Heterogeneity, Endogeneity, Measurement Error and Identification of the Union Wage Impact\textsuperscript{1,2}

Georgios Marios Chrysanthou  
Department of Economics  
Universidad de Alicante  
03080, Alicante, Spain  
E-mail: gmc@ua.es

Version: June, 2014

This paper studies the union wage impact when membership decisions are non-coercive and employers reserve the right of union recognition for bargaining purposes. Using the BHPS during 1995-2009 we find robust evidence of union membership wage premia for the lower male and middle female observed skill groups. The initial estimates indicate negative wage differentials for union members at the upper end of the observed skill distribution that disappear upon measurement error reduction. While union members are negatively selected, individuals from the lower tail of the observed skill distribution are positively selected in union employment. Using the unified analytical expression of the bias resulting from non-classical measurement error and union status endogeneity, the empirical findings indicate a discernible pattern between uncorrected, endogeneity corrected and longitudinal estimates of the union wage effect.

Key Words: unobserved heterogeneity, endogeneity, measurement error, union wage differentials  
Subject Classification: C33, C35, J31, J51

1. INTRODUCTION

Estimation of union wage differentials is an old though yet unresolved issue in labour economics. There is still absence of a consensus regarding whether union wage differentials represent the impact of trade unions, are a product of selectivity of unionised sector employees or indicate insufficient controls of both employer and individual characteristics. Further, in institutional environments where union membership is voluntary and employers reserve the right of union recognition for bargaining purposes it is unclear if positive membership premia, over and above any common coverage premia, can actually exist.

The joint determination of union membership and wage decisions means that in exploring how observationally equivalent employees’ wages differ we cannot ignore

\textsuperscript{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.

\textsuperscript{2}I am especially grateful to Pedro Albarran for his helpful comments and our discussions. I also wish to thank Jeffrey Wooldridge, Francis Vella, Nicholas Cox, Leslie Godfrey, Rasmus Jørgensen, Keith Hartley, Andrew Jones, John Hutton, Wim Koeveets, Angel Leon participants at the Applied Econometrics Workshop of Universidad Carlos III de Madrid, the 24th EALE Annual Conference and the Understanding Society 2013 Research Conference. All errors and omissions are my own.
how the unobserved individual heterogeneity underlying the union membership decision is rewarded. The ubiquitous phrases in literally any paper in the field stress that:

"Union members are different from nonmembers in unobserved ways, biasing your estimates. You should make a selectivity bias correction...simultaneously determine union status and economic outcomes...develop an unobservables model...Use longitudinal data" (Freeman, 1984, p.2).

Longitudinal analyses, however, do not provide a research panacea due to the substantial impact of measurement error in union membership status on longitudinal estimates of the union effect (see Freeman 1984, Card 1996, Swafield, 2001). Additionaly, the standardisation of wage rates implies reduced human capital premia meaning that the probability of joining an establishment recognising unions for wage bargaining purposes is inversely related to individual skill level. Sorting into union/non-union employment is likely to follow a multiple indices rule. Effectively then, successful instrumentation employing either the control function approach, or instrumental variables, becomes problematic (see for instance Farber, 1982; Card, 1996; Heckman et al., 2006). Lastly but not least, there is the issue of the lack of availability of convincing instrumental variable candidates.

Such estimation and identification difficulties, coupled with the complexity of the institutional framework governing union-firm wage bargaining interactions render the empirical evidence on union wage differentials generally inconclusive despite the presence of a large and refined literature.

In this study 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. These legislative changes raise the issue of whether union membership wage premia can actually exist since by law employees within establishments covered by bargaining agreements should receive the same wage independently of individual union membership status. Therefore, we should only observe a coverage differential and union membership should not have an additional impact on wage structure.

In the absence of coercion, unions face a free rider problem given the positive monetary cost of membership in which case unions may offer excludable goods or services to encourage membership (see Olson, 1965). Empirical evidence from both the U.K. and U.S., however, is generally inconclusive and many authors suggested the existence of membership premia over and above the corresponding coverage premia.

Positive union membership wage premia invalidate the free rider problem and raise a free rider puzzle issue instead as to why covered nonmembers abstain from membership and renounce the respective wage premium. Two types of rationalisations were put forward in the literature to explain the existence of positive membership premia in the absence of coercion: selectivity or omitted variables,

3 Longitudinal typically fixed effects (FE), within transformation, estimators are highly susceptible to measurement error since a small percentage of missclassified union membership status changes generally produces large biases given that the true occurrence of movements between unionised and nonunionised employment is low (Card, 1996, p.957).

4 The overall impact was that the 1980s was a decade of a dramatic decline in aggregate union membership and union recognition in the U.K. (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).
and discriminatory behaviour by either unions, employers or both (see Booth and Bryan, 2004).

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 methods to eliminate unobserved individual heterogeneity and account for measurement error (e.g. Swaffield, 2001), longitudinal studies that employ instrumental variables to address the issue of endogeneity and account for measurement error (Hildreth, 2000, Koevets, 2007) and finally, studies that employ a selection on observables approach relying on the conditional independence assumption and do not address the issue of endogeneity (e.g. Blanchflower and Bryson, 2010).

We estimate union wage differentials using longitudinal data from the British Household Panel Survey, henceforth the BHPS. We account for the impact of measurement error in union status while considering changes in union recognition, distinguish between membership and coverage premia, and use an endogeneity correction methodology that is flexible in its treatment of unobserved heterogeneity and models union membership determination. In addition, 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 for wage bargaining purposes sector.

Building upon the contributions of Aigner (1973) and Vella and Verbeek (1998), this study contributes to the literature by obtaining the unified analytical bias expression stemming from endogeneity of union status and non-classical measurement error. It is formally 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.

Consequently, to our knowledge the current study is one of few works in the literature that tackles all of these issues simultaneously building upon the influential contributions of Freeman (1984), Robinson (1989a), Card (1996) and Vella and Verbeek (1998).

We obtain robust estimates of union membership wage differentials for males at the bottom of the observed skill distribution and females in the middle observed skill group of approximately 4.7 and 7.6 percent, respectively. We find weaker evidence of a 2.6 percent membership differential for males at the middle of the observed skill distribution disappearing upon the exclusion of the potentially endogenous socioeconomic group controls. The initial estimates indicate negative differentials for union members in the high observed skill group implying reduced human capital premia due to wage standardisation. These negative differentials disappear by removing the attenuation bias caused by measurement error.

In addition, there is weak evidence of coverage differentials in the case of male nonmembers from the low end of the observed skill distribution indicating free-riding behaviour and reducing the union membership premium to approximately 1.9 percent, only. We also find weak evidence of negative coverage differentials for female nonmembers from the middle of the observed skill distribution effectively acting as a punishment for nonmembers and raising the respective membership differential to approximately 10.5 percent.

While union members are negatively selected, individuals pertaining to the lower tail of the observed skill distribution are positively selected indicating the conflicting interests of employers and employees determining the unobserved differences among
individuals located in unionised and non-unionised establishments. The estimates demonstrate that, mean regression methodologies ignoring distinct selection biases per skill group, as well as, generic fixed effects and instrumental variable estimators restricting unobserved heterogeneity to be individual specific and time invariant are inappropriate.

Finally, using the expression of the bias resulting from non-classical measurement error and union status endogeneity, the estimates indicate a discernible pattern between uncorrected, endogeneity corrected and longitudinal estimates of the union wage effect.

The paper is organised as follows. Section 2 briefly discusses the problem of endogeneity and simultaneity; Sections 3 and 4 outline the econometric model and estimation procedure; Section 5 discusses the issue wage equation parameterisation; Section 6 discusses data and sample selection issues; Section 7 treats the estimation of union effects by predicted wage quantile; Section 8 treats measurement error and identification issues and presents the estimates. Lastly, Section 9 summarises and concludes.

2. THE JOINT DETERMINATION OF UNION STATUS AND UNION WAGE EFFECTS

The general consensus in the earlier literature was that cross-sectional least squares analyses of the union wage effect are contaminated by the selectivity of union members inflating the union wage effect (see Abowd and Farber, 1982; Freeman, 1984). 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: endogeneity correction methods (such as the Heckman, 1979 two-step estimator or instrumental variables) produced an upward adjustment as opposed to OLS, whereas longitudinal (differencing) methods produced a downward adjustment (Robinson, 1989a, p.640). Though, some researchers employing longitudinal data sets attributed the resulting reduction in the union wage effect to fixed effects of higher quality workers present within establishments recognising unions for bargaining purposes this is not consistent with studies employing endogeneity correction methods to deal with such effects (Robinson, 1989a, p.658).

Estimation of union wage differentials using longitudinal data requires controlling for the endogeneity of union membership status. Fixed effects 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).

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

Generic instrumental variables methods constrain unobserved individual heterogeneity to be identical across sectors (i.e. across establishments recognising unions for wage bargaining and those that do not). This enforces the constraint that the ordering of employees’ productivity is sector invariant (termed as hierarchical sorting). The FE 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, p.644; Vella and Verbeek, 1998, pp.165-166).

Further, a comparative advantage structure is precluded, a priori, by either generic instrumental variables or restricted control function estimators. Therefore, variants of the control function approach, other than its restricted form, are preferred in that they convey valuable information about the sorting process of employees within the unionised and non-unionised sector (for a detailed treatment refer to Vella and Verbeek, 1999a).

Estimation of the union wage effects calls for a procedure that explicitly identifies the different sources of endogeneity of union membership status: we ought to employ a methodology that is flexible in its treatment of unobserved individual heterogeneity so that we are able to distinguish among the different economic sorting structures (hierarchical/comparative advantage) into the unionised and non-unionised sectors. As it is demonstrated in Section 8, unbiased estimation of the union wage impact requires both significant measurement error reduction and flexible unobserved heterogeneity treatment.

3. A DYNAMIC MODEL OF UNIONISM AND WAGE DETERMINATION

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 status by exploiting the longitudinal nature of the data.

Equation (1) outlines 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 \(j = \{0, 1\}\) 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.\(^6\)

The unobserved random components of an individual employee’s wage are given by \((\alpha_{j,i}, \varepsilon_{j,it})\) and the usual error component structure assumes \(\alpha_{j,i} \sim iidN(0, \sigma_{\alpha}^2)\)

\(^5\)By the term restricted control function estimators we refer to variants of the control function approach imposing the requirement that on average the best performers in unionised establishments, will also perform better in establishments not recognising unions for bargaining purposes (see Section 4).

\(^6\)Note \(t = 1, ..., T_i\) as opposed to \(t = 1, ..., T\) in eq.(1) and all of the remaining equations reflects the unbalanced nature of the panels employed in the estimations.
and \( \varepsilon_{j,it} \sim iidN(0, \sigma_z^2) \):

\[
\begin{align*}
    w_{j,it} &= \beta\prime_j x_{it} + \varepsilon_{j,it}, \\
    \varepsilon_{j,it} &= \alpha_{j,i} + \varepsilon_{j,it} \\
    t &= 1, \ldots, T; \ i = 1, \ldots, N; \ j = \{0, 1\}
\end{align*}
\] (1)

Gaining employment in an establishment that recognises unions for bargaining purposes is also contingent on the employer’s willingness to hire an individual (see Abowd and Farber, 1982). A limitation of the estimation methodology 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 employer characteristics are captured through industrial classification, establishment size and public sector controls these are not adequate in order to assign any specific effects purely to unobserved heterogeneity.

However, in the subsequent analysis we do study the impact of job and employer changes in association with changes in union membership and changes in union recognition on the part of employers. Furthermore, by employing a multiple indices sorting rule into unionised jobs, we account for distinct biases according to observed skill level and are thus able to examine in greater depth employer-employee interactions reflecting the conflicting interests of the two parties.

The dynamic reduced form model depicting individual union 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 dummy variable \( U_{it} \). The unknown parameters to be estimated are \( (\gamma_1, \gamma_2)\prime \) and the composite error term \( \nu_{it} \) captures the unobserved individual heterogeneity underlying the union membership decision and is decomposed into an individual-specific component \( \theta_i \) and an individual time-specific effect \( \eta_{it} \). 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.\(^7\)

\[
\begin{align*}
    U^*_it &= \gamma_1 x_{it} + \gamma_2 U_{i,t-1} + \nu_{it}, \ \nu_{it} = \theta_i + \eta_{it} \\
    U_{it} &= I(U^*_it > 0) \\
    w_{it} &= w_{j,it} \ \text{if} \ U_{it} = j
\end{align*}
\] (2) (3) (4)

Potential seniority and non-pecuniary benefits can prolong union membership in the long term, irrespective of wage changes, and this introduces state dependence in the model. 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.\(^9\)

\(^7\)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.

\(^8\)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 BHPS based studies (accounting for fixed effects and measurement error) such as Swafield (2001) and Koevets (2007) also employing the hourly earnings measure.

\(^9\)It is not possible to include dynamics coming through the lagged dependent variable in the wage equation. Arellano et al. (1997) propose an alternative estimator, constrained to models with Tobit types of censoring, that permits for lagged latent dependent variables to enter both the primary and reduced form equations linearly.
The random components \((\alpha_{j,i}, \theta_i), (\varepsilon_{j,it}, \eta_{it})\) in equations (1), (2) denote the individual-specific and the individual/time-specific effects 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. More specifically the four covariances \((\sigma_{j,i}, \sigma_{\theta, \theta}, \sigma_{j,\varepsilon, \eta})\) are allowed to be non-zero. 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.\(^{10}\)

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:

\[
\begin{align*}
    w_{it} &= \beta^t x_{it} + \delta U_{it} + \varepsilon_{it} \\
    e_{it} &= U_{it}(\alpha_{1,i} + \varepsilon_{1,it}) + (1 - U_{it})(\alpha_{0,i} + \varepsilon_{0,it})
\end{align*}
\]

The samples used in this study are generally heterogeneous. We need to maintain reasonably large sample sizes given the dynamic nature of the reduced forms (requiring both successive observations and a common start date to facilitate initial conditions estimation) and considering the subsequent sample losses when evaluating measurement error impact. We therefore do not restrict estimations to the traditional male manual employees as for instance in Swaffield (2001) or distinct public and private sector samples as Blanchflower and Bryson (2010). This might question our assumption on homogeneous returns on observables invoked in the estimating wage equation (5).\(^{11}\)

As discussed in Section 7, recognising that selection into union jobs may lead to differing selection biases at different skill levels we construct a skill index and estimate models for three distinct skill groups. This makes our assumption of homogeneous sector returns on observables more plausible and further, given the usage of an independent sample to construct the skill index we avoid the pitfalls and inconsistencies associated with distinctions based on endogenous sample selection mechanisms.

4. 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 \(x_{it}\):

\[
E(w_{it} | x_{it}, U_i) = \beta^t E(x_{it} | x_{it}, U_i) + \delta E(U_{it} | x_{it}, U_i) + E(\alpha_{j,i} | x_{it}, U_i) + E(\varepsilon_{j,it} | x_{it}, U_i)
\]

\(^{10}\)The covariances between the effects in the union/ non-union wage equations are not specified, whereas all remaining covariances are set to zero.

\(^{11}\)Higher-order terms of the latent effects were added to all estimated models to detect departures from normality (see Pagan and Vella, 1989). These were generally insignificant and are only shown when statistically significant in the appropriate estimates according to the sorting mechanism supported by the data.
Estimation of the reduced form, equation (2), provides the estimates of the unobserved individual heterogeneity. This is a dynamic random effects Probit model with a likelihood function:

$$\prod_{i=1}^{N} \prod_{t=1}^{\infty} \Phi \left( \frac{\gamma' \Psi_{it} + \theta_i}{\sigma_\eta} \right) U_{it} \left( \Phi \left( -\frac{\gamma' \Psi_{it} + \theta_i}{\sigma_\eta} \right) ^{1-U_{it}} \right) \left( 1/\sigma_\delta \right) \phi(\theta/\sigma_\delta) d\theta$$

where, $\gamma = (\gamma_1', \gamma_2')'$, $\Psi_{it} = [x_{it}, U_{i,t-1}]$, and $(\Phi, \phi)$ correspond to the cumulative probability and density functions of the standard Normal distribution.\(^{12,13}\)

The inclusion of the lagged union membership variable as a regressor in equation (2) gives rise to the problem of initial conditions (refer to Heckman, 1981a). The presence of individual-specific effects $\theta_i$ invalidates the exogeneity assumption of union status in the first period and the random effects ML estimator in its standard form will be inconsistent (see Heckman, 1981a,b).

We employ Wooldridge’s (2005) solution to the initial conditions problem due to its computational simplicity as opposed to Heckman’s (1981b) estimator. This involves modelling the distribution of the unobserved effect conditional on the initial value and the observed history of strictly exogenous explanatory variables instead of obtaining the joint distribution of all outcomes of the endogenous variables. 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 the reduced form can be estimated by a standard random effects Probit model with the only difference that the explanatory variables at time $t$ are $z_{it} \equiv (1, x_{it}, U_{i,t-1}, U_{i1}, \bar{x})$, (see Wooldridge, 2005).\(^{14,15,16}\)

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

$$E(\alpha_{j,i} | x_{it}, U_i) = \sigma_{j,\alpha \theta} \left[ \frac{T_i}{\sigma_\eta^2 + T_i/\sigma_\theta^2} E(\nu_i | x_{it}, U_i) \right] = \sigma_{j,\alpha \theta} B_i \tag{8}$$

\(^{12}\)Given $U_{it}$ is dichotomous, a normalisation is necessary. Since we are estimating using Stata, the estimated coefficients $\gamma = (\gamma_1', \gamma_2')$ are normalised on $\sigma_\eta$. That is, $\sigma_\eta^2 = 1$ as it is commonly assumed.

\(^{13}\)If the intra-panel correlation coefficient is local to zero a simple pooled probit can be estimated instead. This gives consistent parameter estimates provided the correct parameterisation is used along with the Mundlak (1978)-Chamberlain (1984) specification.

\(^{14}\)Orme (2001) proposes an alternative solution that is an approximation when the correlation between the initial condition and the random effect is small. Due to the highly persistent nature of union membership and the strong correlation between the initial condition and the random effect it produces inconsistent parameter estimates.

\(^{15}\)Arulampalam and Stewart (2009) demonstrate that none of the three estimators dominates the other and once the Mundlak (1978)-Chamberlain (1984) correlated random effects framework is employed the three estimators provide similar results.

\(^{16}\)Fernández-Val and Vella (2011) propose an alternative two-step methodology estimating the reduced form by fixed effects. This avoids the independence axiom and parametric assumptions though exclusion restrictions remain the same as in the random effects variant. Nonlinear fixed effects estimation introduces the incidental parameters problem (Neyman and Scott, 1948) and a bias correction is necessary. Carro (2007) proposes a modified ML estimator though effective bias reduction is achieved when $t \geq 8$. As fixed effects discard observations for which the dependent variable is invariant, introducing the additional restriction that individuals are continuously present for at least eight periods is prohibitive in terms of sample attrition. Using the same data set as Vella and Verbeek (1998) and employing the two-step fixed effects methodology, Fernández-Val and Vella (2011) obtain similar estimates of the union wage differential to those of Vella and Verbeek (1998).


\[ E(\varepsilon_{jt} \mid x_{it}, U_i) = \sigma_{j, \epsilon \eta} \left[ \frac{E(\nu_{it} \mid x_{it}, U_i)}{\sigma^2} - \frac{T_i \sigma^2_\theta}{\sigma^2_\eta (\sigma^2\eta + T_i \sigma^2_\theta)} E(\tilde{\nu}_{it} \mid x_{it}, U_i) \right] = \sigma_{j, \epsilon \eta} B_{it} \]  

(9)

where, \( \nu_{it} = \theta_i + \eta_{it} \) and \( \tilde{\nu} = T_i^{-1} \sum_{t=1}^{T_i} \nu_{it} \).

Using standard results (see Vella, 1998; Vella and Verbeek, 1998) it can be demonstrated that \( E(\nu_{it} \mid x_{it}, U_i) \) corresponds to:

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

(10)

Note that \( E(\eta_{it} \mid x_{it}, U_i, \theta_i) \) in the numerator of eq.(10) is the generalised Probit residual (Gourieroux et al., 1987) and the term in the denominator of eq.(10) is the likelihood contribution for individual \( i \). Given the parameter estimates from the reduced form model \( \psi = (\gamma_1, \gamma_2, \sigma_\theta) \) we can approximate eq.(10) using numerical integration (simulation).\(^{17}\)

Substituting the estimates for \( E(\nu_{it} \mid x_{it}, U_i) \), \( E(\tilde{\nu}_{it} \mid x_{it}, U_i) \), the endogeneity correction terms \( (B_i, B_{it}) \) defined in equations (8) and (9) can be computed and are then added as additional terms in the structural equation to be estimated jointly with \( (\beta', \delta) \) in the second step from conditional moment restrictions such as least squares based on equation (5).\(^{18}\)

Under the null hypothesis of exogeneity \( (\sigma_{j, \epsilon \theta} = \sigma_{j, \epsilon \eta} = 0) \) the conventional standard errors can be used. Otherwise, the standard errors should be adjusted for heteroskedasticity and the inclusion of the endogeneity correction terms (see Newey, 1984; Vella and Verbeek, 1999b).\(^{19}\)

\(^{17}\)The simulated counterpart of eq.(10) is estimated taking \( R \) draws of \( \theta_i^r \) and computing \( f(U_i \mid x_{it}, \theta_i^r) \) i.e. the corresponding log-likelihood for \( i \) conditional on the draw. To provide a better coverage of the integrals we use randomised Halton draws supplementing these with antithetic draws to induce a negative correlation over observations and further improve coverage (see Train, 2003; Cappellari and Jenkins, 2006). Repeating \( R \) times and averaging over replications we obtain the simulated log-likelihood function. Given \( \tilde{\psi} = (\tilde{\gamma}_1, \tilde{\gamma}_2, \tilde{\sigma}_\theta) \) we can simulate \( E(\nu_{it} \mid x_{it}, U_i) \). Taking again \( R \) draws eq.(10) is approximated by

\[ \tilde{\nu}_{it} = \left[ \frac{1}{R} \sum_{r} f(U_i \mid x_{it}, \theta_i^r) \right]^{-1} \left[ \frac{1}{R} \sum_{r} \left[ \theta_i^r + E(\eta_{it} \mid x_{it}, U_i, \theta_i^r) \right] f(U_i \mid x_{it}, \theta_i^r) \right]. \]

The individual specific means, \( E(\tilde{\nu}_{it} \mid x_{it}, U_i) \), are obtained using \( \tilde{\nu}_{it} = T_i^{-1} \sum_{t=1}^{T_i} \tilde{\nu}_{it} \). We use 1000 randomised Halton draws and 1000 antithetic draws, \( R = 2000 \), so that given the sample sizes \( \sqrt{N}/R \to 0 \).

\(^{18}\)We use the Generalised Least Squares (GLS) estimator, as opposed to pooled Ordinary Least Squares (OLS). Under the null of 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.

\(^{19}\)While in balanced panels, GLS is obtained as a simple weighted least squares (WLS), the weights are dependent upon the lengths of the time series per cross-section in the unbalanced panel case. We follow Baltagi’s (2012) suggestion and use the Swamy and Arora (1972) proposal to estimate the standard error estimates while allowing for individual-level clustering (see Baltagi, 2012, pp.182-184). Bootstrapping both stages in order to obtain the appropriate standard errors in the presence of heteroskedasticity and the inclusion of the generated correction terms, provides similar standard errors to the ones obtained using the Swamy and Arora (1972) estimator and
The four covariances \( (\sigma_{j,\alpha \theta}, \sigma_{j,\varepsilon \eta}) \) convey valuable information about the form of sorting into the two sectors (see Vella and Verbeek, 1999a). Note that the \( \theta_i \) are constructed so that their average value for union members is positive while, their average value for nonmembers is negative (see Vella and Verbeek, 1998). 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- and similarly if either covariance between \( (\varepsilon_{j,it}, \eta_{it}) \) is non-zero.

If both covariances \( (\sigma_{0,\alpha \theta}, \sigma_{1,\alpha \theta}) \) are positive, individuals with high values of \( \theta \) are on average the best employees in terms of their endowment of unobserved productivity irrespective of their union membership status (and \textit{vice-versa}). This is termed as hierarchical sorting. A comparative advantage (or positive sorting structure) 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 (\textit{i.e.} \( \sigma_{1,\alpha \theta} > 0, \sigma_{0,\alpha \theta} < 0 \)). Note that \( \sigma_{1,0} \) cannot be estimated directly.

The estimates obtained upon inclusion of the endogeneity correction terms, as well as, their interactions with union status are indicated as \textit{unrestricted sorting estimates} in the estimation results. The respective estimates including only \( (B_i, B_{it}) \) and omitting interactions are indicated as \textit{hierarchical sorting estimates}. Under unrestricted sorting we obtain a sector-invariant and a sector-variant contribution given by the coefficients on \( (B_i, B_{it}) \) and \( (B_iU_{it}, B_{it}U_{it}) \), respectively. For non-union members the latter is equal to zero as \( U_{it} = 0 \).

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 restricted control function estimators. A comparative advantage structure is precluded \textit{a priori} (for a detailed treatment see Vella and Verbeek, 1999a).

5. WAGE EQUATION PARAMETERISATION

The legal framework governing collective bargaining arrangements in the U.K was substantially altered during the 1980s. The introduction of the 1980 and 1982 Employment Acts strengthened the case for claiming unfair dismissal on the basis of an employee’s refusal to enter a closed shop. The 1982 Act dictated that all post-entry closed shops had to be sanctioned via a ballot of the workforce. The successive 1988 and 1990 Employment Acts 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 their unionisation decision 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). Such incentive private excludable goods 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).\footnote{Booth (1985) using a formalised model demonstrates that social customs enforced by a threat accounting for clustering. This is due to the generally low estimated coefficients of the endogeneity correction terms.}

\footnote{Under unrestricted sorting we obtain a sector-invariant and a sector-variant contribution given by the coefficients on \((B_i, B_{it})\) and \((B_iU_{it}, B_{it}U_{it})\), respectively. For non-union members the latter is equal to zero as \( U_{it} = 0 \).}
The process of union affiliation is not as clear in the U.K since no elections are held and there is no legal requirement of union recognition on the part of the employer even when the majority of employees are favourable of union representation. Union recognition is a voluntary act as is union membership upon joining an establishment that recognises unions for bargaining purposes (Hildreth 2000, p.133).

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

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

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

If union status is exogenous then

\[
E(\alpha_{j,i} | x_{it}, m_{it} c_{it}) = E(\epsilon_{j,it} | x_{it}, m_{it} c_{it}) = 0
\]

and 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

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

In the presence of endogeneity eq. (11) provides consistent estimates of the union membership and coverage differentials if an appropriate endogeneity correction procedure is employed. When eq. (11) is used instead of eq. (5) the estimation procedure described in Section 4 remains the same noting that \( j = 1 \) for covered members and \( j = 0 \) for covered nonmembers and uncovered employees. Accordingly, the \( \theta_i \) are constructed so that their average value for covered members is positive while their average value for covered nonmembers and uncovered employees is negative. This construction is based on the assertion that it is high union density at the establishment level that produces an above-average union differential while coverage on its own is insufficient (see Stewart, 1987, Andrews et al., 1998). Profit maximising employers will accordingly seek to employ the most productive members.\footnote{\textsuperscript{22,23,24}}

\footnote{As in the union membership case under unrestricted sorting we obtain a sector-invariant and a sector-variant contribution given by the coefficients on \( (B_i, B_{it}) \) and \( (B_i, m_{it} c_{it}, B_{it}, m_{it} c_{it}) \), respectively. For covered nonmembers and uncovered employees, the latter is equal to zero since in both cases \( m_{it} c_{it} = 0 \).}

\footnote{While it is conventional to consider three groups of prime economic interest, a distinction among uncovered members and uncovered nonmembers could be made leading to four categories. Due to few individuals pertaining to the former category the identification of such a differential is, however, fairly imprecise.}

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

Studies such as Hildreth (2000) and Swa¢ eld (2001) employing longitudinal BHPS data and accounting for fixed effects and measurement error, do find evidence of substantial positive wage differentials for covered members. Booth and Bryan (2004) using cross-sectional data from the linked Workplace Employee Relations Survey (WERS 1998), exploiting the within-establishment variation in wages, conclude that the apparent covered membership wage differential is illusory.25

Booth and Bryan (2004) provide a summary of the rationalisations advanced in the literature for the presence of positive membership wage premia. These can be broadly classified into selectivity or omitted variable arguments and discriminatory behaviour by either unions or employers.

Selectivity or omitted variable advanced explanations are based on the assertion of systematic differences among members and nonmembers in terms of unobserved productivity augmenting attributes. Assuming a positive association among such unobserved characteristics and union membership there will be a positive selection effect accompanied by the resulting upwardly biased covered membership premium.

Discriminatory behaviour arguments, assume that wage premia result from cooperative behaviour among unions and employers. Recognising that union cooperative behaviour can enhance profitability employers might tacitly engage in discriminatory practices such as systematically targeting training programs to union members hence establishing a wage advantage or, alternatively, employers might pay nonmembers from a point lower down the corresponding union wage scale (see Booth and Bryan, 2004, p. 405).26

It is worth noting that, high union density at the establishment produces a higher than average union differential and coverage on its own right is insufficient (see Stewart, 1987). It is therefore important to bear in mind that while individuals doing the same job at the same workplace should in theory earn the same when conducting analyses across establishments as in this study, membership is a closer determinant of a differential than coverage (Andrews et al., 1998, p.453).

Finally, as Hildreth (2000) clarifies we are measuring a wage differential and not a wage difference. Hence, it is indeed conceivable that members might enjoy positive wage premia thus invalidating the free riding problem.

6. DATA AND SAMPLES

We employ the BHPS data set to construct unbalanced panels of male and female employees during 1995-08.27

---

25 Hildreth (2000) finds some evidence of a wage differential for covered nonmembers though this was not generally robust while, Swa¢ eld (2001) generally finds lower positive wage differentials for covered nonmembers as opposed to covered members.

26 Booth et al. (2003) using BHPS data during 1991-96 find that among employees who received training, those in covered establishments enjoyed greater returns to training and higher wage growth.

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 was not asked of employees still in same job as in the previous year- see Section 8.1. 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 for example Arulampalam et al., 2000).

We therefore exclude individuals that were inactive in labour market during 1995, and exclude individuals 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 and individuals employed in the agricultural, forestry and fishing industrial classification or 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. 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 (Swaffield, 2001, p.439).

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 or covered members and covered nonmembers (frequencies and percentages). Tables providing detailed descriptions of the variables included in the reduced form union membership models and primary wage equations are provided in the Appendix.

7. 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. However, some key contributions in the field such as Abowd and Farber (1982) and Card (1996), as well as, the econometric study on endogeneity and heterogeneity by Heckman et al. (2006), suggest that sorting into sectors may follow 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 via the bargaining process implies reduced skill premia, the probability of joining an establishment that recognises unions for wage bargaining is expected to be inversely related to individual skill level. Alternatively, insider-outsider theories predict that highly skilled employees can be viewed as acting as a de facto union on its own since they cannot
be rapidly and costlessly replaced thus being less reliant on unions in order to extract concessions from employers (see Lindbeck and Snower, 1986; Blanchflower et al., 1990).

In effect, while potential candidates for employment within establishments recognising unions for bargaining purposes 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 include 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 (1-18) in Section 8, 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 and higher predicted observed skill quantiles in the male and 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, it is clear that 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 the suggestion of Vella and Verbeek (1998) we added interactions between the endogeneity correction terms, and their higher-order values, with educational controls to the list of covariates in eq. (5). Highest education qualification was chosen as the interaction variable since union wage effects generally differ by education level which can be thought of as a skill measure.

This is an informal first stage assessment of whether essential heterogeneity, in Heckman et al. (2006) terminology, is present. That is if individuals undertake their union membership decision 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. When δ in eq.(5) varies even after controlling for \((x_{it}, U_{it})\), there may be sorting on the gain, \(\text{cov}(\delta, U_{it}) \neq 0\), which is what Heckman et al. (2006) define as the essential heterogeneity model.\(^{28}\)

Inclusion of interactions between endogeneity correction terms and educational controls, generally produces distinct union wage differential estimates. This out-

\(^{28}\)Successful instrumentation to solve the endogeneity problem using either control function methods, modelling levels of conditional means, or the instrumental variables approach, modelling the slope of conditional means, becomes problematic. The single index endogeneity correction will only identify the mean union wage impact if individual decisions are made without knowledge of the corresponding idiosyncratic gain (see Heckman et al., 2006, p.391).
come suggests that a multiple indices rule should instead be employed. Crucially then as there evidence of essential heterogeneity, employing a single index approach will not permit identification of the mean union wage differential.29

Recognising that selection into union jobs leads to differing selection biases at different skill levels, we perform separate estimations for three distinct skill groups. Following Card (1996), we construct a skill index using the predicted wage quantile in the nonunion 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.

As the dynamic reduced form model for union membership calls for consecutive observations, the single index original 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 then form three distinct observed skill groups (referred to as low, middle and high skill) and re-estimate separate dynamic models of unionism and wage determination. The corresponding quantiles vary by gender and estimation sample and are clearly indicated in the Wage Regression tables.30

8. MEASUREMENT ERROR AND IDENTIFICATION

Measurement error issues are frequent in union status variables from longitudinal micro data sets in which event, least squares and fixed effects estimates are biased and inconsistent. Relatively more recent studies, such as Swaffield (2001) and Hirsch (2004), conclude that misclassification of union status causes attenuation in longitudinal (fixed effects) estimates of the union wage effect.

Freeman (1984) and Card (1996) note that since true changes in union status are infrequent in short time periods, observed union status changes are particularly prone to reporting error. As identification of the union wage effect in longitudinal studies relies on the partial correlation of wage changes with respect to changes in union membership status, misclassification bias is of primary concern. The FE estimator is especially problematic 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).31

29The estimation results inclusive of \((B_i, B_{it})\), and educational control interactions, are indicated as essential heterogeneity estimates and appear in all Wage Regressions tables. The interaction terms (not shown in the tables) are jointly statistically significant in all estimated models.

30The number of groups was chosen so as to maintain reasonable sample sizes within each observed skill category. As observations are stratified into quantiles according to the predicted wage cutoff the three groups are unequal in terms of size. Note that some individuals change predicted wage quantile across time. As the dynamic nature of the reduced form requires consecutive observations from the corresponding sample regressions. Estimating the original models using the pooled samples from the three quantiles excluding individuals changing quantiles across time we obtain similar results so there seems to be no issue of introducing additional sample selection problems.

31Griliches and Hauman (1986), under the assumption of stationary and uncorrelated errors of
Since union membership status is dichotomous, we are dealing with a non-standard error of measurement. As we generally find no evidence of union coverage differentials in the single index results we can simplify our analysis by initially assuming that we estimate the union wage differential in its simplest form, i.e. ignoring distinctions between covered members/nonmembers so that observed union status, $U_{it}$, corresponds to either the unconditional membership variable or covered membership.

For tractability, assume union membership is the only explanatory variable (along with a constant term) and let the true union status, $U^*_{it}$, be related to the observed measure via $U_{it} = U^*_{it} + \tau_{it}$ where $\tau_{it}$ is a random measurement error so that the respective wage equation is $w_{it} = \beta_0 + \delta U^*_{it} + \epsilon_{it}$ which corresponds to:

$$
\begin{align*}
w_{it} &= \beta_0 + \delta U_{it} + (\epsilon_{it} - \delta \tau_{it}) \\
\epsilon_{it} &= U_{it}(\alpha_1, i + \epsilon_{1,it}) + (1 - U_{it})(\alpha_0, i + \epsilon_{0,it})
\end{align*}
$$

Following Aigner (1973), denote the observed membership rate in the respective sample by $\hat{\pi} = N^{-1}T_i^{-1} \sum_{i=1}^N \sum_{t=1}^T U_{it}$ so that the observed proportion of nonmembers is $(1 - \hat{\pi})$. Let $\alpha_2$ denote the proportion of union members that have been misclassified as nonmembers according to the observed membership measure. Further, let $\alpha_1$ denote the proportion of individuals that have been erroneously classified as members according to $U_{it}$. The observed membership rate, $\hat{\pi}$, is thus related to the true membership rate, $\pi$, by

$$
\pi = (1 - \alpha_1) \hat{\pi} + \alpha_2 (1 - \hat{\pi}), \quad \alpha_1 = \Pr(U^*_{it} = 0 \mid U_{it} = 1), \quad \alpha_2 = \Pr(U^*_{it} = 1 \mid U_{it} = 0)
$$

or alternatively by

$$
\pi = [\Pr(U^*_{it} = 1 \mid U_{it} = 1)] \hat{\pi} + [\Pr(U^*_{it} = 1 \mid U_{it} = 0)] (1 - \hat{\pi}).
$$

The marginal distributions of $(U^*_{it}, U_{it})$ are Bernoulli with parameters $(\pi, \hat{\pi})$, respectively while the joint distribution of $(U_{it}, \tau_{it})$ is easily obtained by manipulation and is given in the joint frequency table that follows:

<table>
<thead>
<tr>
<th>$\tau_{it}$</th>
<th>$U_{it}$</th>
<th>$f(\tau_{it})$</th>
<th>$f(U_{it})$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$-1$</td>
<td>$\alpha_2 (1 - \hat{\pi})$</td>
<td>$0$</td>
<td>$\alpha_2 (1 - \hat{\pi})$</td>
</tr>
<tr>
<td>$0$</td>
<td>$(1 - \alpha_2) (1 - \hat{\pi})$</td>
<td>$(1 - \alpha_1) \hat{\pi}$</td>
<td>$(1 - \alpha_2) (1 - \hat{\pi}) + (1 - \alpha_1) \hat{\pi}$</td>
</tr>
<tr>
<td>$1$</td>
<td>$0$</td>
<td>$\alpha_1 \hat{\pi}$</td>
<td>$\alpha_1 \hat{\pi}$</td>
</tr>
</tbody>
</table>

measurement demonstrate that the fixed effects (within transformation) under classical errors in variables assumptions, converges to:

$$
\lim_{N \to \infty} \hat{\sigma}^2 = \delta \left[ 1 - T \frac{\sigma_U^2}{\sigma^2} \right].
$$

Thus, the attenuation bias caused by measurement error is exaggerated by the fixed effects estimator since $\sigma_U^2 < \sigma_{\hat{\pi}}^2$. This is due to weak or no serial correlation in the measurement error and high persistence in observed union membership status.
Note that, there is a negative correlation between true union status and the measurement error. Considering the pair \( (U_{it} = 0, \tau_{it} = -1) \) then \( U^*_{it} = 1 \), whereas, if \( (U_{it} = 1, \tau_{it} = 0) \) then \( U^*_{it} = 1 \). On the other hand, when \( (U_{it} = 0, \tau_{it} = 0) \) then \( U^*_{it} = 0 \) and finally, when \( (U_{it} = 1, \tau_{it} = 1) \) then \( U^*_{it} = 0 \). That is when \( U^*_{it} = 1, \tau_{it} \) is either -1 or 0 while when \( U^*_{it} = 0, \tau_{it} \) is either 0 or 1. Using the joint frequency table we obtain:

\[
E(\tau_{it}) = \alpha_1 \hat{\pi} - \alpha_2 (1 - \hat{\pi}) \quad (17)
\]
\[
Var(\tau_{it}) = [\alpha_1 \hat{\pi} + \alpha_2 (1 - \hat{\pi})] - [\alpha_1 \hat{\pi} - \alpha_2 (1 - \hat{\pi})]^2 \quad (18)
\]
\[
E(U_{it}) = \hat{\pi}, \quad Var(U_{it}) = \hat{\pi} (1 - \hat{\pi}) \quad (19)
\]
\[
E(U_{it}\tau_{it}) = \alpha_1 \hat{\pi}, \quad Cov(U_{it}, \tau_{it}) = (\alpha_1 + \alpha_2) \hat{\pi} (1 - \hat{\pi}) \quad (20)
\]

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.

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

\[
\text{plim} \hat{\delta}_{LS} = \delta \left[ 1 - \frac{\text{Cov}(U_{it}, \tau_{it})}{\text{Var}(U_{it})} \right] + \frac{\text{Cov}(U_{it}, \epsilon_{it})}{\text{Var}(U_{it})}. \quad (21)
\]

Recall that the random components \((\alpha_{j,i}, \varepsilon_{j, it}), \theta_{it}, \eta_{it}\) in equations (1, 2) are assumed to be \( iid \) drawings from a multivariate normal distribution, where every effect is potentially correlated with its counterpart of the same dimension, in the other equation. Hence the second source of asymptotic bias in eq.(21) stemming from the endogeneity of union status simply corresponds to the sum of the coefficients of the endogeneity correction terms. Substituting the corresponding quantities using (19, 20) into eq.(21) and given \( \frac{\text{Cov}(U_{it}, \epsilon_{it})}{\text{Var}(U_{it})} = \sigma_{j, \alpha \theta} + \sigma_{j, \epsilon \eta} \) we obtain

\[
\text{plim} \hat{\delta}_{LS} = \delta (1 - \alpha_1 - \alpha_2) + (\sigma_{j, \alpha \theta} + \sigma_{j, \epsilon \eta}), \quad j = 0, 1. \quad (22)
\]

Under the null of exogeneity of union membership status, since the misclassification rates by definition satisfy \((0 < a_m < 1, m = 1, 2)\) the least squares estimate of the union wage effect 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. A consistent estimator for \( \delta \) would be \( \hat{\delta}_{LS} / (1 - \alpha_1 - \alpha_2) \). In the multiple regression model, however, there is no straightforward expression available (see Aigner, 1973, p.53).

Rewriting eq.(22) in terms of conditional probabilities we obtain

\[
\text{plim} \hat{\delta}_{LS} = \delta [1 - \text{Pr}(U^*_{it} = 0 | U_{it} = 1) - \text{Pr}(U^*_{iT} = 1 | U_{it} = 0)] + (\sigma_{j, \alpha \theta} + \sigma_{j, \epsilon \eta}) \quad (23)
\]

which clearly indicates that under null of exogeneity the union membership differential will be attenuated as long as a certain proportion of nonmembers are classified as members, whereas, a certain proportion of members are classified as
Hierarchical sorting restricts the covariances among the random components of the same dimension in the structural and reduced form equations to be sector-invariant and therefore eq. (22) reduces to
\[
\text{plim } \hat{\delta}_{LS} = \delta \left(1 - \alpha_1 - \alpha_2\right) + (\sigma_{\alpha \theta} + \sigma_{\varepsilon \eta}) \cdot \quad (26)
\]
Hence, if the null of exogeneity is rejected the conventional downward bias result will hold if \((\sigma_{\alpha \theta} + \sigma_{\varepsilon \eta}) < 0\) i.e. if negative selection is present or dominant considering both the fixed and time-variant endogeneity correction term coefficients. In the presence or prevalence of positive selection, \((\sigma_{\alpha \theta} + \sigma_{\varepsilon \eta}) > 0\), the conventional downward bias will hold if the attenuation bias from measurement error is greater than the selection bias.

On the other hand, under unrestricted sorting \((\sigma_{0,\alpha \theta} + \sigma_{0,\varepsilon \eta}) \neq (\sigma_{1,\alpha \theta} + \sigma_{1,\varepsilon \eta})\). Thus, measurement error reduction will only raise the uncorrected GLS estimate of \(\delta\) under unrestricted sorting since assuming that \((\alpha_1 = \alpha_2 \approx 0)\) then
\[
\text{plim } \hat{\delta}_{LS} = \delta + (\sigma_{0,\alpha \theta} + \sigma_{0,\varepsilon \eta}) + (\sigma_{1,\alpha \theta} + \sigma_{1,\varepsilon \eta}) \cdot \quad (27)
\]
and comparative advantage requires \((\sigma_{1,\alpha \theta} > 0, \sigma_{0,\alpha \theta} < 0)\) and \((\sigma_{1,\varepsilon \eta} > 0, \sigma_{0,\varepsilon \eta} < 0)\) or alternatively that either \((\sigma_{1,\alpha \theta} > 0, \sigma_{0,\alpha \theta} < 0)\) 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 \((\sigma_{1,\alpha \theta} \approx 0, \sigma_{0,\alpha \theta} \approx 0)\). In other words, the positive contribution of unobserved heterogeneity not eliminated by the uncorrected GLS estimator biases upwards the membership differential.33

Conclusively then, in the absence of other covariates and assuming there are no coverage differential, it has been demonstrated that in the presence of measurement error the conventional attenuation bias result will hold if union status is exogenous or, under hierarchical sorting and negative selection so that the selectivity and measurement error biases reinforce each other.

Under hierarchical sorting and negative selection, measurement error reduction will not raise the uncorrected GLS estimate since the negative endogeneity bias has

---

32 Alternatively let \(q_1 \equiv \text{Pr}(U_{it} = 1| U_{it}^* = 1), q_0 \equiv \text{Pr}(U_{it} = 1| U_{it}^* = 0)\) where \(q_0, (1 - q_1)\) is the probability of a false positive (negative). Since eq.(23) corresponds to
\[
\text{plim } \hat{\delta}_{LS} = \delta \left[\text{Pr}(U_{it}^* = 1| U_{it} = 1) - \text{Pr}(U_{it}^* = 1| U_{it} = 0)\right] + (\sigma_{j,\alpha \theta} + \sigma_{j,\varepsilon \eta}) \quad (24)
\]
and given \(\text{plim } \hat{\pi} = \pi q_1 + (1 - \pi)q_0\) using Bayes’ rule we have
\[
\text{Pr}(U_{it}^* = 1| U_{it} = 1) = \frac{\text{Pr}(U_{it} = 1| U_{it}^* = 1) \text{Pr}(U_{it}^* = 1)}{\text{Pr}(U_{it} = 1)} = \frac{\pi q_1}{\pi q_1 + (1 - \pi)q_0}
\]
and similarly
\[
\text{Pr}(U_{it}^* = 1| U_{it} = 0) = \frac{\pi (1 - q_1)}{\pi (1 - q_1) + (1 - \pi)(1 - q_0)}.
\]
Substituting the relevant quantities in eq.(24) we obtain the following expression that under the null of exogeneity corresponds to the attenuation bias derived by Card (1996, p.959) noting that the term in square brackets is less than 1 if \(q_1 < 1\) and \(\pi \approx \hat{\pi}\):
\[
\text{plim } \hat{\delta}_{LS} = \delta \left[\frac{\pi(q_1 - \hat{\pi})}{\pi(1 - \hat{\pi})}\right] + (\sigma_{j,\alpha \theta} + \sigma_{j,\varepsilon \eta}), \quad j = 0, 1. \quad (25)
\]
33 Recall that since the \(\hat{\theta}_j\) are constructed so that their average value is positive for members (covered members) and negative for nonmembers (covered nonmembers and uncovered employees) in the two model variants, the average contribution of unobserved heterogeneity is positive for all groups under comparative advantage sorting.
not been removed, whereas, under comparative advantage sorting removing measurement error will raise the uncorrected GLS estimate. Finally, provided that the endogeneity underlying union status is individual-specific and fixed, the FE (within transformation) estimate will be increased upon measurement error reduction as the time-invariant endogeneity is eliminated.

While we find no robust evidence of coverage differentials in the single index results, there is some evidence of coverage differentials in the multiple indices results (see Sections 8.2, 8.3 and 8.5, 8.6). In this event in the absence of additional covariates, and assuming the true covered non-membership status 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

\[
\begin{align*}
  w_{it} &= \beta_0 + \delta_1 U_{it} + \delta_2 F_{it} + (e_{it} - \delta_1 \tau_{it} - \delta_2 \xi_{it}) \\
  U_{it} &= m_{it}c_{it}, \quad F_{it} = (1 - m_{it})c_{it}, \quad U_{it} = U^*_it + \tau_{it}, \quad F_{it} = F^*_it + \xi_{it} \\
  e_{it} &= m_{it}c_{it}(\alpha_{1,i} + \varepsilon_{1,it}) + (1 - m_{it}c_{it})(\alpha_{0,i} + \varepsilon_{0,it})
\end{align*}
\]  

(28)

Therefore, as \( N \to \infty \), the \( \text{plim} \) of the pooled (ordinary/generalised) least squares estimator of \( \delta_1 \), assuming that the variances, covariances and misclassification rates are all constant across \( t \) corresponds to

\[
\text{plim} \, \hat{\delta}_{1,LS} = \delta_1 (1 - \alpha_1 - \alpha_2) + \delta_2 \left[ \frac{\text{Cov}(U_{it}, F_{it})}{\text{Var}(U_{it})} - \frac{\text{Cov}(U_{it}, \xi_{it})}{\text{Var}(U_{it})} \right] + (\sigma_{j,\alpha \theta} + \sigma_{j,\xi \eta}).
\]  

(29)

The marginal distributions of \( (F^*_{it}, \, F_{it}) \) are Bernoulli with parameters \((\pi, \hat{\pi})\) and the joint distribution of \( (F_{it}, \xi_{it}) \) can be easily obtained by interchanging \((1 - \hat{\pi})\) with \( \hat{\pi} \) in the joint frequency table of \((U_{it}, \tau_{it})\) since when \( c_{it}=1 \), i.e. the case of interest, then \((1-U_{it}) = F_{it} \).34

It then follows that \( \text{Cov}(U_{it}, F_{it}) = -\text{var}(U_{it}) \) and \( \text{Cov}(U_{it}, \xi_{it}) = -\text{Cov}(F_{it}, \xi_{it}) \) and plugging these into eq.\((29)\) we obtain

\[
\text{plim} \, \hat{\delta}_{1,LS} = \delta_1 (1 - \alpha_1 - \alpha_2) + \delta_2 \left[ \frac{\text{Cov}(F_{it}, \xi_{it})}{\text{Var}(U_{it})} - 1 \right] + (\sigma_{j,\alpha \theta} + \sigma_{j,\xi \eta}).
\]  

(30)

and since \( \text{Cov}(F_{it}, \xi_{it}) = (\alpha_1 + \alpha_2) \hat{\pi}(1 - \hat{\pi}) \) then the \( \text{plim} \) of the least squares estimator of \( \delta_1 \) reduces to

\[
\text{plim} \, \hat{\delta}_{1,LS} = (\delta_1 - \delta_2) (1 - \alpha_1 - \alpha_2) + (\sigma_{j,\alpha \theta} + \sigma_{j,\xi \eta}), \quad j = 0, 1.
\]  

(31)

Similarly, under the assumption that the variances, covariances and misclassification rates are all constant across \( t \) as \( N \to \infty \) the \( \text{plim} \) of the pooled (ordi-

---

34 Using the joint distribution of \((F_{it}, \xi_{it})\) it is clear that we are dealing with non-classical measurement error since:

\[
\begin{align*}
  E(\xi_{it}) &= \alpha_1 (1 - \hat{\pi}) - \alpha_2 \hat{\pi} \\
  \text{Var}(\xi_{it}) &= [\alpha_1 (1 - \hat{\pi}) + \alpha_2 \hat{\pi}] - [\alpha_1 (1 - \hat{\pi}) - \alpha_2 \hat{\pi}]^2 \\
  E(F_{it}) &= (1 - \hat{\pi}), \quad \text{Var}(F_{it}) = \hat{\pi}(1 - \hat{\pi}) \\
  E(F_{it}\xi_{it}) &= \alpha_1 (1 - \hat{\pi}), \quad \text{Cov}(F_{it}, \xi_{it}) = (\alpha_1 + \alpha_2) \hat{\pi}(1 - \hat{\pi})
\end{align*}
\]
nary/generalised) least squares estimator of $\delta_2$ corresponds to
\[
\text{plim}_{N \to \infty} \hat{\delta}_{2,LS} = \delta_2 (1 - \alpha_1 - \alpha_2) + \delta_1 \left[ \frac{\text{Cov}(F_{it}, U_{it})}{\text{Var}(F_{it})} - \frac{\text{Cov}(F_{it}, \tau_{it})}{\text{Var}(U_{it})} \right] + (\sigma_{j,\alpha\theta} + \sigma_{j,\varepsilon\eta})
\]
and given $\text{Cov}(F_{it}, U_{it}) = -\text{var}(F_{it})$ and $\text{Cov}(F_{it}, \tau_{it}) = -\text{Cov}(U_{it}, \tau_{it})$ substituting these relationships in eq. (32) the plim of the least squares estimator of $\delta_2$ simplifies to
\[
\text{plim}_{N \to \infty} \hat{\delta}_{2,LS} = (\delta_2 - \delta_1) (1 - \alpha_1 - \alpha_2) + (\sigma_{j,\alpha\theta} + \sigma_{j,\varepsilon\eta}), \ j = 0, 1.
\]

Using the preceding analysis it is straightforward to obtain the respective versions of eq.(31, 33) under restricted and unrestricted sorting.

Equation (31) is critical in that it indicates that in the presence of a coverage differential then the union membership differential, defined in eq.(12), corresponds to $(\delta_1 - \delta_2)$ and will be attenuated in the presence of measurement error under the null of exogeneity.

If the null of exogeneity is rejected then the uncorrected GLS estimate of the union membership differential will be biased downwards under hierarchical sorting and negative selection so that the measurement error and selectivity biases reinforce each other. Conclusively then, measurement error reduction will only raise the uncorrected GLS estimate of $(\delta_1 - \delta_2)$ under comparative advantage sorting.\(^{35}\)

Given the absence of suitable instrumental variables, correlated with union status and uncorrelated with measurement error, one should employ alternative error reduction strategies. We initially compare the two membership questions to detect potential inconsistencies and further, we restrict samples so that union status changes occur only when corresponding employer and/or job changes are observed while accounting for changes in union recognition.

While our strategy leads to sample reduction, assuming that individuals change status when accompanying employer changes are observed, is in fact a requirement of our exclusion restriction of lagged union membership in the primary wage equation (see Section, 8.4.1). Original unrestricted sample estimates not accounting for measurement error are also presented for comparability.\(^{36}\)

### 8.1. Measurement Error in Union Status Response

The best measure of union membership 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

\(^{35}\)The same holds regarding $(\delta_2 - \delta_1)$ noting that across all estimates $\delta_1 > \delta_2$.

\(^{36}\)Retaining measurement error identified observations via imputation of values according to initial period membership status (as in Swafield, 2001) is clearly not an option due to the relatively large time span of our analysis during which genuine changes in union status are likely to occur.
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 question is our preferred measure of union membership status as it is reported annually without gaps from 1995 onwards.

Since we wish to distinguish among membership premia 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.

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.

As the latter is a broader definition, given that it includes staff associations, 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 our preferred covered membership status variable only for the corresponding years.

8.2. Male Initial Measurement Error Correction Results

We primarily estimate the union wage impact using the continuous membership variable obtained by combining the unconditional "Member of trade union" and conditional upon coverage questions for the years the former was unavailable.

Under the single sorting assumption the estimates (in Table 1) indicate that union members are negatively selected thus biasing downwards the uncorrected GLS estimate. Therefore, employees receiving lower wages, conditioning on their attributes and in the absence of unions, are those more likely to be union members.

Given that the interactions between the correction terms and membership status are

---

37 The results obtained using the entire respective sample under analysis are henceforth referred to as single index sorting estimates and are provided on the left-hand side of the Wage Regressions tables. The economic 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 economic sorting structure, with educational controls). The predicted quantile results, on the right-hand side, are referred to as multiple indices sorting estimates.
statistically insignificant, the single index estimates are consistent with hierarchical sorting.

The statistically significant coefficient on the time-variant effects in both the single index, as well as, the intermediate skill group estimates (in Table 1) imply that FE estimation is inappropriate since the time-varying endogeneity is not eliminated and continues to contaminate our estimates. The difference among the essential and hierarchical estimates indicates that the single index endogeneity correction fails to identify the mean union effect. The appropriate multiple indices estimates indicate that solely union members pertaining to the middle skill group are in receipt of an approximately 4.1 percent wage differential. On the other hand, high skill group union members experience an approximately 2.5 percent reduction in their wages and are therefore in receipt of reduced human capital premia due to wage standardisation in unionised establishments.

| TABLE 1: WAGE REGRESSIONS (UNION MEMBERSHIP), MALE (1995-2009) |
|--------------------------|--------------------------|--------------------------|--------------------------|--------------------------|
| Predicted Quantile       | <=29                     | >29-<=63                 | >63                      |
| GLS                      | GLS-H                    | GLS-U                    | GLS-EM                   |
| Coef. z                  | Coef. t                  | Coef. z                  | Coef. z                  |
| Uit                      | -0.007                   | -0.040                   | -0.001                   |
| Bi                       | -0.038                   | -0.036                   | -0.018                   |
| Bi.Uit                   | 0.067                    | 1.42                     |                            |
| Observations             | 9,599                    | 6,418                    | 2,529                    |
| Members                  | 2,828                    | 2,048                    | 502                      |
| Members (%)              | 29.46                    | 31.91                    | 19.85                    |
| Log-L                    | -1695.4                  | -1695.4                  | 236.4                    |

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

It is worth taking a closer look at the unrestricted estimates under both single and multiple indices sorting assumptions. While the negative coefficient on the individual-specific correction term \(B_i\) indicates that union members are negatively selected, the positive interaction term between \(B_i\) and covered membership means that their membership decision is positively rewarded. Since by construction the average value of the \(B_i\) for covered members is positive while for covered nonmembers and uncovered employees is negative, the average contribution of unobserved heterogeneity on wages is positive for all three groups (i.e. covered members, covered nonmembers and uncovered employees).

As there is evidence of essential heterogeneity, the appropriate multiple indices estimates indicate that unions establish wage premia of approximately 6.2 percent.
only for covered members pertaining to the lower predicted skill group.\(^{38}\)

**TABLE 2: WAGE REGRESSIONS (UNION COVERAGE), MALE (1995-2009)**

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

\(m_{c2}\) 0.054 0.105 0.050 2.87 \(1-0.03\) 3.90 0.081 3.02 0.064 2.86
\(1-m_{c2}\) 0.003 0.21 0.007 0.46 \(-0.001\) -0.001 -0.06 0.004 0.24
\(B_{1}\) -0.010 -2.38 -0.033 -4.06 \(-0.011\) -2.03 -0.030 -4.74
\(B_{2}\) -0.013 -1.49 -0.020 -1.49 \(-0.005\) -0.54 0.012 0.83
\(B_{i}(m_{c2},c)\) 0.051 3.22 0.030 3.26 0.030 3.26 0.016 0.84
\(B_{i}(m_{c2},c)\) -0.012 -0.72 -0.030 -1.50

**TABLE 3: WAGE REGRESSIONS (CORRECTED UNION COVERAGE), MALE (1995-2009)**

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

\(m_{c2}\) 0.043 3.37 0.042 2.45 0.085 3.48 0.076 3.02 0.066 3.03
\(1-m_{c2}\) 0.014 1.21 0.015 1.06 0.002 0.15 0.002 0.13 0.004 0.25
\(B_{1}\) -0.004 -0.42 -0.027 -1.05 \(-0.001\) -0.04 0.027 0.25
\(B_{2}\) -0.031 -3.12 -0.034 -2.36 -0.030 -1.50
\(B_{i}(m_{c2},c)\) 0.046 0.99 0.006 0.30
\(B_{i}(m_{c2},c)\) -0.012 -0.72 -0.030 -1.50

**Notes:**
- **H**: hierarchical, **U**: unrestricted, **EH**: essential heterogeneity test (endogeneity correction terms not shown)
- GLS: Swamy-Aroa estimator with cluster-robust std. errors; FE: cluster-robust std. errors; Log-L: Reduced Form Log-likelihood
- Predicted Quantile: stratification into quantiles according to a predicted wage in the uncovered sector; appropriate estimates shown only

Undertaking the first stage measurement error correction to detect whether an individual is indeed a union and not a staff association member (during the years both membership questions are available) we obtain our preferred covered membership measure.

The single index and intermediate observed skill group estimates (in Table 3) are consistent with a hierarchical sorting structure, union members are negatively selected biasing downwards the uncorrected GLS estimate and endogeneity is time-variant rendering FE estimation inappropriate. Given the differences among the membership differentials obtained using hierarchical and essential heterogeneity specifications, there is presence of essential heterogeneity. The appropriate multiple indices estimates indicate wage premia for covered members pertaining to both the low and middle skill groups of around 4.8 and 3.4 percent, respectively. Lastly, if one considers the coverage differential of around 3.1 percent to be statistically significant at a sufficiently stringent level then, the effective union membership differential for the lower skilled group is reduced to a mere 1.6 percent.

**TABLE 3: WAGE REGRESSIONS (CORRECTED UNION COVERAGE), MALE (1995-2009)**

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

\(m_{c2}\) 0.043 3.37 0.042 2.45 0.085 3.48 0.076 3.02 0.066 3.03
\(1-m_{c2}\) 0.014 1.21 0.015 1.06 0.002 0.15 0.002 0.13 0.004 0.25
\(B_{1}\) -0.004 -0.42 -0.027 -1.05 \(-0.001\) -0.04 0.027 0.25
\(B_{2}\) -0.031 -3.12 -0.034 -2.36 -0.030 -1.50
\(B_{i}(m_{c2},c)\) 0.046 0.99 0.006 0.30
\(B_{i}(m_{c2},c)\) -0.012 -0.72 -0.030 -1.50

**Notes:**
- **H**: hierarchical, **U**: unrestricted, **EH**: essential heterogeneity test (endogeneity correction terms not shown)
- GLS: Swamy-Aroa estimator with cluster-robust std. errors; FE: cluster-robust std. errors; Log-L: Reduced Form Log-likelihood
- Predicted Quantile: stratification into quantiles according to a predicted wage in the uncovered sector; appropriate estimates shown only

\(^{38}\) Note that since we find no robust evidence of statistically significant union coverage differentials \((\delta_2)\) in the single index estimates, we generally treat the total union differential \((\delta_1)\) and the union membership differential \((\delta_1 - \delta_2)\) defined in eq.(11) as the same concept i.e. assuming that \(\delta_2 = 0\).
Union members in the higher observed skill group experience an approximately 3.5 percentage reduction in their wages suggesting reduced human capital premia within unionised establishments due to wage standardisation. However, this effect is not robust and disappears when restricting membership changes to those more likely to experience true changes as a measurement error reduction device (see Section 8.4).

Regarding the relationship among the single index estimates of the union wage effect in this Section, a discernible pattern arises. FE estimation provides the lower bound as it fails to eliminate the negative impact of the time-variant selection terms. By eliminating the negative impact of the unobservables we are effectively adding their contribution to the union wage effect and FE ignore the role of time-varying endogeneity. Hence, corrected GLS corresponds to the upper bound of the union membership effect.

Using uncorrected coverage instead of membership raises the single index uncorrected GLS in Table 2. Recall that the union membership variable suffers from discontinuity and our approach to use union coverage for the years that the unconditional membership question was not asked is a source of measurement error. By using the coverage variable for the entire period instead, we are effectively reducing measurement error and the uncorrected GLS estimate of the membership effect rises since the estimates (in Table 2) are consistent with comparative advantage sorting—see eq. (27). As Table 3 estimates are consistent with hierarchical sorting and negative selection, measurement error reduction lowers the uncorrected GLS estimate of the membership effect in Table 3 as opposed to Table 2—refer to eq. (26).

Reducing measurement error raises the single index FE estimate in Table 2 as it eliminates the negative time-invariant selection effect and time-variant heterogeneity is statistically insignificant. On the other hand, the FE estimate is lower in Table 3, as opposed to Table 2, since the FE estimator fails to eliminate the statistically significant negative time-variant selection.

Turning to the multiple indices results it is clear that usage of the coverage variable (in Table 2) increases the low skilled union membership differential to around 6.2 percent compared to the statistically insignificant baseline estimate of approximately 2.6 (in Table 1). This occurs in that the 2.1 coverage differential (in Table 2) is nearly three times smaller than the corresponding total union differential of 6.2 that effectively coincides with the membership differential due to the statistical insignificance of the coverage premium. However, as revealed by Table 3, further measurement error reduction raises the coverage premium and therefore reduces the effective membership differential for the low predicted skill group to approximately 1.6 percent—see eq. (31).

Correcting the coverage measure increases the intermediate group membership differential in Table 3, noting that accounting for the insignificant coverage premium it nearly converges back to the respective membership differential of Table 1. Lastly, measurement error correction of both the baseline membership and coverage measures, increases the negative high skill group membership differential as the negative coverage differentials are insignificant (refer to Tables 1, 2 and 3).

### 8.3. Female Initial Measurement Error Correction Results

Starting with the baseline estimates, using the continuous membership variable constructed by combining the unconditional and conditional on coverage questions, the negative selection of union members biases downwards the uncorrected GLS
As can be seen the endogeneity underlying union decisions under the single index assumption (in Table 4) is time-variant rendering FE estimation inappropriate. Note that the difference between the hierarchical sorting and essential heterogeneity differentials suggests that a multiple indices rule should be employed instead. Accordingly, the multiple indices estimates indicate differential patterns of selection biases stemming through the time-invariant effects. There is clear evidence of negative selection of union members from the middle and upper skill groups and weaker evidence, in terms of statistical significance, of positive selection occurring at the lower end of the observed skill distribution. Closer inspection of the coefficients of the endogeneity correction terms makes it clear that the endogeneity impact is least at the middle of the observed skill distribution. Finally, the estimates indicate a 5.2 percent premium for union members in the intermediate skill group and provide rather weak evidence of a negative membership differential in the higher group case.

As in the baseline estimates, under the multiple indices sorting assumption we obtain distinct selection biases between the lower and the remaining two observed skill categories. While there is positive selection of union members at the lower end of the observed skill distribution, the selection biases for the middle and higher skill categories are instead negative. Once more, the lowest endogeneity correction term coefficients occur at the middle group. Hence, endogeneity has a more pronounced impact at the extremes of the observed skill distribution. As the two dominant biases at the ends of the observed skill distribution have opposite signs, they approximately offset each other and render the unrestricted and essential het-

<table>
<thead>
<tr>
<th>Predicated Quantile</th>
<th>GLS FE</th>
<th>GLS-H</th>
<th>GLS-U</th>
<th>GLS-EH</th>
<th>GLS-H</th>
<th>GLS-U</th>
</tr>
</thead>
<tbody>
<tr>
<td>&lt;=34</td>
<td>&gt;=34 &lt;=66</td>
<td>&gt;66</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GLS FE</td>
<td>Coef. z</td>
<td>Coef. t</td>
<td>Coef. z</td>
<td>Coef. z</td>
<td>Coef. z</td>
<td>Coef. z</td>
</tr>
<tr>
<td>U_1</td>
<td>0.041</td>
<td>3.39</td>
<td>-0.015</td>
<td>-1.05</td>
<td>0.091</td>
<td>2.83</td>
</tr>
<tr>
<td>B_1</td>
<td>0.000</td>
<td>-0.07</td>
<td>-0.006</td>
<td>-0.34</td>
<td>0.001</td>
<td>1.93</td>
</tr>
<tr>
<td>B_2</td>
<td>-0.047</td>
<td>-3.34</td>
<td>-0.061</td>
<td>-3.48</td>
<td>0.010</td>
<td>-0.36</td>
</tr>
<tr>
<td>B_1U_2</td>
<td>0.005</td>
<td>0.17</td>
<td>0.098</td>
<td>3.23</td>
<td>0.098</td>
<td>3.23</td>
</tr>
<tr>
<td>B_1U_2</td>
<td>0.038</td>
<td>1.36</td>
<td>-0.012</td>
<td>-0.47</td>
<td>0.038</td>
<td>1.36</td>
</tr>
</tbody>
</table>

Observations: 7,056 7,056 4,050 4,050 4,050 885 1,081 1,686
Members: 2,191 2,191 1,457 1,457 1,457 166 347 758
Members (\%): 31.05 31.05 35.98 35.98 35.98 18.76 32.10 44.96

H: hierarchical, U: unrestricted, EH: essential heterogeneity test (endogeneity correction terms not shown)
GLS: Swamy Arora estimator with cluster-robust std. errors; FE: cluster-robust std. errors; Log-L: Reduced Form Log-likelihood
Predicated Quantile: stratification into quantiles according to a predicted wage in the uncovered sector; appropriate estimates shown only
A rationalisation of the distinct biases by skill group is offered by Card (1996) using a two-sided selection model interacting employee and employer criteria in the selection process. Assuming a negative association among union membership premia and location in the observed skill distribution (consistent with our findings), the model predicts that the employee’s selection criterion is more likely to be binding than the employer’s selection criterion given a higher level of observed skill. Highly skilled employees within the unionised sector are thus more likely to have negative values of unobserved individual-specific effects \( (\theta_i) \) yielding a negative selection bias. On the other hand, the employer’s selection criterion is more likely to be binding than the employee’s selection criterion given a low level of observed skill. Hence, among unionised employees with lower levels of observed skill the frequency of higher values of individual-specific effects \( (\theta_i) \) is greater producing a positive selection bias (see Card, 1996, pp.976-978).

Moving to the single sorting estimates obtained using the corrected union coverage measure (in Table 6) are consistent with hierarchical sorting and negative selection biasing downwards the uncorrected GLS membership differential noting the inappropriateness of FE estimation as endogeneity stems from \( B_{it}. \) Given the discrepancy among the unrestricted and essential heterogeneity membership premia there is evidence of essential heterogeneity. As in the remaining female estimates, the appropriate multiple indices estimates indicate distinct biases per observed skill group. There is negative selection for middle and higher skill group union employees and weak evidence, in terms of statistical significance, of positive selection of union members from the lower end of the observed skill distribution. In line with the preceding estimates, the impact of the endogeneity correction terms is least at the middle of the observed skill distribution.

Regarding wage differentials it is only union members from the intermediate skill group that are recipients of a 5.1 percent premium, whereas, the approximately 6.9 percent negative membership differential for high skilled employees diminishes compared to the respective Table 5 estimate (in terms of both magnitude and statistical significance).

---

We now turn to the relationship among the distinct single index estimators. Similarly to the male results the FE estimate of the union wage effect provides the lower bound though in the female models it is additionally local to zero and statistically insignificant. Unlike the male results, in the female estimates there is recurring evidence of differential biases occurring at the lower end of the observed skilled distribution and the remaining two skill groups. Noting that the impact of endogeneity is least at the middle of the observed skill distribution the dominant biases occur at the two extremes. As these opposite dominant effects approximately offset each other, they drive the female FE estimate of the membership differential towards zero. Conclusively then, the corrected GLS estimates correspond to the upper bound of the union wage impact noting that in Table 5 this is only the case with the restricted sorting differential.

Estimation using the uncorrected coverage measure in Table 5, raises the single index uncorrected GLS membership differential compared to the corresponding baseline estimate in Table 4. The observed increase in the uncorrected GLS membership premium occurs in that Table 5 estimates are consistent with comparative advantage sorting- refer to eq. (27).

Undertaking further measurement error reduction (in Table 6), however, reduces the uncorrected GLS estimate of the union membership differential. Referring to eq. (31), it can be seen that the decrease occurs in that the estimates (in Table 6) are consistent with hierarchical sorting and negative selection while additionally the coverage premium increases and becomes statistically significant.

We finally consider the impact of measurement error on the multiple indices estimates. The 4 percent effective membership differential for the middle skill group (in Table 5) coincides with the total differential since the corresponding 2 percent negative coverage differential that would have otherwise increased the membership differential (compared to the respective 5.2 baseline differential in Table 4) is statistically insignificant. Correcting the coverage measure (in Table 6) does raise the effective middle skill membership differential which coincides with the total differential of approximately 5.1 percent- refer to eq. (31).

The weakly significant negative membership differential for high skilled members in the baseline estimates (Table 4) remains negative though reduced in terms of magnitude while its statistical significance increases when using uncorrected coverage (see Table 5). This is further reduced in Table 6 though it is not statistically insignificant.
significant at a sufficiently stringent level. Referring to eq.(27) it can be seen that measurement error reduction reduces the magnitude of the negative premium as the estimates at the upper extreme of the observed skill distribution are consistent with comparative advantage sorting.

8.4. 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 status variable is performed by restricting status changes to those that are 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.

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 other studies employing BHPS data and accounting for measurement error such as Swafield (2001) and Koevets (2007), our time period under analysis is considerably greater. It then stands to reason that we should observe several true changes in union recognition for bargaining purposes on the part of employers.

Unless we additionally account for changes in union recognition, our restriction regarding status changes only when corresponding employer changes are observed would be erroneously discarding some true changes. We therefore re-estimate all models by omitting observations where union status changes were not accompanied by either a job and/or employer changes if and only if a corresponding change in union recognition did not occur- that is membership status changes with accompanying recognition changes are classified as true changes.

8.4.1. Identification and Simultaneity

The assumptions regarding the errors identify all parameters in equation (5) given the non-linear mapping from the reduced form variables to the endogeneity correction terms \( (B_i, B_{it}) \) defined in equations (8) and (9). The imposition of exclusion restrictions is, however, desirable. The exclusion of lagged union status from the empirical counterpart of (eq.5) identifies the equation as long as \( \gamma_1 \) in eq.(2) differs from zero.\(^{40}\)

Reduced form models include political closeness controls conceptually expected to affect unionisation decisions while not impacting on wages and thus excluded

\(^{40}\)Of course, the same argument carries forward regarding estimating eq.(11).
from the structural equations. However, these were not always statistically significant and achieving instrumentation solely on the basis of political controls was unfortunately not possible.

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

Crucially then, restricting samples so 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.

Think of an employee moving to a different establishment. Since the most dominant level of wage bargaining in the U.K. is establishment level negotiation and that the collective agreement length is one year, previous membership status will only affect individual propensity towards unionisation and will not be an integral part of the future employer wage setting decision. In fact, given the current state of U.K. industrial relations, membership will be irrelevant if the new employer does not recognise unions for wage bargaining purposes. 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 and past union status at the previous establishment will play no role whatsoever on current establishment wage rate determination.\(^{41}\)

Considering an individual that changes union status while at the same establishment, provided that the observed change is not due to a coverage recognition change, the assumption of past membership not impacting upon 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 process in particular when employers engage in discriminatory behaviour in return for union cooperation.

A further underlying assumption of our exclusion restriction is that long-term advantages of union employment, whilst generating persistence of union membership status, 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 estimate upon the inclusion of tenure. Inclusion of tenure and its square in our set of controls, reduces the union wage effect in all of our estimates therefore partialling out potential \(U_{i,t-1}\) effects.

8.5. Male Estimates

Prior to analysing the results it should be stressed that comparisons in this Section are made with respect to the estimates obtained using what is termed as the corrected coverage measure (in Table 3). The same holds regarding the respective female estimates (in Section 8.6) that are accordingly compared to Table 6 estimates.

This is done for two reasons. Primarily since corrected coverage is the best possible union membership measure that can be derived from the BHPS dataset

\(^{41}\)For institutional features of wage bargaining in the UK refer to Du Cahu et al. (2008).
being conditional upon coverage and ignoring staff association members. Secondly, the additional measurement error reduction techniques adopted identify potentially erroneous observations by restricting union membership status changes to those more likely to experience changes. Since such measurement error reduction strategies might exclude some true changes, we ought to make comparisons with respect to the estimates obtained using the best possible membership measure.\footnote{A related issue concerns the notably low log-likelihoods of the reduced form models corresponding to the estimates reported in Sections (8.5, 8.6) under both single and multiple indices assumptions. Restricting union status changes, as a means of measurement error reduction, lowers the variation of the dependent variable in the reduced form models thus making it increasingly difficult to model union membership determination by the set of observables.}

### TABLE 7: WAGE REGRESSIONS (UNION CHANGE IF JOB CHANGE, COVERAGE), MALE (1995-2009)

<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 FE GLS-H GLS-U GLS-EM</td>
<td>GLS FE GLS-H GLS-U GLS-EM</td>
<td>GLS FE GLS-H GLS-U GLS-EM</td>
<td></td>
</tr>
<tr>
<td>m-it</td>
<td>Coef. 2 Coef. 1 Coef. 2 Coef. 2</td>
<td>Coef. 2 Coef. 2 Coef. 2 Coef. 2</td>
<td></td>
</tr>
<tr>
<td>(1-m-it)c-it</td>
<td>0.061 4.09 0.093 3.55</td>
<td>0.127 4.76 0.103 3.98</td>
<td>0.089 3.60 0.057 3.34</td>
</tr>
<tr>
<td>Coef. z Coef. t Coef. z Coef. z</td>
<td>Coef. z Coef. t Coef. z Coef. z</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bi-it</td>
<td>-0.013 -3.55 -0.027 -5.05</td>
<td>-0.018 -3.28 -0.031 -4.97</td>
<td>-0.002 -0.19 -0.003 -0.003</td>
</tr>
<tr>
<td>Bi-it.(m-it.c-it)</td>
<td>0.039 3.47</td>
<td>0.035 2.97</td>
<td>0.009 1.02</td>
</tr>
</tbody>
</table>

| Observations | 8,654 | 8,654 | 5,469 | 5,469 | 5,469 | 2,436 | 1,502 | 2,063 |
| Covered Members | 2,213 | 2,213 | 1,473 | 1,473 | 1,473 | 420 | 467 | 505 |
| Covered Nonmembers | 1,402 | 1,402 | 821 | 821 | 821 | 399 | 236 | 315 |
| Covered Nonmembers (%) | 16.20 | 16.20 | 15.05 | 15.05 | 15.05 | 16.38 | 15.71 | 15.27 |
| Log-L | -464.8 | -464.8 | -464.8 | -464.8 | -96.3 | -130.5 | -136.7 |

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

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

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

Upon restricting union membership status changes to those with corresponding job changes the single and multiple indices estimates (in Table 7) indicate negative selection of union members except at the lower end of the observed skill distribution.\footnote{Note that, in the estimates reported in this study the endogeneity correction terms for male employees pertaining to the low predicted skill group generally indicate positive selection though they enter all specifications with statistically insignificant effects.}

The single index results (in Table 7) suggest clear presence of essential heterogeneity and absence of a coverage differential while the appropriate multiple indices estimates indicate union membership premia for the low and middle skill groups in the vicinity of 5.9 and 3.5 percent, respectively.

Additionally accounting for union recognition changes, lowers the low skilled group membership differential to 5.3 while the medium skill differential remains the same (see Table 8). Further, the single index membership premia (in Table 8) are clearly deflated compared to the previous estimates (in Table 7) ignoring recognition changes. Hence, restricting union changes to those with accompanying job changes while ignoring recognition changes overstates the degree of measurement error thus inflating the membership premia. The same pattern generally holds for all remaining single index estimates reported in this study.

In Table 8, as well as the remaining male estimates except in Table 11, there is weak evidence of a coverage differential for male employees located at the lower end of the observed skill distribution. These coverage premia are only statistically significant at the not so stringent 10 percent level noting that the effective union membership differential remains positive though reduced to approximately 2
Since the single indices estimates in Tables (7, 8) are consistent with comparative advantage sorting the uncorrected GLS membership differentials upon additional measurement error reduction are clearly increased, refer to eq. (27), compared to the respective estimate in Table 3. As expected, the FE estimate of the union membership effect is significantly higher in the single index estimates provided in Tables (7, 8) as opposed to Table 3 due to measurement error reduction. Note that the statistically significant impact of the time-variant endogeneity correction terms in the unrestricted sorting estimates (in Table 8) augments the difference between the FE and unrestricted sorting membership premia noting that the same outcome is observed in all estimates accounting for recognition changes in this Section. 45

In effect, the single index results of this Section establish that significant measurement error reduction renders the FE estimate the intermediate bound of the union wage impact, while the uncorrected GLS estimate becomes the lower bound. Therefore, the attenuation bias is indeed exaggerated by the FE estimator.

Regarding the multiple indices estimates (in Tables 7, 8) measurement error reduction raises the lower skill group membership premia since the coverage differentials are reduced compared to Table 3 - refer to eq. (31). Further, the intermediate group membership premia (in Tables 7, 8) are slightly higher than in Table 3 as the average unobserved heterogeneity contribution is positive for all groups- see eq. (27).

Quite importantly, referring once more to eq. (27), the negative differential for highly skilled union members found previously (in Table 3) disappears and becomes statistically insignificant upon additional measurement error reduction since the corresponding estimates (in Tables 7, 8) are consistent with comparative advantage sorting. The same argument applies to the remaining high skill group differentials in Tables (9-12) that are clearly statistically insignificant.

44 The exact union membership differential if $\delta_1 \neq 0, \delta_2 \neq 0$ corresponds to $e^{(\delta_1 - \delta_2)} - 1$. That is, the membership differential for low skilled employees at the not stringent 10% level becomes approximately (2.3, 2.1, 2, 1.9) percent in the estimates given in Tables (8, 9, 10, 12) respectively.

45 In Tables (10, 12) while the contribution of the time-variant endogeneity correction terms is statistically insignificant, their coefficients are markedly higher compared to the respective estimates (in Tables 9, 11) that ignore changes in union recognition.
Therefore, while union members at the upper extreme of the observed skill distribution are negatively selected the sorting mechanism consistent with the estimates indicates that their sorting decision is positively rewarded. That is since the employee’s selection criterion is more likely to be binding at the upper end of the skill distribution (as in Card’s, 1996 two-sided selection model) union members might be more likely to be the worse among their high-skill counterparts (in terms of their endowment of unobserved productivity) but, their sorting decision is compensated.

Finally, we account for an additional potential measurement error source by restricting union status changes to those experiencing employer changes. As predicted by eq.(27) the unrestricted GLS single index membership premia in Tables (9, 10) are higher than in Table 3 since the estimates are consistent with comparative advantage sorting. Moreover, as already discussed, ignoring recognition changes inflates the single index membership premia (in Table 9) with respect to the corresponding estimates (in Table 10).

---


<table>
<thead>
<tr>
<th>Predicted Quantile</th>
<th>GLS</th>
<th>FE</th>
<th>GLS-H</th>
<th>GLS-U</th>
<th>GLS-EM</th>
</tr>
</thead>
<tbody>
<tr>
<td>&lt;=31</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Coef.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>m_{it}</td>
<td>0.057</td>
<td>3.61</td>
<td>0.103</td>
<td>3.15</td>
<td>0.126</td>
</tr>
<tr>
<td>(1-m_{it})_{it}</td>
<td>0.006</td>
<td>0.47</td>
<td>0.004</td>
<td>0.27</td>
<td>-0.004</td>
</tr>
<tr>
<td>B</td>
<td>-0.018</td>
<td>-4.26</td>
<td>-0.027</td>
<td>-5.18</td>
<td></td>
</tr>
<tr>
<td>B_{it}</td>
<td>-0.003</td>
<td>-0.18</td>
<td>-0.009</td>
<td>-0.46</td>
<td></td>
</tr>
<tr>
<td>B_{it} (m_{it}c_{it})</td>
<td>0.034</td>
<td>2.69</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Observations</td>
<td>8,556</td>
<td></td>
<td>5,371</td>
<td></td>
<td>5,371</td>
</tr>
<tr>
<td>Covered Members</td>
<td>2,158</td>
<td>2,158</td>
<td>1,418</td>
<td>1,418</td>
<td></td>
</tr>
<tr>
<td>Covered Members (%)</td>
<td>25.22</td>
<td>25.22</td>
<td>26.40</td>
<td>26.40</td>
<td></td>
</tr>
<tr>
<td>Covered Nonmembers</td>
<td>1,380</td>
<td>1,380</td>
<td>799</td>
<td>799</td>
<td></td>
</tr>
<tr>
<td>Log-L</td>
<td>-277.5</td>
<td>-277.5</td>
<td>-277.5</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

---

Consequent to the presence of essential heterogeneity (in Tables 9, 10) the appropriate multiple indices estimates indicate negative selection of union members except at the lower end of the observed skill distribution.46

The low skill membership premia in Tables (9, 10) rise to approximately 4.9 and 5 percent correspondingly since the coverage premia are reduced with respect to Table 3- see eq.(31). As the middle skill group estimates in Table 9 are consistent with comparative advantage the membership premium rises, compared to Table 3, to approximately 4.2 percent as predicted by eq.(27).

However, the middle skill group premium in Table 10 is reduced (as opposed to Table 3) to around 2.6 percent since there is an increase in the statistically

---

46 The hierarchical sorting estimates for the intermediate group in Table 10 indicate negative selection and an approximately 5.2 percent membership premium though, the endogeneity correction terms are weakly statistically significant.
insignificant coverage premium- refer to eq.(31).

**TABLE 10: WAGE REGRESSIONS (UNION CHANGE IF EMPLOYER CHANGE/ CHANGES IN RECOGNITION, COVERAGE), MALE (1995-2009)**

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>0.056</td>
<td>0.35</td>
<td>0.074</td>
<td>3.07</td>
<td>0.105</td>
<td>3.77</td>
<td>0.080</td>
<td>2.86</td>
<td>0.070</td>
<td>2.63</td>
<td>0.049</td>
<td>2.78</td>
</tr>
<tr>
<td>0.001</td>
<td>0.003</td>
<td>0.18</td>
<td>0.001</td>
<td>0.007</td>
<td>0.029</td>
<td>1.83</td>
<td>0.007</td>
<td>0.73</td>
<td>-0.021</td>
<td>-1.01</td>
<td></td>
</tr>
<tr>
<td>(1-m)_c</td>
<td>0.008</td>
<td>0.68</td>
<td>0.004</td>
<td>0.29</td>
<td>0.001</td>
<td>0.05</td>
<td>0.003</td>
<td>0.18</td>
<td>0.001</td>
<td>0.07</td>
<td>0.029</td>
</tr>
<tr>
<td>B_1</td>
<td>-0.010</td>
<td>-2.71</td>
<td>-0.025</td>
<td>-4.66</td>
<td>-0.031</td>
<td>-4.61</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>B_0</td>
<td>0.001</td>
<td>0.10</td>
<td>-0.022</td>
<td>-1.35</td>
<td>0.008</td>
<td>0.59</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>B_0(m_c)</td>
<td>0.037</td>
<td>3.23</td>
<td>0.060</td>
<td>4.04</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>B_0(m_c)</td>
<td>0.045</td>
<td>1.63</td>
<td>0.012</td>
<td>-0.38</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Observations: 8,701 8,701 5,516 5,516 5,516 2,446 2,628 2,077
Covered Members: 2,227 2,227 1,487 1,487 1,487 420 800 511
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

GLS: Swamy Arora estimator with cluster-robust std. errors; FE: cluster-robust std. errors; Log-L: Reduced Form Log-likelihood
H: hierarchical, U: unrestricted, EH: essential heterogeneity test (endogeneity correction terms not shown)
Predicted Quantile: stratification into quantiles according to a predicted wage in the uncovered sector; appropriate estimates shown only

**TABLE 11: WAGE REGRESSIONS (UNION CHANGE IF EMPLOYER CHANGE, COVERAGE, NO SOCIAL GROUP), MALE (1995-2009)**

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>0.041</td>
<td>2.48</td>
<td>0.100</td>
<td>3.01</td>
<td>0.101</td>
<td>3.59</td>
<td>0.092</td>
<td>3.23</td>
<td>0.084</td>
<td>3.05</td>
<td>0.044</td>
<td>2.52</td>
</tr>
<tr>
<td>0.003</td>
<td>0.21</td>
<td>0.004</td>
<td>0.25</td>
<td>-0.004</td>
<td>-0.27</td>
<td>-0.003</td>
<td>-0.18</td>
<td>-0.005</td>
<td>-0.29</td>
<td>0.024</td>
<td>1.52</td>
</tr>
<tr>
<td>(1-m)_c</td>
<td>0.015</td>
<td>3.40</td>
<td>-0.033</td>
<td>-4.01</td>
<td>-0.020</td>
<td>-3.77</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>B_0</td>
<td>-0.001</td>
<td>-0.05</td>
<td>-0.001</td>
<td>-0.07</td>
<td>-0.020</td>
<td>-2.08</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>B_0(m_c)</td>
<td>0.027</td>
<td>1.87</td>
<td>0.037</td>
<td>2.97</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>B_0(m_c)</td>
<td>0.009</td>
<td>-0.31</td>
<td>0.001</td>
<td>-0.56</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Observations: 8,620 8,620 5,418 5,418 5,418 2,654 1,972 2,040
Covered Members: 2,161 2,161 1,419 1,419 1,419 452 427 475
Covered Members (%): 25.07 25.07 26.96 26.96 26.96 17.03 30.44 24.60
Covered Nonmembers: 1,384 1,384 802 802 802 433 207 307
Covered Nonmembers (%): 15.95 15.95 14.63 14.63 14.63 16.31 15.09 15.06

GLS: Swamy Arora estimator with cluster-robust std. errors; FE: cluster-robust std. errors; Log-L: Reduced Form Log-likelihood
H: hierarchical, U: unrestricted, EH: essential heterogeneity test (endogeneity correction terms not shown)
Predicted Quantile: stratification into quantiles according to a predicted wage in the uncovered sector; appropriate estimates shown only

Socioeconomic group classification is a measure of ability and could contaminate our conclusions concerning the role of unobserved individual heterogeneity. To check the response of the estimated union wage differentials, we exclude the potentially endogenous socioeconomic controls from the previous models restricting union changes to accompanying employer changes.

As expected a priori, inclusion of socioeconomic controls inflates the estimated union membership differentials obtained under both single and multiple indices rules. This is clear comparing the estimates of Table 9 to Table 11 and the estimates of Table 10 to the corresponding estimates of Table 12. Crucially, the middle skill group membership differential becomes statistically insignificant in both set of estimates provided in Tables (11, 12) while the reduced premium obtained upon accounting for recognition changes (in Table 10) becomes local to zero (in Table 12).

Since the estimates in Tables (11, 12) are consistent with hierarchical sorting and negative selection of union members, the unrestricted GLS single index membership premia are reduced (with respect to Table 3) as predicted by eq.(26).

Given the presence of essential heterogeneity, the appropriate multiple indices estimates indicate membership differentials solely for individuals pertaining to low
observed skill group of approximately 4.5 and 4.7 percent in Tables 11 and 12, respectively. Following the exclusion of socioeconomic controls these are deflated compared to the corresponding estimates of 4.9 and 5 percent percent obtained in Tables (9, 10).

Lastly but not least, in our preferred estimates provided in Table 12 there is weak evidence of a coverage premium at the lower end of the observed skill distribution. As already discussed, if one considers that the coverage differential is statistically significant at a sufficiently stringent level, the effective membership premium is consequently reduced to a mere 1.9 percent, approximately.

TABLE 12: WAGE REGRESSIONS (UNION CHANGE IF EMPLOYER CHANGE/CHANGES IN RECOGNITION, COVERAGE, NO SOCIAL GROUP), MALE (1995-2009)

<table>
<thead>
<tr>
<th>Predicted Quantile</th>
<th>GLS</th>
<th>GLS-U</th>
<th>GLS-UE</th>
<th>GLS-UEH</th>
<th>GLS-H</th>
<th>GLS-UH</th>
<th>GLS-UH</th>
</tr>
</thead>
<tbody>
<tr>
<td>Coef. z</td>
<td>Coef. t</td>
<td>Coef. z</td>
<td>Coef. t</td>
<td>Coef. z</td>
<td>Coef. t</td>
<td>Coef. z</td>
<td>Coef. t</td>
</tr>
<tr>
<td>m_{it}</td>
<td>0.041</td>
<td>2.66</td>
<td>0.072</td>
<td>2.95</td>
<td>0.077</td>
<td>2.69</td>
<td>0.064</td>
</tr>
<tr>
<td>(1-m_{it}c_{it}</td>
<td>0.005</td>
<td>0.36</td>
<td>0.004</td>
<td>0.26</td>
<td>0.000</td>
<td>0.00</td>
<td>0.001</td>
</tr>
<tr>
<td>B_{it}</td>
<td>-0.007</td>
<td>-2.03</td>
<td>-0.019</td>
<td>-3.15</td>
<td>-0.001</td>
<td>-2.48</td>
<td>-0.026</td>
</tr>
<tr>
<td>B_{it}(m_{it}c_{it})</td>
<td>0.024</td>
<td>1.90</td>
<td>0.023</td>
<td>1.90</td>
<td>0.006</td>
<td>0.36</td>
<td>-0.023</td>
</tr>
</tbody>
</table>

Observations: 8,766
Covered Members: 2,231
Covered Members (%): 25.45
Covered Nonmembers: 1,392
Covered Nonmembers (%): 15.88
Log L: -550.8

H: Hierarchical, U: unrestricted, EH: essential heterogeneity test (endogeneity correction terms not shown)
GLS: Swamy Aroa estimator with cluster-robust std. errors; FE: cluster-robust std. errors; Log-L: Reduced Form Log-likelihood
Predicted Quantile: stratification into quantiles according to a predicted wage in the uncovered sector; appropriate estimates shown only

8.6. Female Estimates

The most distinguishing feature of the female estimates presented in this Section is the clear evidence of opposing biases occurring at the two extremes of the observed skill distribution. Union members are positively selected among employees located in the lower end of the skill distribution and negatively selected among employees in the upper extreme of the distribution. Following the relevant discussion in Section 8.3, we favour the rationale that the opposing biases reflect the conflicting interests of employers and employees (see Card, 1996). As these opposing effects approximately offset each other, the single index estimates provide generally weak evidence of essential heterogeneity.\(^{47}\),\(^{48}\)

Contrary to the male estimates (in Section 8.5), the positive selection occurring at the lower extreme of the observed skill distribution is statistically significant across the female estimates provided in the present Section. As these two opposing effects offset each other, they drive the FE estimates of the female membership differential towards zero. This outcome contrasts with the male FE estimates of

\(^{47}\) A notable exception are the results in Table 17 where there is strong evidence of essential heterogeneity given the 2.8 percentage discrepancy between the appropriate unrestricted and the essential heterogeneity membership premia.

\(^{48}\) While in Section 8.3 results the impact of the endogeneity correction terms was least in the intermediate skill group estimates, in Section 8.6 there seems to be no endogeneity issue at the middle of the observed skill distribution regarding the first half of the estimates. Similar to Section 8.3 in Table 17, the endogeneity correction terms have their lowest impact at the middle of the skill distribution. Lastly, the estimates in Tables (16, 18) indicate weaker impact of the endogeneity correction terms at the middle of the skill distribution while the significant higher order correction terms suggest departures from normality (possibly due to the heterogeneous nature of the female samples under analysis).
the membership premium which are not far from the appropriate sorting single index endogeneity corrected estimates obtained upon sufficient measurement error reduction (refer to Tables 7-12 in Section 8.5).

Effectively then, regarding all single index estimates in this Section, even upon significant measurement error reduction the FE estimator remains the lower bound of the female union membership effect, the uncorrected GLS estimates lie in the middle and the endogeneity corrected GLS estimate represents the upper bound.

Recall that the estimates presented in this Section are compared against the results obtained using the best possible measure of union membership available in the BHPS dataset (provided in Table 6). Furthermore, regarding the single index estimates in the entire Section two features are common to the respective male results provided in Section 8.5. Firstly, ignoring union recognition changes overstates the degree of measurement error and inflates the single index estimates of the membership premia in all cases. Secondly, the estimates provide no evidence of any coverage premia.

Embarking on a more detailed analysis of the results we begin with the estimates obtained upon restricting union status changes to those with accompanying job changes (ignoring/accounting recognition changes, correspondingly).

Since under the single index assumption the estimates in Tables (13, 14) are consistent with comparative advantage sorting, the uncorrected GLS estimate of the membership premium is significantly higher compared to the respective estimate in Table 6. This can be easily seen by reference to eq. (27) and noting that the coverage premium previously found in the uncorrected GLS single index estimates of Table 6 diminishes and becomes statistically insignificant in Tables (13, 14).

Though the unrestricted and essential heterogeneity single index estimates in both Tables (13, 14) are seemingly comparable the offsetting distinct selection biases obtained under the multiple indices assumption indicate that the inappropriateness of the single index rule. The appropriate multiple indices estimates yield no evidence of coverage differentials and suggest that solely females pertaining to the middle of the observed skill distribution enjoy membership premia of approximately 3.1 and 3.3 percent in Tables (13, 14) respectively.
Note that, the middle skill group premia obtained in Tables (13, 14) are reduced compared to the respective estimate of Table 6. Using eq. (26) and assuming that \( \alpha_1 = \alpha_2 \approx 0 \) due to extensive measurement error reduction, it can be seen that under hierarchical sorting and negative selection the endogeneity corrected estimate is obtained by adding the coefficients of the endogeneity correction terms. Thus, the significant negative selection at the middle of the skill distribution in Table 6, raises the membership premium compared to the respective Table (13, 14) estimates where there is no endogeneity issue at the middle of the skill distribution.

Restricting union status changes to those experiencing employer changes increases the uncorrected GLS membership premia obtained under the single index assumption in Tables (15, 16). Note that while our selected sorting mechanism in Table 15 is the hierarchical one, there is also sufficient evidence in favour of comparative advantage sorting (though the interaction term among the time-invariant endogeneity term and covered membership is not statistically significant). Therefore, as predicted by eq. (27) the uncorrected GLS membership premia in Tables (15, 16) are higher with respect to the corresponding estimates in Tables (13, 14). Using eq. (31) it is clear that significant measurement error reduction raises the uncorrected GLS membership premia in Tables (15, 16) compared to the respective estimate in Table 6 since the coverage premium disappears and the estimates are in line with comparative advantage sorting.

### Table 14: Wage Regressions Union Change if Job Change/Changes in Recognition, Coverage, Female (1995-2009)

<table>
<thead>
<tr>
<th>Predicted Quantile</th>
<th>GLS</th>
<th>FE</th>
<th>GLS-H</th>
<th>GLS-U</th>
<th>GLS-EH</th>
<th>GLS-H</th>
<th>GLS-U</th>
</tr>
</thead>
<tbody>
<tr>
<td>&lt;=39</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>( m_{1,c1} )</td>
<td>0.095</td>
<td>5.30</td>
<td>0.005</td>
<td>0.20</td>
<td>0.133</td>
<td>4.05</td>
<td>0.119</td>
</tr>
<tr>
<td>( [1-m_{1,c1}] )</td>
<td>0.010</td>
<td>0.67</td>
<td>-0.007</td>
<td>-0.35</td>
<td>0.003</td>
<td>0.14</td>
<td>0.003</td>
</tr>
<tr>
<td>( B_{1} )</td>
<td>-0.011</td>
<td>-2.45</td>
<td>-0.025</td>
<td>-3.12</td>
<td>0.020</td>
<td>2.42</td>
<td>-0.045</td>
</tr>
<tr>
<td>( B_{1} (m_{1,c1}) )</td>
<td>-0.006</td>
<td>-0.53</td>
<td>-0.013</td>
<td>-0.83</td>
<td>0.021</td>
<td>0.92</td>
<td>-0.013</td>
</tr>
<tr>
<td>( B_{1} (m_{1,c1}) )</td>
<td>0.032</td>
<td>2.18</td>
<td></td>
<td></td>
<td>0.055</td>
<td>2.67</td>
<td></td>
</tr>
</tbody>
</table>

Observations | 6,314 | 6,314 | 3,309 | 3,309 | 3,309 | 859 | 1,781 | 1,295 |
Covered Members | 1,686 | 1,686 | 971 | 971 | 971 | 137 | 533 | 500 |
Covered Nonmembers (%) | 26.70 | 26.70 | 29.34 | 29.34 | 29.34 | 192 | 470 | 266 |
Covered Nonmembers | 1,475 | 1,475 | 739 | 739 | 739 | 192 | 470 | 266 |
Observations | 6,135 | 6,135 | 3,131 | 3,131 | 3,131 | 834 | 1,763 | 1,194 |
Covered Members | 1,604 | 1,604 | 889 | 889 | 889 | 121 | 519 | 469 |
Covered Nonmembers (%) | 1,434 | 1,434 | 698 | 698 | 698 | 190 | 463 | 240 |

H: Hierarchical, U: unrestricted, EH: essential heterogeneity test (endogeneity correction terms not shown)
GLS: Swamy Arora estimator with cluster-robust std. errors; FE: cluster-robust std. errors; Log-L: Reduced Form Log-likelihood
Predicted Quantile: stratification into quantiles according to a predicted wage in the uncovered sector; appropriate estimates shown only

### Table 15: Wage Regressions Union Change if Employer Change, Coverage, Female (1995-2009)

<table>
<thead>
<tr>
<th>Predicted Quantile</th>
<th>GLS</th>
<th>FE</th>
<th>GLS-H</th>
<th>GLS-U</th>
<th>GLS-EH</th>
<th>GLS-H</th>
<th>GLS-U</th>
</tr>
</thead>
<tbody>
<tr>
<td>&lt;=39</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>( m_{2,c2} )</td>
<td>0.120</td>
<td>5.77</td>
<td>0.007</td>
<td>0.14</td>
<td>0.186</td>
<td>5.36</td>
<td>0.172</td>
</tr>
<tr>
<td>( [1-m_{2,c2}] )</td>
<td>0.010</td>
<td>0.64</td>
<td>-0.006</td>
<td>-0.29</td>
<td>0.005</td>
<td>0.24</td>
<td>0.005</td>
</tr>
<tr>
<td>( B_{2} )</td>
<td>-0.020</td>
<td>-2.91</td>
<td>-0.029</td>
<td>-3.83</td>
<td>0.023</td>
<td>2.37</td>
<td>-0.029</td>
</tr>
<tr>
<td>( B_{2} (m_{2,c2}) )</td>
<td>0.011</td>
<td>0.58</td>
<td>0.006</td>
<td>0.26</td>
<td>0.051</td>
<td>1.68</td>
<td>0.010</td>
</tr>
<tr>
<td>( B_{2} (m_{2,c2}) )</td>
<td>0.035</td>
<td>1.63</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Observations</td>
<td>6,314</td>
<td>6,314</td>
<td>3,309</td>
<td>3,309</td>
<td>3,309</td>
<td>859</td>
<td>1,781</td>
</tr>
</tbody>
</table>
Covered Members | 1,686 | 1,686 | 971 | 971 | 971 | 137 | 533 | 500 |
Covered Nonmembers (%) | 26.70 | 26.70 | 29.34 | 29.34 | 29.34 | 192 | 470 | 266 |
Covered Nonmembers | 1,475 | 1,475 | 739 | 739 | 739 | 192 | 470 | 266 |
Observations | 6,135 | 6,135 | 3,131 | 3,131 | 3,131 | 834 | 1,763 | 1,194 |
Covered Members | 1,604 | 1,604 | 889 | 889 | 889 | 121 | 519 | 469 |
Covered Nonmembers (%) | 1,434 | 1,434 | 698 | 698 | 698 | 190 | 463 | 240 |

H: Hierarchical, U: unrestricted, EH: essential heterogeneity test (endogeneity correction terms not shown)
GLS: Swamy Arora estimator with cluster-robust std. errors; FE: cluster-robust std. errors; Log-L: Reduced Form Log-likelihood
Predicted Quantile: stratification into quantiles according to a predicted wage in the uncovered sector; appropriate estimates shown only
Considering Table 15 the unrestricted and essential heterogeneity single index estimates are nearly identical whereas in Table 16 there is stronger evidence of essential heterogeneity. Nevertheless, as has been already analysed, the offsetting distinct selection biases occurring at the two extremes of the observed skill distribution render the single index estimates unreliable. The appropriate multiple indices estimates in Table 15 indicate a 3.3 membership differential at the middle of the observed skill distribution. As in the estimates provided in Tables (13, 14) using equation (26) it can be seen that the absence of endogeneity at the middle of the skill distribution explains the higher estimated respective premium in Table 6.

However, accounting for changes in union recognition increases markedly the middle skill group membership premium obtained in Table 16. Firstly note the presence of non-normality due to the significant impact of the higher order endogeneity terms which could be attributed to the heterogeneous nature of the female samples that contrary to the male samples are also inclusive of part-time employees. While inclusion of the higher order terms complicates the identification of the sorting mechanism consistent with the estimates, using the statistically significant correction terms it appears that the negative selectivity effects outweigh the positive ones. Bearing this into consideration and using eq.(31) it is straightforward that the increase in the membership premium is owed to the decrease in the coverage premium though it still remains statistically insignificant.

Ignoring the higher order correction terms the middle skill group estimates do not provide any evidence of endogeneity and the corresponding premium is approximately 3.4 percent while the coverage premium is local to zero and statistically insignificant. Once more, the absence of statistically significant negative selection rationalises the lower premium compared to the corresponding Table 6 estimate.

Nevertheless, our preferred estimate for the intermediate skill group premium is the approximately 8 percent differential reported Table 16 for two reasons. Primarily, because the final estimates omitting socioeconomic controls (in Table 18) also provide evidence of departures from normality thus necessitating the inclusion of higher order correction terms. Secondly, since the respective middle group membership premia are closer to the single index essential heterogeneity estimates (in both of Tables 16, 18) in particular when the coverage premia are taken into consideration (raising the effective membership premium to approximately 10.6 percent in Table 16).

<table>
<thead>
<tr>
<th>Predicted Quantile</th>
<th>&lt;= 39</th>
<th>&gt; 39 &lt;= 69</th>
<th>&gt; 69</th>
</tr>
</thead>
<tbody>
<tr>
<td>GLS FE</td>
<td>GLS-H</td>
<td>GLS-U</td>
<td>GLS-U</td>
</tr>
<tr>
<td>Coef. z</td>
<td>0.098</td>
<td>5.31</td>
<td>-0.002</td>
</tr>
<tr>
<td>Coef. t</td>
<td>0.004</td>
<td>0.02</td>
<td>0.001</td>
</tr>
<tr>
<td>(1-mit)</td>
<td>0.018</td>
<td>-0.024</td>
<td>-1.62</td>
</tr>
<tr>
<td>B.</td>
<td>-0.015</td>
<td>-2.81</td>
<td>-0.030</td>
</tr>
<tr>
<td>B.(mit)</td>
<td>0.022</td>
<td>0.07</td>
<td>1.34</td>
</tr>
<tr>
<td>B.(mit)</td>
<td>-0.001</td>
<td>-0.04</td>
<td>-0.003</td>
</tr>
<tr>
<td>B.(mit)</td>
<td>0.042</td>
<td>2.66</td>
<td>-0.014</td>
</tr>
<tr>
<td>B.(mit)</td>
<td>0.002</td>
<td>0.07</td>
<td>0.13</td>
</tr>
<tr>
<td>B.(mit)</td>
<td>-0.003</td>
<td>-1.62</td>
<td>-0.030</td>
</tr>
<tr>
<td>B.(mit)</td>
<td>0.029</td>
<td>1.01</td>
<td>0.001</td>
</tr>
<tr>
<td>B.(mit)</td>
<td>-0.083</td>
<td>-2.73</td>
<td>-0.019</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Observations</th>
<th>6,246</th>
<th>6,246</th>
<th>3,241</th>
<th>3,241</th>
<th>3,241</th>
<th>855</th>
<th>841</th>
<th>1,276</th>
</tr>
</thead>
<tbody>
<tr>
<td>Covered Members</td>
<td>1,654</td>
<td>1,654</td>
<td>939</td>
<td>939</td>
<td>939</td>
<td>135</td>
<td>226</td>
<td>492</td>
</tr>
<tr>
<td>Covered Nonmembers</td>
<td>1,447</td>
<td>1,447</td>
<td>711</td>
<td>711</td>
<td>711</td>
<td>190</td>
<td>213</td>
<td>254</td>
</tr>
<tr>
<td>Covered Nonmembers (%)</td>
<td>37.97</td>
<td>37.97</td>
<td>21.94</td>
<td>21.94</td>
<td>21.94</td>
<td>22.22</td>
<td>25.31</td>
<td>19.93</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>GLS</th>
<th>FE</th>
<th>GLS-H</th>
<th>GLS-U</th>
</tr>
</thead>
<tbody>
<tr>
<td>Reduced Form Log-likelihood</td>
<td>Log-L</td>
<td>-382.0</td>
<td>-382.0</td>
</tr>
</tbody>
</table>

H: hierarchical, U: unrestricted, EH: essential heterogeneity test (endogeneity correction terms not shown)
GLS: Swamy Anora estimator with cluster-robust std. errors; FE: cluster-robust std. errors; Log-L: Reduced Form Log-likelihood
Predicted Quantile: stratification into quantiles according to a predicted wage in the uncovered sector; appropriate estimates shown only
As in the male estimates, exclusion of the potentially endogenous socioeconomic group controls, deflects all of the single index estimates of the union membership premia. This is clearly visible by comparing the respective premia in Tables 15 and 17 and those of Tables 16 and 18. The same occurs regarding the multiple indices estimates with the exception of the intermedia skill group membership premium in Table 17 that has been instead increased. This outcome can be best understood by reference to equation (31). While there is no significant endogeneity impact in the middle skill estimates of Table 15, the corresponding estimates in Table 17 indicate significant negative selection under the hierarchical sorting mechanism. As the exclusion of socioeconomic controls reduces substantially the coverage differential which is additionally statistically significant in Table 17 the union membership premium goes up.

Since under single index sorting the estimates in Tables (17, 18) are consistent with comparative advantage sorting, and as there is no evidence of coverage premium, referring to the appropriate eq.(27) it follows that significant measurement error reduction raises the uncorrected GLS membership premia compared to the respective estimate in Table 6.

### Table 17: Wage Regressions (Union change if Employer change, Coverage, No Social Group), Female (1995-2009)

<table>
<thead>
<tr>
<th>Predicted Quantile</th>
<th>GLS-U</th>
<th>GLS-H</th>
<th>GLS-U</th>
<th>GLS-H</th>
<th>GLS-U</th>
<th>GLS-H</th>
</tr>
</thead>
<tbody>
<tr>
<td>&lt;=40</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>&gt;40 &lt;=68</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>&gt;68</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

As in the male estimates, exclusion of the potentially endogenous socioeconomic group controls, deflects all of the single index estimates of the union membership premia. This is clearly visible by comparing the respective premia in Tables 15 and 17 and those of Tables 16 and 18. The same occurs regarding the multiple indices estimates with the exception of the intermedia skill group membership premium in Table 17 that has been instead increased. This outcome can be best understood by reference to equation (31). While there is no significant endogeneity impact in the middle skill estimates of Table 15, the corresponding estimates in Table 17 indicate significant negative selection under the hierarchical sorting mechanism. As the exclusion of socioeconomic controls reduces substantially the coverage differential which is additionally statistically significant in Table 17 the union membership premium goes up.

### Table 18: Wage Regressions (Union change if Employer change/Changes in Recognition, Coverage, No Social Group) Female (1995-2009)

<table>
<thead>
<tr>
<th>Predicted Quantile</th>
<th>GLS-U</th>
<th>GLS-H</th>
<th>GLS-U</th>
<th>GLS-H</th>
<th>GLS-U</th>
<th>GLS-H</th>
</tr>
</thead>
<tbody>
<tr>
<td>&lt;=39</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>&gt;39 &lt;=68</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>&gt;68</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

As in the male estimates, exclusion of the potentially endogenous socioeconomic group controls, deflects all of the single index estimates of the union membership premia. This is clearly visible by comparing the respective premia in Tables 15 and 17 and those of Tables 16 and 18. The same occurs regarding the multiple indices estimates with the exception of the intermedia skill group membership premium in Table 17 that has been instead increased. This outcome can be best understood by reference to equation (31). While there is no significant endogeneity impact in the middle skill estimates of Table 15, the corresponding estimates in Table 17 indicate significant negative selection under the hierarchical sorting mechanism. As the exclusion of socioeconomic controls reduces substantially the coverage differential which is additionally statistically significant in Table 17 the union membership premium goes up.
The results ignoring changes in recognition (in Table 17) indicate presence of essential heterogeneity and the appropriate multiple indices estimates indicate an approximately 8.2 percent premium solely for union members pertaining to the middle skill group.

As has already been discussed, the observed increase in the middle skill premium is due to a substantial decrease in the coverage differential which is now statistically significant unlike the respective estimates provided in Tables 6 and 15. The middle skill negative coverage premium appears also in Tables 16 and 18 though it is not well defined in terms of statistical significant in particular in the former case. Of course, the important outcome is that there are indications that free-riders are punished instead of receiving coverage premia possibly due to discriminatory behaviour by either unions, employers or both (see Booth and Bryan, 2004).

Finally, turning to our preferred estimates reported in Table 18 the single index results offer little evidence of essential heterogeneity since the appropriate unrestricted sorting and essential heterogeneity estimates are seemingly comparable. However, as this is the product of the offsetting biases occurring at the two extremes of the observed skill distribution one should refer to the multiple indices estimates.

The appropriate multiple indices estimates, in line with the results obtained including socioeconomic controls (in Table 16) indicate presence of non-normality. As in Table 16, the inclusion of the additional higher order endogeneity correction term complicates the identification of the sorting mechanism consistent with the estimates. Nevertheless, using the statistically significant correction terms there seems that negative selection effects prevail. Taking this into consideration and referring to eq. (31), it is clear that extensive measurement error reduction combined with the substantial decrease of the coverage premium produce the increase in the intermediate group membership premium obtained in Table 18. Note that the estimated coverage differential in Table 18 is lower compared to the respective estimates in both Tables 6 and 16 and further its statistical significance has been improved.

Alternatively, ignoring the higher order endogeneity correction terms the appropriate estimates indicate absence of selectivity and yield an approximately 3 percent middle group membership premium while the coverage premium is local to zero and statistically insignificant. Our preferred estimate is the approximately 7.6 membership premium that rises to around 10.5 if the free-rider punishment is considered to be statistically significant at a sufficiently stringent level.

9. SUMMARY AND CONCLUSIONS

This paper studies the union wage impact using the BHPS dataset during 1995-2009. The estimates presented account for the impact of measurement error in union status, consider changes in union recognition, identify both membership and coverage premia, and employ an endogeneity correction methodology that is flexible in the treatment of unobserved heterogeneity.

Furthermore, we test for the presence of essential heterogeneity (in Heckman et al., 2006 parlance) 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 for wage bargaining purposes sector.

Based on the contributions of Aigner (1973) and Vella and Verbeek (1998), the present study contributes to the literature by obtaining the combined analytical
bias expression resulting from endogeneity of union status and non-classical measurement error. It is formally 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.

Therefore, the present study is one of few works in the related literature treating with these issues simultaneously while building upon the influential contributions of Freeman (1984), Robinson (1989a), Card (1996) and Vella and Verbeek (1998).

We obtain robust estimates of union membership wage differentials for males located at the bottom and females at the middle of the corresponding observed skill distributions of approximately 4.7 and 7.6 percent, respectively. Noting that the middle female observed skill group wage cutoff always lies within the bottom observed skill wage interval of the respective male samples, it appears that British unions are solely able to establish positive premia for members pertaining at the lower end of the observed skill distribution.

There is weaker evidence of a 2.6 percent membership differential for males in the intermediate predicted observed skill group that disappears upon omitting the potentially endogenous socioeconomic group controls. The initial estimates indicate negative differentials for union members at the upper extreme of the observed skill distribution implying reduced human capital premia due to wage standardisation. These negative premia disappear upon significant measurement error reduction.

Additionally, we find evidence of a coverage differential for male nonmembers from the bottom of the observed skill distribution suggesting free-riding behaviour and reducing the corresponding membership premium to approximately 1.9 percent, only. Further, there is weak evidence of a negative coverage differential for female nonmembers from the middle of the observed skill distribution effectively acting as a punishment for nonmembers and raising the corresponding membership premium to approximately 10.5 percent.

While union members are negatively selected, individuals from the lower end of the observed skill distribution are positively selected indicating the conflicting interests of employers and employees determining the unobserved differences between individuals located in unionised and non-unionised establishments. The estimates obtained in this study make it clear that, mean regression methodologies ignoring distinct selection biases per skill group and generic fixed effects or instrumental variable methods restricting unobserved heterogeneity to be individual specific and fixed are inappropriate.

Lastly but not least, using the unified expression of the bias resulting from non-classical measurement error and the endogeneity of union status we can obtain a discernible pattern between uncorrected, endogeneity corrected and longitudinal estimates of the union wage effect.

An interesting path of future research would be using matched employer-employee data to study the union wage impact, as well as, the underlying determinants of union membership decisions. At the econometric front, the obvious research direction is the development of fully specified two-step quantile regression models allowing for estimation of longitudinal dynamic models with binary endogenous covariates. In particular, nonlinear panel quantile estimation with predetermined variables still remains a challenge.
REFERENCES


10. 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
O Levels (1987- ) or Certificate of Secondary Education (CSE: 1965-1987)

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

TABLE A2
WAGE REGRESSIONS
(structural form)

DEPENDENT VARIABLE

Gross Average Real Hourly Wage Rate (log of weekly wage divided by usual paid hours including overtime)

EXPLANATORY VARIABLES

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