We study union wage effects when membership is non-coercive and employers reserve the right of union recognition. Using the BHPS during 1995-2009 we find evidence of membership differentials for the lower male and middle female observed skill groups. The coverage differential at the bottom of the observed male skill distribution suggests free riding. While union members are negatively selected, there is evidence of positive selection at the bottom of the observed skill distribution. Using the unified measurement error and endogeneity bias expression, we obtain a discernible pattern between uncorrected, endogeneity corrected and fixed effects estimates of the union wage effect.

Key Words: union wage differentials, measurement error, unobserved heterogeneity, endogeneity

Subject Classification: C33, C35, J31, J51

1. INTRODUCTION

The estimation of union wage differentials is one of the most heavily studied issues in labour economics. While there is an extensive and refined literature, in many cases there is lack of a consensus regarding whether union wage differentials represent the impact of unions, are a product of selectivity or indicate insufficient observable controls.

The joint determination of union membership and wages means that in exploring how observationally equivalent employees’ wages differ we cannot ignore how the unobserved heterogeneity underlying membership decisions is rewarded. Literally every paper in the field stresses that:

"Union members differ from nonmembers in unobserved ways, biasing your estimates...make a selectivity bias correction...simultaneously determine union status and economic outcomes...Use longitudinal data" (Freeman, 1984, p.2).

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 especially grateful to Pedro Albarran for his helpful comments and our discussions. I also wish to thank Leslie Godfrey, Rasmus Jørgensen, Keith Hartley, Andrew Jones, John Hutton, Wim Koevets, Mark Bryan, Angel Leon, participants at the Applied Econometrics Workshop of Universidad Carlos III de Madrid, the 63rd AFSE Annual Meeting, the 24th EALE Annual Conference and the Understanding Society 2013 Research Conference. I gratefully acknowledge financial support from the Spanish Ministry of Education by the grant ECO2013-43119-P and the University of Alicante by the grant GRE13-04.
Longitudinal analyses, however, do not provide a research panacea due to the substantial impact of measurement error in membership status on the estimated wage differentials (see Freeman 1984, Card 1996, Swafield, 2001).

Additionally, as wage rate standardisation implies reduced human capital premia the probability of joining establishments recognising unions for wage bargaining is inversely related to skill level. Sorting into union employment is likely to follow a multiple indices rule in which event, instrumentation using either the control function approach or instrumental variables becomes problematic (see Farber, 1982; Card, 1996; Heckman et al., 2006).

We analyse the impact of British unions during 1995-2009. The British case is of particular interest, given the effective outlaw of pre-entry and post-entry closed shops in the advent of the 1980, 1982, 1988 and 1990 Employment Acts.3

These legislative changes raise the issue of whether membership wage premia can actually exist since by law employees within establishments covered by bargaining agreements should receive the same wage independently of membership status.

In the absence of coercion, unions face a free rider problem given the positive monetary membership cost. Empirical evidence from both the U.K. and U.S., however, is inconclusive and many authors suggested the existence of membership premia in which case there is a free rider puzzle instead i.e. as to why covered nonmembers abstain from membership and renounce the respective premium.

Recent U.K. studies estimating union wage differentials can be classified into four broad categories. Cross sectional studies employing endogeneity correction methods (Bryson 2002, Booth and Bryan 2004, Arulampalam et al., 2009), longitudinal studies using fixed effects and accounting for measurement error (e.g. Swafield, 2001), longitudinal studies using instrumental variables and accounting for measurement error (Hildreth, 2000, Koevets, 2007) and studies relying on the conditional independence assumption (e.g. Blanchflower and Bryson, 2010).

We estimate union wage differentials using British Household Panel Survey (BHPS) data. We account for measurement error, distinguish between membership and coverage premia, and use an endogeneity correction methodology that is flexible in its treatment of unobserved heterogeneity. We allow for distinct selection biases at three different skill levels defined by quantiles of predicted wages obtained from an independent sample of employees in the uncovered sector.

Building upon the contributions of Aigner (1973) and Vella and Verbeek (1998), we obtain the unified measurement error and endogeneity bias expression. It is demonstrated that the conventional attenuation bias arises only under hierarchical sorting and negative selection so that the measurement error and selectivity biases reinforce each other.

We obtain robust estimates of membership differentials at the bottom and middle of the observed male and female skill distributions of approximately 2 and 7.6 percent, respectively noting that the more parsimonious female differential corresponds to 3 percent. There is evidence of free riding at the bottom of the observed male skill distribution given the presence of a 2.7 coverage premium in the selected estimates.

While union members are negatively selected, those located at the bottom of the observed skill distribution are in general positively selected indicating the conflicting

3The overall impact was that the 1980s was a decade of a dramatic decline in aggregate union membership and recognition (see Stewart, 1995, pp.143-145) while the following decades are characterised by an initial period of rapid decline during the 1990s and an ensuing moderation in decline and relative stability of unionisation (see Bryson and Forth, 2010).
interests of employers and employees in determining unobserved differences.

Using the unified bias expression the estimates indicate a discernible pattern between uncorrected, endogeneity corrected and fixed effects (FE) estimates thus, refuting the pessimistic conclusions of Freeman and Medoff (1982) and Lewis (1986).

Sufficient measurement error reduction raises the FE estimate of the male union effect making it the middle bound, whereas, the endogeneity corrected estimate always corresponds to the upper bound. However, the strong offsetting biases occurring at two the extremes of the female observed skill distribution render the FE estimate local to zero.

Sections 2 and 3 outline the model and estimation procedure; Section 4 discusses wage equation parameterisation; Section 5 discusses data issues and sample selection; Section 6 treats estimation by predicted wage quantile; Section 7 treats measurement error and identification issues and presents the estimates and Section 8 concludes.

2. A DYNAMIC MODEL OF UNIONISM AND WAGE DETERMINATION

Estimation of union wage differentials using longitudinal data requires controlling for the endogeneity of membership status. FE and instrumental variables estimators in their generic form (e.g. Hausman and Taylor, 1981 or its GMM generalisation by Arellano and Bover, 1995) assume that this endogeneity is individual-specific and fixed and are thus restrictive in their treatment of unobserved heterogeneity (see Robinson 1989a,b; Vella and Verbeek 1999a).

We adopt the two-step estimation methodology (outlined in Vella, 1998 and generalised by Vella and Verbeek, 1999b) originally employed by Vella and Verbeek (1998) to estimate the union wage impact. This is essentially the panel data variant of Heckman’s (1979) two-step estimator that explicitly identifies the different sources of endogeneity of union membership by exploiting the longitudinal nature of the data.

Equation (1) is the primary wage equation and assumes that individuals sort themselves into union/non-union membership employment on the basis of wages which are determined by observed and unobserved attributes and their respective prices. The potential wage corresponding to individual \( i \) with union membership status \( j \), in time period \( t \) is given by \( w_{j;it} \). A non-union/union member employee is denoted by \( j \in \{0, 1\} \) respectively, \( \beta \) 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 an individual’s wage are \( (\alpha_{j;i}, \varepsilon_{j;it}) \) and it is assumed that \( \alpha_{j;i} \sim iidN(0, \sigma_\alpha^2) \) and \( \varepsilon_{j;it} \sim iidN(0, \sigma_\varepsilon^2) \):

\[
\begin{align*}
  w_{j;it} &= \beta'_{j} x_{it} + e_{j;it}, \quad e_{j;it} = \alpha_{j;i} + \varepsilon_{j;it} \\
  t &= 1, ..., T_i; \quad i = 1, ..., N; \quad j \in \{0, 1\}
\end{align*}
\] (1)

The dynamic reduced form model depicting individual membership decision is given in equation (2). Union employment benefits are captured by the latent variable \( U_{it} \). The union membership status of an individual \( i \) in period \( t \), is indicated by the binary dichotomous variable \( U_{it} \). The unknown parameters to be estimated are \( (\gamma_1', \gamma_2) \) and the composite error term \( \nu_{it} \) captures the unobserved individual heterogeneity underlying the union membership decision. This is decomposed into
an individual-specific component \( \theta_i \) and an individual time-specific effect \( \eta_{it} \) and it is assumed that \( \theta_i \sim iidN(0,\sigma^2_\theta) \), \( \eta_{it} \sim iidN(0,\sigma^2_\eta) \). The logarithm of the gross average real hourly wage rate, denoted by \( w_{it} \), corresponds to the logarithm of weekly real wage divided by usual paid hours including overtime.\(^4,5\)

\[
\begin{align*}
\tilde{U}_{it} &= \gamma_1' x_{it} + \gamma_2 U_{i,t-1} + \nu_{it}, \quad \nu_{it} = \theta_i + \eta_{it} \\
U_{it} &= 1(\tilde{U}_{it} > 0) \\
w_{it} &= w_{j,it} \quad \text{if} \quad U_{it} = j
\end{align*}
\]

(2)

(3)

(4)

Potential seniority and non-pecuniary benefits can prolong membership in the long term, irrespective of wage changes, and this introduces state dependence in the model. The inclusion of lagged union membership status in eq.\(^2\) prevents the error components from incorrectly capturing the dynamics which should be credited to \( U_{i,t-1} \).

Concerning the random components \( (\alpha_{j,i};\theta_i), (\varepsilon_{j,it};\eta_{it}) \) in eq.\(^1\),\(^2\) every effect is potentially correlated with its counterpart of the same dimension in the other equation i.e. covariances \( (\sigma_{j,\alpha\theta}, \sigma_{j,\varepsilon\eta}) \) are allowed to be non-zero.

To estimate the union wage differential we enforce the restriction that the returns to observed characteristics are invariant with respect to both time and union membership status. The wage equation \(^1\) then becomes:

\[
w_{it} = \beta' x_{it} + \delta U_{it} + \epsilon_{it}
\]

\[
\epsilon_{it} = U_{it}(\alpha_{1,i} + \varepsilon_{1,it}) + (1 - U_{it})(\alpha_{0,i} + \varepsilon_{0,it})
\]

(5)

The samples used in this study are heterogeneous. We need reasonably large sample sizes given the dynamic nature of the reduced forms and considering the subsequent sample losses when evaluating measurement error impact. We therefore do not restrict estimations to the traditional male manual employees (as Swaffield, 2001) or distinct public/private sector samples (as Blanchflower and Bryson, 2010). This might question the assumption of homogeneous returns on observables.\(^6\)

Recognising that selection into union jobs may lead to differing selection biases at different skill levels, we estimate models for three distinct skill groups thus rendering the assumption of homogeneous sector returns on observables more plausible (see Section 6).

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 \( x_{it} \):

\(^4\)The CPI-index for the UK, in 2005 consumer prices, is used as a deflator and obtained from the ONS. Estimating all models using the gross average nominal hourly wage rate instead, produced identical wage differentials since models include regional and time controls.

\(^5\)Using hourly as opposed to weekly earnings might produce a slight increase in the union wage effect assuming union employees, on average, tend to work less hours per week (see Andrews et al., 1998). Our choice is based on comparability grounds with similar studies employing BHPS data such as Swaffield (2001) and Kooves (2007).

\(^6\)We use higher-order terms of the latent effects to detect departures from normality (see Pagan and Vella, 1989).
E(w_{it} \mid x_{it}, U_i) = \beta' E(x_{it} \mid x_{it}, U_i) + \delta E(U_{it} \mid x_{it}, U_i) + E(\alpha_{j,i} \mid x_{it}, U_i) + E(\varepsilon_{j,it} \mid x_{it}, U_i)

(6)

Estimation of the reduced form, eq.(2), provides the estimates of the unobserved individual heterogeneity. This is a dynamic random effects probit model.\(^7\)

The inclusion of lagged membership in eq.(2) gives rise to the initial conditions problem (see Heckman, 1981a). We employ Wooldridge’s (2005) solution to the initial conditions problem due to its computational simplicity as opposed to Heckman’s (1981b) estimator. Adopting the Mundlak (1978)-Chamberlain (1984) specification to allow for a correlation between the unobserved effect and the time means of the observed time-varying characteristics, eq.(2) is estimated by a random effects probit where the explanatory variables at t are \(z_{it} \equiv (1, x_{it}, U_{it-1}, U_{i1}, \bar{X}_i)\).

The conditional expectations of the random components in eq.(6) are estimates of the unobserved heterogeneity taking the form of:

\[
E(\alpha_{j,i} \mid x_{it}, U_i) = \sigma_{j,\alpha\theta} \left[ \frac{T_i}{\sigma_\theta^2 + T_i \sigma_\theta^2} E(\tilde{\psi}_i \mid x_{it}, U_i) \right] = \sigma_{j,\alpha\theta} B_i
\]

(7)

\[
E(\varepsilon_{j,it} \mid x_{it}, U_i) = \sigma_{j,\varepsilon\eta} \left[ \frac{E(\psi_{it} \mid x_{it}, U_i)}{\sigma_\eta^2} - \frac{T_i \sigma_\theta^2}{\sigma_\eta^2 (\sigma_\eta^2 + T_i \sigma_\theta^2)} E(\tilde{\psi}_i \mid x_{it}, U_i) \right] = \sigma_{j,\varepsilon\eta} B_{it}
\]

(8)

where, \(\tilde{\psi}_i = T_i^{-1} \sum_{t=1}^{T_i} \nu_{it}\). Given the parameter estimates from the reduced form model \(\psi = (\gamma_1, \gamma_2, \sigma_\theta)\) we estimate \(E(\nu_{it} \mid x_{it}, U_i)\) using numerical integration (simulation)- see Vella, (1998); Vella and Verbeek (1998, 1999b).\(^8\)

Substituting the estimates for \(E(\nu_{it} \mid x_{it}, U_i)\), \(E(\tilde{\psi}_i \mid x_{it}, U_i)\), the endogeneity correction terms \((B_i, B_{it})\) defined in equations (7) and (8) can be computed and are then added as additional terms in the structural equation, with coefficients \(\sigma_{j,\alpha\theta}\) and \(\sigma_{j,\varepsilon\eta}\). These parameters can be estimated jointly with \((\beta', \delta)\) in the second step from conditional moment restrictions such as least squares based on equation (5).\(^9\)\(^,\)\(^10\)

Under exogeneity the conventional standard errors can be used. Otherwise, standard errors should be adjusted for heteroskedasticity and the inclusion of \((B_i, B_{it})-\)

\(^7\)When the intra-panel correlation coefficient is local to zero a simple pooled probit is estimated instead. This gives consistent parameter estimates provided the correct parameterisation is used along with the Mundlak (1978)-Chamberlain (1984) specification.

\(^8\)Using standard results

\[
E(\nu_{it} \mid x_{it}, U_i) = \int [E(\eta_{it} \mid x_{it}, U_i, \theta_i)] f(U_i \mid x_{it}, \theta_i) f(\theta_i) d\theta_i
\]

\[
\int f(U_i \mid x_{it}, \theta_i) f(\theta_i) d\theta_i
\]

where \(E(\eta_{it} \mid x_{it}, U_i, \theta_i)\) is the cross-sectional generalised probit residual from eq.(2) and, the term in the denominator is the likelihood contribution of individual \(i\) in eq.(2).

\(^9\)We use the random effects GLS estimator, as opposed to pooled OLS. Under exogeneity, this is the conventional random-effects estimator and exploits the serial correlation of the composite error (ignored by pooled OLS) within a GLS framework.

\(^{10}\)Fernández-Val and Vella (2011) estimate the reduced form by FE. Effective bias reduction, required by nonlinear FE estimation, is achieved when \(T_i\) is reasonably big and since FE discards observations for which \(U_{it}\) is invariant this is prohibitive in terms of sample attrition. The authors obtain similar estimates of the union differential to Vella and Verbeek (1998).

The four covariances \((\sigma_{j,\theta}, \sigma_{j,\eta})\) convey valuable information about the form of sorting into the two sectors (see Vella and Verbeek, 1999a). By construction the average value of the \(\theta_j\) is positive for union members and negative for nonmembers. For tractability assume that the endogeneity operates via the individual-specific effects \((\alpha_{j,i}, \theta_i)\). If either covariance between \((\alpha_{j,i}, \theta_i)\) is non-zero then the unobserved factors that determine union membership influence wages as well.

If both covariances \((\sigma_0, \sigma_1)\) are positive, individuals with high values of \(\theta\) are on average the best employees in terms of their endowment of unobserved productivity irrespectively of membership status (and vice-versa). This is termed as hierarchical (or restricted) sorting. A comparative advantage (or unrestricted/positive sorting) instead requires that employees perform differently in the two sectors and sort themselves appropriately. This implies a negative association between the relative productivity in the two sectors \((\sigma_{1,0} < 0)\) and demands that the contribution of unobserved heterogeneity raises wages in both sectors (i.e. \(\sigma_{1,\theta} > 0, \sigma_{0,\theta} < 0\)). Note that \(\sigma_{1,0}\) cannot be estimated directly.

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 generic instrumental variables or the FE estimator. A comparative advantage structure is precluded \textit{a priori} (see Vella and Verbeek, 1999a).

4. WAGE EQUATION PARAMETERISATION

The introduction of the 1988 and 1990 Employment Acts in the U.K., prohibited all means to enforce a closed shop and respectively rendered post and pre-entry closed shops illegal (see Stewart, 1995, pp.143-145).

Given the effective outlaw of closed shops, individuals undertake unionisation decisions on the basis of wages, individual preferences and non-pecuniary benefits. In the absence of coercion unions may offer excludable goods or services to encourage membership (see Olson, 1965).12

If all employees at covered establishments receive a fixed union negotiated mark-up, a coverage differential should arise whilst membership should not be expected to have an additional influence upon pay structure. Arguably then, in absence of coerced membership trade unions face a free rider problem.

Nevertheless, at the establishment level high union density produces a higher than average union differential and coverage on its own right is insufficient (see Stewart, 1987). At the individual level, while employees doing the same job at the same workplace should in theory earn the same when conducting analyses across establishments, membership is a closer determinant of a differential than coverage (Andrews et al., 1998, p.453).

11While in balanced panels, GLS is obtained as a simple WLS, the weights are dependent upon the lengths of the time series per cross-section in the unbalanced panel case. We use the Swamy and Arora (1972) estimator of the standard errors while allowing for individual-level clustering (see Baltagi, 2012, pp.182-184). We estimate bootstrap standard errors using 100 replications over both stages adjusted for clustering on individuals. Due to the nonlinear nature of the reduced form and the numerical integration required to obtain the endogeneity correction terms in each replication this procedure is computationally intensive.

12These include protection against unfair dismissal, discrimination, grievance procedures, pension plans advice and the implementation of well-defined dismissal arrangements in recessionary periods (see Arulampalam and Booth, 2000, p.291).
Booth and Bryan (2004) provide a summary of the rationalisations of positive membership wage premia advanced in the literature. These can be broadly classified into selectivity or omitted variable arguments and discriminatory behaviour by either unions/employers (e.g. systematically targeting training programs to members or, pay nonmembers from a point lower down the corresponding union wage scale).

Hildreth (2000) and Swafield (2001) employing BHPS data and accounting for fixed effects and measurement error, find evidence of positive differentials for covered members. However, Booth and Bryan (2004) using cross-sectional data from the linked WERS (1998) conclude that the apparent covered membership wage differential is illusory.\(^\text{13}\)

In all respects when estimating union wage effects in the U.K. it is imperative to account for three distinct classes of employees namely, covered members, covered nonmembers and uncovered employees. Following Andrews \textit{et al.}(1998), the appropriate parameterisation of eq.(5) when observations on both membership, \(m\), and coverage, \(c\), are available is

\[
\begin{align*}
  w_{it} &= \beta' x_{it} + \delta_1 m_{it} c_{it} + \delta_2 (1 - m_{it}) c_{it} + e_{it} \\
  e_{it} &= m_{it} c_{it} (\alpha_{1,i} + \varepsilon_{1,it}) + (1 - m_{it} c_{it})(\alpha_{0,i} + \varepsilon_{0,it})
\end{align*}
\]

where the unobserved components of an employee’s wage are given by \((\alpha_{j,i}, \varepsilon_{j,it})\) where \(\varepsilon_{j,it} = \alpha_{j,i} + \varepsilon_{j,it}\) and \(j = 1\) if \(m_{it} c_{it} = 1\), \(j = 0\) if \(m_{it} c_{it} = 0\).

If union status is exogenous, the total union differential is given by \(\delta_1\), the union membership differential can be defined as

\[
E(w_{it} | m_{it} = 1, c_{it} = 1; x_{it}) - E(w_{it} | m_{it} = 0, c_{it} = 1; x_{it}) = \delta_1 - \delta_2
\]

and the union coverage differential can be defined as

\[
E(w_{it} | m_{it} = 0, c_{it} = 1; x_{it}) - E(w_{it} | c_{it} = 0; x_{it}) = \delta_2.
\]

When eq.(9) is used instead of eq.(5) the estimation procedure described in Section 3 remains the same noting that, by construction the average value of the \(\theta_t\) is positive for covered members \((j = 1)\) and negative for covered nonmembers and uncovered employees \((j = 0)\).\(^\text{14}\)

5. DATA AND SAMPLES

We employ the BHPS data set to construct unbalanced panels of male and female employees during 1995-08.\(^\text{15}\)

The period choice was data driven due to the continuous availability of the key union membership (conditional upon coverage) variable since during 1992-94 this

---

\(^{13}\)Hildreth (2000) finds some evidence of a differential for covered nonmembers though this was not generally robust while, Swafield (2001) generally finds lower differentials for covered nonmembers as opposed to covered members.

\(^{14}\)Since \(w_{it}\) is in logarithmic form the exact differentials in each case are given by \((e^{\delta_1} - 1)\), \(l = 1, 2\).

was not asked of employees still in same job as in the previous year- see Section 7.1. The samples were selected according to 1995 characteristics and include individuals that were active in the labour market during 1995, do not have any missing relevant information and are below 65 years old in 2008.

We use unbalanced panels allowing individuals to exit the sample but do not allow individuals to enter the sample ex post 1995. This selection mechanism is necessary given the dynamic nature of the reduced form models. The inclusion of lagged union membership in the reduced form requires both consecutive observations and a common start date in order to facilitate the estimation of initial conditions (see Arulampalam et al., 2000).

We therefore exclude individuals that were inactive in labour market during 1995, and those that exit post-1995 and subsequently re-enter. Employing balanced panels instead, given the large time period under analysis and the subsequent restrictions for measurement error and identification purposes, is prohibitive in terms of sample attrition.

Due to few observations and lack of variation, we exclude the self-employed, individuals pertaining to the agricultural, forestry and fishing industrial classification and those belonging to the agricultural and own account socioeconomic groups.

Part-time male employees are excluded since the small gains in terms of sample size are more than outweighed by the costs of a potential increase in the heterogeneity of the male samples. In all respects, part-time controls were statistically insignificant in all male estimates.

Given the considerable number of distinct samples used in this study, we are unable to provide descriptive statistics for all explanatory variables. All Wage Regression Tables report the number of observations, number of union members/covered members and covered nonmembers. Tables providing detailed descriptions of the variables included in both estimation stages are provided in the Appendix.

6. UNION EFFECTS BY POSITION IN THE WAGE DISTRIBUTION

We initially treat the probability of union membership as a function of the single index appearing in the reduced form equation. Some key contributions in the field such as Abowd and Farber (1982) and Card (1996), suggest that sorting into sectors follows a multiple indices rule.

Provided collective bargaining produces higher standardised wages, profit maximising employers faced with a pool of potential employees, will systematically undertake their selection decisions so as to minimise production costs (Abowd and Farber, 1982, p.355).

However, since the standardisation of wage rates implies reduced skill premia, the probability of joining an establishment recognising unions for wage bargaining is expected to be inversely related to individual skill level.

In effect, while potential candidates for employment in establishments recognising unions are likely to be possessing relatively low levels of human capital, profit maximising employers will aim to hire those having the exact opposite attributes. It is precisely these opposing aims underlying the bargaining strategies of the two parties that eventually determine the union wage differential.

Thus the union sector is predicted to have a higher proportion employees from the middle of the observed skill distribution, and relatively lower proportions of employees from either tail (see Card, 1996, p.977). The descriptive statistics, provided at the bottom of the Wage Regression Tables in Section 7, verify the tendency for
union members to be drawn from the middle of the observed skill distribution thus indicating that single index estimation may well be inappropriate.

The highest percentage of covered members (union members in the baseline models) is consistently found at the middle/higher predicted observed skill quantiles in the male/female samples, respectively. Noting that the predicted wage cutoff for the highest female quantile always lies in the intermediate wage interval of the corresponding male samples, the highest proportion of union members is indeed drawn from the middle of the observed skill distribution regarding both genders.

If the appropriate controls are included in the index an incorrect selection model will only assign incorrect weights for covariates in the construction of the index in which case, a higher-order polynomial of the single index will in part capture the true random effects (Vella and Verbeek, 1998, p.176).

Following Vella and Verbeek (1998) we added interactions between \((B_1, B_{1t})\) and their higher-order values, with educational controls to the list of covariates in eq.(5/9). Highest education qualification was chosen since union wage effects generally differ by education level which can be thought of as a skill measure.\(^{16}\)

This is a first stage assessment of whether essential heterogeneity, in Heckman et al. (2006) terminology, is present. That is if individuals undertake their membership decisions with partial or full knowledge of the impact of their (unobserved to ourselves) characteristics then our endogeneity correction strategy will fail to identify the mean treatment effect.

If the union wage effect varies in the population even after conditioning upon the set of explanatory variables then the response distribution cannot generally be captured by a single number. Even if our interest lies in the mean of the distribution, an additional complication distinct from selection bias arises when there is sorting on the gain which is what Heckman et al. (2006) define as the essential heterogeneity model.

As the estimates in Sections (7.3, 7.4) indicate presence of essential heterogeneity, we perform separate estimations for three distinct skill groups. Following Card (1996), we construct a skill index using the predicted wage quantile in the non-union sector employing an independent sample.

To construct the skill index we only use individuals that were not used in the estimation panels and are not employed at covered establishments, that is recognising unions for bargaining purposes, so that our skill index is unaffected from potentially distortionary union effects on pay structure in the covered sector.

Since estimation of eq.(2) calls for consecutive observations, the initial single index estimates discard individuals having discontinuous labour market spells. Using those individuals with discontinuous labour market spells, we estimate flexible wage equations applying GLS to the pooled sample of uncovered employees. The predicted wage from these regressions represents an index of observed individual skill that is free from union distortionary effects upon wages.

According to the predicted wage quantile we form three distinct observed skill groups and re-estimate separate dynamic models of unionism and wage determination.

\(^{16}\)The estimations inclusive of educational control interactions with \((B_1, B_{1t})\) are indicated as essential heterogeneity estimates in the Wage Regressions Tables. The interaction terms (not shown) are jointly statistically significant in all estimated models except in Table 2 (see Section 7.3).
7. MEASUREMENT ERROR AND IDENTIFICATION

Measurement error issues are frequent in union status variables from longitudinal micro data sets, rendering least squares and FE estimates biased and inconsistent. Since true changes in union status are infrequent over short time periods, observed changes are particularly prone to reporting error—see Freeman (1984), Card (1996).

As identification of the union wage effect in longitudinal studies relies on the partial correlation of wage changes with respect to membership changes, misclassification bias is of primary concern. The FE estimator is particularly affected since the error of measurement is generally a serially uncorrelated noise while observed membership status is serially correlated across time (see Hirsch, 2004, p. 243; Griliches and Hausman, 1986).

Since membership status is dichotomous, we are dealing with a non-standard error of measurement. Assume initially that we estimate the union wage differential in its simplest form so that observed membership, \( U_{it} \), corresponds to either the unconditional membership variable or covered membership.

Following Aigner (1973) for tractability assume union membership is the only explanatory variable. Let true union status, \( U^*_t \), be related to the observed measure via

\[
U_{it} = U^*_t + \tau_{it}
\]

where \( \tau_{it} \) is a random measurement error so that the respective wage equation \( w_{it} = \beta_0 + \delta U^*_t + \epsilon_{it} \) corresponds to:

\[
w_{it} = \beta_0 + \delta U_{it} + (\epsilon_{it} - \delta \tau_{it})
\]

(12)

Denote the observed membership/nonmembership rates by \( \hat{\pi} = \left( \sum_{i=1}^{N} T_i \right)^{-1} \sum_{i=1}^{N} \sum_{t=1}^{T_i} U_{it}, \) \( (1 - \hat{\pi}) \). Let \( \lambda_2 \) denote the proportion of union members that have been misclassified as nonmembers and \( \lambda_1 \) denote the proportion misclassified as members, according to \( U_{it} \). The observed membership rate, \( \hat{\pi} \), is thus related to the true membership rate, \( \pi \), by

\[
\pi = (1 - \lambda_1) \hat{\pi} + \lambda_2 (1 - \hat{\pi}), \quad \lambda_1 = \Pr(U^*_t = 0 | U_{it} = 1), \lambda_2 = \Pr(U^*_t = 1 | U_{it} = 0)
\]

(13)

The marginal distributions of \((U^*_t, U_{it})\) are Bernoulli with parameters \((\pi, \hat{\pi})\), respectively while the joint distribution of \((U_{it}, \tau_{it})\) is easily obtained by manipulation:

<table>
<thead>
<tr>
<th>( \tau_{it} )</th>
<th>0</th>
<th>1</th>
<th>( f(\tau_{it}) )</th>
</tr>
</thead>
<tbody>
<tr>
<td>-1</td>
<td>( \lambda_2 (1 - \hat{\pi}) )</td>
<td>0</td>
<td>( \lambda_2 (1 - \hat{\pi}) )</td>
</tr>
<tr>
<td>0</td>
<td>( (1 - \lambda_2) (1 - \hat{\pi}) )</td>
<td>( (1 - \lambda_1) \hat{\pi} )</td>
<td>( (1 - \lambda_2) (1 - \hat{\pi}) + (1 - \lambda_1) \hat{\pi} )</td>
</tr>
<tr>
<td>1</td>
<td>0</td>
<td>( \lambda_1 \hat{\pi} )</td>
<td>( \lambda_1 \hat{\pi} )</td>
</tr>
<tr>
<td>( f(U_{it}) )</td>
<td>( (1 - \hat{\pi}) )</td>
<td>( \hat{\pi} )</td>
<td>1</td>
</tr>
</tbody>
</table>

There is a negative correlation between true union status and the measurement error since when \( U^*_t = 1 \), \( \tau_{it} \) is either -1 or 0 while when \( U^*_t = 0 \), \( \tau_{it} \) is either 0 or 1. Using the joint frequency table:

\[
E(\tau_{it}) = \lambda_1 \hat{\pi} - \lambda_2 (1 - \hat{\pi})
\]

(14)

\[
Var(\tau_{it}) = [\lambda_1 \hat{\pi} + \lambda_2 (1 - \hat{\pi})] - [\lambda_1 \hat{\pi} - \lambda_2 (1 - \hat{\pi})]^2
\]

(15)

\[
E(U_{it} \tau_{it}) = \lambda_1 \hat{\pi}, \quad Cov(U_{it}, \tau_{it}) = (\lambda_1 + \lambda_2) \hat{\pi} (1 - \hat{\pi})
\]

(16)
Contrary to classical errors in variables assumptions, the measurement error does not have a zero mean, it is negatively correlated with true union status and it is also correlated with observed union status.

Note that measurement error in union status also produces inconsistent estimates of the reduced form parameters given that \( U_{it} \) is the dependent variable (see Hausman et al., 1998). However, since the endogeneity correction terms \((B_1, B_{it})\) are non-linear functions of the reduced form variables, the inconsistency in \( \psi = (\gamma_1, \gamma_2, \sigma_\theta) \) is diluted via the non-linear mapping.

Assuming that the variances, covariances and misclassification rates are all constant across \( t \), noting that GLS and OLS use the same orthogonality assumptions between covariates and the composite error term, in micropanels as \( N \to \infty \), the \( \text{plim} \) of the least squares estimator reduces to

\[
\text{plim} \, \hat{\delta}_{LS} = \delta \left[ 1 - \frac{\text{Cov}(U_{it}, \tau_{it})}{\text{Var}(U_{it})} \right] + \left[ \frac{\text{Cov}(U_{it}, e_{it})}{\text{Var}(U_{it})} \right]. \tag{17}
\]

The second source of asymptotic bias in eq.(17) stems from union status endogeneity and given \( \frac{\text{Cov}(U_{it}, e_{it})}{\text{Var}(U_{it})} = \sigma_{j,a\theta} + \sigma_{j,\varepsilon\eta} \) we obtain

\[
\text{plim} \, \hat{\delta}_{LS} = \delta \left( 1 - \lambda_1 - \lambda_2 \right) + (\sigma_{j,a\theta} + \sigma_{j,\varepsilon\eta}), \quad j = 0, 1. \tag{18}
\]

Under the null of exogeneity assuming, \((0 < \lambda_m < 1, m = 1, 2)\), \( \hat{\delta}_{LS} \) will be attenuated. In theory, knowledge of the misclassification rates and assuming that these are time-invariant, one could correct for the least squares bias in the simple regression model.

Hierarchical sorting restricts \((\sigma_{j,a\theta}, \sigma_{j,\varepsilon\eta})\) to be sector-invariant and therefore eq.(18) reduces to

\[
\text{plim} \, \hat{\delta}_{LS} = \delta \left( 1 - \lambda_1 - \lambda_2 \right) + (\sigma_{a\theta} + \sigma_{\varepsilon\eta}). \tag{19}
\]

Hence, the conventional downward bias result will hold if \((\sigma_{a\theta} + \sigma_{\varepsilon\eta}) < 0 \) i.e. if negative selection is present or dominant. In the presence or prevalence of positive selection, \((\sigma_{a\theta} + \sigma_{\varepsilon\eta}) > 0 \), the conventional downward bias will hold if the attenuation bias from measurement error is greater than the selection bias.

Under unrestricted sorting \((\sigma_{0,a\theta} + \sigma_{0,\varepsilon\eta}) \neq (\sigma_{1,a\theta} + \sigma_{1,\varepsilon\eta}) \). Thus, measurement error reduction will raise the uncorrected \( \hat{\delta}_{LS} \) under unrestricted sorting since assuming \( \lambda_1 \approx 0 \) and \( \lambda_2 \approx 0 \) then

\[
\text{plim} \, \hat{\delta}_{LS} = \delta + (\sigma_{0,a\theta} + \sigma_{0,\varepsilon\eta}) + (\sigma_{1,a\theta} + \sigma_{1,\varepsilon\eta}) \tag{20}
\]

and comparative advantage requires \((\sigma_{1,a\theta} > 0, \sigma_{0,a\theta} < 0) \) and \((\sigma_{1,\varepsilon\eta} > 0, \sigma_{0,\varepsilon\eta} < 0) \) or alternatively that either \((\sigma_{1,a\theta} > 0, \sigma_{0,a\theta} < 0) \) while \((\sigma_{1,\varepsilon\eta} \approx 0, \sigma_{0,\varepsilon\eta} \approx 0) \) or \((\sigma_{1,\varepsilon\eta} > 0, \sigma_{0,\varepsilon\eta} < 0) \) while \((\sigma_{1,a\theta} \approx 0, \sigma_{0,a\theta} \approx 0) \). In other words, the positive contribution of unobserved heterogeneity biases upwards the membership differential.

In the presence of a coverage differential, assuming absence of additional covariates, if true covered non-membership status \( F_{it}^* \) is related to the observed measure via \( F_{it} = F_{it}^* + \xi_{it} \) where \( \xi_{it} \) is a random measurement error, the corresponding model is

\[
\begin{align*}
w_{it} & = \beta_0 + \delta_1 U_{it} + \delta_2 F_{it} + (e_{it} - \delta_1 \tau_{it} - \delta_2 \xi_{it}) \tag{21} \\
U_{it} & = m_{it} c_{it}, \quad F_{it} = (1 - m_{it}) c_{it}, \quad U_{it} = U_{it}^* + \tau_{it}, \quad F_{it} = F_{it}^* + \xi_{it} \\
e_{it} & = m_{it} c_{it}(\alpha_{1,t} + \varepsilon_{1,it}) + (1 - m_{it} c_{it})(\alpha_{0,t} + \varepsilon_{0,it})
\end{align*}
\]
As \( N \to \infty \), the \( \text{plim} \) of \( \hat{\delta}_{1,LS} \), assuming that the variances, covariances and misclassification rates are all constant across \( t \) corresponds to

\[
\text{plim}_{N \to \infty} \hat{\delta}_{1,LS} = \delta_1 (1 - \lambda_1 - \lambda_2) + \delta_2 \left[ \frac{\text{Cov}(U_{it}, F_{it})}{\text{Var}(U_{it})} - \frac{\text{Cov}(U_{it}, \xi_{it})}{\text{Var}(U_{it})} \right] + (\sigma_{j,\alpha \theta} + \sigma_{j,\varepsilon \eta}).
\]

(22)

The marginal distributions of \( F_{it}^a, F_{it} \) are Bernoulli with parameters \((\pi, \tilde{\pi})\) and the joint distribution of \( (F_{it}, \xi_{it}) \) can be easily obtained by interchanging \((1 - \tilde{\pi})\) with \( \tilde{\pi} \) in the joint frequency table of \((U_{it}, \tau_{it})\) since when \( c_{it} = 1 \), \( i.e. \) the case of interest, then \((1-U_{it}) = F_{it} \).

Thus, \( \text{Cov}(U_{it}, F_{it}) = -\text{var}(U_{it}) \) and \( \text{Cov}(U_{it}, \xi_{it}) = -\text{Cov}(F_{it}, \xi_{it}) \) and plugging these into eq.(22) we obtain

\[
\text{plim}_{N \to \infty} \hat{\delta}_{1,LS} = \delta_1 (1 - \lambda_1 - \lambda_2) + \delta_2 \left[ \frac{\text{Cov}(F_{it}, \xi_{it})}{\text{Var}(U_{it})} - 1 \right] + (\sigma_{j,\alpha \theta} + \sigma_{j,\varepsilon \eta})
\]

(23)

and since \( \text{Cov}(F_{it}, \xi_{it}) = (\lambda_1 + \lambda_2) \tilde{\pi}(1 - \tilde{\pi}) \) then the \( \text{plim} \) of \( \hat{\delta}_{1,LS} \) becomes

\[
\text{plim}_{N \to \infty} \hat{\delta}_{1,LS} = (\delta_1 - \delta_2) (1 - \lambda_1 - \lambda_2) + (\sigma_{j,\alpha \theta} + \sigma_{j,\varepsilon \eta}) \quad j = 0, 1.
\]

(24)

Similarly, assuming that variances, covariances and misclassification rates are all constant across \( t \) as \( N \to \infty \) the \( \text{plim} \) of \( \hat{\delta}_{2,LS} \) is

\[
\text{plim}_{N \to \infty} \hat{\delta}_{2,LS} = \delta_2 (1 - \lambda_1 - \lambda_2) + \delta_1 \left[ \frac{\text{Cov}(F_{it}, U_{it})}{\text{Var}(F_{it})} - \frac{\text{Cov}(F_{it}, \tau_{it})}{\text{Var}(F_{it})} \right] + \left[ \frac{\text{Cov}(F_{it}, e_{it})}{\text{Var}(F_{it})} \right]
\]

(25)

and substituting \( \text{Cov}(F_{it}, U_{it}) = -\text{var}(F_{it}), \text{Cov}(F_{it}, \tau_{it}) = -\text{Cov}(U_{it}, \tau_{it}) \) and \( \frac{\text{Cov}(F_{it}, e_{it})}{\text{Var}(F_{it})} = \frac{\text{Cov}(U_{it}, e_{it})}{\text{Var}(U_{it})} \) in eq.(25) we obtain

\[
\text{plim}_{N \to \infty} \hat{\delta}_{2,LS} = (\delta_2 - \delta_1) (1 - \lambda_1 - \lambda_2) - (\sigma_{j,\alpha \theta} + \sigma_{j,\varepsilon \eta}) \quad j = 0, 1.
\]

(26)

Using the preceding analysis it is straightforward to obtain the respective versions of eq.(24, 26) under restricted and unrestricted sorting.

Equation (24) indicates that in the presence of a coverage differential and measurement error, then the membership differential corresponds to \( (\delta_1 - \delta_2) \) and will be attenuated under exogeneity.

If the null of exogeneity is rejected then the uncorrected \( \hat{\delta}_{1,LS} \) will be biased downwards under hierarchical sorting and negative selection so that the measurement error and selectivity biases reinforce each other. Hence, measurement error reduction will only raise the uncorrected \( \hat{\delta}_{1,LS} \) under either comparative advantage or hierarchical sorting and positive selection.

Considering eq.(26) measurement error reduction will increase the uncorrected \( \hat{\delta}_{2,LS} \) under hierarchical sorting and negative selection of union members so that unobserved heterogeneity impacts positively upon the wages of covered nonmembers.

Given the absence of suitable instrumental variables, correlated with union status and uncorrelated with measurement error, we employ alternative error reduction strategies.
7.1. Measurement Error in Union Status Response

The best membership measure in the BHPS could potentially be obtained via the question "Are you currently a member of: Trade Unions" in the Social and Interest Group Membership section. The respective variable is "Member of trade union".

Unfortunately, this was only asked every other year after the fifth wave of the survey (1995-96) as data depositors adopt the conventional notion of little movement in and out of organisations and infrequent changes in membership status.

The BHPS contains an alternative membership question named "Member of workplace union". This is derived via the question asked conditionally following a positive response to the question regarding union recognition phrased as: "Is there a trade union, or a similar body such as a staff association, recognised by your management for negotiating pay or conditions for the people doing your sort of job in your workplace?".

The resulting variable is termed as "Union or staff association at workplace" and is our measure of union coverage. Following this response, a positive reply to "Are you a member of this trade union/association?" is recorded as membership of workplace union noting that it includes "in-house" staff associations and excludes employers’ organisations.

The union recognition and corresponding conditional membership questions were not asked during 1992-94 of employees still in same job as in the previous year. Nevertheless, the conditional union status variable referred to as covered membership, is our preferred membership status measure as it is reported annually without gaps from 1995 onwards.

Since we wish to distinguish among membership and coverage differentials using longitudinal methods requiring continuous observations, we employ the time framework spanning 1995 up to the end of the BHPS in 2009 given our choice of union status variable. The broad nature of the definition, including staff associations, is a source of measurement error.

A first stage correction of the covered membership status variable can be achieved using the two alternative membership variables to detect inconsistencies and reconstruct the covered membership status variable accordingly (see Swaffield, 2001).

If an individual responded yes to whether he/she is a trade union member, in the unconditional narrower membership question in the Social and Interest Group section, he/she should also respond positively to the conditional question of whether he/she is a member of a trade union or similar recognised by management for wage bargaining purposes.

A negative reply in the unconditional membership question and a positive reply in the broader membership case means that the individual is actually a member of a staff association and not a trade union and is therefore classified as a measurement error. Since the narrower membership status variable is only available every other year ex post 1995, we can correct the covered membership variable only for the corresponding years.\footnote{Provided measurement error attenuates the union wage effect, failure to correct the covered membership variable across all years under analysis (by excluding staff association members) might slightly depress the estimated differentials. On the other hand, as discussed in footnote 5, using hourly instead of weekly earnings may slightly increase the union wage differential. To the extent that the two aforementioned effects cancel each other out, the estimated union wage differentials should remain unaffected.}
7.2. Restricting Union Status Changes to those with either Job and/or Employer Changes or Changes in Union Recognition

A second stage correction of the covered membership variable is performed by restricting changes to those more likely to have experienced true changes i.e. those with either accompanying job and/or employer changes provided no corresponding change in union recognition on the part of the employer is observed (see Swaffield, 2001).

Using the length (in days) of the current labour market spell of the individual, noting that not all job changes are employer changes, we can identify those individuals changing union status while remaining at the same job- a potential source of measurement error.

Current labour market spell length on its own sake, however, is insufficient as for those changing jobs while at the same establishment the corresponding labour market spell length is reset to zero. If we wish to restrict union status changes to those with a higher probability of experiencing true changes, we ought to identify those that change status while changing employers and not solely jobs. The labour force status code variable, from the job history files of the BHPS, permits identifying individuals doing a different job at the same employer in the previous period and hence, account for the additional source of potential measurement error by restricting union status changes to those that change employers.

Unlike Swaffield (2001) or Koevets (2007), our time period under analysis is considerably greater meaning that we should observe several true changes in union recognition on the part of employers.

Unless we account for recognition changes, our restrictions regarding membership status changes could be discarding some true changes. Therefore we re-estimate models excluding observations where union status changes were not accompanied by either job and/or employer changes iff a corresponding change in union recognition did not occur- i.e. membership status changes with accompanying recognition changes are classified as true changes.\textsuperscript{18}

7.2.1. Identification and Simultaneity

The assumptions regarding the errors identify all parameters in eq.\( (5/9) \) given the non-linear mapping from the reduced form variables to \((B_i, B_{it})\). The imposition of exclusion restrictions is, however, desirable. The exclusion of \(U_{it-1} \) from the empirical counterpart of eq.\( (5/9) \) identifies the equation as long as \( \gamma_1 \neq 0 \) in eq\.(2).

Reduced form models additionally include political closeness controls hypothesised to affect unionisation decisions while not impacting on wages. However, as these were not always statistically significant, achieving instrumentation solely on the basis of political controls was not possible.

It can be argued that, while the lagged value of membership status affects the unionisation decision it does not have a significant effect on the current wage. This occurs in that membership status may capture movement costs not specific to union employment since employees are only likely to change status if they change jobs (Vella and Verbeek, 1998, p.167).

\textsuperscript{18}Restricting union changes to those with accompanying job/employer changes while ignoring recognition changes overstates the degree of measurement error hence inflating the membership premia. This pattern generally holds across all single index male/female estimates.
Therefore, restricting samples so that union status changes are accompanied by job and/or employer changes is not only a measurement error reduction device but is in fact imposed by our model assumptions.

Consider an employee changing establishments. Since the dominant form of wage bargaining in the U.K. is establishment level negotiation, and given that collective agreement length is one year, past membership status will only affect individual unionisation propensity and will not be an integral part of the future employer’s wage setting decision. In fact, membership will be irrelevant if the new employer does not recognise unions. If, on the other hand, the new employer does recognise unions a member’s wage rate will correspond to the current establishment union bargained wage.

Considering an individual changing union status while at the same establishment, if the observed change is not due to a recognition change, the assumption of past membership not impacting on current wage determination becomes less plausible. In this event, past membership status might be an integral part of an employer’s current period wage rate decision particularly when engaging in discriminatory behaviour.

An underlying assumption of our exclusion restriction is that long-term advantages of union employment, whilst generating membership persistence, do not have a significant impact on wages. Length of tenure is the inverse function of quits that are expected to be lower in unionised establishments due to benefits, higher wages and union voice (see Hirsch 2004, pp.240-241). To the extent that tenure captures accumulated skills an indicator of whether union sector long-term benefits are appropriately accounted for, is the response of the union wage differential upon inclusion of tenure. Including tenure and its square, reduces the union differentials in all estimates thus partiaIIing out potential $U_{i,t-1}$ effects.

### 7.3. Male Estimates

We initially estimate the union wage impact using the membership variable obtained by combining the unconditional "Member of trade union" and the conditional upon coverage membership questions for the years the former was unavailable.\(^\text{19}\)

\begin{table}[h]
\centering
\begin{tabular}{ccccccccc}
\hline
& & & & & & & & \\
\text{Predicted Quantile} & <=29 & >29<-63 & >63 \\
\hline
GLS & GLS-H & GLS-U & GLS-EH & GLS & GLS-H & GLS-U & GLS-U \\
\hline
\hline
$U_{it}$ & 0.031 & 2.99 & 0.026 & 2.02 & 0.092 & 4.47 & 0.081 & 4.05 & 0.071 & 4.59 \\
$B_{i}$ & -0.007 & -0.20 & -0.039 & -0.49 & 0.000 & -0.02 & -0.038 & -3.80 & -0.036 & -3.33 & -0.018 & -2.45 \\
$B_{i}, U_{it}$ & 0.066 & 0.49 & 0.003 & 0.14 \\
\hline
\hline
Observations & 9,599 & 9,599 & 6,418 & 6,418 & 6,418 & 6,418 & 6,418 & 6,418 & 6,418 & 6,418 & 6,418 & 6,418 & 6,418 \\
Members & 2,828 & 2,828 & 2,048 & 2,048 & 2,048 & 2,048 & 2,048 & 2,048 & 2,048 & 2,048 & 2,048 & 2,048 & 2,048 \\
Log-L & -1695.4 & -1685.4 & -1685.4 & -1685.4 & -236.4 & -433.5 & -665.0 \\
\hline
\end{tabular}
\caption{WAGE REGRESSIONS (UNION MEMBERSHIP), MALE (1995-2009)}
\end{table}

\(^\text{19}\)The sorting structure consistent with the single index estimates appears in bold letters as are the corresponding essential heterogeneity estimates (obtained by interacting the endogeneity correction terms, included in the appropriate sorting structure, with educational controls).
The single index estimates (in Table 1) are in line with hierarchical sorting and negative selection biasing downwards the uncorrected GLS estimate. Thus, employees receiving lower wages, conditioning on their attributes and in the absence of unions, are those more likely to be union members.

The statistically significant $B_{it}$ coefficients render FE estimation inappropriate as the time-varying endogeneity is not eliminated. The difference among the essential and hierarchical estimates suggests that single index estimation fails to identify the mean effect. The appropriate multiple indices estimates indicate an approximately 4.1 percent membership differential at the middle of the observed skill distribution while at the upper end, members experience a 2.5 percent wage reduction implying reduced human capital premia following wage standardisation.

Employing the covered membership variable (in Table 2) provides no evidence of coverage differentials. The single index estimates provide only weak indication of negative selection via $B_{it}$. Referring to eq. (24), it is clear that measurement error reduction raises the uncorrected GLS single index membership differential.

The multiple indices estimates indicate an approximately 6.2 percent total union differential at the lower end of the observed skill distribution, only. This coincides with the union membership differential due to the statistical insignificance of the coverage premium.\footnote{In all male estimates, except Table 2, the endogeneity correction terms at the bottom of the observed skill distribution indicate positive selection though they enter all specifications with statistically insignificant effects.}

Using corrected covered membership, the single index and intermediate observed skill group estimates (in Table 3) are consistent with hierarchical sorting, members are negatively selected and endogeneity is time-variant rendering FE inappropriate. Given the difference among the hierarchical and essential heterogeneity membership differentials, there is evidence of sorting on the gain.

The appropriate multiple indices estimates yield an approximately 4.8 and a 3.4 percent total union differential for the low and middle observed skill groups, respectively. In the former case, the statistically significant coverage differential reduces the membership premium to a mere 1.7 percent whereas, in the latter case the total union and membership differentials coincide. Members at the top of the

\begin{table}[h]
\centering
\begin{tabular}{lcccccc}
\hline
& & & & & & \\
 & GLS & FE & GLS-H & GLS-U & GLS-EH & \\
\hline
$\mu_{it}$ & Coef. & 0.054 & 4.10 & 0.050 & 2.87 & 0.104 & 4.63 & 0.082 & 3.65 & 0.081 & 4.13 \\
& $t$ & 0.003 & 0.21 & 0.007 & 0.46 & -0.001 & -0.05 & 0.003 & 0.27 & \\
$(1-m_{it})c_{it}$ & Coef. & -0.011 & -1.41 & -0.020 & -1.58 & \\
& $t$ & 0.052 & 0.87 & \\
$B_{it}$ & Coef. & 0.003 & 0.21 & 0.007 & 0.46 & -0.001 & -0.05 & \\
& $t$ & -0.011 & -1.41 & -0.001 & -0.05 & \\
$B_{it}(m_{it}c_{it})$ & Coef. & 0.003 & 0.21 & 0.007 & 0.46 & -0.001 & -0.05 & \\
& $t$ & -0.011 & -1.41 & -0.001 & -0.05 & \\
$B_{it}(m_{it}c_{it})$ & Coef. & 0.003 & 0.21 & 0.007 & 0.46 & -0.001 & -0.05 & \\
& $t$ & -0.011 & -1.41 & -0.001 & -0.05 & \\
\hline
Observations & 9,599 & 9,599 & 6,418 & 6,418 & 6,418 & 2,529 & 1,699 & 2,481 \\
Covered Members & 2,917 & 2,917 & 2,119 & 2,119 & 2,119 & 508 & 604 & 776 \\
Covered Members (%) & 30.39 & 30.39 & 33.02 & 33.02 & 33.02 & 20.09 & 35.55 & 33.28 \\
Covered Nonmembers & 1,412 & 1,412 & 890 & 890 & 890 & 388 & 243 & 351 \\
Log-L & -1227.6 & -1227.6 & -1227.6 & \\
\hline
\end{tabular}
\caption{WAGE REGRESSIONS (COVERED MEMBERSHIP), MALE (1995-2009)}
\end{table}

\begin{equation}
(\text{24})
\end{equation}
distribution experience an approximately 3.5 percentage wage reduction following wage standardisation.

Since single index estimates in Table 3 indicate hierarchical sorting and negative selection, noting the increase in the insignificant coverage differential, measurement error reduction lowers the uncorrected GLS membership differential—see eq. (19, 24).

<table>
<thead>
<tr>
<th>Predicted Quantile</th>
<th>&lt;=29</th>
<th>&gt;29-&lt;=63</th>
<th>&gt;63</th>
</tr>
</thead>
<tbody>
<tr>
<td>GLS Coef.</td>
<td>2</td>
<td>Coef.</td>
<td>2</td>
</tr>
<tr>
<td>GLS-H Coef.</td>
<td>2</td>
<td>Coef.</td>
<td>2</td>
</tr>
<tr>
<td>GLS-U Coef.</td>
<td>2</td>
<td>Coef.</td>
<td>2</td>
</tr>
<tr>
<td>GLS-EH Coef.</td>
<td>2</td>
<td>Coef.</td>
<td>2</td>
</tr>
</tbody>
</table>

Concerning the multiple indices results, using covered membership raises the membership differential relative to the statistically insignificant baseline estimate (in Table 1) to approximately 6.2 percent at the bottom of the observed skill distribution (see Table 2). Additional measurement error reduction, increases the now statistically significant coverage premium to approximately 3.1 percent (in Table 3) thereby reducing the effective membership differential at the lower end of the distribution to approximately 1.7 percent—see eq. (24).

Upon restricting union status changes to those with accompanying job or recognition changes, the estimates (in Table 4) indicate negative selection of members except at the lower end of the observed skill distribution.

<table>
<thead>
<tr>
<th>Predicted Quantile</th>
<th>&lt;=30</th>
<th>&gt;30-&lt;=65</th>
<th>&gt;65</th>
</tr>
</thead>
<tbody>
<tr>
<td>GLS Coef.</td>
<td>2</td>
<td>Coef.</td>
<td>2</td>
</tr>
<tr>
<td>GLS-H Coef.</td>
<td>2</td>
<td>Coef.</td>
<td>2</td>
</tr>
<tr>
<td>GLS-U Coef.</td>
<td>2</td>
<td>Coef.</td>
<td>2</td>
</tr>
<tr>
<td>GLS-EH Coef.</td>
<td>2</td>
<td>Coef.</td>
<td>2</td>
</tr>
</tbody>
</table>

The single index estimates provide evidence of essential heterogeneity and absence of a coverage differential noting that the membership premia in Table 4 are
higher than the respective premia in Table 3.\(^{21}\)

Considering the appropriate multiple indices results, while at the lower end of the observed skill distribution the total union differential is estimated to be approximately 5.3 percent, the presence of a 2.9 percent coverage premium reduces the membership premium to 2.4 percent. Given the absence of a coverage premium at the middle of the distribution, the corresponding membership premium is approximately 3.5 percent.

Since the single index estimates in Table 4 are consistent with comparative advantage sorting and as the coverage premium is local to zero, the uncorrected GLS membership differential is higher compared to the respective estimate in Table 3- see eq.\((20, 24)\).

Note the increase in the FE estimate of the membership effect in Table 4 which is markedly higher compared to the respective estimates in Tables (1-3). The same holds regarding the FE membership premia reported in Tables (5,6) noting that, the statistically significant impact of the \(B_{it}\) in the appropriate unrestricted sorting estimates of Table 4 augments the difference relative to the FE premium.

Effectively, significant measurement error reduction renders the FE estimate the intermediate bound of the male union membership differential. Therefore, the attenuation bias is indeed exaggerated by the FE estimator.

Regarding the multiple indices estimates (in Table 4) measurement error reduction raises the lower skill group membership premium since the coverage differential is reduced compared to Table 3 (see eq.\((24)\)), whereas, the middle group membership premium is slightly higher.

The negative differential for highly skilled members found in Tables (1,3) is reduced in magnitude and becomes statistically insignificant. This occurs in that the average wage contribution of unobserved heterogeneity is positive since the estimates at the upper end of the distribution are consistent with comparative advantage sorting- the same holds regarding Tables (5,6). Therefore, while union members at the upper end of the observed skill distribution are negatively selected, their sorting decision is positively rewarded.

### TABLE 5: WAGE REGRESSIONS (UNION CHANGE IF EMPLOYER CHANGE/ CHANGES IN RECOGNITION), MALE (1995-2009)

<table>
<thead>
<tr>
<th>Predicated Quantile</th>
<th>GLS FE</th>
<th>GLS-H</th>
<th>GLS-U</th>
<th>GLS-EH</th>
</tr>
</thead>
<tbody>
<tr>
<td>&lt;=30</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>&gt;30 &lt;=65</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>&gt;65</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Coef. z</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Coef. t</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Coef. z</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Coef. t</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Coef. z</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Coef. t</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Coef. z</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Coef. t</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

| Observations        | 8,701  | 8,701 | 5,516 | 5,516  | 5,516  | 5,516  | 2,446 | 2,628 | 2,077 |
| Covered Members (%) | 25.59  | 25.59 | 26.96 | 26.96  | 26.96  | 17.17  | 30.44 | 24.60 |
| Covered Nonmembers (%) | 1,388  | 1,388 | 807   | 807    | 807    | 399    | 427   | 312  |
| Covered Nonmembers (%) | 15.95  | 15.95 | 14.63 | 14.63  | 14.63  | 16.31  | 16.25 | 15.02 |

Log-L: -357.2 -357.2 -357.2 -108.8 -156.5 -159.8

<table>
<thead>
<tr>
<th>Predicated Quantile</th>
<th>GLS FE</th>
<th>GLS-H</th>
<th>GLS-U</th>
</tr>
</thead>
<tbody>
<tr>
<td>&lt;=30</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>&gt;30 &lt;=65</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>&gt;65</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

H: hierarchical, U: unrestricted, EH: essential heterogeneity test (endogeneity correction terms not shown)

GLS: Swamy Anora estimator with individual level cluster-robust std.errors; FE: cluster-robust std.errors; Log-L: Reduced Form Log-likelihood

GLS-H, GLS-U, GLS-EH: Bootstrap std. errors using 100 replications over both estimation stages adjusted for clustering on individuals

Predicted Quantile: stratification into quantiles according to a predicted wage in the uncowed sector; appropriate estimates shown only

\(^{21}\) Restricting union status changes lowers the variation of the dependent variable in reduced form models thus producing notably lower log-likelihoods.
Restricting union status changes to those with accompanying employer or recognition changes, raises the uncorrected single index GLS membership premium in Table 5 compared to Table 3 (see eq. (20, 24)) though it is slightly deflated relative to Table 4. Nevertheless, as the additional measurement error reduction strategies in Tables (4,5) could be excluding some true changes, comparisons should be made relative to the estimates in Table 3 using the best possible measure of membership status that can be obtained from the BHPS.  

Measurement error reduction raises the low skill membership premium (in Table 5) to approximately 2.1 percent since the coverage premium is reduced compared to Table 3—see eq. (24). As negative selection becomes insignificant, the intermediate membership premium is reduced to approximately 2.6 percent.

Socioeconomic group is a measure of ability and could contaminate our conclusions concerning the role of unobserved heterogeneity. To check the response of the estimated differentials, we exclude socioeconomic controls from the models restricting union changes to accompanying employer or recognition changes. As expected, inclusion of socioeconomic controls inflates the membership differentials obtained under both single and multiple indices rules (refer to Tables 5,6).

<table>
<thead>
<tr>
<th>TABLE 6: WAGE REGRESSIONS (UNION CHANGE IF EMPLOYER CHANGE/CHANGES IN RECOGNITION, NO SOCIAL GROUP), MALE (1995-2009)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Predicted Quantile</td>
</tr>
<tr>
<td>GLS FE</td>
</tr>
<tr>
<td>Coef.</td>
</tr>
<tr>
<td>0.041</td>
</tr>
<tr>
<td>0.005</td>
</tr>
<tr>
<td>-0.007</td>
</tr>
<tr>
<td>0.000</td>
</tr>
<tr>
<td>0.006</td>
</tr>
<tr>
<td>Observations</td>
</tr>
<tr>
<td>Covered Members</td>
</tr>
<tr>
<td>Covered Members (%)</td>
</tr>
<tr>
<td>Covered Nonmembers</td>
</tr>
<tr>
<td>Covered Nonmembers (%)</td>
</tr>
<tr>
<td>Covered Nonmembers (%)</td>
</tr>
<tr>
<td>Log-L</td>
</tr>
<tr>
<td>H: hierarchical, U: unrestricted, EH: essential heterogeneity test (endogeneity correction terms not shown)</td>
</tr>
<tr>
<td>GLS: Swamy Arora estimator with individual level cluster-robust std.errors; FE: cluster-robust std.errors; Log-L: Reduced Form Log-likelihood</td>
</tr>
<tr>
<td>GLS-H, GLS-U, GLS-EH: Bootstrap std. errors using 100 replications over both estimation stages adjusted for clustering on individuals</td>
</tr>
</tbody>
</table>

Given the presence of essential heterogeneity, we turn to the multiple indices estimates. Crucially, the membership differential at the middle of the observed skill distribution disappears as it is local to zero and becomes statistically insignificant (see Table 6).

Considering the low skill membership differential, it slightly falls from 2.1 percent (in Table 5) to approximately 2 percent in our preferred estimates (in Table 6) as the total union differential decreases more than the coverage premium. Compared to Table 3, measurement error reduction raises the low skill membership premium—see eq. (24).

Finally, as in all other male estimates using the corrected covered membership variable, there is evidence of free riding at the lower end of the observed skill distribution given the presence of an approximately 2.7 percent coverage premium.

---

22 The same holds for the respective female models that should be compared to Table 9 estimates.
7.4. Female Estimates

Starting with the estimates employing the constructed membership variable, there is negative selection while the endogeneity under the single index assumption stems from $B_{it}$ rendering FE inappropriate. The difference between the data consistent hierarchical sorting and essential heterogeneity differentials suggests that a multiple indices rule should be employed (see Table 7).

Accordingly, the multiple indices estimates indicate differential patterns of selection biases stemming from $B_i$. Members at the middle of the observed skill distribution are negatively selected, whereas, there is weak evidence of positive selection at the lower end. There is an approximately 5.2 percent membership differential at the middle of the observed skill distribution while at the upper end, there is a negative differential of approximately 2.7 percent indicative of reduced human capital premia.

Employing the covered membership variable provides no evidence of coverage differentials (see Table 8). Given the discrepancy among the unrestricted and essential heterogeneity estimates the appropriate estimates indicate negative membership...
differentials of approximately 7.5 and 7.6 percent at the lower and upper extreme of the observed skill distribution, respectively, and a weakly significant 4 percent membership premium at the middle. 

Crucially, the multiple indices estimates indicate positive selection at the bottom, whereas, at the middle and upper end of the observed skill distribution members are negatively selected. The differently signed selection effects approximately offset each other, rendering the FE estimate local to zero— an outcome generally holding across all female estimates. 

The single sorting estimates in Table 9 are consistent with hierarchical sorting and negative selection, noting the inappropriateness of FE as the endogeneity is time-variant. The appropriate multiple indices estimates indicate negative selection and an approximately 5.1 percent membership differential at the middle of the observed skill distribution which rises relative to Table 8. Lastly, the negative membership differential at the upper end of the distribution found in Tables (7,8) becomes statistically insignificant. 

| TABLE 9: WAGE REGRESSIONS (CORRECTED COVERED MEMBERSHIP), FEMALE (1995-2009) |
|-----------------|-----------------|-----------------|-----------------|-----------------|
|                | GLS FE          | GLS-H           | GLS-U           | GLS-EH          |                 |
| Predicated Quantile | <=34            | >34<=66         | >66             |                 |                 |
|                   | Coef. z         | Coef. t         | Coef. z         | Coef. t         |                 |
| m_{it}            | 0.061 3.95      | -0.016 -0.78    | 0.114 4.36      | 0.106 3.03      | 0.096 4.09      |
| E(m_{it}c)_{it}   | 0.036 2.54      | 0.009 0.53      | 0.024 1.57      | 0.021 1.37      | 0.025 1.66      |
| B_{it}            | -0.010 -0.27    | -0.013 -0.17    | -0.012 -2.64    |                 |                 |
| B_{it}(m_{it}c_{it}) | -0.049 -3.39 | -0.061 -3.29 | -0.069 -3.89 |                 |                 |
|                   | 0.002 0.01      |                 |                 |                 |                 |
|                   | 0.034 1.29      |                 |                 |                 |                 |
| GLS              | GLS-H           | GLS-U           | GLS-EH          |                 |                 |
| Observations      | 7,052           | 7,052           | 4,044           | 4,044           | 4,044           |
| Covered Members   | 2,142           | 2,142           | 1,426           | 1,426           | 1,426           |
| Covered Members (%) | 30.37           | 30.37           | 35.26           | 35.26           | 35.26           |
| Covered Nonmembers | 1,693           | 1,693           | 955             | 955             | 955             |
| Covered Nonmembers (%) | 24.01          | 24.01           | 23.62           | 23.62           | 23.62           |
| Observations      | 2,310 2,308     | 1,077 975       |                 |                 |                 |
| log L             | -1356.3         | -1356.3         | -1356.3         | -1356.3         | -1356.3         |

H: hierarchical, U: unrestricted, EH: essential heterogeneity test (endogeneity correction terms not shown)
GLS: Swamy-Aora estimator with individual level cluster-robust std. errors; FE: cluster-robust std. errors; Log L: Reduced Form log-likelihood
GLS-H, GLS-U, GLS-EH: Bootstrap std. errors using 100 replications over both estimation stages adjusted for clustering on individuals
Predicted Quantile: stratification into quantiles according to a predicted wage in the uncovered sector; appropriate estimates shown only

Using covered membership (in Table 8) raises the single index uncorrected GLS membership differential as the estimates are consistent with comparative advantage sorting. The respective Table 9 membership differential, however, falls since the estimates are in line with hierarchical sorting and negative selection and the coverage differential increases in magnitude and becomes statistically significant (see eqs. (19, 20, 24)).

Upon restricting union status changes to those with either job/employer or recognition changes, a distinguishing feature of opposing biases at the two extremes of the observed skill distribution arises (see Tables 10-12). While at the lower end union members are positively selected, at the upper end there is negative selection. As these opposing effects approximately offset each other, the single index estimates provide weaker evidence of essential heterogeneity while the FE estimates of the membership differential are driven towards zero.\textsuperscript{23}

\textsuperscript{23}Card (1996) proposes a two-sided selection model interacting employee and employer criteria. Given a negative association between the membership differential and location in the observed skill distribution, the employee’s (employer’s) selection criterion is more likely to be binding than the employer’s (employee’s) selection criterion given a higher (lower) level of observed skill. Union members with high (low) levels of observed skill are thus more likely to have negative (positive) values of $\theta_i$ yielding a negative (positive) selection bias (see Card, 1996, pp.976-78).

21
As the single index estimates (in Table 10) indicate comparative advantage sorting and the coverage differential found in Table 9 is reduced in magnitude and becomes insignificant, the uncorrected GLS estimate of the membership differential rises significantly (see eq. (20, 24)).

The offsetting selection biases render the unrestricted and essential heterogeneity single index estimates seemingly comparable. Nevertheless, the appropriate multiple indices estimates suggest that solely females pertaining to the middle of the observed skill distribution enjoy a membership premium of approximately 3.3 percent. This is reduced compared to Table 9 due to the absence of negative selection and the reduction in the magnitude of the insignificant negative coverage differential- see eq.(24).

### Table 10: Wage Regressions (Union Change if Job Change/Changes in Recognition), Female (1995-2009)

<table>
<thead>
<tr>
<th>Predicted Quantile</th>
<th>GLS-H</th>
<th>GLS-U</th>
<th>GLS-EH</th>
</tr>
</thead>
<tbody>
<tr>
<td>&lt;=39</td>
<td>0.095</td>
<td>0.530</td>
<td>0.005</td>
</tr>
<tr>
<td>&gt;=49</td>
<td>0.200</td>
<td>0.213</td>
<td>0.333</td>
</tr>
<tr>
<td>&gt;=69</td>
<td>0.333</td>
<td>0.296</td>
<td>0.154</td>
</tr>
<tr>
<td>B</td>
<td>-0.005</td>
<td>-2.61</td>
<td>-0.030</td>
</tr>
<tr>
<td>B[0,0.05]</td>
<td>-0.012</td>
<td>-0.17</td>
<td></td>
</tr>
</tbody>
</table>

### Table 11: Wage Regressions (Union Change if Employer Change/Changes in Recognition), Female (1995-2009)

<table>
<thead>
<tr>
<th>Predicted Quantile</th>
<th>GLS-H</th>
<th>GLS-U</th>
<th>GLS-EH</th>
</tr>
</thead>
<tbody>
<tr>
<td>&lt;=39</td>
<td>0.098</td>
<td>0.531</td>
<td>-0.002</td>
</tr>
<tr>
<td>&gt;=49</td>
<td>0.146</td>
<td>0.253</td>
<td>0.035</td>
</tr>
<tr>
<td>&gt;=69</td>
<td>0.200</td>
<td>0.309</td>
<td>0.004</td>
</tr>
</tbody>
</table>

Restricting union status changes to those accompanied by employer or recognition changes, raises the uncorrected GLS membership differential (in Table 11) compared to Table 9 since the coverage premium disappears and the estimates are in line with comparative advantage sorting- see eq.(20, 24).

### Table 11: Wage Regressions (Union Change if Employer Change/Changes in Recognition), Female (1995-2009)

<table>
<thead>
<tr>
<th>Predicted Quantile</th>
<th>GLS-H</th>
<th>GLS-U</th>
<th>GLS-EH</th>
</tr>
</thead>
<tbody>
<tr>
<td>&lt;=39</td>
<td>0.098</td>
<td>0.531</td>
<td>-0.002</td>
</tr>
<tr>
<td>&gt;=49</td>
<td>0.146</td>
<td>0.253</td>
<td>0.035</td>
</tr>
<tr>
<td>&gt;=69</td>
<td>0.200</td>
<td>0.309</td>
<td>0.004</td>
</tr>
</tbody>
</table>

The weak evidence of non-normality at the middle of the distribution could be attributed to the heterogeneous nature of the female samples that include part-time employees. The appropriate multiple indices estimates indicate a membership
differential of approximately 3.5 percent at the middle of the distribution which is notably lower than the 8 percent differential obtained upon inclusion of the higher-order correction terms.\textsuperscript{24}

As in the male estimates, excluding socioeconomic group controls, deflates the membership differentials under both single and multiple indices assumptions (refer to Tables 11,12). Since under the single index assumption Table 12 estimates indicate comparative advantage sorting, measurement error reduction raises the uncorrected membership premium compared to Table 9 (see eq.20, 24).

Lastly, the single index results offer weaker evidence of essential heterogeneity as the unrestricted and essential heterogeneity estimates are seemingly comparable. Provided this is the product of the offsetting biases occurring at the two extremes of the distribution one should refer to the multiple indices estimates.

At the middle of the distribution there is evidence of non-normality noting that the p-value for the joint significance of the endogeneity correction terms included is 0.054. The dominant form of endogeneity at the middle of the distribution stems from the time-invariant selection terms. Taking this into consideration, measurement error reduction combined with the increase of the negative (though insignificant) coverage premium raise membership differential at the middle of the distribution (in Table 12) to approximately 7.6 percent which is markedly higher compared to Table 9- see eq.24).

Ignoring the higher order correction terms the appropriate estimates indicate absence of selectivity and yield a more parsimonious estimate of an approximately 3 percent membership differential at the middle of the observed skill distribution (as the estimate not rounded to three significant figures is 0.0294388).

8. SUMMARY AND CONCLUSIONS

We study the union wage impact using BHPS data during 1995-2009. We account for the impact of measurement error in union status, consider changes in union recognition, identify both membership and coverage premia, and employ an

\textsuperscript{24}Since the p-value for the joint significance of \((B_i, B_{it})\) and their higher-order functions is 0.164, both sets of estimates are only included for comparability with Table 12.
endogeneity correction methodology that is flexible in the treatment of unobserved heterogeneity.

Furthermore, we test for the presence of essential heterogeneity and allow for distinct selection biases at three different skill levels defined by predicted wage quantiles obtained from an independent sample of employees in the uncovered sector.

Based on the contributions of Aigner (1973) and Vella and Verbeek (1998), we obtain the unified measurement error and endogeneity bias expression. It is demonstrated that the conventional attenuation bias arises only under hierarchical sorting and negative selection so that the measurement error and selectivity biases reinforce each other.

We obtain robust estimates of union membership wage differentials at the bottom and middle of the observed male and female skill distributions of approximately 2 and 7.6 percent, respectively. However, the more parsimonious female differential corresponds to 3 percent. Further, there is evidence of free riding at the bottom of the male observed skill distribution given the presence of a 2.7 coverage premium.

While union members are negatively selected, those at the lower end of the observed skill distribution are in general positively selected indicating the conflicting interests of employers and employees in determining unobserved differences.

Using the unified bias expression, we obtain a discernible pattern between uncorrected, endogeneity corrected and FE estimates of the union wage effect.

Extensive measurement error reduction raises significantly the FE estimate of the male union wage effect making it the intermediate bound, whereas, the endogeneity corrected estimate always corresponds to the upper bound. Unlike the male estimates, the strongOffsetting biases occurring at the two extremes of the female observed skill distribution render the FE estimate local to zero.

The obvious future research challenge is the development of a two-step quantile regression model for estimation of longitudinal dynamic models with binary endogenous covariates.

REFERENCES


9. APPENDIX: VARIABLES AND DEFINITIONS

VARIABLES AND DEFINITIONS

TABLE A1
UNION MEMBERSHIP DETERMINATION MODELS, DYNAMIC RANDOM EFFECTS (reduced form)

DEPENDENT VARIABLE

Trade Union Membership

EXPLANATORY VARIABLES

Initial Condition: Trade Union Membership Status in 1995
Lagged Trade Union Membership Status (previous year)

Age (age at date of interview)
Married or in Civil Partnership
In Full-Time Employment (included in female models only)
Job Tenure (Current Year - year started current job)
Public Sector Employee
Training Incidence

Regional Controls

London (Inner/Outer) and rest of South East
South West
Scotland
Wales
Northwest (Greater Manchester, Merseyside, rest of Northwest England)
Northeast (South/West Yorkshire, rest of Yorkshire and Humber, Tyne and Wear, rest of North England)
East Anglia
Midlands (base group): East Midlands, West Midlands conurbation, rest of West Midlands)

Highest Academic Qualification Controls

University Qualification: First Degree or Higher
Vocational Qualifications (HND, HNC, Teaching)
A Level Qualifications

Establishment Size Controls

More than 500 Employees at Current Job
100-499 Employees at Current Job
25-99 Employees at Current Job
Less than 25 Employees at Current Job (base group)
VARIABLES AND DEFINITIONS

TABLE A1 (continued)
UNION MEMBERSHIP DETERMINATION MODELS, DYNAMIC RANDOM EFFECTS
(reduced form)

Political Party Closest to

Conservative
Labour
Liberal Democrats (post 1988: merger of LD & SDP)
Scottish National Party, The Party of Wales, Green Party, Other Parties, none (base group)

Socioeconomic Group Controls

Professional, Manager or Employer
Foreman or Skilled Manual (incl. Armed forces)
Semi-skilled Manual
Unskilled
Intermediate/Junior non-manual, Personal Service (base group)

Industrial Classification Controls

Extraction of Minerals & Manufacture of Metals, Minerals, Chemicals
Metal Goods, Engineering & Vehicles Industries
Other Manufacturing Industries
Construction
Distribution (wholesale, retail), Hotels & Catering, Repairs
Transport & Communication
Banking, Finance, Insurance, Business Services & Leasing
Public Administration, Sanitary, Education, Research & Development
Energy and Water Supplies (base group)

Notes:

1. All models include time dummies
2. All models are inclusive of time averages of time-variant characteristics
3. Variables treated as time invariant due to minimal or no within variation:
   highest academic qualification, regional, socioeconomic group, industrial classification controls

Data Source

## VARIABLES AND DEFINITIONS

**TABLE A2**

### WAGE REGRESSIONS

*(structural form)*

<table>
<thead>
<tr>
<th><strong>DEPENDENT VARIABLE</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td>Gross Average Real Hourly Wage Rate (log of weekly wage divided by usual paid hours including overtime)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th><strong>EXPLANATORY VARIABLES</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td>Covered Union Member (Union Member in baseline models using the unconditional membership question)</td>
</tr>
<tr>
<td>Covered Non-Union Member</td>
</tr>
<tr>
<td>Age (age at date of interview)</td>
</tr>
<tr>
<td>Squared Age</td>
</tr>
<tr>
<td>Married or in Civil Partnership</td>
</tr>
<tr>
<td>In Full-Time Employment (included in female models only)</td>
</tr>
<tr>
<td>Job Tenure (Current Year - year started current job)</td>
</tr>
<tr>
<td>Squared Job Tenure</td>
</tr>
<tr>
<td>Public Sector Employee</td>
</tr>
<tr>
<td>Training Incidence</td>
</tr>
<tr>
<td>Regional, Highest Academic Qualification, Establishment Size</td>
</tr>
<tr>
<td>Socioeconomic Group, Industrial Classification Controls</td>
</tr>
</tbody>
</table>

*(same as in reduced form models)*

### Notes:

1. All models include time dummies
2. Hierarchical, Unrestricted Sorting, Essential Heterogeneity models include respective endogeneity correction terms

### Data Sources