UNIVERSIDAD CARLOS III DE MADRID working papers

Working Paper 08-22
Economic Series (14)
May 2008

Departamento de Economía
Universidad Carlos III de Madrid
Calle Madrid, 126
28903 Getafe (Spain)
Fax (34) 916249875

Estimating Union Wage Effects in Great Britain During 1991-2003[1,2]

Georgios Marios Chrysanthou
Department of Economics
Universidad Carlos III de Madrid
28903, Madrid, Spain
E-mail: gchrysan@eco.uc3m.es

Abstract

Using a dynamic model of unionism and wage determination we find that the unobserved factors that influence union membership also affect wages. The estimates suggest that UK trade unions still play a non-negligible, albeit diminishing, role in wage formation. It appears that the greater impact of un observables in determining individual union propensity concerning the second period under analysis, versus past unionisation experience, implies that those remaining in unions during (1997-2002) gain most from their sorting decision. The significant contribution of unobserved heterogeneity renders the total union wage differential highly variable across individuals. The endogeneity correction procedure employed yields a discernible pattern of the estimated union wage effect relative to OLS and Fixed effects. This is in line with Robinson (1989a) and Vella and Verbeek (1998) and refutes the pessimistic conclusions reached by Freeman and Medoff (1982) and Lewis (1986) that endogeneity correction methodologies do not contribute to our understanding of the union wage effect puzzle.

Keywords: union status, union wage effects, unobserved heterogeneity, dynamic model of unionism and wage determination.

JEL Classification: C33, J31, J51

1 The views presented in this paper are the author´s and do not reflect those of the BHPS data depositors, namely, the Institute for Social and Economic Research at the University of Essex, U.K.

2 I am grateful to Jeffrey Wooldridge, Mark Stewart, Pierre-Carl Michaud, Francis Vella, Rasmus Jørgensen, Bertel Schjerning, Alicia Pérez Alonso, Nicholas Cox, Gianluigi Vernasca, John Bone, Federico Crudu, Leslie Godfrey, Keith Hartley, John Hutton, Andrew Jones and Eduardo Anderson Stanciole for their valuable help. All errors and omissions are my own.

1. INTRODUCTION: THE JOINT DETERMINATION OF UNION STATUS AND UNION WAGE EFFECTS

This paper is concerned with the estimation of union wage differentials in the United Kingdom during 1991-2003 using data from the British Household Panel Survey (BHPS). We are essentially exploring the question of how observationally equivalent employees' pecuniary rewards diverge in union and non-union employment.

It is hard to conceptualise a realistic scenario in which the unobserved determinants of the sorting decision regarding the two sectors would not affect wages. Therefore, we cannot readily ignore to encompass how the unobserved individual heterogeneity underlying the union/non-union decision is rewarded in the respective sectors (see for instance Robinson, 1989a; Vella and Verbeek, 1998).

The primary complication arises because union sector employees differ from non-union sector employees in unobserved ways, unobserved individual heterogeneity, thus resulting in biased estimates of the union wage effect. As Freeman (1984) notes the ubiquitous phrases in literally any paper in the field stress that:

"You should make a selectivity bias correction...simultaneously determine union status and economic outcomes...develop an unobservables model...Use longitudinal data" (p.2).

The general consensus by many researchers in the past was that cross-sectional OLS analyses of the union wage effect are contaminated by the selectivity[3] of the union sector employees that leads to a positive correlation between the error term and union membership status hence, inflating the union wage effect (see Abowd and Farber, 1982; Freeman, 1984).

Longitudinal estimates of the union wage effect are biased downward and are expected to be lower than the corresponding cross-sectional estimates reflecting the presence of unobserved heterogeneity bias. Of course, longitudinal estimators of the union wage impact do not provide a research panacea. This is due to the potentially substantial impact of measurement or misclassification error of the union membership variable on longitudinal estimates of the union effect (see Freeman, 1984).[4]

While there was little disagreement that union membership status is not exogenous (e.g. Freeman, 1984; Duncan and Leigh, 1985; Robinson, 1989a) authors such as Freeman and Medoff (1982) and Lewis (1986) have reached the pessimistic conclusion that there is no discernible pattern to the estimates of the union wage impact, many were considered to be suspiciously high or low, and endogeneity correction methodologies have contributed little to our understanding of the union wage differential puzzle (Robinson, 1989a, p.640).

Upon summarising the then existing literature Robinson (1989a) concludes that the outcome is conflicting: cross-sectional methods (such as Inverse Mills Ratio and Instrumental Variables) produced an upward adjustment as opposed to OLS, whereas longitudinal (differencing) methods produced a downward adjust-

[3] The selectivity argument assumes that given that unions are able to establish wage premia within the unionised sector, profit maximising employers will select the most productive workers.

[4] In fact, the union membership variable constructed for the twelve years of the survey suffers from a certain degree of discontinuity (refer to Chrysanthou, 2007). The two alternative questions available in the survey do allow us to construct a continuous measure of union membership. The individual responses from the two alternative membership questions and the union coverage outcomes can be compared in order to detect any potential measurement error and its impact on the union wage effects (see Swaffield, 2001).

ment (Robinson, 1989a, p.640). Though, some researchers employing longitudinal data sets attribute the resulting reduction in the union wage effect to fixed effects of higher quality workers present within the unionised sector this is not consistent with the cross-section studies employing IMR or IVE methods to deal with such effects (Robinson, 1989, p.658).

Given a comparative advantage interpretation of the unobserved individual heterogeneity effects, the endogeneity corrected estimates of the union wage effect can actually be higher than their uncorrected OLS counterparts. The selectivity argument (e.g. Abowd and Farber, 1982; Freeman, 1984) will only produce an upward bias for OLS under a hierarchical notion of unobserved individual heterogeneity (Robinson, 1989a, p.665).

Estimation of union wage differentials using longitudinal data requires controlling for the endogeneity of union membership status. Fixed effects and Instrumental Variables estimators assume that this endogeneity is individual-specific and fixed and are thus restrictive in their treatment of unobserved individual heterogeneity.[5]

The Fixed effects estimator can provide consistent estimates of the union wage impact to the extent that the individual heterogeneity that triggers the endogeneity of union status runs solely through the individual fixed effects which are further constrained to be equal in the two sectors (see Robinson, 1989a; Vella and Verbeek, 1998).

Instrumental Variables procedures (e.g. Hausman and Taylor, 1981) can also be put into use although these estimation methodologies, in their generic form, do not account for endogeneity operating through the other error components. Further, Instrumental Variables methods constrain unobserved individual heterogeneity to be identical across sectors. This enforces the constraint that the ordering of employees' productivity in each sector is invariant to sector (see Vella and Verbeek, 1998, 1999a).

In response to this, we employ the estimation procedure offered by Vella and Verbeek (1998, 1999b) to explicitly identify the different sources of endogeneity of union membership status via the decomposition of the endogeneity underlying union membership status into an individual-specific component and an individual/time-specific effect.

This is a two-step estimation methodology of a simultaneous error component model with an endogenous explanatory variable that is also the basis of the selection rule. It extends the existing cross-sectional estimators (e.g. Lee, 1978 and Heckman, 1979) via the exploitation of the panel nature of the data to separate the type of unobserved heterogeneity causing the endogeneity/selection bias. By introducing dynamics it enables the isolation of individual effects from state dependence (see Vella and Verbeek, 1999b, p.245).

A major advantage of this approach is its relative computational simplicity, as opposed to maximum likelihood, and although two-step procedures are in general inefficient (see Newey, 1987) this method does provide initial consistent estimators for a limited information framework (LIML) approach so that asymptotically efficient estimators can be obtained at one iteration (see Vella and Verbeek, 1999b, pp.240-241).

In this paper we wish to establish whether UK trade unions still play a role in wage formation following the introduction of the successive (1980, 1982, 1988 and

[5]For a detailed discussion on this issue refer to Robinson (1989a,b) and Vella and Verbeek (1999a).

1990) Employment Acts targeted towards weakening their bargaining strength. Further, we are interested in exploring the role of unobserved individual heterogeneity in the union/non-union decision and the manner in which it is rewarded in the two sectors. Finally, we wish to investigate whether the economic sorting structure governing the entry into the two sectors provides a discernible pattern of the endogeneity corrected estimates relative to the uncorrected estimates of the union wage effect.

The empirical results indicate that trade unions in the UK still play a non-negligible, albeit diminishing, role in wage formation. The estimated union wage differentials are consistent with the recent estimates of Blanchflower and Bryson (2007) using data from the Labour Force Survey (LFS). While during (1991-1996) the average contribution of unobserved individual heterogeneity to the wages of male and female union members, under the sorting structure supported by the estimates, is negative the corresponding contribution is positive in the (1997-2002) estimates. Further, the significant contribution of unobserved heterogeneity renders the total union wage differential highly variable across individuals. Finally, the endogeneity correction procedure employed yields a discernible pattern of the estimated union wage impact relative to OLS and Fixed effects. This is in line with Robinson (1989a) and Vella and Verbeek (1998) and refutes the pessimistic conclusions reached by Freeman and Medoff (1982) and Lewis (1986) in their influential surveys.

The rest of the paper is organised as follows. Section 2 outlines the econometric model; Section 3 the estimation procedure; Section 4 analyses the estimation results; and finally Section 5 concludes.

2. A DYNAMIC MODEL OF UNIONISM AND WAGE DETERMINATION

Equation (1) outlines the primary wage equation and assumes that individuals sort themselves into their preferred sector (union/non-union) on the basis of wages which are determined by observed and unobserved attributes and their respective prices. The potential wage corresponding to individual i employed in sector j, in time period t is given by $w_{j,it}$. The non-unionised and unionised sectors are denoted by $j = \{0,1\}$ respectively, β is an unknown parameter vector and x_{it} is the conventional vector of personal and industrial characteristics which is also inclusive of time dummies. The unobserved random components of the employee's wage are given by $(\alpha_{j,i}, \varepsilon_{j,it})$ and the usual error component structure assumes $\alpha_{j,i} \sim iidN(0, \sigma^2_\alpha)$ and $\varepsilon_{j,it} \sim iidN(0, \sigma^2_\varepsilon)$:

$$\begin{aligned} w_{j,it} &= \beta'_{j,t} x_{it} + \alpha_{j,i} + \varepsilon_{j,it} \\ t &= 1, ..., T;\ i = 1, ..., N;\ j = \{0,1\} \end{aligned} \tag{1}$$

Employment within a unionised establishment is also contingent on the employer's willingness to hire him/her (see Abowd and Farber, 1982). The major limitation of the estimation methodology employed here is that it does not sufficiently control for employer characteristics while on the other hand, individual employees' attributes are allowed to be an integral part of the employer's decision making process. While in the estimated models presented employer attributes are captured through the industrial classification dummies and establishment size controls these are not adequate in order to assign any specific effects purely to unobserved heterogeneity (see Vella and Verbeek, 1998, p.164).

The dynamic reduced form model depicting the decision of an individual to join either the unionised or non-unionised sector is provided in equation (2). The benefits of employment within the unionised sector are captured by the latent variable U_{it}^*. The union membership status of an individual i in period t, is indicated by the dummy variable U_{it}.

The unknown parameters to be estimated are $(\gamma_1', \gamma_2)'$ and the composite error term ν_{it} captures the unobserved individual heterogeneity underlying the union membership decision and is decomposed into an individual-specific component ϑ_i and an individual time-specific effect η_{it} and w_{it} is the logarithm of the gross average hourly wage:[6]

$$\begin{aligned} U_{it}^* &= \gamma_1' x_{it} + \gamma_2 U_{i,t-1} + \vartheta_i + \eta_{it} \\ U_{it} &= I(U_{it}^* \rangle 0) \\ w_{it} &= w_{j,it} \;\; if \;\; U_{it} = j \end{aligned} \tag{2}$$

Denote $ej_{,it} = \alpha_{j,i} + \varepsilon_{j,it}$ and $\nu_{it} = \vartheta_i + \eta_{it}$, let ν_i be a T vector of ν_{it} and $x_i = [x_{i1}, ..., x_{iT}]'$. Assuming that:

$$\nu_i \mid x_i \sim iidN(0, \sigma_\vartheta^2 ii' + \sigma_\eta^2 I) \tag{3}$$

$$E\{ej_{,it} \mid x_{i,}\nu_i\} = \tau_1 \nu_{it} + \tau_2 \bar{\nu}_i \tag{4}$$

where $\bar{\nu}_i = \frac{1}{T}\sum_{t=1}^{T} \nu_{it}$, i is a T dimensional vector of ones and (τ_1, τ_2) are unknown constants to be estimated.

Expression (3) enforces normality and a strict error components structure on the reduced form model for union membership[7] and precludes any form of autocorrelation in η_{it} while equation (4) permits heteroskedasticity and autocorrelation in $\varepsilon_{j,it}$ but imposes the strict exogeneity[8] of x_{it} (see Vella and Verbeek, 1999b, pp.242-244).

Potential seniority and non-pecuniary benefits can be sufficiently strong motives for individuals to remain within the unionised sector, irrespective of wage changes, and this introduces state dependence in the model (see Vella and Verbeek, 1998). The inclusion of a lagged union membership status variable in the reduced form model prevents the error components from incorrectly capturing the dynamics which should be credited to lagged union membership. Regrettably, it is not possible to include dynamics coming through the lagged dependent variable in the wage equation as well. An alternative estimator proposed by Arellano *et al* (1997) permits for lagged latent dependent variables to enter both the primary and reduced form equations linearly. Nevertheless this is constrained to models with Tobit types of censoring (see Vella and Verbeek, 1999b, p.242).

The random components $(\alpha_{j,i}, \theta_i), (\varepsilon_{j,it}, \eta_{it})$ in equations $(1, 2)$ denote an individual-specific effect, and an individual/time-specific effect, respectively. It is assumed that these are independently and identically distributed drawings from a multivariate normal distribution, where every effect is potentially correlated with its counterpart, of the same dimension, in the other equation (Krishnakumar, 1996, p.230). More

[6] Logarithm of weekly wage divided by usual paid hours (including overtime).

[7] Note that testing for non-normality in the reduced form model for union membership can be quite difficult computationally.

[8] The errors are assumed to be independent of future and lagged values of x_{it}.

specifically the four covariances ($\sigma_{j,\alpha\theta}$, $\sigma_{j,\varepsilon\eta}$) are allowed to be non-zero.[9] These covariances indicate that the random components in the wage equation are potentially correlated with the random components in the union membership equation and this is precisely what produces the potential endogeneity of union membership status in the primary equation.

The covariances convey valuable information about the form of sorting into the two sectors (see Vella and Verbeek, 1999a). Note that the $(\theta_i; \forall\ i = 1, ..., N)$ are constructed so that their average value for union employees is positive while their average value for non-union employees is negative. For tractability assume that the endogeneity is taken to operate purely via the individual-specific effects $(\alpha_{j,i}, \theta_i)$. In the case that either covariance between (α, θ) is non-zero then the unobserved factors that determine union membership influence wages as well.

A hierarchical sorting structure occurs when both covariances are positive ($\sigma_{1,0} \succ 0$) so that individuals with high values of θ are, on average, the best employees in terms of their endowment of unobserved productivity, irrespective of whether they are located in the union or non-union sector. A comparative advantage or positive sorting structure occurs when employees perform differently in the two sectors and sort themselves appropriately ($\sigma_{1,0} \prec 0$). This implies a negative association between the relative productivity in the two sectors and demands that the contribution of unobserved individual heterogeneity raises wages in both sectors (i.e. $\sigma_{1,\alpha\theta} \rangle\ 0, \sigma_{0,\alpha\theta} \langle\ 0$). One should bear in mind that $\sigma_{1,0}$ cannot be estimated directly.

Note that solely a degenerate hierarchical structure, imposing perfect correlation between sector-specific skills, can meet the strict and restrictive requirement of the equality of the two covariances imposed by either of the Instrumental Variables or the restricted Control Function estimators. A comparative advantage structure is precluded *a priori* (see Vella and Verbeek, 1999a).

To estimate the union wage differential we enforce the restriction that the returns to observed characteristics are both time and sector invariant. The wage equation (*eq*.1) then becomes:

$$\begin{aligned} w_{it} &= \beta' x_{it} + \delta U_{it} + e_{it} \\ e_{it} &= U_{it}(\alpha_{1,i} + \varepsilon_{1,it}) + (1 - U_{it})(\alpha_{0,i} + \varepsilon_{0,it}) \end{aligned} \tag{5}$$

3. ESTIMATION PROCEDURE

Following Vella and Verbeek (1998, 1999b) we start with equation (5) which is made conditional on the t-dimensional vector U_i, and the vector of exogenous variables x_{it}:

$$E(w_{it} \mid x_{it}, U_i) = \beta' E(x_{it} | \ x_{it}, U_i) + \delta E(U_{it} | \ x_{it}, U_i) + E(\alpha_{j,i} | \ x_{it}, U_i) + E(\varepsilon_{j,it} | \ x_{it}, U_i) \tag{6}$$

Estimation of the reduced form model for union membership (*eq*.2), provides the estimates of the unobserved individual heterogeneity. This is a dynamic Random effects Probit model with a likelihood function:

$$\prod_i \int \prod_t \Phi\left(\frac{\gamma' \Psi_{it} + \theta_i}{\sigma_\eta}\right)^{U_{it}} \Phi\left(-\frac{\gamma' \Psi_{it} + \theta_i}{\sigma_\eta}\right)^{1-U_{it}} \frac{1}{\sigma_\theta} \phi(\theta/\sigma_\theta) d\theta \tag{7}$$

[9] The covariances between the effects in the union/ non-union wage equations are not specified, whereas all remaining covariances are set to zero.

where, $\gamma = (\gamma_1', \gamma_2)'$, $\Psi_{it} = [x_{it}, U_{i,t-1}]$, and (Φ, ϕ) correspond to the cumulative probability and density functions of the standard Normal distribution.

The inclusion of the lagged union membership variable as a regressor in ($eq.$2) gives rise to the problem of initial conditions (refer to Heckman, 1981a). The initial conditions problem occurs when the initial value of the dependent variable is correlated with unobserved individual heterogeneity. The presence of individual-specific effects θ_i clearly invalidates the assumption of exogeneity of union membership status in the first period of the survey.

Heckman's solution to the initial conditions problem approximates the reduced form marginal probability of the initial state by a Probit function which has as its argument all of the available pre-sample information on the exogenous variables (Heckman, 1981b, p.188).

We employ Stewart's (2006) Stata command, -redprob-, to implement Heckman's (1981b) estimator of the dynamic Random effects Probit model for union membership.[10]

The conditional expectations in equation (6), the estimates of the unobserved heterogeneity, can be expressed as:[11]

$$
\begin{aligned}
E(\alpha_{j,i}|x_{it}, U_i) &= \sigma_{j,\alpha\theta}\left\{\frac{T}{\sigma_\eta^2 + T\sigma_\theta^2}E(\bar{\nu_i}|\ x_{it}, U_i)\right\} = \sigma_{j,\alpha\theta}C_i \\
E(\varepsilon_{j,it}|x_{it}, U_i) &= \sigma_{j,\varepsilon\eta}\left\{\sigma_\eta^{-2}E(\nu_{it}|\ x_{it}, U_i) - \frac{T\sigma_\theta^2}{\sigma_\eta^2(\sigma_\eta^2 + T\sigma_\theta^2)}E(\bar{\nu_i}|\ x_{it}, U_i)\right\} = \sigma_{j,\varepsilon\eta}C_{it}
\end{aligned} \tag{8,9}
$$

The endogeneity correction terms (C_i, C_{it}) are added as additional terms in the equation of primary interest to be estimated jointly with (β', δ) in the second step from conditional moment restrictions such as least squares based on (5). Under the null hypothesis of no endogeneity ($\sigma_{j,\alpha\theta} = \sigma_{j,\varepsilon\eta} = 0$), the conventional standard errors can be used. Otherwise, the standard errors should be adjusted for heteroskedasticity and for the inclusion of the endogeneity correction terms (see Vella and Verbeek, 1999b, pp.259-260).

4. ESTIMATION RESULTS

We employ balanced panels of employees aged between 16-65 during (1991-1996 & 1997-2002) consisting of a full-time male employees' sample and a full sample of female employees (full-time and part-time). Part-time male employees are excluded since the small gains in terms of sample size are more than outweighted by the costs of a potential increase in the heterogeneity of the male samples. The female samples can provide a comparison group against the male sample that could potentially suffer from selectivity bias. Note that the former is also prone to sample selection bias caused by the labour market participation decision (see Swaffield, 2001, p.439).

The estimated models presented in Tables (2-5) in Appendix III, include a set of personal and industrial characteristics. All models, except the Fixed effects (within)

[10] Employing Orme's (2001) two-step methodology the correlations between the initial condition and the random effect were all in excess of 0.8 in the light of which the inherent heteroskedasticity of the residual component produces inconsistent parameter estimates. Wooldridge's (2005) estimator, extended by a standard probit estimator for the initial period to enable comparability with Heckman's (1981b) estimator, produced an inferior log-likelihood since it doesn't incorporate pre-sample information (for a detailed treatment refer to Chrysanthou, 2007).

[11] See Appendix I.

estimates, are also inclusive of a set of time dummies. The set of explanatory variables included in the models were jointly statistically significant in every reported specification. The descriptive statistics for the variables are provided in Table 1, Appendix III while those of the endogeneity correction terms are given in Tables (6-9), Appendix III.

The unconditional male union wage differential is approximately 4.5% and 4% in the (1991-1996) and (1997-2002) samples, respectively, while, the corresponding unconditional female union wage differentials are approximately 14% and 11.9%. Further, the descriptive statistics Tables reveal that females earn approximately 20.6% less than their male counterparts during (1991-1996) while, they earn approximately 14.4% less during (1997-2002). Male union membership during (1997-2002) has fallen by approximately 15.3% compared to the (1991-1996) male membership level. The female union membership rate on the other hand, has actually risen by approximately 9.5% so that by the last cross-section of the survey female unionisation rates converged to the level of male unionisation rates at a marginally higher percentage.

Prior to embarking on the analysis of the estimated results a couple of issues need to be addressed. Regarding the issue of identification, the non-linear mapping from the reduced form union membership variables to the endogeneity correction terms identifies all parameters in the wage equation (5). Exclusion restrictions are of course desirable though one should be cautious with remaining consistent with the economics of trade union membership (see Vella and Verbeek, 1999a).

The exclusion of lagged union membership status from the empirical counterpart of (eq.5) identifies the equation as long as γ_1 differs from zero. Note that it is assumed that the long-term advantages of union employment, whilst generating persistence of union membership status, do not have a significant impact on wages. Thus, lagged union membership status is only expected to have a minor effect on current wages (see Vella and Verbeek, 1998, p.167).

Vella and Verbeek treat industry of employment as exogenous since their data come from the Youth (male) sample National Longitudinal Survey. Citing the results of Jovanovic and Moffitt (1990) and Topel and Ward (1992) they argue that although workers generally sort themselves into jobs on the basis of industry-specific skills younger workers experience multiple job changes prior to entering longer term employment arrangements. Occupational status on the other hand, being a measure of ability, is treated as endogenous as it can contaminate the conclusions concerning the role of unobserved individual heterogeneity (see Vella and Verbeek 1998, pp.168-169).

However, the samples used in this study are fairly heterogeneous and consist of employees aged between 16-65. Thus, assuming that industry of employment is exogenous using the argument in the preceding paragraph is not as straightforward. Including occupational controls reduces the coefficients on the fixed individual effects in most estimated models suggesting that a component of unobserved heterogeneity underlying union status is correlated with occupational classification. Nevertheless, to the extent that this was not found to affect the union wage differentials or the conclusions reached with regard to the role of the endogeneity correction terms we opted to include occupational controls in the models.

Given that the industrial and occupational classifications overlap to a certain extent, our exclusion decision of either set of controls in the reduced form union membership models was based on the convergence problems incurred in the maximum likelihood estimators (due to the given degree of collinearity). Concerning the exclusion of either set of controls in the primary wage regressions this was based on the detection of non-normality. Higher-order terms of the latent effects, squared correction terms and their interactions with union status, were included in the models in order to detect non-normality (see Pagan and Vella, 1989). Exclusion of either set of the industrial and occupational classification controls was based on minimising the joint statistical significance of the higher-order terms. These are reported only in the models where normality was an issue. The statistically significant higher order terms in most of the endogeneity corrected estimates in Tables (2-5) indicate non-normality, an outcome which was not unexpected given the fairly heterogeneous nature of the samples.

As expected *a priori*, including the occupational controls does change the coefficients of several personal characteristics variables since these variables are bound to be correlated with occupational status. One should bear in mind, however, that the set of personal and industrial variables included in the models only serve in controlling for some observed heterogeneity and are not the focus of this study. The latent effects in the estimated wage regressions are linear functions of the corresponding latent effects in the reduced form models for union membership. Hence, the inclusion of the endogeneity correction terms complicates the interpretation of the coefficients on explanatory variables other than union membership.

The standard errors in the models inclusive of the endogeneity correction terms need to be adjusted to account for the structure of the errors and the two-step nature of the estimation procedure using the formulae provided in Newey (1984). Otherwise the resulting standard errors will be deflated and significance levels overstated (see Heckman, 1979).

While we use heteroskedasticity robust standard errors, in conjunction with all estimation methodologies reported in Tables (2-5), we fail to account for the inclusion of the generated regressors in the second step due to the complexity entailed in computing the asymptotic covariance matrix of the second step estimator. Obtaining boostrap standard errors via re-sampling the data is an attractive alternative though this needs to be carried out for both estimation stages. Unfortunately, the computational time involved in estimating the first stage Random effects Probit models for union membership, using Stewart's (2006) maximum likelihood estimator to account for the initial conditions problem, renders the boostrapping alternative prohibitive.[12] Nevertheless since the generated endogeneity correction terms are only jointly, as opposed to unilaterally, statistically significant the impact on the standard errors in the reported estimates should not be substantial.[13]

4.1. The Wage Regressions Under Hierarchical Sorting

To determine the sorting structure consistent with the data we begin with the hierarchical sorting estimates that assume that the best employees, in terms of

[12] The first stage maximum likelihood estimators require between 52-120 hours each, depending on the model, to provide convergent results using a computer with normal capacity. The number of iterations ranges from 7-14. Combining this with the boostrap option requires more than 72 hours for a single iteration. The only options reducing computational time to reasonable levels would be to either conduct a pilot study with very few boostrap repetitions, in which case the asymptotic validity of the results is questionable, or exclude the vast majority of the explanatory variables in the first stage regressions which would result in model misspecification.

[13] For instance, Vella and Verbeek (1998) report that failing to account for the inclusion of strongly statistically significant generated regressors in the second stage results in an underestimation of the standard errors that ranges between 1-11%.

their unobserved productivity, will receive a higher remuneration irrespective of whether they are located in the union or non-union sector. Essentially then, under this assumption the endogeneity correction terms are restricted to be invariant to sector.

Regarding the (1991-1996) male and (1997-2002) female models the negative and statistically significant coefficients on the individual-specific correction terms suggest that employees who receive lower wages, upon conditioning on their attributes and in the absence of trade unions, are those more likely to be trade union members (refer to Tables 2, 5, respectively). The statistically significant and negative coefficient on the individual time-variant effect in the (1991-1996) female estimates is also indicating that lower paid individuals, in the absence of trade unions, have a higher propensity towards unionisation *ceteris paribus* (see Table 4).

However, concerning the (1997-2002) male estimates the positive and statistically significant coefficient on the individual-specific effect, implies that male employees who receive higher wages, controlling for their characteristics and in the absence of trade unions, are those more likely to be trade union members (refer to Table 3).

Note that apart from the (1991-1996) female estimates (in Table 4) the individual time-variant effects were not found to be unilaterally statistically significant. Thus, it seems that with the exception of the (1991-1996) female estimates, the union effects are mainly due to the individual-specific effects. Nevertheless, all selection terms were generally found to be jointly statistically significant and this is suggestive of selectivity bias.

The statistically significant coefficients on the time-variant individual effects in the (1991-1996) female estimates imply that Fixed effects estimation is inappropriate as the time varying endogeneity is not eliminated and continues to contaminate the resulting estimates.

In all reported models the estimated union effect under Fixed effects assumptions is notably low (refer to Tables 2-5).[14] This is consistent with the general consensus in the union literature that the longitudinal estimates of the union wage effect are lower than the cross-sectional estimates (e.g. Freeman, 1984; Robinson, 1989a) and in line with the empirical findings of authors such as Jakubson (1991) and Vella and Verbeek (1998).

While, of course, there is no doubt regarding the question of the appropriateness of the Fixed effects approach in the female (1991-1996) model it could be argued that joint, as opposed to unilateral, statistical significance in the remaining models weakens the argument. One should note, however, that even in the case whereby unobserved heterogeneity is individual-specific, Fixed effects impose the invariance of heterogeneity rewards to sector and further that covariance $\sigma_{\alpha\theta}$ is constant across sector (see Vella and Verbeek, 1998, pp.171-172).

The outcome that higher paid employees are less likely to seek union employment, apart from the (1997-2002) male estimates, is in line with insider-outsider theories (Lindbeck and Snower, 1986) suggesting that groups of highly skilled employees can be seen as acting as a *de facto* union on its own since they cannot be rapidly and costlessly replaced (see Blanchflower *el al*, 1990). It is also consistent with Abowd and Farber's (1982) argument that the standardisation of wage rates

[14] The only exception are the (1997-2002) male estimates (see Table 3) where the Fixed effects estimator provides a similar estimate of the union effect to the hierarchical sorting estimate. This is due to the very low statistical significance of the individual time-variant effects in both the hierarchical and unrestricted estimates.

via the bargaining process implies that workers with a high degree of human capital invested in themselves would be less prone to unionisation as it would entail reduced human capital premia. Robinson (1989a) and Vella and Verbeek (1998) in their studies also reach the conclusion that there is no empirical evidence that better employees are chosen from a queue to join unions.

The result regarding the (1997-2002) male employees' remuneration and union membership propensity could be explained by resorting to the reduced form estimates of union membership determination (see Chrysanthou, 2007). The union status determination models reveal that while male employees in white collar occupations during (1991-1996) display a greater degree of union aversion, highly skilled male employees within the public sector are more likely to be union members. This outcome is reinforced by the respective (1997-2002) estimates suggesting that the public sector, which still remains heavily unionised, does not seem to share the fate of the British manufacturing industry in terms of rapid deunionisation. Moreover, the reduced form estimates suggest that the role of habit persistence (i.e. the behavioural effect of remaining in unions due to experiencing unionisation in the past) has diminished relative to that of unobserved heterogeneity that appears to play an even more important role in determining individual unionisation propensity in the second period estimates. Thus, it seems plausible that the greater proportion of highly productive male employees remaining in unions during the second period, do so because of unobserved factors such us solidarity as opposed to maximising their earnings.

Finally, though on the basis of individual behaviour it seems reasonable to argue that less productive employees might be more prone towards unionisation in order to appropriate some share of monopoly rents, it appears improbable that profit maximising employers would hire them (see Abowd and Farber, 1982; Vella and Verbeek, 1998, p.173). Since in modelling wage determination and the union membership decision we do not explicitly incorporate the employers' role, at a minimum we ought to investigate the impact on the estimates upon the relaxation of the assumption that unobserved heterogeneity is equally valued in each sector.

4.2. The Wage Regressions Under Unrestricted Sorting

To test whether the random components are differentially rewarded in the two sectors we modify the models by allowing the endogeneity correction terms to vary by sector. To investigate how the random components are rewarded in each sector more rigorously we need to examine the average wage contribution of the endogeneity correction terms. This is obtained by multiplying the mean values[15] of the endogeneity correction terms (C_i, C_{it}) by the corresponding estimates of covariances ($\sigma_{j,\alpha\theta}$, $\sigma_{j,\varepsilon\eta}$).[16]

The unrestricted sorting structure estimates for males during (1991-1996), in Table 2, reveal that the sign of the coefficient on the fixed individual effect remains negative as in the hierarchical structure estimates, though now it is weakly statistically significant. Given the statistical insignificance of the interaction between the fixed effect and union status the dominant effect determining the sorting structure comes from the time-variant effects. The corresponding return to the time-variant

[15] Provided in Tables (6-9), Appendix III.

[16] Under unrestricted sorting we obtain a sector-invariant and a sector-variant contribution given by the coefficients on (C_i, C_{it}) and ($\mathrm{C}_i.\mathrm{U}_{it}$, $\mathrm{C}_{it}.\mathrm{U}_{it}$) respectively. For non-union employees the latter is equal to zero as $\mathrm{U}_{it} = 0$.

effect for union and non-union employees is negative, ($\sigma_{1,0} \succ 0$). The descriptive statistics in Table 6 reveal that, the average contribution of the time-variant effects to the wages of union members is negative while the average contribution to the wages of non-union members it is positive. This indicates that upon incorporating the role of unobserved heterogeneity union members receive a lower wage irrespective of the sector they are located and thus sorting into union and non-union employment seems to be done on the basis of a hierarchical structure.

Considering the (1997-2002) male estimates under unrestricted sorting (see Table 3) the appropriate return to the fixed individual effect for union and non-union employees is positive, ($\sigma_{1,0} \succ 0$). The time-variant effects remain statistically insignificant as in the respective hierarchical sorting estimates. The mean values of the endogeneity correction terms (in Table 7) imply that the average contribution of the individual-specific effects to the wages of union members is positive whereas the corresponding return to the wages of non-union members is negative. Thus, union members receive a higher wage as a result of unobserved heterogeneity, regardless of the sector they are located in, an outcome that suggests that sorting into the two sectors is performed on the basis of a hierarchical structure. Note the negative coefficient on the interaction between the individual fixed effects and union status; while the positive coefficient on the fixed individual effect suggests that individuals with characteristics typically associated with higher wages are those more likely to be in unions, the negative sector-variant fixed effect (though statistically significant at the not so stringent 10% level) implies that the standardisation of wage rates within the union sector entails reduced human capital premia for highly skilled employees (see Abowd and Farber, 1982). This is in line with the conclusion reached in the restricted sorting estimates of Section 4.1, that the larger proportion of highly productive male employees remaining in unions in the second period, are likely to do so due to unobserved factors such us solidarity as opposed to maximising their earnings.

With regard to the female (1991-1996) unrestricted sorting structure estimates (in Table 4) given the statistical insignificance of the fixed individual effect the dominant effect determining the sorting structure consistent with the estimates stems from the time-variant effects. The return to the time-variant effect is negative for union and non-union employees and thus, ($\sigma_{1,0} \succ 0$). According to the respective mean values for the endogeneity correction terms (see Table 8) the average contribution of the time-variant effects to the wages of union members is negative whereas the corresponding contribution to the wages of non-union members is positive. Therefore unobserved individual heterogeneity impacts negatively on the wages of union members, irrespective of the sector they sort themselves into. Hence, we cannot reject the hierarchical sorting specification.

Finally, the unrestricted sorting estimates for females during (1997-2002: see Table 5) are consistent with a comparative advantage sorting specification since the appropriate return to the dominant fixed individual effect for union members is positive whereas for non-union it is negative. Hence, sector-specific skills are negatively correlated, ($\sigma_{1,0} \prec 0$). Note the negative coefficient on the individual-specific effect. This indicates that employees that are the recipients of lower wages, upon conditioning on their attributes and in the absence of trade unions, are those most prone towards unionisation. The positive coefficient on the interaction between the fixed effect and union membership, however, implies that such individuals perform relatively better in the union sector. In other words, employees perform differently within the two sectors and sort themselves appropriately so that the average contribution of unobserved heterogeneity to the wages of union and non-union members is positive (refer to Table 9). Therefore, those with a relatively higher propensity towards unionisation benefit most from sorting themselves into the union sector. It is plausible that employers, faced with higher union wages and an increasing female unionisation rate, select those females that are more likely to perform better in the union sector from the pool.

Interestingly then while during (1991-1996) the average contribution of unobserved heterogeneity to the wages of union members, under the sorting structure consistent with the estimates, is negative the respective contribution in the (1997-2002) estimates is positive. In modelling union membership determinants we found evidence that regarding union membership status, in the second period under analysis, the role of habit persistence has diminished relative to that of unobserved heterogeneity (see Chrysanthou, 2007). Thus, concerning the male estimates it seems that the higher proportion of highly qualified employees remaining in the union sector during the second period, do so due to unobserved determinants such us solidarity as opposed to maximising their earnings. Regarding the female estimates, the results suggest that females that are the recipients of lower wages, controlling for their characteristics and in the absence of unions, have a relatively higher propensity towards unionisation thoughout the entire period under analysis. The greater effect of unobserved heterogeneity versus habit persistence during the second period reverses the sign of average contribution to union members' wages from negative, in the first period, to positive so that those persisting in unions benefit most from sorting themselves into the union sector.

While the restriction that the random components are equally rewarded in the two sectors is only rejected in the female (1997-2002) estimates we do not believe that estimation of the remaining models by Instrumental Variables or restricted Control Function estimators is appropriate. Primarily since the structure of sorting into the union and non-union sectors cannot be known *a priori* we see no reason why one should preclude a comparative advantage sorting structure by employing these estimators.

Further, one should bear in mind that solely a degenerate hierarchical structure, imposing perfect correlation between sector-specific skills, can meet the strict and restrictive requirement of the equality of the two covariances imposed by Instrumental Variables and variants of the restricted Control Function estimators. These estimators do not only require a hierarchical structure, but also that sector-specific skills are perfectly correlated (see Vella and Verbeek, 1999a, p.477).

According to the sorting structure supported by the estimates the male union wage effect[17] has fallen from 9.2% in the (1991-1996) estimates to 5.8% in (1997-2002) which is a sizeable decrease of approximately 37% (refer to Tables 2, 3). The respective female union wage effect has also fallen from 17.6% to 11.5% which is a significant decrease of approximately 35% (see Tables 4, 5). The corresponding decreases in the union wage differentials are not surprising given the successive legislative changes targeted towards weakening the bargaining strength of UK trade unions (for a detailed account refer to Stewart, 1995). The overall effect of the (1980, 1982, 1988 and 1990) Employment Acts was that the past quarter century was an era of a dramatic decline in aggregate trade union membership and union recognition in the UK. Nevertheless, the estimated union wage differentials imply

[17] Given that the dependent variable is the natural logarithm of the gross average hourly wage the union wage differential is obtained by $e^{\delta} - 1$.

that unions still play a non-negligible role in wage formation.

The outcome that female employees receive higher union differentials is consistent with the expectation that unions raise the wages of those who would have been relatively lower paid in their absence. Further, provided that unionism is assumed to be a normal good an employee's demand for unionism is positively related to his/her union wage differential (see Chang and Lai, 1997, p.121). Therefore, the individuals that are lower paid, in the absence of unions, for discriminatory reasons are more likely to be prone towards unionisation (see for instance Heywood, 1990). This is in line with the finding that, in the period under analysis, while male union membership has been declining female union membership has been rising albeit not markedly (see Chrysanthou, 2007).

The estimated union wage effects are consistent with the recent estimates of Blanchflower and Bryson (2007) using data from the Labour Force Survey (LFS).[18] The corresponding uncorrected OLS union wage differentials of Blanchflower and Bryson (2007), using personal and industrial plus occupational controls, are approximately (7.9%, 7.1%) for males and (11.2%, 8.3%) for females during (1993-2000) and (2001-2006), respectively.[19] The uncorrected OLS union wage differentials provided in Tables (2-5) are approximately (9%, 11.7%) for males and (13.5%, 12.1%) for females during (1991-1996) and (1997-2002) correspondingly. The evident continuous decline in union wage differentials explains the relatively higher uncorrected OLS estimates from the present study that uses data from an earlier period. Note that the (1997-2002) uncorrected OLS male union wage differential is upwardly biased and this could explain the higher male union effect for the second period (2001-2006) reported by Blanchflower and Bryson (2007) as opposed to our preferred endogeneity corrected estimate of 5.8% for the (1997-2002) period.

Under the sorting structure consistent with the estimates the average contribution of unobserved individual heterogeneity[20] to the wages of male union and non-union members during (1991-1996) is approximately -4.1% and 1.8% respectively while the corresponding figures for the (1997-2002) estimates are 3.1% and -2.8%. The respective figures for female union and non-union members during (1991-1996) are approximately -2.9% and 0.7% while in the (1997-2002) estimates the average contribution of unobserved individual heterogeneity to the wages of female union members is approximately 3.5% and for non-union members 4.1%. The significant contribution of unobserved heterogeneity renders the total union wage differential highly variable across individuals.

OLS differentials denote the difference between the average wage of those that are actually in the union sector and the average of those employees that are located in the non-union sector *ceteris paribus.* Endogeneity corrected differentials on the other hand come from a random assignment to sector (see Robinson, 1989a, p.660).[21] Given a non-hierarchical interpretation of the unobserved individual heterogeneity effects, endogeneity corrected estimates of the union wage differential

[18] Note that Blanchflower and Bryson (2007) fail to account for the endogeneity of union status.

[19] Refer to the first column of Table 4, heading C, (Blanchflower and Bryson, 2007, p.17).

[20] This amounts to multiplying the estimated coefficient on the dominant latent effect, determining the sorting structure, in Tables (2-5) by the respective mean value for union and non-union members provided in the descriptive statistics Tables (6-9) for the endogeneity correction terms. Concerning the calculations with respect to the female estimates for (1991-1996) and (1997-2002) see also footnotes 14, 12 correspondingly.

[21] Note that there are no individuals who are randomly assigned. Random assignment refers to controlling for the unobservables underlying the unionisation decision which are correlated with the wage rate (see Vella and Verbeek, 1999a, p.474).

may be higher than the uncorrected pooled cross-sectional OLS estimates (Robinson, 1989a, p.665).

This is, of course, at odds with the *orthodox* view that profit maximising establishments within a unionised sector will select the best employees thus inflating the OLS estimate of the union wage differential (e.g. Abowd and Farber, 1982; Freeman, 1984). However, this will only result in an upward bias for OLS under a hierarchical sorting structure of omitted ability when the average contribution of unobserved individual heterogeneity to the wages of union members is positive as is the case with the (1997-2002) male estimates in this study. Note that while we do recognise that the estimation methodology used fails to explicitly account for employers' behaviour we favour the interpretation that the greater proportion of highly productive male employees remaining in unions during the second period do so because of unobserved determinants such us solidarity as opposed to maximising their earnings, rather than the selectivity argument.

As the average contribution of unobserved heterogeneity to the wages of male and female union members during (1991-1996) is negative, so that they receive lower wages irrespective of the sector they are located in, this reverses the direction of the OLS bias from positive to negative despite that the data are consistent with a hierarchical structure. Under the comparative advantage sorting structure consistent with the (1997-2002) female estimates the endogeneity correction methodology produces a slightly lower, rather than higher according to Robinson (1989a), union effect relative to OLS. This stems from the fact that though the average contribution of unobserved individual heterogeneity to the wages within both sectors is positive, the respective contribution to the wages of non-union members is higher compared to that of union members.

Recalling that the estimated union effect under Fixed effects assumptions is markedly low, so that it provides the lower bound, the endogeneity correction procedure employed yields a discernible pattern of estimates of the union wage differential relative to OLS and Fixed effects. Hence, it does contribute to our understanding of the union wage impact puzzle. This conflicts with Freeman and Medoff (1982) and Lewis (1986) and is in line with Robinson (1989a) and Vella and Verbeek (1998).

5. SUMMARY AND CONCLUSIONS

Using data from the BHPS (1991-2003) we employ a dynamic model of union status and wage determination to estimate the wage impact of UK trade unions. The findings suggest that the unobserved factors that influence union membership also affect the wage impact of unions.

It is worth noting that the major restriction of the estimation methodology used is that we fail to adequately control for employer characteristics while on the other hand, individual employees' attributes are allowed to be an integral part of the employer's decision making process. In the models presented employer attributes are captured through the industrial classification dummies and establishment size controls but, these are not sufficient so as to assign any specific effects purely to unobserved heterogeneity (see Vella and Verbeek, 1998, p.164). Further, it is likely that the paucity of employer controls might be biasing upwards the union wage impact since unionised establishments tend to pay more for reasons not directly linked to union membership (see Blanchflower and Bryson, 2007, p.7).

According to the sorting structure supported by the estimates the male union wage effect has fallen from 9.2% in the (1991-1996) estimates to 5.8% in (1997-2002), whereas the corresponding female union wage differential has fallen from 17.6% to 11.5%. The higher union wage differentials accruing to female members are consistent with the expectation that unions raise the wages of those who would have been relatively lower paid in their absence. Furthermore, the significant contribution of unobserved heterogeneity renders the total union wage differential highly variable across individuals.

The observed decrease in the union wage impact is in agreement with our expectations following the successive legislative changes (1980, 1982, 1988 and 1990 Employment Acts) targeted towards weakening the bargaining strength of trade unions in the UK. The estimated union wage differentials imply that unions still play a non-negligible, though diminishing, role in wage formation. Further, the estimated union wage effects are consistent with the recent estimates of Blanchflower and Bryson (2007) using data from the Labour Force Survey (LFS).

The endogeneity correction procedure employed reveals that during (1991-1996) the average contribution of unobserved heterogeneity to the wages of union members, under the sorting structure supported by the estimates, is negative whereas the corresponding contribution in the (1997-2002) estimates becomes positive.

The reduced form models for union status determination suggest that concerning union membership in the second period under analysis the role of habit persistence (i.e. the behavioural effect of remaining in unions due to experiencing unionisation in the past) has diminished relative to the impact of unobserved heterogeneity. Therefore, regarding the male estimates it seems plausible that the higher proportion of highly skilled employees remaining in the union sector in the second period, do so due to unobserved determinants such us solidarity as opposed to maximising their earnings. Concerning the female estimates, the results indicate that recipients of lower wages, controlling for their attributes and in the absence of unions, are more likely to seek union employment during the whole period under analysis. The greater impact of unobserved heterogeneity versus habit persistence during the second period reverses the sign of the average contribution to union members' wages from negative, in the first period, to positive so that those remaining in unions benefit most from their sorting decision.

The finding that higher paid employees are less prone towards unionisation, excluding the (1997-2002) male estimates, is in accordance with insider-outsider theories (Lindbeck and Snower, 1986) indicating that groups of highly qualified employees do not need to rely on unions to gain higher wages since they can be valuable assets to employers (see Blanchflower *el al*, 1990). It is also consistent with Abowd and Farber's (1982) argument that the standardisation of wage rates through the bargaining process suggests that those with higher levels of human capital would be less likely to join unions as it would entail reduced human capital premia. Moreover, this is in line with the conclusion reached by Robinson (1989a) and Vella and Verbeek (1998) that there is no supporting empirical evidence that more able employees are selected from a queue to join unions.

The endogeneity correction procedure employed yields a discernible pattern relative to OLS and Fixed effects and therefore contributes to our understanding of the union wage differential puzzle. This refutes the pessimistic conclusions reached by Freeman and Medoff (1982) and Lewis (1986) and is in agreement with Robinson (1989a) and Vella and Verbeek (1998).

The estimated union effect under Fixed effects assumptions is notably low and provides the lower bound; an outcome which is consistent with the general consensus in the union effects literature that the longitudinal estimates of the union wage impact are lower than the cross-sectional estimates (e.g. Freeman, 1984; Robinson, 1989a; Jakubson, 1991; Vella and Verbeek, 1998).

We conclude that OLS estimates of the union wage effect will be biased upwards solely under a hierarchical sorting structure of omitted ability given a positive average contribution of unobserved individual heterogeneity to the wages of union members as is the case with the (1997-2002) male estimates. Since the average contribution of unobserved heterogeneity to the wages union members during (1991-1996) is negative, so that they receive lower wages regardless of the sector they sort themselves into, this reverses the direction of the OLS bias from positive to negative even though the data are consistent with a hierarchical sorting structure. While under the comparative advantage sorting structure consistent with the (1997-2002) female estimates the endogeneity correction procedure produces a slightly lower, as opposed to higher, union effect relative to OLS this does not conflict with Robinson (1989a). This occurs in that, while the average contribution of unobserved heterogeneity to the wages in both sectors is positive, the corresponding contribution to the wages of union members is lower to that of non-union members.

Future work could involve investigating the impact of measurement error in union membership status on the union wage differential estimates. Nevertheless, though an interesting exercise, proving that OLS and Fixed effects estimators bound the true impact of unions on wages when both estimators produce biased results might not be an attractive research direction.

The probability of employment within the union sector is treated as a function of the single index (of observed and unobserved characteristics) appearing in the reduced form union membership model. Given that we employ fairly heterogeneous samples it is likely that our methodology is restrictive in that sorting into union employment is expected to follow a multiple indices rule (see Abowd and Farber, 1982; Card, 1996; Vella and Verbeek, 1998). Provided that the union wage effect may vary with an employee's skill level, and that the selection process into the union sector may give rise to differing selection biases at different skill levels, future work should entail estimating separately the models for distinct skill groups (see Card, 1996, p.969). This, however, will be rather costly in terms of sample sizes.

Finally, another interesting extension could entail comparing the estimates to the recent bias-corrected estimator for nonlinear panel data offered by Fernandez-Val and Vella (2007) that estimates the reduced form model by Fixed effects procedures in order to obtain estimates of the time-variant heterogeneity driving the selectivity bias. This estimator does not use the independence axiom of the random effects, with respect to the explanatory variables, and avoids parametric assumptions for the unobserved individual heterogeneity.

6. APPENDIX I: DERIVATION OF THE ENDOGENEITY CORRECTION TERMS

Following Vella and Verbeek (1998, 1999b) we use the assumption of joint normality to derive the expectation of $e_{j,it}$ in ($eq.5$) conditional on the composite error ν_{it} from the reduced form model and the vector of exogenous variables x_{it}. Employing the standard formulae for the conditional expectation of normally distributed vectors:

$$E(\alpha_{j,i}|x_{it},\nu_i) = \sigma_{j,\alpha\theta}\left\{\frac{T}{\sigma_\eta^2+T\sigma_\theta^2}\bar{\nu}_i\right\} \quad (A_1)$$

$$E(\varepsilon_{j,it}|x_{it},\nu_i) = \sigma_{j,\varepsilon\eta}\left\{\sigma_\eta^{-2}\nu_{it}-\frac{T\sigma_\theta^2}{\sigma_\eta^2(\sigma_\eta^2+T\sigma_\theta^2)}\bar{\nu}_i\right\} \quad (A_2)$$

To obtain the conditional expectations, given the t-dimensional vector U_i, we replace ν_i in ($eq.A_1, A_2$) by their conditional expectations given U_i. Using the first law of iterated expectations:

$$\begin{aligned} E(\nu_{it}|x_{it},U_i) &= E(\vartheta_i+\eta_{it}|x_{it},U_i) = E_{\vartheta_i}[\vartheta_i+E\{\eta_{it}|x_{it},U_i,\vartheta_i\}] \quad (A_3)\\ &= \int_{-\infty}^{\infty}[\vartheta_i+E(\eta_{it}|x_{it},U_i,\vartheta_i)]f(\vartheta_i|x_{it},U_i)d\vartheta_i \end{aligned}$$

where $f(\vartheta_i|x_{it},U_i)$ is the conditional density of ϑ_i.

Given the assumption of the strict exogeneity of x_{it} the conditional distribution of ϑ_i is:

$$f(\vartheta_i|x_{it},U_i) = \frac{f(U_i|x_{it},\vartheta_i)f(\vartheta_i)}{f(U_i|x_{it})} \quad (A_4)$$

Substituting expression A_4 into A_3 we arrive at:

$$\begin{aligned} E(\nu_{it}|x_{it},U_i) &= \frac{1}{f(U_i|x_{it})}\int_{-\infty}^{\infty}[\vartheta_i+E(\eta_{it}|x_{it},U_i,\vartheta_i)]f(U_i|x_{it},\vartheta_i)f(\vartheta_i)d\vartheta_i = \quad (A_5)\\ &= \frac{1}{\displaystyle\int f(U_i|x_{it},\vartheta_i)f(\vartheta_i,x_{it})d\vartheta_i}\int_{-\infty}^{\infty}[\vartheta_i+E(\eta_{it}|x_{it},U_i,\vartheta_i)]f(U_i|x_{it},\vartheta_i)f(\vartheta_i)d\vartheta_i \end{aligned}$$

where we have used $f(U_i|x_{it}) = \int f(U_i|x_{it},\vartheta_i)f(\vartheta_i,x_{it})d\vartheta_i$ and $E\{\eta_{it}|x_{it},U_i,\vartheta_i\}$ denotes the cross-sectional generalised Probit residual (see Gourieroux *et al*, 1987) obtained from the first step estimates of the reduced form model:

$$E(\eta_{it}|x_{it},U_i,\vartheta_i) = (2U_i-1)\sigma_\eta\left\{\frac{\phi\{(2U_i-1)(\gamma'\Psi_{it}+\theta_i)/\sigma_\eta\}}{\Phi\{(2U_i-1)(\gamma'\Psi_{it}+\theta_i)/\sigma_\eta\}}\right\} \quad (A_6)$$

7. APPENDIX II: ESTIMATION OF THE ENDOGENEITY CORRECTION TERMS

The term in the denominator in expression A_5 is the likelihood contribution for individual i. Given the parameter estimates from the reduced form model $\delta_1 = (\gamma_1, \gamma_2, \sigma_{\vartheta_i})$ we can approximate expression A_5 using quadrature methods or simulation (numerical integration).

To estimate the simulated counterpart of A$_5$ we obtain R draws of ϑ_i^r from $f(\vartheta_i|\sigma_{\vartheta_i}) = N(0, \sigma_{\vartheta_i})$ and compute the corresponding log-likelihood for individual i conditional on the draw:

$$f(U_i|x_{it}, \vartheta_i^r) = \prod_{t=1}^{T} f(U_{it}|x_{it}, \vartheta_i^r) \tag{A$_7$}$$

To provide a better coverage of the integrals we use randomised Halton draws since the asymptotic properties of simulation-based estimators are obtained under the assumption of randomness.[22] Supplementing these with antithetic draws[23] we induce a negative correlation over observations thus, further improving the coverage (see Train, 2003; Cappellari and Jenkins, 2006). The procedure is repeated R times and averaging over these replications we obtain the simulated log-likelihood function:

$$f(U_i|x_{it}, \vartheta_i^r) = \ln \frac{1}{R} \sum_{r}^{R} f(U_i|x_i, \vartheta_i^r) \tag{A$_8$}$$

Estimators obtained by maximising likelihoods that are approximated by simulation techniques are known as MSL estimators and have the same large-sample properties as maximum likelihood, if the number of repetitions used to approximate the integral grows at a higher rate than the square root of the number of observations in the sample ($\sqrt{N}/R \longrightarrow 0$), (see Drukker, 2006, p.153). We use 500 randomised Halton draws and 500 antithetic draws hence, $R = 1000$. Given the sample sizes, N, of the estimated models presented $\sqrt{N}/R \longrightarrow 0$.

Train (2003, p.233) claims that 100 Halton draws provide similar estimates to 1000 pseudorandom draws though the former result in lower standard deviations of the parameter estimates. However, having estimated the models with 100, 500 and 600 Halton draws, with and without antithetics, we found that the standard deviations of the parameters with fewer Halton draws are not unambiguously lower and this questions the validity of simplistic definitive statements about the trade-off between estimation time and accuracy.[24]

Given the estimates from the reduced form model we can simulate the expression for $E(\nu_{it}|x_{it}, U_i)$.

[22] Randomised Halton draws have identical properties of coverage over observations as Hatlon draws, however, they are not systematic at least in the same manner pseudorandom numbers are random (see Train, 2003, pp.234-235).

[23] Note that, Halton sequences already induce a negative correlation over observations. It is customary however to supplement them with antithetic draws to further reduce the variance of the simulator.

[24] Capellari and Jenkins (2006, p.174) reach a similar conclusion while Train (2003) himself underlines the need for further research on the properties of Halton sequences in simulation-based inference.

Taking again R draws from $f(\vartheta_i|\sigma_{\vartheta_i})$ expression A_5 is approximated by:

$$\tilde{\nu_{it}} = \frac{1}{\frac{1}{R}\sum_{r}^{R} f(U_i|x_{it}, \vartheta_i^r)} \frac{1}{R}\sum_{r}^{R}[\vartheta_i^r + E(\eta_{it}|x_{it}, U_i, \vartheta_i^r)]f(U_i|x_{it}, \vartheta_i^r) \qquad (\mathrm{A}_9)$$

The individual specific means $E(\bar{\nu_i}|\ x_{it}, U_i)$ can be computed using: $\bar{\nu}_i = \frac{1}{T_i}\sum_{t=1}^{Ti} \tilde{\nu_{it}}$. Substituting the estimates for $E(\nu_{it}|x_{it}, U_i)$ and $E(\bar{\nu_i}|\ x_{it}, U_i)$ into $(eq.8, 9)$ we obtain the endogeneity correction terms (C_i, C_{it}).

References

Abowd, J. M. and Farber, H. S. (1982). 'Job Queues and the Union Status of Workers', *Industrial and Labor Relations Review*, Vol. 35, pp. 354-367.

Arellano, M., Bover, O. and Labeaga, J.M. (1997). 'Autoregressive Models with Sample Selectivity for Panel Data', Paper 9706, *Centro de Estudios Monetarios Y Financieros*, Madrid.

Blanchflower, D. G., Oswald, A. J. and Garrett, M. D. (1990). 'Insider Power in Wage Determination', *Economica*, Vol. 57, pp. 143-170.

Blanchflower, D. G. and Bryson, A. (2007). 'The Wage Impact of Trade Unions in the UK Public and Private Sectors', *IZA*, Discussion Paper, No.3055.

Brown, C. and Medoff, J. (1989). 'The Employer Size-Wage Effect', *Journal of Political Economy*, Vol. 97, pp.1027-1059.

Cameron, A. C. and Trivedi, P. K. (2005). '*Microeconometrics: Methods and Applications*', Cambridge University Press.

Cappellari, L. and Jenkins, S. P. (2006).'Calculation of Multivariate Normal Probabilities by Simulation, with Applications to Maximum Simulated Likelihood Estimation', *The Stata Journal*, Vol. 6(2), pp.156-189.

Card, D. (1996). 'The Effect of Unions on the Structure of Wages: A Longitudinal Analysis', *Econometrica*, 64(4), pp.957-979.

Chang, J. and Lai, C. (1997). 'Union Membership and Employment Dynamics with Endogenous Union Density', *Economics Letters*, Vol. 57, pp. 119-125.

Chrysanthou, G. M. (2007). 'Determinants of Trade Union Membership in Great Britain During 1991-2003', Discussion Paper 2007/01, *University of York*.

Disney, R., Gosling, A. and Machin, S. (1996). 'What Has Happened to Union Recognition in Britain', *Economica*, Vol.63, pp. 1-18.

Drukker, D. M. (2006). 'Maximum Simulated Likelihood: Introduction to a Special Issue', *The Stata Journal*, Vol. 6(2), pp.153-155.

Duncan, G. and Leigh, D. (1985). 'The Endogeneity of Union Status: An Empirical Test', *Journal of Labor Economics,* 3, pp. 385-402.

Fernandez-Val, I. and Vella, F. (2007). 'Bias Corrections for Two-Step Fixed Effects Panel Data Estimators', *IZA,* Discussion Paper Series No.2690.

Freeman, R. B. (1984). 'Longitudinal Analyses of the Effects of Trade Unions', *Journal of Labor Economics,* 2, pp. 1-26.

Freeman, R. B, Medoff, L. J. (1984). '*What Do Unions Do?*', New York: Basic Books.

Gould, W., Pitblado, J. and Sribney, W. (2003). 'Maximum Likelihood Estimation with Stata', *Stata Press*.

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

Hausman, J. A . and Taylor, W.E. (1981). 'Panel Data and Unobservable Individual Effects', *Econometrica,* 41(2), pp. 263-80.

Heckman, J. J. (1979). 'Sample Selection Bias as a Specification Error', *Econometrica*, 47(1), pp.153-62.

Heckman, J. J. (1981a). 'Statistical Models for Discrete Panel Data', in Manski, C. F., and McFadden, D. L., Editors, '*Structural Analysis of Discrete Data and Econometric Applications*', Cambridge: The MIT Press.

Heckman, J. J. (1981b). 'The Incidental Parameters Problem and the Problem of Initial Conditions in Estimating a Discrete Time-Discrete Data Stochastic

Process', in Manski, C. F. and McFadden, D. L., Editors, '*Structural Analysis of Discrete Data and Econometric Applications*', Cambridge: The MIT Press.

Heywood, J.S. (1990). 'Who Queues for a Union Job?', *Industrial Relations*, Vo.29, pp. 99-127.

Jakubson, G. (1991). 'Estimation and Testing of the Union Wage Effect Using Panel Data', *Review of Economic Studies*, 58, pp. 971-991.

Jovanovic, B., and Moffitt, R. (1990). 'An Estimate of a Sectorial Model of Labor Mobility', *Journal of Political Economy*, 98, pp.827-852.

Krishnakumar, J. (1996). 'Simultaneous Equations: Simultaneous Error Component Models with Censored Endogenous Variables' in Mátyás, L., Sevestre, P. (Eds.), *'The Econometrics of Panel Data: A Handbook of the Theory with Applications'*, 2nd rev. ed., Kluwer Academic Publishers.

Lee, L. (1978). 'Unionism and Wage Rates: A Simultaneous Equations Model with Qualitative and Limited Dependent Variables', *International Economic Review,* 19, pp. 415-433.

Lewis, H. G. (1986). 'Union Relative Wage Effects: A Survey', Chicago: *University of Chicago Press.*

Lindbeck, A. and Snower, D. (1986) 'Wage Setting, Unemployment and Insider-Outsider Relations', *American Economic Review*, Papers and Proceedings, 76, pp. 235-239.

Maire, D. and Schjerning, B. (2007). 'Earnings, Uncertainty, and the Self-Employment Choice', Discussion Paper 2007-04, *CEBR.*

Newey, W. K. (1984). 'A Method of Moments Interpretation of Sequential Estimators', *Economics Letters*, 14, pp. 201-206.

Newey, W. K. (1987). 'Efficient Estimation of Limited Dependent Variable Models with Endogenous Explanatory Variables', *Journal of Econometrics*, Vol. 36 (3), pp. 231-250.

Orme, C. D. (2001). 'Two Step Inference in Dynamic Non-Linear Panel Data Models', School of Economic Studies, *University of Manchester*, mimeo.

Pagan, A. and Vella, F. (1989). 'Diagnostic Tests for Models Based on Unit Record Data: A Survey', *Journal of Applied Econometrics*, 4, pp.29-60.

Robinson, C. (1989a). 'The Joint Determination of Union Status and Union Wage Effects: Some Tests of Alternative Models', *Journal of Political Economy*, 97, pp. 639-667.

Robinson, C. (1989b). 'Union Endogeneity and Self-selection', *Journal of Labor Economics*, 7(1), pp.106-12.

Stewart, M. B. (1995). 'Union Wage Differentials in an Era of Declining Unionization', *Oxford Bulletin of Economics & Statistics*, Vol. 57, pp.143-66.

Stewart, M. B. (2006). '-redprob- A Stata Program for the Heckman Estimator of the Random Effects Dynamic Probit Model', *University of Warwick*, mimeo.

Stewart, M. B. (2007). 'The Inter-related Dynamics of Unemployment and Low-Wage Employment', *Journal of Applied Econometrics*, Vol. 22(3), pp. 511-531.

Swaffield, J. (2001). 'Does Measurement Error Bias Fixed-Effects Estimates of the Union Wage Effect?', *Oxford Bulletin of Economics & Statistics*, Vol. 63, pp. 437-457.

Topel, R. H. and Ward, M.P. (1992). 'Job Mobility and the Careers of Young Men', *Quarterly Journal of Economics*, 107, pp.439-79.

Train, K. E. (2003). *'Discrete Choice Methods with Simulation'*, Cambridge University Press.

University of Essex, *Institute for Social and Economic Research*, British Household Panel Survey; Waves 1-12, 1991-2003 [computer file]. Colchester, Essex: UK Data Archive [distributor], SN: 4967, June, 2004.

Vella, F. and Verbeek, M. (1998). 'Whose Wages Do Unions Raise? A Dynamic Model of Unionism and Wage Rate Determination for Young Men', *Journal of Applied Econometrics*, 13(2), pp.163-183.

Vella, F. and Verbeek, M. (1999a). 'Estimating and Interpreting Models with Endogenous Treatment Effects', *Journal of Business & Economic Statistics*, 17, pp. 473-478.

Vella, F. and Verbeek, M. (1999b). 'Two-step Estimation of Panel Data Models with Censored Endogenous Variables and Selection Bias', *Journal of Econometrics*, 90, pp. 239-263.

Wooldridge, J. M. (2005). 'Simple Solutions to the Initial Conditions Problem in Dynamic, Nonlinear Panel Data Models with Unobserved Effects', *Journal of Applied Econometrics*, Vol. 20, pp.39-54.

8. APPENDIX III: TABLES

TABLE 1

Descriptive Statistics

	1991-1996				1997-2002			
Gender	**Male**		**Female**		**Male**		**Female**	
Variable	**Mean**	**Std. Dev.**	**Mean**	**Std.Dev.**	**Mean**	**Std. Dev.**	**Mean**	**Std.Dev.**
Log of Hourly Wage	**2.073**	**0.464**	**1.746**	**0.472**	**2.264**	**0.466**	**1.995**	**0.461**
Trade Union Membership	**0.4099**	**0.492**	**0.3173**	**0.465**	**0.3471**	**0.476**	**0.3474**	**0.476**
Log (1+Potential Experience)	3.223	0.634	3.166	0.715	3.558	0.456	3.544	0.467
Marital Status	0.692	0.462	0.683	0.465	0.665	0.472	0.642	0.479
Full-Time Employment	_	_	0.672	0.469	_	_	0.709	0.454
Maternity Leave	_	_	0.011	0.104	_	_	0.013	0.113
Black (Caribbean, African, Other)	0.001	0.035	0.005	0.068	0.002	0.047	0.006	0.079
Asian (Indian, Pakistani, Chinese, Other)	0.014	0.116	0.008	0.090	0.013	0.113	0.013	0.111
Other Ethnic Minority Group	0.009	0.093	0.005	0.068	0.007	0.080	0.005	0.072
Inner/ Outer London and R of South East	0.300	0.458	0.300	0.458	0.269	0.444	0.282	0.450
South West	0.097	0.296	0.070	0.255	0.094	0.291	0.095	0.293
Midlands	0.167	0.373	0.158	0.365	0.192	0.394	0.172	0.377
Scotland	0.078	0.269	0.097	0.296	0.078	0.268	0.089	0.285
Wales	0.053	0.224	0.047	0.211	0.056	0.230	0.050	0.219
North West	0.110	0.313	0.112	0.316	0.102	0.303	0.111	0.315
North East	0.150	0.358	0.175	0.380	0.170	0.376	0.160	0.366
East Anglia	0.044	0.205	0.041	0.198	0.040	0.195	0.041	0.198
University Degree or Higher	0.158	0.365	0.109	0.312	0.169	0.375	0.133	0.340
HND, HNC, Teaching	0.082	0.274	0.074	0.263	0.087	0.281	0.078	0.268
A Levels	0.239	0.426	0.155	0.362	0.267	0.442	0.225	0.418
O Levels or CSE	0.333	0.471	0.426	0.495	0.345	0.475	0.409	0.492
Fair, Poor, V Poor Self-Assessed Health	0.157	0.364	0.199	0.399	0.181	0.385	0.206	0.404
Workforce >500	0.213	0.410	0.156	0.363	0.206	0.405	0.173	0.379
Workforce 100-499	0.303	0.460	0.228	0.419	0.297	0.457	0.213	0.409
Workforce 25-99	0.271	0.445	0.275	0.447	0.248	0.432	0.271	0.445
Workforce <25	0.213	0.409	0.340	0.474	0.249	0.432	0.343	0.475
Industrial Classification Dummies								
Agriculture, Forestry & Fishing	0.016	0.125	0.006	0.080	0.006	0.078	0.005	0.067
Energy and Water Supplies	0.051	0.220	0.007	0.085	0.025	0.155	0.007	0.083
Extraction of Minerals & Manufacture of Metals	0.052	0.221	0.018	0.134	0.053	0.225	0.014	0.116
Metal Goods, Engineering & Vehicles Industries	0.156	0.363	0.050	0.219	0.138	0.345	0.030	0.171
Other Manufacturing Industries	0.122	0.328	0.067	0.250	0.151	0.359	0.062	0.241
Construction	0.044	0.204	0.006	0.076	0.055	0.227	0.006	0.079
Distribution, Hotels & Catering (Repairs)	0.116	0.320	0.177	0.382	0.122	0.327	0.185	0.388
Transport & Communication	0.091	0.288	0.030	0.171	0.105	0.307	0.036	0.187
Banking & Finance	0.134	0.341	0.153	0.360	0.143	0.350	0.154	0.361
Public Administration, Education, Other	0.219	0.414	0.484	0.500	0.202	0.401	0.502	0.500
Occupational Classification Dummies								
Professional Occupations	0.124	0.330	0.107	0.310	0.098	0.297	0.106	0.308
Managers & Administrators	0.180	0.384	0.097	0.296	0.196	0.397	0.120	0.325
Associate Professional & Technical	0.108	0.311	0.124	0.329	0.110	0.312	0.141	0.349
Clerical & Secretarial	0.099	0.298	0.333	0.471	0.081	0.273	0.302	0.459
Craft & related	0.182	0.386	0.026	0.161	0.202	0.401	0.021	0.143
Personal & Protective Service	0.069	0.253	0.117	0.322	0.059	0.235	0.131	0.338
Sales	0.033	0.179	0.076	0.266	0.035	0.183	0.090	0.287
Plant & Machine Operatives	0.152	0.359	0.037	0.188	0.158	0.365	0.030	0.170
Other Occupations	0.054	0.227	0.081	0.273	0.063	0.243	0.059	0.235
Number of Observations	**4818**		**5172**		**5538**		**5760**	
Source: BHPS (1991-2003), ISER, Essex, SN:4967, June 2004								

TABLE 2

Wage Regressions (1991-1996), Males

	OLS		Fixed Effects		Hierarchical		Unrestricted	
Variable	**Coef.**	**t**	**Coef.**	**t**	**Coef.**	**t**	**Coef.**	**t**
Trade Union Membership	0.086	7.51	0.037	2.88	0.088	4.19	0.068	1.96
Potential Labour Market Experience	0.123	7.16	0.422	12.41	0.146	7.06	0.137	6.51
Squared Experience	0.000	-2.91	0.000	13.00	0.000	-2.75	0.000	-2.53
Marital Status	0.143	12.35	0.009	0.40	0.135	10.80	0.135	10.59
Black(Carribean, African, Other)	-0.035	-0.36	_	_	0.006	0.05	0.012	0.11
Asian(Indian, Pakistani, Chinese, Other)	0.036	0.90	_	_	0.056	1.30	0.058	1.34
Other Ethnic Minority Group	-0.056	-1.14	_	_	-0.040	-0.73	-0.037	-0.67
Inner/ Outer London and R of South East	0.200	12.57	0.032	0.39	0.175	10.33	0.178	10.41
South West	0.125	5.99	0.103	1.27	0.087	3.77	0.089	3.79
Scotland	0.022	1.02	-0.114	-0.83	-0.004	-0.18	0.002	0.10
Wales	0.038	1.59	0.042	0.40	0.049	1.91	0.046	1.73
North West	0.080	3.92	-0.168	-1.15	0.090	4.01	0.089	3.86
North East	0.088	4.73	0.246	1.82	0.084	4.21	0.086	4.30
East Anglia	0.092	3.31	-0.007	-0.07	0.064	2.11	0.067	2.21
University Degree or Higher	0.360	13.71	0.161	1.58	0.346	12.42	0.347	12.36
HND, HNC, Teaching	0.216	8.76	0.046	0.42	0.225	8.24	0.221	7.76
A Levels	0.180	9.48	0.204	1.92	0.185	8.99	0.184	8.89
O Levels or CSE	0.106	6.39	0.089	1.07	0.116	6.45	0.115	6.30
Fair, Poor, V Poor Self-Assessed Health	-0.018	-1.34	-0.004	-0.37	-0.026	-1.83	-0.028	-1.99
Workforce >500	0.172	10.14	0.033	1.77	0.186	9.91	0.171	8.36
Workforce 100-499	0.118	7.43	0.031	1.89	0.125	7.28	0.110	5.95
Workforce 25-99	0.086	5.31	0.023	1.47	0.084	4.80	0.074	4.17
Agriculture, Forestry & Fishing	-0.356	-6.86	-0.256	-3.57	-0.311	-5.80	-0.308	-5.78
Extraction of Minerals & Manufacture of Metals	-0.244	-8.34	-0.069	-1.52	-0.235	-7.46	-0.233	-7.35
Metal Goods, Engineering & Vehicles Industries	-0.281	-11.45	-0.099	-2.30	-0.261	-9.78	-0.253	-9.38
Other Manufacturing Industries	-0.280	-10.73	-0.089	-1.94	-0.264	-9.31	-0.261	-9.15
Construction	-0.239	-7.60	-0.133	-2.78	-0.226	-6.55	-0.223	-6.45
Distribution, Hotels & Catering (Repairs)	-0.414	-14.61	-0.189	-4.08	-0.391	-12.94	-0.390	-12.79
Transport & Communication	-0.291	-10.48	-0.141	-2.97	-0.281	-9.45	-0.276	-9.20
Banking & Finance	-0.140	-4.95	-0.146	-3.35	-0.133	-4.43	-0.133	-4.38
Public Administration, Education, Other	-0.258	-10.53	-0.128	-2.67	-0.235	-8.96	-0.241	-8.86
Professional Occupations	0.399	13.31	-0.022	-0.53	0.398	12.02	0.410	11.91
Managers & Administrators	0.497	17.26	0.002	0.05	0.470	14.51	0.491	14.02
Associate Professional & Technical	0.400	13.77	-0.004	-0.10	0.382	11.81	0.397	11.83
Clerical & Secretarial	0.136	4.80	-0.023	-0.60	0.110	3.50	0.119	3.69
Craft & related	0.157	5.96	-0.015	-0.38	0.152	5.26	0.157	5.42
Personal & Protective Service	0.283	8.72	-0.004	-0.08	0.269	7.57	0.271	7.64
Sales	0.251	5.72	0.001	0.02	0.228	5.05	0.248	5.34
Plant & Machine Operatives	0.120	4.57	0.003	0.08	0.111	3.80	0.119	4.01
Ci	_	_	_	_	-0.085	-3.23	-0.072	-1.79
Cit	_	_	_	_	-0.005	-0.30	-0.196	-3.42
Ci*Union Membership	_	_	_	_	_	_	0.097	1.06
Cit*Union Membership	_	_	_	_	_	_	0.188	2.20
Ci Squared	_	_	_	_	0.230	5.48	0.219	3.57
Cit Squared	_	_	_	_	0.009	0.86	-0.122	-3.24
(Ci*Union Membership)Squared	_	_	_	_	_	_	-0.117	-1.24
(Cit*Union Membership)Squared					_	_	0.148	2.12
Constant	1,145	18.98	0.164	1.11	1,083	14.98	1,109	14.69
R Squared	0.455		_		0.455		0.457	

Source: BHPS (1991-2003), ISER, Essex, SN:4967, June 2004

TABLE 3

Wage Regressions (1997-2002), Males

	OLS		Fixed Effects		Hierarchical		Unrestricted	
Variable	**Coef.**	**t**	**Coef.**	**t**	**Coef.**	**t**	**Coef.**	**t**
Trade Union Membership	0.111	10.47	0.055	3.59	0.056	2.60	0.067	2.61
Potential Labour Market Experience	0.158	5.67	0.730	12.45	0.157	4.94	0.158	4.96
Squared Experience	0.000	-7.11	0.000	9.92	0.000	-5.98	0.000	-6.01
Marital Status	0.199	19.22	0.082	4.44	0.178	15.59	0.179	15.70
Black(Carribean, African, Other)	0.222	2.21	_	_	0.201	1.68	0.211	1.80
Asian(Indian, Pakistani, Chinese, Other)	-0.055	-1.14	_	_	-0.076	-1.39	-0.069	-1.25
Other Ethnic Minority Group	-0.023	-0.41	_	_	0.043	0.65	0.019	0.29
Inner/ Outer London and R of South East	0.146	10.36	0.096	0.77	0.153	10.09	0.151	9.99
South West	0.129	6.76	0.085	0.81	0.125	6.04	0.129	6.22
Scotland	0.053	2.64	-0.069	-0.50	0.046	2.04	0.051	2.24
Wales	0.002	0.09	0.057	0.67	-0.019	-0.83	-0.016	-0.63
North West	0.052	2.98	-0.074	-0.67	0.026	1.22	0.032	1.40
North East	0.002	0.13	0.032	0.34	-0.008	-0.48	-0.003	-0.18
East Anglia	0.046	1.85	0.047	0.22	0.028	1.03	0.034	1.23
University Degree or Higher	0.417	20.41	0.004	0.11	0.453	19.90	0.444	19.44
HND, HNC, Teaching	0.286	13.41	0.171	1.70	0.327	13.35	0.314	12.97
A Levels	0.170	10.57	-0.021	-0.28	0.175	9.92	0.174	9.89
O Levels or CSE	0.117	8.02	-0.157	-1.85	0.117	7.37	0.118	7.40
Fair, Poor, V Poor Self-Assessed Health	-0.053	-4.09	-0.001	-0.09	-0.042	-3.00	-0.044	-3.11
Workforce >500	0.201	13.57	0.032	1.71	0.165	9.41	0.173	9.55
Workforce 100-499	0.149	10.92	0.027	1.55	0.113	6.66	0.122	6.99
Workforce 25-99	0.067	4.93	0.011	0.71	0.032	2.09	0.038	2.45
Professional Occupations	0.411	16.15	0.062	1.94	0.397	13.89	0.403	13.97
Managers & Administrators	0.473	20.78	0.125	3.92	0.467	18.42	0.468	18.27
Associate Professional & Technical	0.428	17.01	0.066	2.00	0.417	14.81	0.419	14.75
Clerical & Secretarial	0.128	5.32	0.032	0.96	0.125	4.57	0.127	4.60
Craft & related	0.228	10.72	0.092	2.84	0.226	9.47	0.228	9.44
Personal & Protective Service	0.252	8.94	0.119	1.94	0.240	7.68	0.245	7.74
Sales	0.304	9.51	0.095	2.26	0.293	8.45	0.292	8.41
Plant & Machine Operatives	0.152	7.03	0.046	1.48	0.148	6.10	0.150	6.13
Ci	_	_	_	_	0.176	3.22	0.276	3.22
Cit	_	_	_	_	0.004	0.27	0.013	0.69
Ci*Union Membership	_	_	_	_	_	_	-0.273	-1.64
Cit*Union Membership	_	_	_	_	_	_	-0.019	-0.67
Ci Squared	_	_	_	_	-0.595	-2.81	_	_
Cit Squared	_	_	_	_	0.002	0.11	_	_
Constant	0.949	11.27	-1.108	-5.64	1.055	10.81	1.043	10.60
R Squared	0.452		_		0.440		0.440	

Source: BHPS (1991-2003), ISER, Essex, SN:4967, June 2004

TABLE 4

Wage Regressions (1991-1996), Females

	OLS		**Fixed Effects**		**Hierarchical**		**Unrestricted**	
Variable	**Coef.**	**t**	**Coef.**	**t**	**Coef.**	**t**	**Coef.**	**t**
Trade Union Membership	0.127	11.42	0.018	1.30	0.162	7.22	0.202	5.57
Potential Labour Market Experience	0.067	4.80	0.405	14.10	0.044	2.69	0.038	2.34
Squared Experience	0.000	-5.29	0.000	12.85	0.000	-3.90	0.000	-3.81
Marital Status	0.045	4.05	0.022	1.24	0.032	2.66	0.018	1.31
Full-Time Employment	0.017	1.35	-0.119	-5.57	0.001	0.05	-0.020	-1.19
Maternity Leave	-0.353	-3.64	-0.402	-4.95	-0.366	-3.81	-0.379	-3.97
Black(Carribean, African, Other)	0.113	2.48	_	_	0.084	1.69	0.057	1.12
Asian(Indian, Pakistani, Chinese, Other)	0.108	2.30	_	_	0.102	1.93	0.082	1.51
Other Ethnic Minority Group	0.082	1.13	_	_	0.120	1.47	0.140	1.71
Inner/ Outer London and R of South East	0.175	11.45	0.085	1.10	0.173	10.52	0.185	10.84
South West	0.002	0.09	-0.024	-0.20	0.014	0.55	0.024	0.93
Scotland	0.091	4.63	0.406	2.67	0.074	3.55	0.057	2.63
Wales	0.073	3.23	0.154	0.88	0.060	2.44	0.035	1.30
North West	0.080	4.22	0.204	0.81	0.076	3.73	0.071	3.48
North East	-0.023	-1.46	0.038	0.22	-0.033	-1.98	-0.053	-2.88
East Anglia	-0.024	-0.90	0.737	3.29	-0.029	-0.99	-0.018	-0.59
University Degree or Higher	0.258	10.00	0.066	1.05	0.253	9.04	0.238	7.91
HND, HNC, Teaching	0.192	8.00	0.004	0.07	0.179	6.76	0.168	5.92
A Levels	0.083	4.73	-0.021	-0.48	0.084	4.50	0.067	3.24
O Levels or CSE	0.052	4.13	-0.063	-1.52	0.051	3.79	0.043	3.02
Fair, Poor, V Poor Self-Assessed Health	-0.041	-3.56	-0.007	-0.65	-0.042	-3.33	-0.044	-3.51
Workforce >500	0.165	11.01	0.035	1.94	0.145	8.86	0.122	7.00
Workforce 100-499	0.136	10.17	0.040	2.60	0.118	7.99	0.107	7.15
Workforce 25-99	0.077	6.33	0.015	1.06	0.064	4.89	0.057	4.28
Agriculture, Forestry & Fishing	-0.249	-3.26	-0.062	-0.64	-0.199	-2.57	-0.186	-2.34
Extraction of Minerals & Manufacture of Metals	-0.064	-1.09	0.074	1.00	-0.027	-0.46	-0.004	-0.07
Metal Goods, Engineering & Vehicles Industries	-0.171	-3.36	0.045	0.68	-0.156	-3.05	-0.143	-2.73
Other Manufacturing Industries	-0.232	-4.40	-0.008	-0.12	-0.218	-4.08	-0.205	-3.80
Construction	-0.170	-2.46	0.049	0.57	-0.168	-2.51	-0.165	-2.34
Distribution, Hotels & Catering (Repairs)	-0.364	-7.22	0.007	0.11	-0.357	-7.06	-0.350	-6.79
Transport & Communication	-0.176	-3.07	-0.019	-0.23	-0.166	-2.91	-0.152	-2.63
Banking & Finance	-0.108	-2.20	0.012	0.18	-0.111	-2.26	-0.097	-1.92
Public Administration, Education, Other	-0.182	-3.72	0.019	0.29	-0.189	-3.87	-0.187	-3.76
Professional Occupations	0.582	20.81	0.088	1.62	0.596	19.43	0.586	18.94
Managers & Administrators	0.572	21.98	0.111	2.23	0.601	21.64	0.616	21.64
Associate Professional & Technical	0.472	21.26	0.090	1.81	0.479	19.87	0.469	19.56
Clerical & Secretarial	0.274	15.44	0.077	1.58	0.287	14.88	0.294	14.95
Craft & related	0.102	2.85	0.107	1.60	0.093	2.36	0.069	1.82
Personal & Protective Service	0.122	6.04	-0.010	-0.21	0.136	6.15	0.136	6.17
Sales	0.119	4.73	-0.031	-0.62	0.126	4.48	0.148	4.98
Plant & Machine Operatives	0.001	0.02	0.060	0.83	0.016	0.46	0.019	0.56
Ci	_	_	_	_	0.013	0.54	0.103	1.62
Cit	_	_	_	_	-0.081	-5.58	-0.243	-5.01
Ci*Union Membership	_	_	_	_	_	_	-0.226	-1.86
Cit*Union Membership	_	_	_	_	_	_	0.158	2.27
Ci Squared	_	_	_	_	_	_	-0.059	-0.55
Cit Squared	_	_	_	_	_	_	-0.094	-2.75
(Ci*Union Membership)Squared	_	_	_	_	_	_	0.099	0.73
(Cit*Union Membership)Squared	_	_	_	_	_	_	0.113	1.80
Constant	1.126	17.27	-0.087	-0.64	1.253	17.76	1.324	17.12
R Squared	0.509		_		0.516		0.518	

Source: BHPS (1991-2003), ISER, Essex, SN:4967, June 2004

TABLE 5

Wage Regressions (1997-2002), Females

	OLS		Fixed Effects		Hierarchical		Unrestricted	
Variable	**Coef.**	**t**	**Coef.**	**t**	**Coef.**	**t**	**Coef.**	**t**
Trade Union Membership	0.114	11.58	0.049	3.55	0.148	7.17	0.109	3.97
Potential Labour Market Experience	0.015	0.55	0.660	10.13	0.008	0.25	0.009	0.27
Squared Experience	0.000	-3.53	0.000	8.84	0.000	-2.96	0.000	-3.04
Marital Status	0.044	4.30	0.064	3.00	0.040	3.19	0.047	3.59
Full-Time Employment	0.021	1.94	-0.074	-4.02	0.033	2.47	0.038	2.72
Maternity Leave	-0.209	-2.99	-0.240	-3.75	-0.201	-2.69	-0.202	-2.70
Black(Carribean, African, Other)	-0.087	-3.00	_	_	-0.075	-2.23	-0.025	-0.58
Asian(Indian, Pakistani, Chinese, Other)	-0.139	-3.24	_	_	-0.132	-2.75	-0.131	-2.75
Other Ethnic Minority Group	-0.022	-0.35	_	_	-0.027	-0.38	-0.007	-0.10
Inner/ Outer London and R of South East	0.171	12.10	0.033	0.56	0.160	9.05	0.153	8.49
South West	0.055	3.28	0.031	0.39	0.041	2.18	0.037	1.94
Scotland	0.087	4.47	0.305	1.77	0.081	3.79	0.087	4.04
Wales	0.073	2.96	0.113	1.19	0.085	3.02	0.105	3.59
North West	0.061	3.56	0.125	0.87	0.055	2.91	0.053	2.77
North East	-0.005	-0.31	-0.007	-0.15	-0.001	-0.06	0.005	0.29
East Anglia	-0.027	-1.10	-0.042	-0.28	-0.039	-1.40	-0.046	-1.63
University Degree or Higher	0.315	14.43	0.380	3.73	0.345	14.24	0.358	14.18
HND, HNC, Teaching	0.232	10.48	0.312	2.53	0.258	10.15	0.268	10.33
A Levels	0.115	6.90	0.135	1.29	0.145	7.67	0.151	7.85
O Levels or CSE	0.087	6.24	0.089	1.14	0.097	6.26	0.098	6.35
Fair, Poor, V Poor Self-Assessed Health	-0.047	-4.19	-0.005	-0.45	-0.048	-3.80	-0.048	-3.86
Workforce >500	0.165	12.33	0.059	2.75	0.170	10.85	0.176	11.04
Workforce 100-499	0.166	12.96	0.056	3.22	0.171	11.65	0.176	11.83
Workforce 25-99	0.101	8.18	0.036	2.51	0.097	7.12	0.099	7.29
Professional Occupations	0.419	14.66	0.069	1.80	0.398	12.37	0.405	12.54
Managers & Administrators	0.444	16.95	0.095	2.73	0.436	15.35	0.434	15.22
Associate Professional & Technical	0.400	16.37	0.086	2.33	0.394	14.56	0.396	14.62
Clerical & Secretarial	0.223	10.32	0.018	0.52	0.212	8.97	0.212	8.97
Craft & related	0.100	2.93	0.002	0.03	0.115	3.12	0.115	3.12
Personal & Protective Service	0.034	1.44	-0.045	-1.28	0.028	1.08	0.030	1.14
Sales	-0.088	-3.51	-0.108	-2.80	-0.090	-3.27	-0.092	-3.32
Plant & Machine Operatives	0.031	0.97	-0.030	-0.61	0.021	0.60	0.022	0.62
Ci	_	_	_	_	-0.115	-2.42	-0.282	-3.13
Cit	_	_	_	_	-0.027	-1.59	-0.024	-1.22
Ci*Union Membership	_	_	_	_	_	_	0.451	2.65
Cit*Union Membership	_	_	_	_	_	_	0.020	0.79
CiSquared	_	_	_	_	0.375	2.12	_	_
CitSquared	_	_	_	_	-0.001	-0.05	_	_
Constant	1.336	16.71	-1.240	-5.94	1.344	13.81	1.318	13.39
R Squared	0.465		_		0.461		0.462	

Source: BHPS (1991-2003), ISER, Essex, SN:4967, June 2004

TABLE 6

Descriptive Statistics, Endogeneity Correction Terms (1991-1996), Males

	Union Members		Non-Union Members	
Latent Effect	Mean	Std. Dev.	Mean	Std. Dev.
Ci	0.481	0.229	-0.213	0.252
Cit	0.265	0.402	-0.093	0.391
Ci Squared	0.284	0.207	0.109	0.139
Cit Squared	0.232	0.582	0.162	0.544

TABLE 7

Descriptive Statistics, Endogeneity Correction Terms (1997-2002), Males

	Union Members		Non-Union Members	
Latent Effect	Mean	Std. Dev.	Mean	Std. Dev.
Ci	0.178	0.090	-0.161	0.097
Cit	0.182	0.385	-0.056	0.393
Ci Squared	0.040	0.040	0.035	0.030
Cit Squared	0.181	0.421	0.158	0.368

TABLE 8

Descriptive Statistics, Endogeneity Correction Terms (1991-1996), Females

	Union Members		Non-Union Members	
Latent Effect	Mean	Std. Dev.	Mean	Std. Dev.
Ci	0.438	0.239	-0.263	0.201
Cit	0.355	0.434	-0.089	0.364
Ci Squared	0.249	0.229	0.110	0.122
Cit Squared	0.314	0.567	0.140	0.422

TABLE 9

Descriptive Statistics, Endogeneity Correction Terms (1997-2002), Females

	Union Members		Non-Union Members	
Latent Effect	Mean	Std. Dev.	Mean	Std. Dev.
Ci	0.206	0.125	-0.147	0.116
Cit	0.240	0.398	-0.091	0.402
Ci Squared	0.058	0.060	0.035	0.033
Cit Squared	0.216	0.410	0.170	0.397