once* earning wages.[5] Stacking the time-series observations for the *i*th

---

[5] The *Within*, or fixed effects dummy variable estimator, is obtained by transforming all variables into deviations from individual time means and then applying OLS, so that T_i must be greater than 1. The *Between* estimator is obtained by applying OLS to the time means of all variables for each individual, so that T_i must be greater than 0, which requires removal of all individuals who were unemployed for the full duration of the sample. Also, any time invariant variables for individuals

observation as $X_i = \left(x_{it_i}, \ldots, x_{iT_i}\right)$ and $Y_i = \left(y_{it_i}, \ldots, y_{iT_i}\right)$ the FGLS estimates are obtained as

$$\hat{\beta}_{FGLS} = \left(\sum_{i=1}^{N^*} X_i' \hat{\Omega}_i^{-1} X_i\right)^{-1} \left(\sum_{i=1}^{N^*} X_i' \hat{\Omega}_i^{-1} Y_i\right), \tag{4.10}$$

with covariance matrix

$$\operatorname{Var}\left(\hat{\beta}_{FGLS}\right) = \left(\sum_{i=1}^{N^*} X_i' \hat{\Omega}_i^{-1} X_i\right)^{-1}. \tag{4.11}$$

4 The Data

The Australian Longitudinal Survey (ALS) provides one of the few panel data sets available in Australia. The data are taken from the ALS, years 1985 – 1988 and have been described in detail by, for example, Miller, 1989; Chapman and Smith; 1993 and; Körösi *et al*, 1993. The data consist of young persons whose age was between 16 and 24 at the inception of the survey.

An important aspect in analysing panel data sets is that of endogenous or exogenous attrition. Following Harris (1996), attention is restricted to "low" education males (defined as Year 11 or below).[6] Most of the data were entered as zero-one indicator (dummy) variables. The most notable exception was the calculation of the replacement ratio. The replacement ratio proxies a reservation wage, and is calculated as the ratio of the unemployment benefit entitlement to expected earnings. In this

(for example, gender) must be discarded and that under the null hypothesis the correction terms are insignificant and therefore are also excluded from X_i.

6 Splitting the sample avoids any sample selection bias as a result of perceived endogenous attrition. In addition to the pursuit of parsimony, the subsequent analysis focuses on the lower educated males due to the difficulties associated with adequately defining an experience variable for females and, importantly, *a priori* it is expected that these individuals are more likely to be affected by industry differences. Results for the other demographics yield broadly similar conclusions, and are available on request from the authors.

paper, the unemployment benefit is calculated using personal characteristics to determine the size of the unemployment benefit an individual is entitled to. We do not estimate other entitlements. Earnings are proxied by averaging weekly earnings for the low education male sub-sample.[7] In accordance with existing knowledge of the relationship between age and the incidence of unemployment (see, for example Miller, 1989), experience was entered in exponential form as the survey year minus the year the individual left school. Due to this parameterisation, a negative relationship between experience and employment/wages is expected.

5 Results

5.1 Cross-section estimates

The cross-section and panel estimates of the employment outcome and wage equation are given in tables 1 and 2.

5.1.1 Employment Outcomes

Employment does not appear to be unduly affected by whether an individual resides in metropolitan or regional areas, although there may be some weak evidence of a positive influence on the employment outcome for rural-dwelling men (Table 1, Column 4). As one would expect, completing year 10 or 11 has a positive impact on the employment probability in the majority of the cross-section results, with the coefficient on education to year 11 being approximately double that of year 10. Being of Western origin also has a positive impact on employment probabilities (see also Miller, 1989).

The partner's employment status made a significant positive contribution to individual employment only in the 1985 sample. Experience was a significant determinant of employment prospects only in the 1988 sample, and had the expected negative sign

---

[7] However, this effectively renders the effect of the ratio to be dominated by movements in the unemployment benefit entitlement.

(as mentioned in the data description section). A common negative influence on employment prospects across samples was having some kind of disability.

5.1.2 The Wage Equation

Previous Australian studies (Borland and Suen, 1990 and Körösi *et al*, 1993) have indicated that males working in the mining sector are likely to enjoy significantly positive wage effects. This is unsurprising, given that in November 1988 around 90 per cent of people working in the mining sector were males and total weekly earnings were around 38 per cent higher than their nearest competitor.[8] Indeed, the cross-sectional results presented in this paper give much the same answer (with the exception of the 1985 sample).

In contrast to most other studies, a relatively limited number of other industries were significant. Individuals were relatively better off (in terms of wages) if they worked in the transport or services industries, but relatively worse off if they worked in the sales or finance industries (Table 2). In the following section, these unobservable differences are accounted for by utilising the panel nature of the data.

As one would expect, pay per hour tended to increase with experience, although at a decreasing rate (the 1988 results were the exception here). Wage outcomes did not appear to be unduly affected by whether an individual resided in metropolitan or regional areas. In line with evidence cited elsewhere (see for example, Christie, 1992; Körösi *et al*, 1993) unions had a significant positive effect on the wage rates of males. Somewhat surprisingly, schooling up to year 11 only had a positive impact on wages in the case of the 1985 example. As expected, and managers and professionals yielded occupation specific benefit for wages, and apprentices experienced worse wage outcomes than the sample as a whole in the 1985 and 1987 samples.

The inverse Mills ratio accounts for the selection bias in the wage equation. The significant negative coefficient on this variable in the 1988 sample suggests that,

---

8 Employment numbers are from ABS Cat. No. 6203.0, *The Labour Force, Australia*. Earnings figures are from ABS Cat. No. 6302.0, *Average Weekly Earnings, Australia*.

given the characteristics of low education males, actual wage offers are relatively homogenous compared to their reservation wages.[9]

5.2 Panel estimates

This subsection presents the panel estimates for both the employment outcomes and the wage equation.

5.2.1 Employment Outcomes

Similarly to the cross section results, whether or not an individual has completed year 10 or 11 has a positive impact on the employment probability (Tables 1 and 2). The partner's workforce status also made a strong contribution, which may explain in part the negative (albeit insignificant) sign on the married variable; that is, a two income family is less likely to require one of the individuals to find employment as compared to a one income family. Experience also continued to exert a positive influence on employment in the panel sample. Individual characteristics that appeared to lessen the probability of employment included residing in a rural area and living rent-free. Two of the time dummies—the macroeconomic proxies—were weakly significant.

5.2.2 The Wage Equation

In direct contrast to the results outlined above, the panel estimates suggest that there are *no* significant industry specific wage benefits. Even the mining industry—where one might expect a premium for compensating differentials—was found to be insignificant.[10]

This is an important finding. Typically, the existence of industry wage premiums has been used as an argument against competitive wage determination (see, for example,

---

9 See Ermisch and Wright (1994) for an exposition on why a negative inverse Mills ratio is not necessarily a cause for concern.

10 Using robust standard errors (see the Appendix). A test for $\sigma_{\varepsilon\mu} = \sigma_{\eta u} = 0$ seems to suggest that the employment and wage equations are in fact independent.

Haskel and Martin, 1991; Kim, 1998). The fact that no premiums were found tends to invalidate this assertion. However, there is still the potential for non-competitive wage theories to explain the determination of wages in Australia given that trade unions and macroeconomic variables were also found to be significant determinants of wages.

An interesting result is that there was some evidence of racial disadvantage in the wage rates that individuals received. This is important given that the model controls for unobserved heterogeneity. As expected pay per hour tended to increase at a decreasing rate with experience and schooling up to year 11 had a positive impact on pay per hour. Belonging to a trade union retained its strong influence on wages, even after moving to panel estimation. In contrast to the cross-section results, managers no longer received a pay premium for their position, with professionals yielding the only occupation specific benefit for wages. This would suggest that the unobserved characteristics of managers account for the premiums witnessed in the cross-section results. The macroeconomic time dummies were all positive and significant (unlike the results for the employment equation).

6 Conclusion

Typically wage equations estimated using cross-sectional data appear to provide evidence of significant inter-industry wage premiums. This paper suggested that this significance could be spurious, the result of not adequately controlling for unobserved individual heterogeneity. This premise was illustrated using an Australian panel data set, whereby wage equations were estimated separately on both the single cross-sections and also on the pooled data. A sample selection model was utilised to jointly model employment probabilities and wages conditional on employment to allow for unobserved heterogeneity. Although this is computationally burdensome, it is very important to avoid biased estimates and erroneous policy inference. Due to a combination of a lack of knowledge and of the complexities involved, this procedure has found very few applications in the literature and as yet none in an Australian context. The results suggest that on separate cross-sections, industry effects may be erroneously significant, but not so once unobserved heterogeneity has been controlled for. Thus applied researchers and policy makers need to take care in concluding that significant industry wage effects exist based solely on cross-sectional results.

Table 1. Employment Equation: Cross Section and Panel Estimates

	1985		1986		1987		1988		1985-88	
Constant	0.11	(0.143)	-0.04	(-0.040)	2.18	-0.964	-0.27	(-0.161)	-0.10	(-0.101)
Experience	0.06	(1.465)	0.05	(1.237)	0.03	-0.604	-0.10	(-1.707)**	-8.12	(-2.727)*
Experience2	--	--	--	--	--	--	--	--	--	--
Married	-0.35	(-0.661)	0.55	(0.751)	1.60	(0.825)	-1.38	(-0.838)	-0.17	(-0.370)
Separated	4.41	(0.001)	3.39	(0.002)	0.65	-0.922	-0.58	(-1.134)	-0.10	(-0.197)
City dwelling	-0.02	(-0.097)	0.16	(0.759)	0.14	-0.7	-0.06	(-0.265)	0.11	-0.747
Rural dwelling	0.50	(1.394)	0.70	(1.779)**	0.38	-0.703	0.47	-0.773	-0.44	(-2.207)*
Buying accommodation	4.20	(0.004)	0.56	(0.715)	0.68	-0.542	0.47	-0.551	0.58	-1.469
Rent-free accommodation	-0.78	(-1.386)	0.34	(0.436)	-0.67	(-0.575)	-0.57	(-0.683)	-0.66	(-1.684)**
Renting accommodation	0.18	(0.346)	0.63	(0.840)	-0.44	(-0.382)	-0.34	(-0.463)	0.25	-0.667
Western origin	0.64	(1.745)**	0.45	(0.818)	0.79	(1.899)**	0.87	(2.056)*	0.75	-1.476
Year 10	0.42	(2.056)*	0.42	(1.907)**	0.57	(2.570)*	0.29	-1.215	0.51	(2.253)*
Year 11	0.80	(3.355)*	0.99	(3.586)*	0.92	(3.229)*	0.78	(2.444)*	0.68	(2.645)*
Partner's employment status	1.68	(1.897)**	0.77	(1.463)	1.40	(1.118)	1.09	(1.615)	0.48	(2.078)*
Replacement ratio	-1.97	(-1.425)	-2.29	(-0.991)	-7.40	(-1.050)	4.97	-0.829	-0.64	(-0.422)
Disabled	-0.36	(-1.624)	-0.66	(-2.725)*	-0.66	(-2.665)*	-0.74	(-2.507)*	-0.14	(-0.783)
Number of children	0.05	(0.202)	-0.01	(-0.034)	-0.13	(-0.690)	0.03	-0.148	0.07	-0.473
1986	--	--	--	--	--	--	--	--	-0.19	(-0.871)
1987	--	--	--	--	--	--	--	--	-0.64	(-1.673)**
1988	--	--	--	--	--	--	--	--	-0.90	(-1.688)**
Average experience	--	--	--	--	--	--	--	--	9.36	(3.155)*
ρ	--	--	--	--	--	--	--	--	0.68	(18.689)*

t-statistics in (.). *, ** Denotes significant at 5% and 10% (2-tailed) level respectively.

Table 2. Wage Equation Correcting for Selectivity: Cross Section and Panel Estimates

	1985			1986			1987			1988			1985-88		
Constant	1.02	(7.327)*	[7.146]*	1.05	(5.743)*	[5.303]*	1.54	(10.758)*	[10.871]*	1.83	(10.771)*	[10.482]*	4.63	(5.109)*	[1.651]**
Experience	0.19	(9.201)*	[9.381]*	0.20	(7.281)*	[7.464]*	0.13	(4.836)*	[4.961]*	0.03	(0.868)	[0.866]	8.15	(3.537)*	[1.216]
Experience2	-0.01	(-5.771)*	[-5.879]*	-0.01	(-5.133)*	[-5.270]*	-0.01	(-3.464)*	[-3.558]*	0.00	(-0.250)	[-0.250]	-12.67	(-7.066)*	[-2.481]*
City dwelling	-0.05	(-1.411)	[-1.433]	0.02	(0.468)	[0.466]	0.04	(1.425)	[1.439]	0.03	(1.062)	[1.009]	-0.06	(-0.428)	[-0.298]
Rural dwelling	0.02	(0.421)	[0.425]	0.03	(0.425)	[0.406]	0.07	(1.398)	[1.382]	0.05	(0.798)	[0.738]	-0.06	(-0.275)	[-0.149]
Western origin	0.04	(0.441)	[0.426]	0.07	(0.704)	[0.695]	-0.02	(-0.180)	[-0.183]	0.08	(0.803)	[0.780]	0.98	(1.810)**	[2.015]*
Year 10	0.06	(1.215)	[1.239]	-0.02	(-0.345)	[-0.318]	-0.01	(-0.316)	[-0.320]	0.01	(0.299)	[0.284]	0.38	(1.468)	[1.167]
Year 11	0.16	(2.865)*	[2.906]*	0.03	(0.311)	[0.278]	0.02	(0.386)	[0.390]	0.02	(0.272)	[0.256]	0.83	(2.887)*	[2.532]*
Disabled	0.00	(-0.078)	[-0.078]	-0.02	(-0.319)	[-0.279]	-0.01	(-0.270)	[-0.271]	0.03	(0.596)	[0.563]	0.12	(0.676)	[0.480]
Trade union	0.09	(3.096)*	[3.177]*	0.15	(4.567)*	[4.695]*	0.10	(3.877)*	[3.997]*	0.10	(3.667)*	[3.743]*	0.66	(5.847)*	[2.503]*
Apprentice	-0.16	(-3.477)*	[-3.599]*	-0.10	(-0.569)	[-0.587]	-0.39	(-3.499)*	[-3.646]*	-0.06	(-1.347)	[-1.360]	0.14	(0.755)	[0.395]
Mining	0.12	(1.064)	[1.101]	0.37	(2.910)*	[3.019]*	0.26	(2.668)*	[2.759]*	0.19	(1.977)*	[2.028]*	2.13	(4.473)*	[1.184]
Manufacturing	0.03	(0.519)	[0.537]	0.10	(1.531)	[1.575]	0.02	(0.388)	[0.398]	0.06	(1.111)	[1.142]	0.24	(1.096)	[0.742]
Construction and Utilities	0.11	(1.595)	[1.644]	-0.02	(-0.228)	[-0.235]	0.06	(0.890)	[0.915]	0.09	(1.411)	[1.449]	0.22	(0.855)	[0.561]
Sales	-0.06	(-0.999)	[-1.033]	0.00	(-0.063)	[-0.065]	-0.13	(-2.333)*	[-2.393]*	-0.01	(-0.235)	[-0.240]	-0.02	(-0.064)	[-0.045]
Transport	0.17	(2.483)*	[2.564]*	0.09	(0.995)	[1.024]	0.00	(-0.004)	[-0.004]	-0.07	(-0.802)	[-0.807]	0.45	(1.442)	[0.946]
Communication	0.15	(1.428)	[1.480]	0.12	(1.032)	[1.063]	0.09	(0.763)	[0.789]	0.08	(0.574)	[0.597]	0.49	(0.976)	[0.722]
Finance	-0.13	(-1.620)	[-1.667]**	0.05	(0.452)	[0.466]	-0.03	(-0.406)	[-0.417]	0.00	(0.019)	[0.019]	0.25	(0.717)	[0.492]
Services	0.14	(2.150)*	[2.218]*	0.04	(0.569)	[0.585]	-0.08	(-1.356)	[-1.392]	-0.02	(-0.285)	[-0.292]	0.08	(0.315)	[0.182]
Managers	0.03	(0.441)	[0.455]	0.20	(2.759)*	[2.844]*	0.16	(2.647)*	[2.714]*	0.15	(2.627)*	[2.663]*	0.62	(2.743)*	[1.350]
Professionals	0.11	(0.628)	[0.649]	0.11	(0.649)	[0.673]	0.00	(0.008)	[0.008]	0.26	(2.807)*	[2.869]*	1.07	(2.308)*	[2.308]*
Clerks and salespersons	0.04	(1.154)	[1.187]	0.03	(0.670)	[0.690]	0.03	(0.938)	[0.965]	0.02	(0.674)	[0.685]	0.05	(0.333)	[0.261]
Inverse Mills ratio	0.07	(1.002)	[0.980]	-0.12	(-0.640)	[-0.550]	-0.10	(-1.077)	[-1.080]	-0.25	(-2.064)*	[-1.967]*	--	--	--
1986	--	--	--	--	--	--	--	--	--	--	--	--	0.34	(2.890)*	[2.925]*
1987	--	--	--	--	--	--	--	--	--	--	--	--	1.14	(8.872)*	[7.876]*
1988	--	--	--	--	--	--	--	--	--	--	--	--	1.70	(11.675)*	[8.596]*
$\sigma_{\varepsilon\mu}$	--	--	--	--	--	--	--	--	--	--	--	--	-0.01	(-0.237)	[-0.007]
$\sigma_{\eta\upsilon}$	--	--	--	--	--	--	--	--	--	--	--	--	0.00	(-1.094)	[-0.007]

t-statistics in (.), robust t-statistics in [.]. *, ** Denotes significant at 5% and 10% (2-tailed) level respectively.

References

Bell, D. and Ritchie, F. (1996) *Female earnings and gender differentials in Great Britain 1977-1990*, University of Stirling. Dept. of Economics Discussion paper in economics No. 96/9.

Borland, J. and Suen, A., (1990) ‘The Determinants of Individual Wages in Australia: Competitive and Non-Competitive Influences’, *The Australian Economic Review*, 4th Quarter, pp. 33-44

Butler, J.S. and Moffit, R. (1982) ‘A Computationally Efficient Quadrature Procedure for the One Factor Multinomial Probit Model’, *Econometrica*, Vol. 50, no. 3, pp. 761-764.

Chamberlain, G. (1980) ‘Analysis of Covariance with Qualitative Data’, *Review of Economic Studies*, 47, pp. 225 – 38.

-- (1984) ‘Panel Data’ in Griliches, Z. and Intriligator, M., (eds.) *Handbook of Econometrics*, Vol. II, pp. 1247 – 1318.

Chapman, B.J. and Smith, P.N. (1993), ‘Predicting the Long-Term Unemployed: A Primer for the Commonwealth Employment Service’, Centre for Economic Policy Research, Australian National University, discussion paper 285.

Christie, V. (1992) ‘Union Wage Effects and the Probability of Union Membership’, *The Economic Record*, Vol. 68, pp. 43 – 56.

Crossley, T. (1998) *What can we learn from displaced worker data about the returns to tenure?* ANU. Faculty of Economics and Commerce Working papers in economics and econometrics No. 346

Dickens, W. and Katz, L. (1987) “Inter-industry wage differences and industry characteristics” in Lang, K and Leonard, S. *Unemployment and the structure of labour markets*, Basil Blackwell.

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

Harris, M. N. (1996) ‘Modeling the Probability of Youth Unemployment in Australia’, *The Economic Record*, Vol. 72, pp. 118 – 29.

Haskel, J. and Martin, C. (1991) ‘Non-Competitive Wage Determination, Firms and the Inter-Industry Wage Structure’, QMW University of London Department of Economics Working Paper no. 235.

—— (1992) *Capacity and competition: empirical evidence on UK panel data,* Queen Mary and Westfield College. Dept. of Economics Paper No. 251

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

Hsiao, C. (1985), ‘Benefits and Limitations of Panel Data’, *Econometric Reviews*, Vol. 4, pp. 21-174.

Hsiao, C. (1986), *Analysis of Panel Data*, Cambridge University Press, Cambridge, England.

Kao, C., Polachek, S. W., and Wunnava, P. V., (1994) ‘Male-Female Wage Differentials in Taiwan: A Human Capital Approach’, *Economic Development and Cultural Change,* 42(2), pp. 351-74.

Keane, M., Moffitt, R. and Runkle, D. (1988), ‘Real Wages over the Business Cycle: Estimating the Impact of Heterogeneity with Micro Data’, *Journal of Political Economy,* Vol. 96, pp. 1232-1266.

Kim, D. I. (1998) ‘Reinterpreting Industry Premiums: Match-Specific Productivity’, *Journal of Labor Economics,* Vol. 16 No. 3, pp. 479-504.

Körösi, G., Rimmer, R. J. and Rimmer, S. M. (1993) *Contributions from Gender and Unions to Earnings Differences Among Young Australians: The Analysis of a Panel,* Discussion Paper 28/93, School of Economics and Commerce, La Trobe University.

Krueger, A. and Summers, L. (1988) “Efficiency wages and the inter-industry wage structure” *Econometrica,* 56, March, pp. 259-93

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

Mátyás, L. and Sevestre, P. (eds) (1996), *The Econometrics of Panel Data*, Kluwer Academic Publishers, The Netherlands.

Miller, P.W. (1989), ‘The Structure of Aboriginal and Non-Aboriginal Youth Unemployment’, *Australian Economic Papers*, Vol. 28, no. 52.

Miller, P.W. and Mulvey, C., (1995) ‘Unions, Firm Size and Wages’, University of Western Australia Organisational and Labour Studies Discussion Paper no. 6/95.

Miller, P.W. Mulvey, C. and Neo, M.L., (1997) ‘Unions and Wages: A Further Analysis’ University of Western Australia Organisational and Labour Studies Discussion Paper no. 1/97.

Moghadam, R. (1990) *Wage Determination: An Assessment of Returns to Education, Occupation, Region and Industry in Great Britain*, London School of Economics Centre for Economic Performance Discussion Paper 8, August.

Murphy, K. and Topel, K. (1987) “Unemployment, risk and earnings: testing for equalizing wage differences in the labor market” in Lang, K and Leonard, S. *Unemployment and the structure of labour markets,* Basil Blackwell.

Nijman, T. and Verbeek, M. (1992), ‘Nonresponse in Panel Data: The Impact on Estimates of a Life Cycle Consumption Function’, *Journal of Applied Econometrics*, Vol. 7, pp. 243-257.

Samwick, A. (1997) *Discount rate heterogeneity and social security reform*, NBER Working paper series No. 6219.

Stegwee, M. G. (1990) *Efficiency Wages, Unemployment and Productivity in U.K. Manufacturing*, Oxford Applied Economics Discussion Paper 89, February.

Stern, S. (1991) *Discrete choice models with unobserved heterogeneity: demand for transportation by elderly and disabled people*, University of Virginia. Thomas Jefferson Centre for Studies in Political Economy Discussion paper No. 240

Vella, F. and Verbeek, M. (1996), 'Two-Step Estimation of Simultaneous Equation Panel Data Models with Censored Endogenous Variables', Centrum voor Economische Studien, Katholieke Universiteit Leuven, Discussion Paper nr. 96.02.

Vella, F. and Woodbridge, G. (1993) 'A Single Factor Explanation of Inter-Industry Wage Differentials' University of Melbourne Department of Economics Research Paper no. 380.

Appendix

Estimation of Standard Errors with Endogenous Participation

If the selection rule is found to be non-ignorable, that is selection is endogenous, the FGLS standard errors are invalid as they ignore the fact that the correction terms have been "generated" and that the conditional distribution of μ_i and u_{it} is heteroscedastic. Correct standard errors can be obtained by using the general formulas in Vella and Verbeek, 1996. Defining: $\hat{\gamma}^*$ as the maximum likelihood estimates of $\gamma^* = (\gamma, \rho)$ from the first stage probit equation, with asymptotic covariance matrix V_2; G_i as

$$G_i = \frac{\partial(X_i'\beta)}{\partial\beta'} = X_i';$$

$$M_N = \frac{1}{N}\sum_{i=1}^{N} E\left\{\left[G_i, A_i^*\right]'\left[G_i, A_i^*\right]\right\};$$

$$V_N = \frac{1}{N}\sum_{i=1}^{N} E\left\{\left[G_i, A_i^*\right]'\Omega_i^*\left[G_i, A_i^*\right]\right\};$$

where Ω_i^* is the covariance matrix of e_i under the alternative hypothesis of a non-ignorable selection rule and;

$$D_N = \frac{1}{N}\sum_{i=1}^{N} E\left\{\left[G_i, A_i^*\right]'\left[\frac{\partial A_i^*\sigma}{\partial\gamma^*}\right]\right\}.$$

The covariance matrix of the second stage estimator is

$$V_1 = \lim_{N\to\infty}\frac{1}{N} M_N^{-1}\left(V_N + D_N V_2 D_N'\right) M_N^{-1}.$$

where expectations are replaced by sample moments and parameters by consistent estimates. An autocorrelation and heteroscedastic consistent estimator of V_N is

$$V_N = \frac{1}{N}\sum_{i=1}^{N}\left\{\left[\hat{G}_i, \hat{A}_i^*\right]'\hat{e}_i\hat{e}_i'\left[\hat{G}_i, \hat{A}_i^*\right]\right\},$$

where $\hat{e}_i$ is the T_i vector of FGLS residuals.