The Union Membership Wage-Premium Puzzle:

Is there a Free Rider Problem?

Alison L Booth and Mark L Bryan*

ISER
University of Essex
Wivenhoe Park, Colchester CO4 3SQ
United Kingdom

7th June 2001

*Corresponding author
markb@essex.ac.uk
Tel: +44 (0)1206 874683
Fax: +44 (0)1206 873151

Abstract

Economists have, at least since Olson (1965), suggested that there is a free rider problem associated with labour union membership. The reason is that union-set wages are available to all workers covered by unions irrespective of whether or not they are union members, and - given that there are costs to membership – workers will only join if they are coerced or offered incentive excludable goods. And yet empirical research for both the US and for Great Britain has shown that there is a substantial union membership wage premium amongst private sector union-covered workers. An implication is that the free rider hypothesis is therefore irrelevant, since these studies reveal significant economic gains in the form of higher wages for union members. Using rich data from a new linked employer-employee survey for Britain, we show that this is not the case. While estimates assuming exogenous membership do indeed suggest there is a union membership wage premium of a similar order of magnitude to that found in other studies, we demonstrate that - with appropriate instruments based on theory and with additional controls - this wage premium vanishes.

JEL Classification: J51, J52
Keywords: member/non-member covered wage premium, employer-employee data

Acknowledgements: We thank Wiji Arulampalam, René Böheim, Lorenzo Cappellari, John Forth, Marco Francesconi, Andrew Hildreth and David Metcalf for helpful comments, and the Leverhulme Trust for financial support under Award F/00213C "Work-related Training and Wages of Union and Non-union Workers in Britain". Part of this paper was written while Booth was visiting the Australian National University, whose hospitality is gratefully acknowledged. The authors acknowledge the Department of Trade and Industry, the Economic and Social Research Council, the Advisory, Conciliation and Arbitration Service and the Policy Studies Institute as the originators of the 1998 Workplace Employee Relations Survey, and the Data Archive at the University of Essex as the distributor of the data. None of these organisations bears any responsibility for the authors' analysis and interpretation of the data.

Introduction

Economists have, at least since Olson (1965), suggested that there is a free rider problem associated with organisations such as labour unions, since the union-set wage is a public good applying to all workers in the union-covered sector regardless of their individual union status.[1] Given that there are monetary or psychic costs to membership, workers will therefore behave like rational economic agents faced with a public good and take a free ride on union membership – unless they are coerced into joining, or offered incentive excludable goods.

However empirical research for both the US and Great Britain has shown that there is a substantial union membership wage premium amongst private sector union-covered workers (see *inter alia* Blakemore, Hunt and Kiker, 1986; Hildreth, 2000; Budd and Na, 2000). These studies take explicit account of membership endogeneity using a variety of techniques, and find large statistically significant member/non-member wage effects. The conclusion of those studies is that there are substantial economic gains in the form of higher wages for union members. One implication reached by these authors is that a 'rethinking of the "free rider" literature is warranted' (Budd and Na, 2000:804), since the union-set wage appears not to be a collective good.

Using data from a new linked employer-employee survey with a particularly rich set of industrial relations variables, we show that the free rider hypothesis is not dead – at least for Britain. While our estimates assuming exogenous membership reveal that there is indeed

[1] In the US, federal laws stipulate that union-covered members receive the same wage as uncovered members. In other countries, unions typically do not try to negotiate lower wages for covered non-union workers, perhaps because such activity encourages under-cutting by non-union workers. If otherwise identical non-union members are paid less than union members, then ultimately the credibility and survival of the union will be undermined, and the union driven out of existence, as firms substitute the cheaper non-member labour for costlier union members. We return to this issue later.

a union membership wage premium of a similar order of magnitude to that found in other recent studies, we demonstrate that when account is taken of membership endogeneity - with appropriate instruments whose selection is guided by relevant theory - this wage premium vanishes.

The remainder of the paper is set out as follows. In Section I we briefly outline the theoretical and institutional background, while in Section II we describe our data source, the linked employer-employee data from the 1998 Workplace Employee Relations Survey (WERS). The empirical results are presented in Section III, and our conclusions drawn in the final section.

I. Background

Why Join a Union?

The free rider problem is only relevant to labour unions where individuals can exercise choice, that is, where there is no coercion. This is the case in both Britain and a proportion of US states where closed or union shops are illegal. From the late 1980s, legislation in Britain effectively outlawed closed shop arrangements, while in the US right-to-work laws prohibit union shops in some 20 states.[2] In Britain, the proportion of private-sector union covered workers who are also members currently stands at 57%, while in US right-to-work states it stands at 87%.[3]

According to Olson (1965), a reason why workers might join a union in the absence of coercion lies in the fact that unions may offer incentive excludable goods or services to

---

[2] See Boeri, Brugavini and Calmfors (2001) for details of the Europe-wide situation.
[3] The British figure is calculated from WERS 98 data on workplaces where management reports that trade unions are recognised for the purposes of negotiating pay and conditions. The figure for the US is based on Current Population Survey data 1983-93, reported by Budd and Na (2000).

their members. There is considerable evidence that historically friendly society benefits have been important in attracting workers into unions (see for example Boyer, 1988). More recent examples of incentive excludable goods include legal and pensions advice, reputation from complying with a group norm of membership, and grievance and promotions procedures.[4] However, it is not easy to find appropriate proxies for such excludable goods in available data sets, an issue to which we shall return later.

What might explain the member/non-member covered wage premium?

In spite of the free rider problem, there is evidence from some empirical studies of a positive member/non-member wage differential for covered workers. Although Jones (1982) using the National Longitudinal Survey (NLS) finds a very small member/non-member covered worker wage premium, subsequent investigation typically finds quite large effects - see for example the studies using the NLS by Blakemore et al (1986), and Hunt, Kiker and Williams (1987); the more recent studies using the Current Population Survey (CPS) by Hundley (1993), Schumacher (1999), and Budd and Na (2000); and Hildreth (2000) who uses the British Household Panel Survey (BHPS).

Reasons advanced to explain this observed wage premium fall into two broad categories. The first includes selectivity or omitted variable explanations, while the second argues for discriminatory behaviour on the part of the union, the firm, or both.

According to the first broad category, covered and non-covered workers may differ systematically in some unobservable but productivity-augmenting characteristic. To the extent that this is positively correlated with union status and with wages, then the estimated

[4] Booth (1985) and Naylor (1989) focus on social custom sanctioned by loss of reputation for non-complying individuals. Booth and Chatterji (1995) emphasise grievance procedures. Willman (1990) and Booth (1991) look in detail at what unions actually provide.

coefficient to union membership status in a covered wage equation will be upward-biased: there will be a positive selection effect. For example, Budd and Na (2000) suggest that workers who unionise are those who are more motivated or more prepared to stay with that firm and invest in firm-specific human capital. The longer period of human capital investment can be controlled for if tenure is observed, but in so far as motivation is unobserved it will bias the estimated membership coefficient. Another hypothesis is that workers who unionise may be more conscientious and work harder, and consequently both earn higher wages and participate in union governance or purchase union-negotiated life insurance. Note, however, the contrary (and more usual) argument that it is workers of lower unobserved ability who have a greater incentive to combine in order to protect their wages, whereas higher ability workers may do better on individual merit. Though they have opposite implications for the direction of any bias, both arguments highlight the need to control for unobserved heterogeneity.

A related explanation is that only permanent workers will face pressure from union shop stewards to join. Younger temporary or probationary workers will not be targeted in the same way. Therefore there will be positive omitted variable bias to the union status variable in the covered wage equation. This explanation is easier to control for, however, since information about temporary work and job tenure is available in most individual-level surveys. Finally, it may be that non-members are systematically found in firms with weaker unions that are unable to negotiate higher wages, and this might explain any member/non-member covered wage effect. This too is either an omitted variable argument (since individual-level surveys – unlike the linked survey used in our analysis – have no controls for union power) or a selection argument (to the extent that less productive workers are unable to find a job with a strong union firm).

The second broad set of explanations advanced in the empirical literature for the observed positive member/non-member wage differential relates to discriminatory behaviour by unions or employers. For example, Blakemore *et al* (1986) suggest that the wage premium might be the result of cooperation between the union and the firm. The essence of this argument is that a union's cooperative behaviour can increase the size of the surplus to be shared between workers and the firm.[5] Since both parties gain from union cooperation, then the firm will be willing to assist or to turn a blind eye to union's discriminatory behaviour in ensuring that its members are paid more than non-members. For example, the firm might target training programs systematically towards one group, thereby conferring on that group a wage advantage. As another example firms might, in return for union cooperation, attempt to pay non-members less, or if equal pay laws preclude this, pay from a point lower down the union wage scale.

Notice that the conventional explanation of why unions do not want to see members and covered non-members paid different rates – to avoid under-cutting by non-members - disappears in this cooperative scenario. This is because the firm is now unlikely to substitute cheaper non-member labour for more costly union members. If it were to do so, the cooperative behaviour of the union – which is held to increase the overall surplus and reduce labour turnover - would be withdrawn, and the firm would be made worse off.[6]

Of course the puzzle with such explanations for the member/non-member covered wage premium is that they do not make clear why union non-members do not take appropriate action to improve their lot – by for example joining the union to obtain higher

[5] For a summary of the quite rich body of literature arguing conditions under which unions can enhance efficiency, see Freeman and Medoff (1984) and Booth (1995: pp66-71).
[6] The available empirical evidence shows that while union presence reduces labour turnover, the impact of unions on productivity and profitability is typically negative.

wages. Therefore these explanations are logically not very convincing as an explanation of the wage premium.[7] This is especially the case since, in some of the studies outlined above, the wage premium actually increases once appropriate care has been taken of the potential selectivity or omitted variable bias to the union status variable in the covered wage equation. Why non-members in the covered sector abstain from membership then becomes even more of a puzzle, since the economic gains are so large.

Against this background, the purpose of our paper is to exploit data from a new linked employer-employee survey with a particularly rich set of industrial relations variables that allow us to control for training incidence, tenure and union power, as well as to proxy the quality of incentive excludable goods.

To the extent that these proxies explain individual union membership but are uncorrelated with the unexplained component of wages, they offer potential instruments in models that attempt to control for union status endogeneity in covered workers' wage equations. We also address a further possible source of bias in estimates of the membership premium: that coverage may be measured with more error than membership, resulting in a greater downward bias of the coverage coefficient relative to membership coefficient. We now turn to a description of the data source.

II. The Data

We use the new linked employer-employee data from the 1998 Workplace Employee Relations Survey (WERS 98), the first comprehensive survey of its kind for Britain. This is a

[7] Hildreth (2000) acknowledges this point in discussing his 30 percent average member/non-member wage premium, and suggests instead that there remains the selectivity issue.

nationally representative survey of workplaces with ten or more employees, covering the private and public sectors.[8] The sample of workplaces was obtained through a process of stratified random sampling, with over-representation of larger workplaces and some industries necessitating the use of weights in analysis of these data (see Forth and Kirby, 2000, for details).[9]

We use the three linked cross-sectional components of the 1998 survey: the management interview questionnaire, the worker representative interview questionnaire, and the individual self-completion Survey of Employees questionnaire of 25 employees randomly selected at each workplace (or all employees in smaller workplaces). The management interview was carried out face-to-face with the most senior workplace manager responsible for personnel or employee relations (see Cully et al, 1999). Interviews were conducted in 2191 workplaces over the period October 1997 and June 1998, with a response rate of 80.4%. Where relevant and permitted by management, additional interviews were carried out with worker representatives. This occurred in 947 workplaces, representing a response rate of 82% of eligible cases. The Survey of Employees was distributed to the 1880 workplaces where management permitted it, with a response rate of 64% (representing 28237 employees). These data too are weighted in order to account first for the probability of an employee's workplace being selected and secondly for the probability of the employee's own selection (which is greater in smaller workplaces). We use the individual level responses in this paper, to which workplace characteristics have been linked. All our analyses use the employee-level weights.

[8] Although note that it excludes agriculture, fishing, mining, private households with employed persons and extra-territorial organisations.

[9] The oversampled industries are the SIC 92 major groups covering electricity, gas and water, construction, hotels, financial services, and other community services.

Our estimating sub-sample is all private sector men and women who are union-covered and employed in workplaces with at least 25 employees, and with complete information on the variables of interest.[10] WERS 98 contains much more detailed information on the union coverage status of employees' workplaces than is usual in individual-level data. In particular we exploit the responses of management regarding the extent of union involvement in negotiating pay and conditions to define three progressively narrower sets of covered employees. Our broadest definition covers all employees at workplaces where management responded that any trade union (TU) is officially recognised for negotiating pay and conditions for any section of the workforce.[11] The other definitions use management classifications of the way that pay is set for each occupational group in the workplace. Three of the eight possible responses involve collective bargaining (CB).[12] From these we identify employees at workplaces where there is collective bargaining at any level for any occupational group (the second definition) and a subset of employees whose own occupational group is covered by collective bargaining at any level (the third definition). Our broad estimating group comprises 3299 full-time male workers and 1989 full-time and part-time female workers.[13] [14] The second group is substantially smaller (2923 men and 1578 women) and the third group is narrower still (2168 men and 1224 women).

These different sub-samples enable us to address issues of measurement error. It is commonly suggested (see for example Jones, 1982) that in individual–level surveys union

[10] We lose 10% of our employees in the covered sector by dropping these smaller establishments. We also estimated all our models retaining workers in these smaller workplaces, and the main results of the paper are unchanged.
[11] Because we use data from the worker representative questionnaire, these workplaces must also have allowed a worker representative to be questioned.
[12] Respectively at industry, organisation and workplace level. The other categories are: set by management at higher level, set by management at workplace, individual negotiations, some other way (e.g. review body), and none of the preceding.
[13] Full-time workers are defined as those whose usual total weekly hours exceed 30.
[14] The sample of part-time men was unduly small and it is well-known that these individuals tend to be very heterogeneous.

coverage is measured with more error than membership. The reason is that while employees are aware of their own membership status, they are less likely to know the exact role played by the union, especially if they are non-members who play no part in union governance. Therefore the coverage effect may be biased towards zero, giving rise to an apparent premium to membership.

In this paper we argue that management respondents are more likely to have first-hand knowledge of the pay determination process. Basing the coverage indicators on their responses should thus reduce measurement error. Furthermore the detailed responses in the management questionnaire enable us to focus on the employees most directly affected by collective bargaining agreements. For example, our second coverage definition eliminates those workplaces where unions are formally recognised but appear to take no part in negotiating pay (see Millward et al, 2000, chapter 5 for a discussion of this phenomenon). The figures cited above show that this discards about 15% of the total sample. And our third group restricts the sample to those occupational groups actually covered by collective bargaining. These sharper definitions of coverage should reduce the spurious premium from noisy coverage indicators.

The individual survey asks respondents ‘Are you a member of a trade union or staff association?’ and we use responses to this question to construct our individual union membership variable. Table 1 shows the mean membership within each of the three coverage groups by gender and manual/non-manual occupation.

[Table 1 around here]

Table 1 shows how the incidence of coverage decreases as the coverage definition is tightened. This is particularly striking for full-time non-manual men, where only about half of individuals in workplaces covered by collective bargaining are actually covered themselves (more specifically their occupational groups are covered). Membership levels within covered groups are highest for male manual workers where only around 20% of employees appear to free-ride. Although membership density is highest for the most restrictive definition of coverage, there does not appear to be a pattern (perhaps unexpectedly) between the other two groups.

III. The Estimates for Private Sector Men and Women

Let wages in the covered sector be determined by:

$$w_i = \mathbf{x}_i' \boldsymbol{\beta} + \gamma U_i + \varepsilon_i \tag{1}$$

where w_i is the natural logarithm of hourly wages for the ith union-covered worker, $\mathbf{x}_i$ is a vector of exogenous variables determining wages, $\boldsymbol{\beta}$ is the associated parameter vector, U_i is a dummy variable equal to one if individual i is a union member and zero otherwise, and γ is a parameter. Measurement error and unobservable influences on wages are captured by ε_i assumed i.i.d. $N(0, \sigma^2)$, where σ is a parameter. Note that the equation is only specified over the covered sector. Initial estimations indicated there were statistically significant differences between the $\boldsymbol{\beta}$ vectors in the covered and uncovered sectors and this finding is consistent with several other studies, for example Blakemore et al (1986), Budd and Na (2000) and Hildreth (1999).

The vector $\mathbf{x}_i$ contains individual variables assumed to influence human capital formation (including training incidence in the last 12 months, age, tenure, highest educational qualification, race, occupation and marital status) as well as workplace characteristics like size and the proportions of female and part-time workers.[15] We also

[15] There is no information on individuals' work experience.

control for 1-digit industry and region of workplace, as well as imperfect competition in the product market and a low unemployment to vacancy ratio in the labour market. Finally we include dummies indicating the presence of a *de facto* closed shop and multiple trade unions to proxy union power.[16] Table A1 in the Appendix defines the variables and their source.

The dependent variable is the log of hourly wages. The hourly wage is constructed from information on gross earnings and the total number of hours usually worked from the employee questionnaire.[17] The analysis is complicated by the fact that, although the hours data are continuous, gross earnings are banded. Thus we do not observe the hourly wage but only the upper and lower bounds of this wage that are particular to each individual. We follow the approach of Stewart (1983) where, based on a distributional assumption about ε_i, it is possible to write down the contribution of each individual observation to a likelihood function. In the Stata software package the complex survey version of this interval regression command allows each likelihood contribution to be weighted to account for the selection probability. See Appendix B of Forth and Millward (2000) for a log likelihood function for this model.

If the regressors in (1) are uncorrelated with the error term, the estimates of $\boldsymbol{\beta}$ and γ will be unbiased. Table 2 presents estimates of the wage equation under this exogeneity assumption for the broad sample of workers covered by trade unions. The estimated $\boldsymbol{\beta}$ coefficients generally conform to expectations: there are positive returns to age and tenure and non-manual employees enjoy good returns to education. The age profile is less pronounced for women as age is probably a deficient measure of human capital (since it

[16] Although closed shops are illegal in Britain, some managers reported the presence of pre- or post-entry closed shops or that management recommended that workers join the union. We combined responses to these questions into the single '*de facto* closed shop' variable.
[17] Both standard hours and overtime hours are available. The effective number of hours worked was adjusted for respondents who worked paid overtime on the assumption of a 0.5 overtime premium.

misses periods spent out of the labour force). All the subgroups exhibit a positive tenure profile: controlling for other factors, wages are about 10% higher after 10 years in the workplace. Turning to the estimates of the membership premium, the results suggest that a male manual worker who is a union member earns exp(0.096) – 1 = 10% more than a comparable non-member. This coefficient is well defined with a t-statistic of 3.4. Similarly a female non-manual union member earns around 7% more than a comparable non-member (t = 2.9). For the other two subgroups there is no evidence of a membership premium. The positive premia are of similar magnitude to the ordinary least squares (OLS) estimates of other studies. For example, using pooled British data over 1991-94 Hildreth (2000) found a membership premium of 7% for full-time men and women.

[Table 2 around here]

Our results use the full set of control variables, several of which were unavailable in the studies discussed above – for example in Budd and Na (2000) - and which may be correlated with both the wage and membership status according to the theoretical discussion of section I. To investigate whether or not the membership effect is partly explained by omission of these regressors, we re-estimated the equations without tenure and the dummies for temporary or fixed-term contracts, closed shop and multiple trade unions. The resulting estimates of membership premia are shown in table 3, with the previous estimates included for comparison. In all cases the coefficient increases, with a change of approximately 3.5-4 percentage points for the two subgroups with statistically significant premia. Although the increase itself is not statistically significant, it does suggest that estimates which omit these variables are likely to produce upwardly biased membership coefficients.

[Table 3 around here]

Using the full set of controls, we also estimated wage equation (1) over the two more narrowly defined groups of workers, those in workplaces where any occupational group was covered by collective bargaining and those whose own occupational group was covered. We argued earlier that these more precise definitions were likely to give cleaner measures of union influence and therefore a membership coefficient with less upward bias. Column (1) of table 5 shows the estimated membership coefficients for each sub-sample. For manual men the results provide some weak support for this argument: the coefficient falls from 0.096 in the TU recognition sample to 0.072 in the sample covered by any collective bargaining. However, in the narrowest sample the coefficient rises slightly to 0.083. Furthermore these differences are not statistically significant, thus no firm inference can be drawn. Similarly for non-manual women, the coefficient remains at about 7% for all three coverage definitions. There is little evidence that measurement error explains part of the membership premium.

Next we allow for violation of the exogeneity assumption. As discussed above there are several reasons why membership status might be correlated with unobservable determinants of wages. In this case the estimate of γ will be biased. To control for endogeneity we extend the model to incorporate the membership decision as a reduced-form equation:

$$\begin{aligned} U_i^* &= \mathbf{z}_i'\boldsymbol{\theta} + \upsilon_i \\ U_i &= 1 \text{ if } U_i^* > 0 \\ U_i &= 0 \text{ otherwise} \end{aligned} \qquad (2)$$

where U_i* is the latent propensity of individual i to join the union, $\mathbf{z}_i$ is a vector of variables influencing the decision to join, $\boldsymbol{\theta}$ is the associated parameter vector and υ_i is an error term. If υ_i is correlated with ε_i then unobservables which influence the membership decision also influence the wage and U_i is correlated with ε_i. We control for this possibility by a two-step estimation method in which (2) is estimated first and the resulting coefficient vector $\hat{\boldsymbol{\theta}}$ is used to calculate predicted union status for inclusion in the second-stage wage equation. See Maddala (1983) (section 8.8) and Greene (2000) (pp.133-137) for a discussion of this type of model. The method produces consistent but not efficient estimates (which would require joint estimation of (1) and (2)); however, it is more amenable to alternative distributional assumptions. For example, below we assume that υ_i is normally distributed (and estimate (2) as a probit), but we also tested the sensitivity of the results to assuming a logistic distribution in the first stage (our conclusions are not altered).

The second-stage equation is estimated by interval regression:

$$w_i = \mathbf{x}_i'\boldsymbol{\beta} + \gamma\hat{U}_i + \xi_i \tag{3}$$

where ξ_i is an assumed normally distributed error term and, under the normality assumption on υ_i, $\hat{U}_i = \Phi(\mathbf{z}_i'\hat{\boldsymbol{\theta}})$, where $\Phi(.)$ is the standard normal cumulative density function (CDF). When estimating (3) it is necessary to adjust the standard errors to account for the additional sampling error introduced by using an estimate of $\boldsymbol{\theta}$ rather than its true value. The correction was derived by Murphy and Topel (1985) and is illustrated in Greene (2000). Appendix B details the method used.

The equation is identified by the non-linearity of $\hat{U}_i$. However as discussed by Vella (1998) such identification is often weak. To strengthen it we use instruments from the data, which theory suggests affect the membership decision but are unrelated to the unobserved determinants of wages, i.e. they are included in **z** but excluded from **x.** These instruments are proxies for the quality of incentive excludable goods provided by the union. We derive two of them from answers to questions in the employee questionnaire about the individual's opinion of the union's service: whether or not the union takes notice of members' problems and complaints, and whether or not it would best represent the individual in making a complaint about work. (See the variable definitions in Appendix A for the precise form of the questions.) We chose these instruments because they emphasise the quality of the individual's relationship with the union rather than the union's role as a pay negotiator. We rejected other questions that more pointedly refer to the union's influence on pay setting and its relations with management, because the responses are likely to be correlated with wages. Two other instruments are derived from the worker representative questionnaire: a dummy equal to one if the representative uses newsletters or mailings to communicate with members and a dummy equal to one if union meetings are held at least once a month.

Table 4 shows estimates of the probit membership equation (2) and the wage equation (3) for individuals in workplaces with trade union recognition. Measures which might proxy union power are generally good predictors of membership. For example the coefficient on multiple unions is positive (though the relationship is not statistically significant for manual women). For manual men the presence of closed-shop type arrangements is strongly associated with membership ($t = 4.49$), although the relationship is much weaker for the other groups (and actually negative, but statistically insignificant at 5%, for non-manual men). Across all four subgroups the probability of membership monotonically increases with tenure. Turning to the instruments, all except the frequent meeting dummy for manual men have the expected positive sign and several are highly significant . The statistics given in the second last row of table 4 show that in each selection equation the hypothesis of joint insignificance is rejected at better than the 0.01% level by a Wald test.

[Table 4 around here]

The corresponding wages equations estimates show that endogenising the membership status barely changes most of the control variable coefficients. With regard to the membership coefficients, those for non-manual men and manual women remain small and insignificant. However while the coefficient for manual men is similar in magnitude to the exogenous equation, it is now much less precisely estimated (and is not significantly different from zero at the 5% level). For non-manual women, the membership coefficient has now changed sign and become small and statistically insignificant. On the assumption that our instruments are valid, these results cast doubt on the existence of a membership wage premium, particularly for female non-manual workers. To gain some confidence in the instruments, we then conducted an additional investigation in the spirit of the test of overidentifying restrictions used by Ermisch and Francesconi (2000). This involved adding the instruments to the wage equation (relying on identification by the non-linearity of $\hat{U}_i$) and testing for their joint significance. The resulting statistics, given in the bottom row of table 4, suggest that the instruments do not explain any of the unobserved variation in wages and are thus valid.

How robust are our results to estimation using the two other alternative definitions of union coverage? We have already commented on the membership effects obtained from estimating the exogenous wage equations (reported in column (1) of table 5) and we now consider the membership coefficients obtained when equation (3) was estimated over the two other coverage samples. The resulting coefficients - shown in column (2) of table 5 - do not indicate any significant differences across the different coverage sub-samples, as was also found for the exogenous estimates.

[Table 5 around here]

In order to explore further the source of the observed membership premia, we now make explicit use of the linked workplace-employee nature of the data set. Although our *de facto* closed shop and multiple union controls should capture union power, there could still be an unobserved component of workplace-level union influence. A concentration of union members in workplaces where this unobserved effect was strong would give rise to a positive premium. To explore this possibility we therefore modify the model to specify a workplace-specific effect capturing unobserved union power:

$$w_{ij} = \mathbf{x}'_{ij}\boldsymbol{\beta} + \gamma U_{ij} + \phi_j + \varepsilon_{ij} \tag{4}$$

where the subscript *ij* indicates individual *i* in workplace *j* and ϕ_j is the workplace-specific effect. Note that ε_{ij} is assumed to be an idiosyncratic disturbance uncorrelated with membership status, that is we are assuming that any correlation between membership and the unobserved component of wages is at the workplace level rather than the individual level. This assumption is tested below.

This modified framework offers the possibility of new instruments defined as the deviations of variables from their workplace means. For a variable g_{ij}, define this deviation as $\Delta g_{ij} = g_{ij} - \bar{g}_j$, where $\bar{g}_j$ is the mean over all individuals in workplace *j*. The validity of Δg_{ij} as an instrument only requires that it be uncorrelated with the idiosyncratic part ε_{ij} of the composite error, since by construction it is uncorrelated with the workplace-specific effect ϕ_j.[18] We have already implicitly made this assumption for the private good instruments above; using them as deviations simply relaxes the assumption of non-correlation with ϕ_j.[19]

[18] It is then uncorrelated with the composite error since $E[\Delta g_{ij}(\phi_j + \varepsilon_{ij})] = E_j[\phi_j E_i[\Delta g_{ij}]] + E[\Delta g_{ij}\,\varepsilon_{ij}] = 0 + 0 = 0$, where $E_{i(j)}$ is the expectation over i (j). Of course it must also be correlated with membership.

[19] Note that the (within workplace) individual-invariant instruments (use of newsletters or mailings and frequent union meetings) cannot be used.

Therefore the two-step procedure described above can be repeated using these instruments (the correlation between errors implied by the presence of ϕ_j is controlled for by using robust standard errors).

In addition, the assumption embodied in the new model that membership is uncorrelated with ε_{ij} implies that the deviation of membership from its workplace mean (ΔU_{ij}) can be included in the instrument set. However, because the sign of ΔU_{ij} is a perfect predictor of U_{ij}, the first-stage equation cannot be estimated by probit. Instead it is specified as a linear probability model estimated by ordinary least squares (OLS). This is analogous to estimating the first stage of two-stage least squares and gives rise to a membership prediction expressed as a linear combination of exogenous (to the wage equation) variables, which is therefore orthogonal to the wage equation error. An overidentifying restrictions test can be performed (relying on the other instruments for identification) to test the assumption that ΔU_{ij} is uncorrelated with ε_{ij}.

Column (3) of table 5 shows the membership coefficients using the deviations of the individual-varying private good proxies as instruments. They are highly significant in the first-stage estimations (not reported) and overidentifying restrictions tests indicate that they are uncorrelated with the wage equation residuals. For all sub-samples except manual men the point estimates are very close to those in column (2) (though note a more negative, and almost significant, coefficient for manual women whose own occupation group is covered by collective bargaining). For manual men, the point estimates are substantially lower. Since the deviation instruments are uncorrelated by construction with the workplace effect, the drop in the point estimates may indicate a correlation between the instruments used as levels and ϕ_j. If true, this is an example of the low power of the overidentifying restrictions test.

Column (4) of table 5 shows the coefficients estimated using the deviation of membership from the workplace mean in addition to the private goods proxies. The overidentifying restrictions tests (not reported) indicate that ΔU_{ij} is a valid instrument for all subgroups except non-manual women. Given that the predictive power of ΔU_{ij} is so strong, the results deserve particular attention. For male manual workers the membership point estimates are very small. It appears that high paying workplaces (those with strong unions according to our model) are also those with high membership levels. For non-manual men and manual women there is no evidence of a non-zero wage premium, consistent with the previous estimates. For non-manual women the failure of the overidentifying restrictions test indicates that ΔU_{ij} is correlated with the idiosyncratic disturbance ε_{ij}, and therefore we must treat the estimates of a positive membership premium with extreme caution. Instead we prefer the previous estimates of an insignificant premium. The result that ΔU_{ij} is an invalid instrument for non-manual women suggests that the selection process operates at least partly at the individual level.

Conclusion

The free-rider hypothesis rests on the assumption that any union-bargained wage is available to all covered workers whether or not they are members. Empirical studies that have found a substantial membership wage premium imply, on the contrary, that workers have a positive incentive to join the union and therefore the free rider problem disappears. Using a new data source representing British workplaces and their employees, we have shown that union-covered private sector male manual workers and female non-manual workers who are union members do indeed earn 7-9% more than comparable non-members. However, once unobservable influences on wages are controlled for by exploiting the workplace-employee linked nature of the data and the availability of good instruments, we find that the premium is statistically insignificant. For manual men the observed premium appears to be driven by a concentration of members in high-paying workplaces (which we hypothesise is due to unobserved union power). For non-manual women it seems that individuals with unobserved characteristics that increase their wages tend to select into membership. There could also be a workplace effect at work but our methods do not enable us to distinguish it. The results highlight the different possible selection mechanisms and the need to examine workplace level pay determination in more detail. They also suggest that – for Britain at least – it is premature to suggest that ‘a rethinking of the “free rider” literature is warranted’. Moreover the puzzle noted in the literature surveyed at the start of this paper – as to why non-members in the covered sector abstain from membership when the economic gains from membership were so large - is not a puzzle, at least for our data. Individuals are therefore not behaving irrationally in their union membership decisions.

Appendix A

Table A1: Variable definitions and sources

Employee questionnaire	
log (hourly wage)	Upper and lower bounds for each individual are given by log (earnings bound / [standard hours + p*overtime hours]) where earnings bounds define the weekly earnings band selected by the respondent. p=1.5 for individuals paid for overtime and 1 otherwise. Hours data are continuous.
Trade union member	=1 if member of a trade union (TU) or staff association
Fixed term or temporary	=1 if job is temporary or for a fixed term
Tenure dummies	Derived from banded responses. Total years working at this workplace.
Age dummies	Derived from banded responses.
Education dummies	Highest educational qualification. Derived from categorical responses. Omitted category is CSE or equivalent.
Vocational qualification	=1 if holds any recognised vocational qualifications (e.g. NVQ)
Voc qual missing	=1 if information on vocational qualification is missing
Training incidence	=1 if received any employer-financed training away from normal place of work in last 12 months
Black	=1 if black (Caribbean, African or other)
Indian subcontinent	=1 if Indian, Pakistani or Bangladeshi
Non-white	=1 if non-white
Occupational dummies	Derived from categorical responses (1-digit SOC). Omitted category other occupations.
Health problems	=1 if have long-standing health problems or disabilities which limit activity at work, home or in leisure.
Married	=1 if living with spouse or partner
Dependent children	=1 if have dependent children under 5 years
Children info missing	=1 if information on children is missing
Part time	=1 if work less than 30 hours per week
Work mainly women	=1 if respondent's type of work is mainly done by women
Work only women	=1 if respondent's type of work is only done by women
Female work	=1 if respondent's type of work is mainly or only done by women
TU good rep for complaint	=1 if respondent replies "Trade union" to "Ideally, who do you think would best represent you in dealing with managers here if [you] wanted to make a complaint about working here?" Other categories are Myself, Another employee, Somebody else.
TU takes notice	=1 if respondent strongly agrees with "Unions/staff associations here take notice of members' problems and complaints". Other categories are Agree, Neither agree nor disagree, Disagree, Strongly Disagree, Don't know.
Management questionnaire	
Industry dummies	Derived from 1-digit SIC92 codes
Closed shop	=1 if any employees in workplace have to be union members to get or keep their jobs; or if management strongly recommends membership
Multiple unions	=1 if more than one union is recognised by management for negotiating pay and conditions for any section of the workforce in this establishment
UK org size dummies	Derived from size of UK organisation if workplace part of larger organisation
UK org size missing	=1 if UK org size missing and workplace part of larger organisation
Imperfect competition	=1 if 5 or fewer competitors or organisation dominates product market
Domestic market	=1 if market for main product is local, regional or national
Employee profile questionnaire – completed by management respondent before interview	
Size dummies	Derived from total number of employees at workplace
Proportion part-time	Proportion of employees working less than 30 hours per week at workplace
Proportion female	Proportion of female employees at workplace
Proportion manual	Proportion of employees in manual occupations at workplace
Worker representative questionnaire	

Regular TU meetings	=1 if general union meetings held at least once a month. Question: "In the last 12 months how often have you or other representatives…called a general meeting with the employees that you represent at this workplace?"
TU newsletters / mailings	= 1 if representative uses newsletters / mailings to communicate with employees represented.
Additional data	
Tight labour market	=1 if the unemployment to vacancy rate ratio in travel-to-work-area ≤ 3
Regional dummies	Derived from Government Office Regions

Appendix B

Adjustment of standard errors in two-step estimation

Let the log likelihood function for each step of the estimation be given by:

$$L_w = \sum_{i=1}^{N} l_{iw}(\boldsymbol{\beta}, \gamma \mid w_i^U, w_i^L, \mathbf{x}_i, f(\mathbf{z}_i, \hat{\boldsymbol{\theta}})) \qquad (B1)$$

$$L_u = \sum_{i=1}^{N} l_{iu}(\boldsymbol{\theta} \mid U_i, \mathbf{z}_i) \qquad (B2)$$

where L_w is the log likelihood of the second-step wage equation and L_u is the log likelihood of the first-step union membership equation. l_{iw} and l_{iu} are the corresponding log likelihood contributions of individual i in the sample of N individuals, w_i^U and w_i^L are respectively the observed upper and lower bounds of individual i's unobserved log hourly wage, and f is a function which generates an additional regressor from the first stage estimates. Other symbols are as defined in the text. Note that it is not necessary to specify a joint distribution for w_i^U, w_i^L and U_i.

Murphy and Topel (1985) show that the asymptotic covariance matrix based on estimating (B1) must be corrected to account for the use of $\hat{\boldsymbol{\theta}}$ rather than $\boldsymbol{\theta}$, as follows (notation as in Greene (2000)):

$$\mathbf{V}_w^{adj} = \mathbf{V}_w + \mathbf{V}_w[\mathbf{C}\mathbf{V}_u\mathbf{C}' - \mathbf{R}\mathbf{V}_u\mathbf{C}' - \mathbf{C}\mathbf{V}_u\mathbf{R}']\mathbf{V}_w$$

where $\mathbf{V}_w$ and $\mathbf{V}_u$ are the asymptotic covariance matrices based on (B1) and (B2) respectively and $\mathbf{C}$ and $\mathbf{R}$ are given by:

$$\mathbf{C} = E\left[\left(\frac{\partial L_w}{\partial \boldsymbol{\beta}^*}\right)\left(\frac{\partial L_w}{\partial \boldsymbol{\theta}'}\right)\right], \quad \mathbf{R} = E\left[\left(\frac{\partial L_w}{\partial \boldsymbol{\beta}^*}\right)\left(\frac{\partial L_u}{\partial \boldsymbol{\theta}'}\right)\right]$$

where $\boldsymbol{\beta}^* = (\boldsymbol{\beta}', \gamma)'$.

Estimates of $\mathbf{V}_w$ and $\mathbf{V}_u$ are given by the software and estimates of $\mathbf{C}$ and $\mathbf{R}$ can be calculated as:

$$\hat{\mathbf{C}} = \sum_{i=1}^{N}\left(\frac{\partial l_{iw}}{\partial \hat{\boldsymbol{\beta}}^*}\right)\left(\frac{\partial l_{iw}}{\partial \hat{\boldsymbol{\theta}}'}\right), \quad \hat{\mathbf{R}} = \sum_{i=1}^{N}\left(\frac{\partial l_{iw}}{\partial \hat{\boldsymbol{\beta}}^*}\right)\left(\frac{\partial l_{iu}}{\partial \hat{\boldsymbol{\theta}}'}\right)$$

Note that factors of 1/N have been omitted from these summations since they ultimately cancel with the terms of N implicit in the estimates of $\mathbf{V}_w$ and $\mathbf{V}_u$. The terms in the expression for $\hat{\mathbf{R}}$ are the score vectors which can easily be obtained from the software. In the expression for $\hat{\mathbf{C}}$ the first term is again a score vector while the second term can be expressed in terms of this score vector, $\mathbf{x}_i$, $f(\mathbf{z}_i, \hat{\boldsymbol{\theta}})$, $\mathbf{z}_i$, $\hat{\gamma}$ and $\hat{\boldsymbol{\theta}}$ by using the chain rule.

Note that if the first-stage is estimated by OLS an analogous expression involves the product of the residuals and $\mathbf{z}_i$ instead of the score vector from the first-stage.

In our data, an additional complication is the use of weights and the non-independence (clustering) of errors for observations within the same workplace. Each of the *N* terms in the above summations was therefore weighted appropriately, and the covariance matrices based on weighted hessians (produced by the software) were used for $\mathbf{V}_w$ and $\mathbf{V}_u$. The resulting standard errors were then adjusted to account for clustering by applying a correction factor equal to the ratio of the robust (i.e. accounting for dependent errors) to non-robust standard error obtained using conventional estimation (i.e. without applying any two-step correction).

As a partial check on whether any bias was introduced by these final adjustment procedures, all equations were re-estimated without weights or clustering. The two-step standard error corrections were very similar and none of our conclusions is altered.

References

Blakemore, Arthur E, Janet C Hunt, and B.F. Kiker (1986), 'Collective Bargaining and Union Membership Effects on the Wages of Male Youths,' *Journal of Labor Economics*, 4, April, 193-211.

Boeri, Tito, Agar Brugavini and Lars Calmfors (2001) (eds.) *The Role of the Unions in the Twenty-first Century*, forthcoming Oxford University Press.

Booth, Alison L. (1985), 'The Free Rider Problem and a Social Custom Model of Trade Union Membership,' *Quarterly Journal of Economics*, 100, February, 253-261

Booth, Alison L. (1995), *The Economics of the Trade Union*. Cambridge: Cambridge University Press.

Booth, Alison L. and Monojit Chatterji, (1995) "Union Membership and Wage Bargaining when Membership is not Compulsory" *The Economic Journal*, March, 345-360.

Boyer George R (1988) 'What Did Unions Do in Nineteenth Century Britain?' *Journal of Economic History* 48(2), 319-332.

Budd, John W and In-Gang Na, (2000) 'The Union Membership Wage Premium for Employees Covered by Collective Bargaining Agreements', *Journal of Labor Economics*,18(4), 783-807.

Cully, Mark, Stephen Woodland, Andrew O'Reilly and Gill Dix (1999) *Britain at Work*, London: Routledge.

Ermisch John and Marco Francesconi (2000), 'Educational Choice, Families and Young People's Earnings', *Journal of Human Resources* 35(1), 143-76.

Forth, John and Simon Kirby (2000) *Guide to the Analysis of the Workplace Employee Relations Survey 1998*,London: NIESR, WERS98 Data Dissemination Service (http://www.niesr.ac.UK org /niesr/wers98).

Forth, John and Neil Millward (2000) 'The Determinants of Pay Levels and Fringe Benefit Provision in Britain,' Discussion Paper No. 171, National Institute for Economic and Social Research, London.

Freeman, Richard B and James L Medoff (1984) *What Do Unions Do?* New York: Basic Books.

Greene W. H. (2000), *Econometric Analysis*, 4th Edition, Prentice-Hall International, Inc.

Hildreth, Andrew K.G. (1999), 'What Has Happened to the Union Wage Differential in Britain in the 1990s?', *Oxford Bulletin of Economics and Statistics*, 61, 1, 5-31.

Hildreth, Andrew K.G. (2000), 'Union Wage Differentials for Covered members and Nonmembers in Great Britain' *Journal of Labor Research* 21, Winter, 133-47

Hunt Janet C, BF Kiker and Glyn C Williams (1987) ‘The Effect of Union Type on Member-Nonmember Wage Differentials’, *Journal of Labor Research* 8, 59-65, Winter.

Hundley, G. (1993), ‘Collective Bargaining Coverage of Union Members and Nonmembers in the Public Sector’, *Industrial Relations*, vol. 32, no.1, 72-93

Jones, E. (1982), ‘Union/Nonunion Differentials: Membership or Coverage?’, *Journal of Human Resources*, XVII, 2, 276-285.

Maddala G. S. (1983), *Limited-Dependent and Qualitative Variables in Econometrics*, *Econometric Society Monographs no.3*, Cambridge University Press.

Metcalf, David, Kirstine Hansen and Andy Charlwood (2000) ‘Unions and the Sword of Justice: Unions and Pay Systems, Pay Inequalities, Pay Discrimination and Low Pay’, Discussion paper no. DP0452, Centre for Economic Performance, April.

Millward, Neil., Alex Bryson and John Forth (2000), *All Change at Work*, London: Routledge

Murphy, K. M. and R. H. Topel (1985), ‘Estimation and Inference in Two-Step Econometric Models’, *Journal of Business and Economic Statistics*, vol. 3, no. 4, 370-379.

Naylor, Robin (1989) ‘Strikes, Free Riders and Social Customs,’ *Quarterly Journal of Economics*, 104(4), pp.771-786.

Olson, Mancur, Jr. (1965) *The Logic of Collective Action*, Cambridge Massachusetts: Harvard University Press.

Stewart, Mark B. (1983), ‘On Least Squares Estimation when the Dependent Variable is Grouped’, *Review of Economic Studies*, 50(4), 737-753

Schumacher, E. J. (1999), ‘What Explains Wage Differences Between Union Members and Covered Nonmembers’, *Southern Economic Journal*, 65(3), 493-512

Vella, Francis (1998), ‘Estimating Models with Sample Selection Bias: A Survey’, *Journal of Human Resources*, 33(1), 127-169

Willman, Paul (1990) ‘The Financial Status and Performance of British Trade Unions, 1950-1988’, *British Journal of Industrial Relations*, 28 (3), 313-327.

Table 1: Union coverage and membership density within covered groups

		(1) TU recognition		(2) CB (any occ group)		(3) CB (own occ group)	
	N	Covered	Member	Covered	Member	Covered	Member
Male, manual (FT)	3859	0.63	0.79	0.55	0.82	0.46	0.83
Male, non-manual (FT)	3920	0.49	0.49	0.45	0.48	0.24	0.61
Female, manual (FT/PT)	3051	0.43	0.48	0.34	0.51	0.26	0.53
Female, non-manual (FT/PT)	4143	0.42	0.42	0.38	0.39	0.26	0.49

Notes: number of observations N is unweighted but coverage and membership means are weighted by selection probability. Membership mean is density within covered group, e.g. 79% of manual men covered by TU recognition are members. Column (1) is employees in workplaces where union is recognised for negotiating pay and conditions; column (2) is employees in workplaces where collective bargaining is practised at any level for any occupational group; column (3) is employees in workplaces where collective bargaining is practised at any level for employee's own occupational group. Estimation samples are smaller than implied by coverage means due to missing observations.

Table 2: Wage equations with exogenous membership - employees in workplaces with TU recognition.

	Male manual		Male non-manual		Female manual		Female non-man	
	Mean	Coeff	Mean	Coeff	Mean	Coeff	Mean	Coeff
log (hourly wage)	1.789		2.284		1.377		1.943	
Trade union member	0.802	0.096**	0.561	-0.019	0.488	0.009	0.484	0.071**
		(0.028)		(0.022)		(0.025)		(0.025)
Closed shop	0.113	0.064	0.066	-0.027	0.08	-0.04	0.064	0.051
		(0.049)		(0.044)		(0.044)		(0.050)
Multiple unions	0.599	0.142**	0.669	0.082*	0.254	0.08	0.493	0.071*
		(0.033)		(0.032)		(0.048)		(0.031)
Fixed term or temporary	0.033	-0.023	0.037	-0.054	0.045	0.015	0.043	-0.121
		(0.056)		(0.069)		(0.056)		(0.067)
Tenure 1-2 yr	0.066	-0.003	0.096	0.021	0.125	0.039	0.124	0.073
		(0.043)		(0.047)		(0.043)		(0.056)
Tenure 2-5 yr	0.165	0.093*	0.214	0.051	0.258	0.122**	0.24	0.05
		(0.038)		(0.033)		(0.036)		(0.041)
Tenure 5-10 yr	0.224	0.076	0.208	0.068*	0.213	0.122**	0.236	0.073
		(0.039)		(0.030)		(0.036)		(0.038)
Tenure >10 yr	0.459	0.092*	0.367	0.091**	0.244	0.097*	0.265	0.115**
		(0.037)		(0.034)		(0.046)		(0.037)
Age 20-24	0.042	0.327**	0.05	0.059	0.08	0.021	0.079	0.089
		(0.110)		(0.118)		(0.054)		(0.080)
Age 25-29	0.099	0.457**	0.13	0.192	0.11	0.143*	0.205	0.138**
		(0.095)		(0.119)		(0.058)		(0.068)
Age 30-39	0.289	0.472**	0.315	0.371**	0.275	0.132*	0.316	0.267**
		(0.092)		(0.121)		(0.056)		(0.067)
Age 40-49	0.274	0.419**	0.315	0.408**	0.237	0.177**	0.226	0.276**
		(0.094)		(0.122)		(0.057)		(0.068)
Age 50-59	0.236	0.427**	0.167	0.532**	0.19	0.136*	0.15	0.273**
		(0.095)		(0.121)		(0.063)		(0.066)
Age ≥ 60	0.041	0.341**	0.015	0.662**	0.028	0.226**	0.009	0.200
		(0.105)		(0.150)		(0.084)		(0.106)
O-level or equivalent	0.254	0.046*	0.21	0.102**	0.258	0.079**	0.401	0.092**
		(0.019)		(0.032)		(0.029)		(0.032)
A-level or equivalent	0.085	0.064*	0.23	0.146**	0.094	0.086*	0.197	0.158**
		(0.031)		(0.038)		(0.041)		(0.033)
First degree or equivalent			0.328	0.228**			0.154	0.236**
				(0.036)				(0.049)
Postgrad degree or equiv			0.084	0.24**			0.061	0.341**
				(0.052)				(0.114)
First or postgrad degree	0.025	0.128**			0.03	0.106		
		(0.043)				(0.082)		
Vocational qualification	0.49	0.035	0.498	-0.049	0.204	-0.01	0.3	-0.015
		(0.019)		(0.026)		(0.029)		(0.020)
Constant		1.121**		1.398**		1.361**		1.206**
		(0.175)		(0.136)		(0.131)		(0.106)
Observations		1834		1465		779		1210
Pseudo R2		0.16		0.2		0.24		0.19

Notes: The dependent variables are the upper and lower bounds of the log (hourly wage). Estimation is by interval regression with weighting for selection probability and robust standard errors to allow for correlation of errors within workplaces. Means are unweighted. Mean of log (hourly wage) is calculated from midpoints of bands. Asymptotic standard errors are in parentheses. Other controls are 1-digit occupation and industry, workplace size, region, training incidence, race, presence of health problems, marital status, dependent children, UK organisation size if workplace not independent, work done mainly/only by women, proportions of part-time, female and manual workers, indicators of imperfect product market competition, domestic product market and tight labour market, and part-time status. * significant at 5% level; ** significant at 1% level

Table 3: Membership premium with full and reduced set of controls

	Male manual	Male non-man	Female man	Female non-man
Full controls	0.096**	-0.019	0.009	0.071**
	(0.028)	(0.022)	(0.025)	(0.025)
Reduced controls	0.135**	0.003	0.027	0.096**
	(0.028)	(0.018)	(0.026)	(0.025)
Joint significance of excluded variables χ^2 (7)	38.5 [0.000]	49.0 [0.000]	36.5 [0.000]	15.2 [0.034]

Notes: samples are identical to table 2. Excluded controls are tenure, closed shop (or strong management recommendation to join union), multiple trade unions, and fixed or temporary contract. Asymptotic standard errors in parentheses. p-values in square brackets. * significant at 5% level; ** significant at 1% level

Table 4: Membership equations and wage equations with endogenous membership - employees in workplaces with TU recognition

	Male manual		Male non-manual		Female manual		Female non-man	
	Member	Wages	Member	Wages	Member	Wages	Member	Wages
Trade union member		0.091		0.008		0.017		-0.035
		(0.066)		(0.057)		(0.065)		(0.065)
Closed shop	1.055**	0.064	-0.539	-0.024	0.639	-0.042	0.175	0.056
	(0.235)	(0.049)	(0.301)	(0.036)	(0.344)	(0.045)	(0.330)	(0.053)
Multiple unions	0.354**	0.143**	0.504**	0.078**	0.185	0.079	0.444**	0.081*
	(0.137)	(0.034)	(0.154)	(0.027)	(0.223)	(0.047)	(0.153)	(0.033)
Fixed term or temporary	-0.593*	-0.023	-0.430	-0.050	-0.258	0.016	-0.435	-0.139*
	(0.252)	(0.056)	(0.285)	(0.057)	(0.369)	(0.057)	(0.282)	(0.067)
Tenure 1-2 yr	0.344	-0.003	0.248	0.019	-0.336	0.040	0.508*	0.084
	(0.229)	(0.043)	(0.246)	(0.039)	(0.196)	(0.043)	(0.199)	(0.058)
Tenure 2-5 yr	0.573**	0.094*	0.742**	0.046	0.163	0.121**	0.756**	0.068
	(0.176)	(0.039)	(0.208)	(0.035)	(0.201)	(0.038)	(0.183)	(0.043)
Tenure 5-10 yr	0.665**	0.077	0.866**	0.061	0.577**	0.120**	0.991**	0.098*
	(0.189)	(0.041)	(0.242)	(0.036)	(0.220)	(0.039)	(0.192)	(0.041)
Tenure >10 yr	0.742**	0.093*	1.351**	0.081*	0.932**	0.095	1.250**	0.148**
	(0.206)	(0.043)	(0.262)	(0.040)	(0.227)	(0.052)	(0.177)	(0.045)
Age 20-24	-0.561	0.325**	-0.100	0.061	-0.615	0.022	0.369	0.092
	(0.505)	(0.106)	(0.526)	(0.122)	(0.362)	(0.052)	(0.769)	(0.078)
Age 25-29	-0.475	0.455**	-0.021	0.193	0.046	0.143*	0.393	0.138
	(0.420)	(0.091)	(0.525)	(0.119)	(0.307)	(0.058)	(0.715)	(0.072)
Age 30-39	-0.235	0.470**	-0.099	0.372**	-0.157	0.133*	0.697	0.275**
	(0.420)	(0.088)	(0.521)	(0.120)	(0.327)	(0.056)	(0.690)	(0.070)
Age 40-49	-0.317	0.417**	0.148	0.407**	-0.119	0.178**	0.697	0.284**
	(0.428)	(0.090)	(0.529)	(0.120)	(0.350)	(0.057)	(0.719)	(0.070)
Age 50-59	-0.520	0.424**	0.241	0.531**	0.184	0.136*	0.611	0.279**
	(0.441)	(0.093)	(0.524)	(0.121)	(0.400)	(0.063)	(0.719)	(0.068)
Age ≥ 60	-0.655	0.338**	-1.275	0.671**	-0.430	0.227**	0.813	0.217*
	(0.585)	(0.103)	(0.817)	(0.131)	(0.453)	(0.083)	(0.813)	(0.096)
O-level or equivalent	-0.099	0.046*	-0.416*	0.105**	-0.297	0.079**	0.168	0.094**
	(0.148)	(0.020)	(0.173)	(0.028)	(0.180)	(0.029)	(0.168)	(0.033)
A-level or equivalent	0.216	0.064*	-0.468**	0.149**	-0.426	0.086*	0.321	0.165**
	(0.212)	(0.031)	(0.170)	(0.029)	(0.248)	(0.041)	(0.231)	(0.035)
First degree or equivalent			-0.822**	0.233**			0.295	0.244**
			(0.209)	(0.031)			(0.257)	(0.052)
Postgrad degree or equiv			-0.525*	0.245**			0.304	0.351**
			(0.259)	(0.043)			(0.269)	(0.117)
First or postgrad degree	-0.409	0.127**			-0.316	0.107		
	(0.261)	(0.045)			(0.381)	(0.084)		
Vocational qualification	-0.131	0.035	0.288*	-0.051**	-0.216	-0.009	-0.337**	-0.026
	(0.106)	(0.019)	(0.121)	(0.018)	(0.152)	(0.030)	(0.136)	(0.022)
TU good rep for complaint	1.321**		1.076**		1.182**		1.210**	
	(0.116)		(0.129)		(0.128)		(0.125)	
TU takes notice	0.845**		0.426		0.680**		0.314**	
	(0.225)		(0.246)		(0.245)		(0.228)	
Regular TU meetings	-0.053				0.690			
	(0.212)				(0.431)			
TU newsletters/mailings			0.297*				0.244	
			(0.138)				(0.152)	
Constant	-0.314	1.124	-2.439**	1.398**	-1.126	1.358	-2.565**	1.205**
	(0.551)	(0.175)	(0.691)	(0.147)	(0.651)	(0.134)	(0.827)	(0.109)
Observations	1834	1834	1465	1465	779	779	1210	1210

Pseudo R2	0.41	0.15	0.36	0.21	0.42	0.24	0.35	0.19
Exc. Instr Wald Chi2(3)	133.9		86.5		92.8		101.7	
	[0.000]		[0.000]		[0.000]		[0.000]	
OverID Wald Chi2(3)		2.81		3.05		2.89		0.56
		[0.423]		[0.384]		[0.409]		[0.905]

Notes: Dependent variables are membership dummy in membership equation and log (hourly wage) in wage equation. 2-step estimation is by probit (first stage) and interval regression (second stage) with weighting for selection probability and robust standard errors to allow for correlation of errors within workplaces. Asymptotic standard errors are in parentheses, p-values in square brackets. Other controls are 1-digit occupation and industry, workplace size, region, training incidence, race, presence of health problems, marital status, dependent children, UK organisation size if workplace not independent, work done mainly/only by women, proportions of part-time, female and manual workers, indicators of imperfect product market competition, domestic product market and tight labour market, and part-time status. * significant at 5% level; ** significant at 1% level

Table 5: Estimated membership coefficients

Sub-sample	Sample size	(1) Exogenous	(2) Endogenous	(3) Endogenous	(4) Endogenous
Male manual					
TU recognition	1834	0.096**	0.091	0.050	0.031
		(0.028)	(0.066)	(0.076)	(0.035)
CB (any occupational group)	1597	0.072**	0.070	0.027	0.001
		(0.025)	(0.064)	(0.068)	(0.030)
CB (own occupational group)	1363	0.083**	0.106	0.095	0.029
		(0.028)	(0.078)	(0.080)	(0.032)
Male non-manual					
TU recognition	1465	-0.019	0.008	0.008	-0.023
		(0.022)	(0.058)	(0.080)	(0.026)
CB (any occupational group)	1326	-0.024	0.011	0.011	-0.022
		(0.024)	(0.086)	(0.092)	(0.029)
CB (own occupational group)	805	0.026	0.081	0.081	0.035
		(0.026)	(0.080)	(0.080)	(0.035)
Female manual					
TU recognition	779	0.009	0.017	0.017	-0.012
		(0.025)	(0.066)	(0.065)	(0.029)
CB (any occupational group)	569	0.015	-0.015	-0.046	-0.029
		(0.031)	(0.068)	(0.063)	(0.030)
CB (own occupational group)	441	0.022	-0.075	-0.127	-0.014
		(0.035)	(0.061)	(0.065)	(0.033)
Female non-manual					
TU recognition	1210	0.071**	-0.035	-0.030	0.061*
		(0.025)	(0.064)	(0.077)	(0.030)
CB (any occupational group)	1009	0.078**	-0.091	-0.090	0.067
		(0.028)	(0.081)	(0.091)	(0.034)
CB (own occupational group)	783	0.070**	-0.095	-0.064	0.056
		(0.027)	(0.083)	(0.089)	(0.031)

Notes:
Column (1): membership is assumed exogenous.
Column (2): instruments for membership are private good proxies.
Column (3): instruments for membership are deviations from workplace means of private good proxies (excluding communication by mailings/newsletters and frequent union meetings).
Column (4): instruments for membership are private good proxies and deviation from workplace mean membership.
Asymptotic standard errors in parentheses. * significant at 5% level; ** significant at 1% level.