'The Size of the Union Membership Wage Premium in Britain's Private Sector'

Alex Bryson[†]

Abstract

The paper estimates the union wage premium in Britain's private sector in 1998, after nearly two decades of union decline. It examines the performance of the linear estimator alongside a semi-parametric technique (propensity score matching (PSM)) – hitherto unused in the wage premium literature - which shares the same identifying assumption, namely that selection into membership is captured with observable data. Results using the two techniques are compared, and reasons for differences in results are identified and discussed. By altering the information set entering estimation the paper shows the sensitivity of OLS and PSM results to data quality.

Key words: trade unions, wage premium, treatment effect, matching, propensity score

JEL classification: C14, C81, J31, J51

---

† Policy Studies Institute, 100 Park Village East, London NW1 3SR and Research Associate at the Centre for Economic Performance, London School of Economics; e-mail: a.bryson@psi.org.uk. This paper was funded under the Leverhulme Trust's 'Future of Trade Unions in Modern Britain' Programme and the University of Westminster's Regent Street Polytechnic Trust. I thank participants at a seminar at the Università Cattolica, Milan (February 2002), the Conference on Employee and Employee Data, Aarhus School of Business (Feb. 28th – March 1st 2002), and the Third WERS98 Users Group Seminar at NIESR, London (November 2001), and Professor Jeff Smith, University of Maryland, for useful comments on earlier drafts of this paper. The author acknowledges 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 data, and the Data Archive at the University of Essex as the distributor of the data. None of these organizations or individuals bears any responsibility for the author's analysis and interpretations of the data.

I. Introduction

This paper addresses the question: how much of the wage differential between union members and non-members is attributable to union membership, and how much is due to differences in personal, job and workplace characteristics across members and non-members? The question is prompted by two recent developments, one substantive, and one methodological.

The substantive development is the recent decline in the union membership premium. Studies for the United States and Britain have traditionally found union members' earnings to be 10-20% higher than non-members'. However, for the US, Hirsch and Schumacher (2001) and Hirsch and Macpherson (2000) find the wage premium has been declining in the private sector for some time. Blanchflower and Bryson (2003) show the premium has declined in both countries since the mid-1990s. Machin's (2001) analysis of longitudinal data from the British Household Panel Survey (BHPS) indicates that, although there was a wage gain for people moving into union jobs in the early 1990s, this had disappeared by the late 1990s. This paper contributes to the body of knowledge about the size of the union membership premium in Britain by the late 1990s.

The methodological development is the advent of new data and relatively new estimation techniques permitting a fresh look at the nature of the union wage premium. Over the last quarter century, there have been over 30 studies estimating the union wage premium in Britain. The majority of these capture the union effect through a union membership dummy using simple linear regression techniques. Increasing concern about the possible endogeneity of the union dummy due to the non-random nature of membership, and a belief that the selection processes governing membership are usually unobservable to the analyst, has spawned a growing literature that utilises structural models which explicitly tackle unobserved heterogeneity through methods adjusting for selection bias. These methodological developments reflect those apparent in other strands of the econometrics literature, most notably the evaluation of labour market programmes (Heckman et al., 1999). However, concerns about the functional form assumptions underpinning these alternative techniques, difficulties in testing their identifying assumptions, and concerns that these techniques introduce their own sources of bias, have led some to conclude, as did H.G. Lewis (1986) nearly two decades ago, that the least-biased estimate of the union wage premium is the OLS. This may account for the fact that OLS remains the standard estimation technique in the wage premium literature in the United States (see, for example, Blanchflower and Bryson, 2003; Hirsch and Schumacher, 2002). Yet most analysts rarely discuss the key identifying assumption underpinning the linear regression model, namely that selection into membership is captured with observable data. Nor do they discuss the fact that, for this key identifying assumption to be plausible, one must be able to control for all characteristics affecting both union status and wages. This requires very informative data.

In this paper, the performance of the linear estimator is re-examined. The paper offers two main contributions to the previous literature. First, linear estimation is considered alongside another technique, hitherto unused in the union wage premium literature, which shares the same causal identification assumption. This semi-parametric technique, known as Propensity Score Matching (PSM), offers some notable advantages relative to linear estimation. Results using the two techniques are compared, and reasons for differences in results are identified and discussed. Second, it exploits a development in data quality, namely the advent of nationally representative data linking employers and employees in the 1998 British Workplace Employee Relations Survey (WERS). These linked data have advantages over data collected solely from employees – which tend to dominate the literature - since they provide more accurate and detailed information regarding the nature of the workplace. By altering the information set entering the estimation – first utilising individual-level data only, and then introducing workplace-level data – the paper shows the sensitivity of OLS and PSM results to data quality.

Using OLS and the full set of individual and workplace-level data, the membership premium varies between 5% and 12%, depending on worker type and model specification. However, using identical data, PSM estimates indicate no significant membership premia, suggesting the linear functional form assumption results in an upward bias in the OLS estimates with these data. In both OLS and PSM estimates, the addition of workplace-level data to the individual data substantially reduces the size of any membership effect, suggesting that some of the union effect attributed to membership in analyses based on individual or household data is actually related to the 'better' paying workplaces that members enter. Failure to take account of this results in upwardly biased estimates of the membership effect.

The paper is organised in the following way. Section II discusses the economics of the union wage premium and reviews ways in which biases in estimating the membership premium have been tackled. Section III introduces the data. Section IV outlines the estimation strategy. Section V presents results and Section VI discusses the implications of the findings and draws some conclusions.

II. The economics of the union wage premium

There are two ways unions can affect wages in the economy (Farber, 2001). The first is the direct effect on the wages of workers in jobs where wages are set through collective bargaining. This may affect non-members' and members' wages. The second level is the impact that the presence of unions has in the economy: this can change the level and distribution of wages generally. In theory, these general equilibrium effects may both raise and reduce the level of aggregate wages in the economy. Spillover effects, in which union bargaining sets wages resulting in labour shedding, leading to increased labour supply in the non-union sector, will reduce non-union wages. On the other hand, where non-union employers raise wages to keep unions out (the 'threat' effect) the union-non-union wage premium will diminish. Since it is not possible to observe the counterfactual (wages in the absence of unions) this union general equilibrium effect is not easily estimable. The union-non-union wage differential (the wage premium), defined as

$$\Delta = \frac{W_u - W_n}{W_n}, \tag{1}$$

is estimable because the wages of members (W_u) and non-members (W_n) are observable. Provided differentials are small, this expression is usefully approximated by

$$\Delta \approx \Delta_u - \Delta_n, \tag{2}$$

which says that the measured union wage premium is approximately equal to the difference in the proportional effects of unions on the union and non-union wage. The union wage premium in equation (1) can be usefully approximated by the difference in log wages, implying that

$$\Delta \approx \ln(W_u) - \ln(W_n). \tag{3}$$

The union wage premium may reflect the direct effect of unions on the wages of unionised workers, and the offsetting effects on non-union workers, but the difference is broadly interpreted as the effect of unions on wages. This is the wage premium estimated in this paper, where *u* denotes union membership.

Potential biases in the estimation of the wage premium. The membership differential is often attributed to the rent-seeking behaviour of unions who, through negotiation with employers, are able to procure a wage premium for their members. However, studies also find a membership premium even among workers whose pay is set through collective bargaining ('covered workers') (for the US see Hunt et al., 1987; Budd and Na, 2000; for Britain see Hildreth, 2000). In explaining this puzzling phenomenon, some have argued that employers may conspire to pay lower wages to covered non-members than to members in return for union co-operation, since this may increase the size of the surplus to be shared between workers and the firm (eg. Blakemore et al., 1986). However, even if this sort of collusion occurs in some cases, it seems unlikely that this could account for the size of membership differentials identified in the literature. Since there appear to be no obvious mechanisms by which members should command higher wages than similar non-members other than coverage, the membership premium may be accounted for by unobserved differences between members and non-members which boost members' relative earnings. Biases in estimates of the union membership premium may be accounted for by data deficiencies and, in particular, the paucity of employer controls in the household and employee data sets often used to generate them. The membership premium could be accounted for, at least in part, if the sorts of employer where unionisation is most likely are also better payers than non-unionised employers. This might occur if unions target their organising efforts on employers with larger rents to share, or if training-intensive employers wishing to maximise returns to their training investment contract with a union to lower quit rates. Workplace heterogeneity may also help explain the premium among covered workers if the union differential is positively correlated with union density since the conditional probability of high density given membership is higher than that given coverage. This deficiency in employer controls is addressed directly in this paper with linked employer-employee data from the Workplace Employee Relations Survey 1998 (WERS). As well as information on individual employees' union membership, WERS contains rich information on the employer, including workplace-level union density and pay bargaining arrangements for occupations within the workplace.

A second possible source of bias in the estimates of union membership effects on wages is the potential endogeneity of union status if membership is governed by a selection process. Following Farber (2001), there are two selection processes. The first is 'worker choice' in which workers only choose membership if the union wage is greater than the wage available to the individual outside the union. It is often assumed that workers with a lower underlying earning capacity have more to gain from membership than higher quality workers, in which case this selection process will understate the union wage premium. The second selection process arises through 'queuing' since not all workers desiring union employment can find union jobs (see Bryson and Gomez, 2002a for empirical validation of this model in Britain). Under this model, union employers may choose the best of the workers among those desirous of a union job. This employer selection implies a positive bias in the union premium but, a priori, it is not clear whether this bias is greater or less than the negative bias implied by worker selection. Thus, causal inference is problematic because, where workers who become members differ systematically from those who do not become members in ways which might affect their earnings independent of membership, we can not infer the non-union wage for union members simply by comparing union members' wages with those of non-members.

Tackling biases in identifying the causal effect of membership on the union wage premium. Broadly speaking, four methods have been used in the union wage premium literature to account for endogenous selection into membership status. Two (fixed effects modelling (eg. Hildreth, 1999; 2000) and identification through repeat observations on wages of individuals who change union status (eg. Freeman, 1984; Machin 2001)) require longitudinal data. Panel estimates are particularly prone to misclassification and measurement error which tend to result in estimates of the impact of unions that are *downward* biased (Robinson, 1989; Swaffield, 2001). The other two approaches are viable with cross-sectional data, and can be labelled methods tackling either 'selection on unobservables' or 'selection on observables'.

Methods accounting for selection on unobservables dominate the literature. In these studies, selection bias is addressed by modelling union status determination simultaneously with earnings and estimating an econometric model that takes account of the simultaneity (Robinson, 1989). The virtue of these techniques is that they seek to control directly for the presence of unobserved correlation between union membership and wages, thus purging estimated effects from the bias induced by unobserved heterogeneity. However, they rely on exclusion restrictions whereby variables assumed to affect union status have no direct effect on earnings. These assumptions are difficult to test and most data sets lack suitable instruments for identification (Lewis, 1986; Blanchflower, 1987). In addition, these techniques invoke functional form assumptions that seem arbitrary, often assuming that the errors in the earnings and union status functions are jointly normal. These drawbacks may explain why simultaneous equation methods tend to produce large and unstable estimates. The advent of new data offering better instruments, coupled with new work on testing the identifying assumptions underpinning IV methods, has resulted in more persuasive IV estimation of the union wage premium very recently. Booth and Bryan's (2001) IV estimation of the membership premium among covered workers is of particular note because it uses the same data set at this paper. Their results and how they relate to this paper are discussed in Section Six.

In his review of the literature, H. Gregg Lewis (1986) concluded that, due to the deficiencies of simultaneous equation and panel techniques, the most appropriate way to estimate the impact of unions on wages is using simple linear regression. He suggests OLS may produce an upper bound estimate of the true impact of unions because 'such estimates suffer from upward bias resulting from the omission of control variables correlated with the union status variable' (Lewis, 1986: 9). His assumption is that some of the wage premium attributed to union membership is, in fact, attributable in part to the characteristics of members, their jobs and their employers which would give them higher wages than non-members in any case. In practice, bias in cross-sectional OLS estimation due to unobserved heterogeneity may both upwardly or downwardly bias the 'true' impact (Farber, 2001; Robinson, 1989, Blanchflower and Bryson, 2003).[1] Either way, if there is endogenous selection the membership mark up estimated using standard cross-sectional regression

[1] Wessels (1994) has cogently argued that OLS will not necessarily provide an upper limit estimate. His argument is that, provided a union has a certain degree of bargaining power, union-won increases in the wage that lead the firm to hire more able workers will be followed by further union actions to raise the wage. Knowing this and given repeated bargains, the firm will not necessarily hire more able workers. In practice, Robinson (1989) shows OLS produces results which lie somewhere between the upper bound set by IV and Inverse Mills Ratio estimation and the lower bound set by panel estimates.

techniques ‘can be interpreted as the average difference in wages between union and non-union workers, but it can not be interpreted as the effect of union membership on the wage of a particular worker’ (Farber, 2001: 11).

Recently greater efforts have been made to capture selection effects using observable data. This is due, in part, to evidence from the evaluation literature indicating that controlling for bias due to observable characteristics is more important than controlling for the bias due to unobservables (Heckman et al., 1998). The regression coefficient for the union membership dummy in an OLS can be interpreted as the causal effect of union membership on wages if the variables entering the regression equation account fully for endogenous selection into membership status. This requires very informative data. This paper assess the sensitivity of results to this assumption by varying the information set entering the estimation – first utilising individual-level data only, and then introducing workplace-level data.

A second way of controlling for bias on observables is the semi-parametric statistical matching approach known as propensity score matching (PSM) (Rosenbaum and Rubin, 1983; Heckman et al., 1999) which compares wage outcomes for unionised workers with ‘matched’ non-unionised workers.[2] The method shares the causal identification assumption of the OLS in that it yields unbiased estimates of the treatment impact where differences between individuals affecting the outcome of interest are captured in their observed attributes (often referred to as the conditional independence assumption, or CIA). However, matching has three distinct advantages relative to regression in identifying an unbiased causal impact of membership on wages. First, it is semi-parametric, so it does not require the assumption of linearity in the outcome equation. Second, it leaves the individual

[2] A full explanation of the approach is described in the appendix.

causal effect completely unrestricted so heterogeneous treatment effects are allowed for and no assumption of constant additive treatment effects for different individuals is required. Effects for sub-groups can be estimated by running the match on sub-populations. This is particularly important when estimating the wage premium since empirical research shows the returns to membership differ across worker types in both the US and Britain (Blanchflower and Bryson, 2003). Section V compares OLS and PSM estimates for six sub-groups: covered and uncovered workers; manual and non-manual workers; and men and women.

Thirdly, matching estimators highlight the problem of common support and thus the short-comings of parametric techniques which involve extrapolating outside the common support. Matching eliminates two of the three sources of estimation bias identified by Heckman, Ichimura, Smith and Todd (1998): the bias due to difference in the supports of X in the treated and control groups (failure of the common support condition) and the bias due to the difference between the two groups in the distribution of X over its common support. The other source of bias is due to selection on unobservables. This highlights the importance of the conditional independence assumption since, if this holds, selection on unobservables ceases to be a problem. The appropriateness of the conditional independence assumption is dependent on having data that account for selection into membership. Section 2.1 indicates selection may occur through employee and employer choice, implying that data should be rich in employee and employer data to capture these selection processes. This paper takes advantage of the very informative linked employer-employee data available in WERS. As in the case of the OLS estimates, the sensitivity of results to data quality is assessed by altering the information set entering estimation.

III. Data

The paper uses linked employer-employee data from the Workplace Employee Relations Survey 1998 (WERS). Appropriately weighted, it is a nationally representative survey of workplaces in Britain with 10 or more employees covering all sectors of the economy except agriculture (Airey *et al*, 1999). The analysis exploits two elements of the survey. The first is the management interview, conducted face-to-face with the most senior workplace manager responsible for employee relations. Interviews were conducted in 2,191 workplaces between October 1997 and June 1999, with a response rate of 80%. The second element is the survey of employees where a management interview was obtained. Self-completion questionnaires were distributed to a simple random sample of 25 employees (or all employees in workplaces with 10-24 employees) in the 1,880 cases where management permitted it. Of the 44,283 questionnaires distributed, 28,237 (64%) usable ones were returned.

The sample of workplaces is a stratified random sample with over-representation of larger workplaces and some industries (Airey, et al, 1999). Employees' probability of selection for the survey is a product of the probability of their workplace being selected and the probability of the employee's own selection. To extrapolate from the analyses to the population from which the employees were drawn (namely employees in Britain in workplaces with 10 or more employees) analyses are weighted using the employee weights. The weighting scheme used compensates for sample non-response bias which was detected in the employee survey (Airey *et al.*, 1999: 91-92). All estimation in the paper accounts for the complex sample design, that is, sampling weights, clustering and stratification.

The estimating sub-sample is all private sector employees with complete information on the variables used in the analysis. The estimation of the union membership premium for the whole private sector contains 10,384 non-members and 4,147 members. Membership is derived from individual employees' response to the question: 'Are you a member of a trade union or staff association?'

The paper also exploits the bargaining coverage information in WERS based on management classifications of the way pay is set for each occupational group in the workplace. The eight possible responses include collective bargaining at industry, organisation or workplace-level. The information is used to identify employees whose own occupational group at their workplace is covered by collective bargaining at any level. This sample comprises 1,531 non-members and 2,653 members. Coverage and membership are less highly correlated in Britain than in the United States because there is less pressure on employees to become members where there is a coverage agreement (Hildreth, 2000: 133-134). In the private sector in WERS, 64% of union members belong to a covered occupation, while 13% of non-members are covered. Around one-third (35%) of covered employees are non-members.

The dependent variable. The dependent variable is log gross hourly wages. This is derived from banded earnings data by taking the mid-point of the respondent's earnings band and dividing this by continuous hours worked. The earnings band for the top-coded highest earners is closed by introducing an upper ceiling that is 1.5 times the lower band. The hours denominator used includes overtime hours.[3]

[3] The question asks: 'How many hours do you usually work each week, including any overtime or extra hours?'

Union membership wage differentials are usually higher when measured in terms of hourly earnings than when measured in weekly earnings because many analysts find union workers work fewer hours per week than non-union workers, on average (Andrews et al, 1998). This proves not to be the case in WERS: mean hours worked per week are 39.4 for members and 35.4 for non-members. This is because the incidence of part-time working is much higher among non-members: part-timers were usually excluded in previous analyses (Green, 1988; Blackaby et al., 1991; Andrews et al., 1996). This issue is further investigated by estimating the independent effect of membership and union recognition on log hours worked per week. The model accounts for a substantial amount of the variance in hours (r-squared = 0.51). The membership effect is positive but not significant (0.38, t = 0.84) and workplace union recognition has a significant negative effect (-1.53, t = 3.17).[4] So, although there is an hours differential in the data, it is not attributable to membership per se, and suggests that measurement of the membership differential should vary little whether measured in weekly or hourly wages. Nevertheless, since the division of banded weekly hours by continuous hours worked introduces measurement error into the hourly earnings dependent variable, sensitivity of results to the use of the weekly banded wage is reported in Section Four.

IV. Estimation strategy

This section describes the OLS and PSM estimation techniques in turn. A technical appendix gives more detail on the theory behind PSM.

[4] Full models are available from the author.

4.1: Ordinary Least Squares regression

The standard 'union effect' estimates (Lewis, 1986) are OLS estimates of equation (3) which take the form

$$\ln W_i = X_i ß + dU_i + e_i , \qquad (4)$$

where $\ln W_i$ is the log wage of individual *i*, *X* is a vector of worker, job and workplace characteristics, *U* is a dummy variable indicating union membership, and e is a random component. The parameter d represents the average proportional difference in wages between union and non-union workers adjusted for worker and workplace characteristics, and it is the regression-adjusted analogue of Δ. However, since membership is non-random there is likely to be non-zero correlation between membership and the error term e. This arises because an individual's membership decision is probably based on characteristics that may also affect earnings. If so, and it is not possible to control for all characteristics affecting *W* and U_i simultaneously, some correlation between the error term and membership can be expected. Thus, for d to capture the causal effect of membership on wages, *X* must contain all variables determining both wages and membership status. The assumption is that individuals who are the same in the observable dimension X_i but make different choices on whether or not to be a member do not differ on the average in the unobserved dimension e_i. In this case equation (4) can be viewed as a form of regression-based linear matching (Blundell and Costa-Dias, 2002). If identification is accomplished, the parameter estimated is the effect of membership on members, that is, the gain from moving a member with a given set of attributes from non-membership to membership. This is what

Lewis (1986:11) termed the 'wage gap' and, in the evaluation literature, is known as the 'treatment on the treated' parameter.[5]

The whole private sector analysis captures an average return to membership, irrespective of whether the individual – member or non-member – is covered by collective bargaining. Results are also presented for sub-groups where previous research (Blanchflower and Bryson, 2003) indicates the premium is likely to differ markedly, namely covered and uncovered workers, men and women, manuals and non-manuals. In this way regression analyses can identify heterogeneous effects of union membership. The other advantage of this approach is that it allows coefficients on all covariates to differ by sub-group.

To assess the additional value of linked employer-employee data relative to the individual-level only data available to most analysts, two model specifications are presented. The first contains data collected solely from the employee respondents. Variables entering these models are demographics (gender, marital status, dependent children, health, ethnicity, academic and vocational qualifications) and job-related (occupation, nature of contract, part-time working, if overtime required or voluntarily undertaken, occupational gender segregation at the workplace). Region and local labour market conditions are included too, since the analyst should usually be able to match these into individual-level data. These models also include whether the employee thinks there is a union on-site. The 'individual plus workplace-level' models incorporate all of the above information, but replace the employee data on the presence of a union with employer data on whether the employer recognises a union for pay bargaining, plus union density at the workplace. One can obtain

[5] If a linear specification is assumed with common coefficients for members and non-members then no common support requirement is needed to estimate the treatment on the treated parameter (Blundell and Costa Dias, 2002).

the 'full' union-non-union wage differential by adding together the membership effect and the coverage effect. The other data added from the managerial respondent include workforce composition (by gender, occupation, hours worked) and workplace characteristics (size, activity, industry, ownership). The additional value of the workplace-level data is tested by comparing the fit of models and by estimating the joint significance of the workplace-level variables.

Although the linked employer-employee data provide much of the requisite information to control for selection into membership, it is arguable that some data are missing. For example, there are no data on motivation which, it has been argued, is positively correlated with membership and the desire to invest in workplace-specific human capital, thus raising wages (Budd and Na, 2000). The data set does contain workplace tenure and the amount of employer-provided training undertaken, both of which may be correlated with this tendency. However, because these variables may be influenced by membership itself, and are thus endogenous with respect to membership, their incorporation in the OLS – and the estimation of the propensity score - could undermine the interpretability of estimated effects (Heckman, LaLonde and Smith, 1999). They are therefore excluded from the reported estimates. Although the absence of data on motivation may violate the conditional independence assumption, the absence of workplace tenure and employer-provided training would only bias estimates if they influenced both membership and wages. It is certainly the case that longer workplace tenure is independently associated with higher earnings (Bryson, 2002) and an increased likelihood of union membership.[6] But the empirical

[6] Two-thirds (66%) of members had been working at the workplace for at least five years, compared with 37% of non-members. Conversely, 39% of non-members had been at their workplace for under two years compared with only 7% of members. Regression analysis revealed an independent association between membership and tenure (results available from the author).

literature suggests that membership increases tenure by reducing the likelihood of voluntary quits, consistent with the theory that unions provide a 'voice' alternative to quitting for dissatisfied workers (Freeman and Medoff, 1984). There is little reason to believe that longer tenure might lead to membership since, unlike the United States (Budd and Na, 2000) there are no institutional factors that increase the likelihood of joining the union after the end of a probationary period. So, in the British context, workplace tenure can be omitted without biasing estimates. Similarly, it is hard to see how employee take-up of training can influence union membership – unless, that is, non-members are discriminated against by employers or unions who ensure privileged access for members, whereupon poorly trained non-members may have an incentive to join the union. In fact, the distribution of days spent in employer-provided training was nearly identical across members and non-members.[7] So, in spite of the independent effect it had on employees' earnings (Bryson, 2002), there is little empirical or theoretical justification for its inclusion in the regression or propensity score estimation. In any event, the sensitivity analyses reported later show the inclusion of training and tenure make little difference to the results.

4.2: Empirical implementation of matching

This section describes the empirical implementation of propensity score matching in WERS to yield an unbiased estimate of membership's effects on the wages of union members.

Since the propensity to be a union member is unknown, the first task in matching is to estimate the propensity to be a union member. This is done with a probit estimating a

[7] The measure of employer provided training is employees' responses to the question: 'During the last 12 months, how much training have you had, either paid for or organised by your employer? Include only training away from your normal place of work, but it could be on or off the premises.'

(0,1) variable identifying individuals' union membership status. The conditional independence assumption requires that all variables influencing membership and wages should be included in the estimate.[8] The choice of variables is informed by previous empirical work (Bryson and Gomez, 2002a) and the theory underpinning the worker choice and queuing models of membership discussed earlier. Variables entering the model are demographics (age, gender, marital status, health, ethnicity, qualifications), job-related (occupation, nature of contract, hours worked, gender segregation), workforce composition (by age, gender, occupation, hours worked), workplace (size, activity, industry, ownership, location) and local labour market conditions. As with the OLS analysis, the sensitivity of the PSM results to the inclusion and exclusion of employer data is tested.

The model estimating the probability of union membership for the whole private sector using individual-level and workplace-level data is presented in Appendix Table A1.[9] Among non-members, the predicted probability of union membership ranges from .0001 to .9827, with a mean of .13 and a median of .05. Among members, the predicted probability ranges from .0025 to .9981, with a mean of .64 and median of .73. The distributions of propensity scores are presented in Graph A1 in the technical appendix: although non-members' scores are bunched in the lower quartile of the distribution, they nevertheless offer support for members throughout the distribution.

---

[8] Variables that affect neither membership nor wages are clearly irrelevant. If a variable influences membership but not wages, there is no need to control for differences between members and non-members because wages are unaffected. Conversely, if a variable influences only wages, there is no need to control for it since it will not be significantly different between members and their matched comparators. This just leaves variables that affect membership and wages.

[9] The equivalent models for the sub-group analyses are available from the author. In what follows, the matching process is illustrated with detail from the estimates for the whole private sector. Identical information for all the matching undertaken is provided in the notes reporting results in Table 2.

As discussed in the technical appendix, matching operates by constructing, for those participants with support, a counterfactual from the non-participants. There are a number of ways of defining this counterfactual using the propensity score. This paper uses nearest neighbour that involves taking each treated individual (member) and identifying the non-treated individual (non-member) with the most similar propensity score. The matches were made with replacement so that, in some cases, a non-treated individual provides the closest match for a number of treated individuals, whereupon they feature in the comparison group more than once.[10] To ensure the quality of matches a tolerance is set when comparing propensity scores. This entails imposing a 0.002 caliper: where the propensity score of a treated individual falls beyond this bound for a near comparator, the treated individual remains unmatched. This means of enforcing common support results in the discarding of 138 members from the analysis, that is, 3.3% of all members. Thus, the sub-group of members for whom it is not possible to estimate the membership premium is very small.

The advantage of nearest neighbour matching is that the match is as good as it is possible to achieve in the sense that the bias across the treatment and comparison groups is minimised. In this analysis, the matches are very close: the mean difference in propensity scores between treated individuals and their matched comparators is .0003, and ranges between 0 and .0019. However, nearest neighbour matching disregards potentially useful information by not considering any matches of slightly poorer quality. Over-reliance on a reduced number of observations can result in effects being less precisely identified. Of the 10,384 non-members who could potentially have been matched to the 4,009 members with common support, 1,584 were used as matched comparators. In 58% of cases these matched comparators have a match weight of 1 because they are matched to a single treated case. The

[10] Dehijia and Wahba (1999) find that allowing the non-treated to be used more than once as comparators

largest weight is 49, and in only 59 cases is a non-member used as a match for 10 or more members. The mean match weight for non-members is 2.53.

The PSM estimates of mean wages are for the population from which the sample was drawn, taking account of the complex survey design when comparing mean differences across members and matched non-members. In these population estimates, the survey sample weight of each treatment group member is applied to the corresponding matched comparator(s) (Frölich et al., 2001:12). Hence, population estimates of the union membership differential are based on a weight incorporating both the matching weight and sampling weight. Population differences in mean earnings between members and their non-member comparators also account for variance arising from sample stratification and clustering.

To be effective, matching should balance characteristics across the treatment and comparison groups. Appendix Table A2 presents comparisons of the means in the characteristics used to match members and non-members, as well as a measure of the 'distance' of the marginal distributions of relevant characteristics in both groups (Rosenbaum and Rubin, 1985). For a given covariate, the standardised difference after matching is defined as the difference of the sample means in the treated and matched non-treated subsamples as a percentage of the square root of the average of the sample variances in the treated and non-treated groups (Sianesi, 2001). Overall, the quality of the match seems good, the mean absolute standardised bias for all covariates being -.35. Standardised bias for each variable tends to range from -7% to +7%, and only once does it exceed 9%. Although achieving a reasonable balance on the *X*'s entering the participation equation is an

---

improves the performance of the match.

indicator of how good the match is on observables, it cannot provide an indication as to whether the CIA is plausible.

Having matched on the propensity score, the mean impact of union membership is estimated as the mean difference in the outcomes of the matched pairs.

V. Results

Table 1 reports OLS estimates of the union membership wage premium in the private sector as a whole and for coverage, gender and occupational sub-groups. Using individual-level data only, the estimated membership premium for the whole private sector is 15% (the exponentiated coefficient for union member in column 1). Column 2 introduces data collected from the employer. The employer variables are jointly significant and improve the model fit. What is more, they reduce the membership premium by over half to 6.1%. This pattern, whereby the premium estimated using individual-level only data is substantially reduced with the introduction of workplace-level data, is repeated across the sub-group analyses. The impact of workplace-level data is particularly marked for men and manual workers. These findings suggest that, at least in the British case, OLS estimates of the membership premium based on individual-level and household survey data are upwardly biased because some of the positive wage effect attributed to membership is actually due to members being employed at better paying workplaces. There are many possible reasons why union workplaces might be better payers than non-union workplaces. As noted earlier, unions may target organising efforts on employers with the biggest rents to share. 'Better' employers may chose to unionise to create stable firm-employer conditions conducive to investment in human capital. Alternatively, if union members are 'better' workers than their non-member counterparts in ways unobservable to the analyst but observable to employers, members may be able to sort themselves into the best employers, or may be chosen by the best employers.[11]

These analyses for the whole private sector condition on whether the individual is located in a workplace where the employer engages with a union in pay bargaining, as well as on union density at the workplace where the employee works.[12] However, if the biggest component of any membership premium is that generated by collective bargaining, the premium should be much smaller where the sub-sample consists solely of workers in covered occupations. In general, all these workers should benefit from pay bargaining, unless employers discriminate between members and non-members. In fact, the OLS estimates for covered occupations presented in columns 3 and 4 differ little from those for the whole private sector. Again, the size of the premium falls substantially once account is taken of workplace heterogeneity, but it remains sizeable and statistically significant at 6.7%. On this evidence, the membership premium among covered workers, evident in other recent studies using individual-level data only (for the United States, Schumacher, 1999, Budd and Na, 2000; and for Britain, Hildreth, 2000), persists having accounted for workplace heterogeneity.

Intriguingly, the 14% of employees who are members in uncovered occupations receive a similar membership premium of 5.7% when the OLS is run with individual and

---

[11] Using French linked employer-employee data, Abowd, Kramarz and Margolis (1999) find high-wage firms have more productive workforces and higher profits than other firms. It may be that the unionised sector in Britain comprises such firms.

[12] The full union-nonunion wage differential combining membership and coverage is obtained by exponentiating the sum of the membership and workplace-level union recognition coefficients exp(.059 + -.018) = 4.2%. The union recognition dummy is never statistically significant in the equations presented in Table 1, but wages rise with union density. Appendix Table ?? presents the full model for the whole private sector.

workplace-level controls.[13] However, almost three-quarters (71%) of these members are located in workplaces where other workers have their pay set through collective bargaining. This suggests that these members benefit from the spillover effects of collective bargaining at their workplace.[14]

Whole economy estimates for the United States suggest little difference in the membership premium for men and women (Blanchflower and Bryson, 2003). In Britain, there is a much bigger premium for women but there is no difference in the premium by gender once a public sector control is added (Blanchflower and Bryson, 2003). Contrary to expectations, comparing columns 5 for men and 7 for women, the membership premium in the private sector in WERS is much higher among men than it is among women when estimated on individual-level data only. However, controlling for workplace heterogeneity in columns 6 and 8, the premia are virtually identical for men and women at 6.2%. This implies that, among members, men tend to sort (or are sorted by employers) into better paying workplaces than women.

The most striking evidence that union membership effects are heterogeneous comes from analyses by broad occupation. Running analyses for manual and non-manual employees separately, results confirm those from other studies in showing a larger membership premium among manual workers (Booth, 1995; Forth and Millward, 2002). Indeed, with the introduction of workplace controls, non-manual workers are the only group of workers for whom the OLS estimates do not produce a statistically significant membership premium.

[13] The uncovered worker estimates using OLS and PSM are available from the author on request.
[14] Forth and Millward (2002) find evidence of such spillover effects in their analysis of WERS.

Table 1
Estimated coefficients on union membership dummy from hourly pay equations

	(1) whole private		(2)covered occupations		(3) men		(4) women		(5) manual		(6) non-manual	
Pay	Individual	Individual + workplace	Individual	Individual + workplace	Individual	Individual + workplace	Individual	Individual + workplace	Individual	Individual + workplace	Individual	Individual + workplace
Union member	.140	.059	.125	.065	.165	.060	.098	.061	.204	.075	.055	.019
	(6.34)	(4.58)	(4.83)	(3.16)	(5.95)	(3.33)	(4.90)	(3.33)	(7.12)	(4.08)	(3.09)	(1.17)
Sample size	14531	14531	4184	4184	7974	7974	6559	6559	5406	5406	9125	9125
R squared	.533	.579	.502	.563	.533	.5921	.4761	.5167	.4336	.4941	.5627	.6115
F-stat	45,978= 132.55	64,959= 113.79	45,281= 59.62	63,263= 59.75	44,907= 81.37	63,888= 78.32	44,905= 72.11	63,886= 69.61	40,735= 40.67	59,716= 41.85	41,912= 120.08	60,893= 111.08
Model p-value	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
Workplace P-value		0.0000		0.0000		0.0000		0.0000		0.0000		0.0000

Notes. (1) OLS estimates of log gross hourly wages that use survey stratification weights and account for the presence of repeated observations on the same establishment. (2) Asymptotically robust t-ratios in parentheses. (3) Individual-level data models contain the following controls, all of which are dummies unless otherwise stated: gender, age (6 dummies), highest academic qualification (5 dummies), vocational qualifications, health problem, non-white, dependent children, married or living as married, occupation (9 dummies), permanent contract, part-timer, overtime hours required, overtime hours worked voluntarily, gender segregation for employee's job at the workplace (5 dummies), region (11 dummies), unemployment in local labour market below 5%, respondent says union on-site, respondent doesn't know if union is on-site. (4) Individual + workplace-level data models contain all variables in (3) except the two union dummies plus: employer recognition of union for pay bargaining, union density at workplace (continuous), % working part-time, % employees who are women, % manual employees (3 dummies), workplace size (6 dummies), single-establishment organisation, foreign ownership, industry (9 dummies). (5) In the sub-group analysis of covered occupations (model (2)) union recognition is dropped because it is another coverage variable. (6) Workplace P-value is reporting significance of the adjusted Wald test for the joint significance of the workplace-level data when added to the individual-level only data.

Table 2
Mean percentage hourly wage premium for union members using propensity score matching

Pay	(1) whole private		(2)covered occupations		(3) men		(4) women		(5) manual		(6) non-manual	
	Individual	Individual + workplace	Individual	Individual + workplace	Individual	Individual + workplace	Individual	Individual + workplace	Individual	Individual + workplace	Individual	Individual + workplace
Union Member	8.9 (16.2)	-1.5 (5.1)	2.8 (11.5)	-1.0 (5.8)	11.1 (16.1)	2.1 (6.2)	1.3 (10.4)	-3.3 (10.1)	17.7 (21.4)	6.9 (5.8)	3.6 (4.8)	-1.0 (4.7)
Sample size	6092	5593	3405	3231	3711	3446	2227	1998	2567	2269	3421	3102
F-stat	1,823= 5.11	1,742=.15	1,318=.44	1,307=.06	1,644= 5.72	1,576=.24	1,599=.15	1,512=.95	1,489= 11.85	1,432= 1.86	1,701=.77	1,581=.11
P-value	.0240	.7018	.5063	.8045	.0170	.6275	.7025	.3296	.0006	.1733	.3799	.7443
Bootstrap	4.1 - 14.0	- 6.9 – 3.7	-3.7 – 9.7	-8.2 – 6.0	5.5 – 16.9	-4.6 – 9.5	-4.9 – 7.7	-10.2 – 3.3	11.2– 24.6	-0.7 – 15.3	-2.0 – 10.0	-6.8 – 4.2
Support loss	0.8%	3.3%	3.3%	5.1%	3.3%	5.6%	4.7%	8.5%	3.0%	11.8%	2.3%	3.4%
Mean dif.	.0002	.0003	.0004	.0005	.0003	.0004	.0002	.0004	.0004	.0005	.0001	.0003
Mean NM wt.	2.10	2.53	3.06	3.54	2.46	2.96	1.55	1.85	3.01	3.40	1.61	2.06
Max. NM wt.	23	49	24	51	24	53	10	13	22	60	10	15
Bias	.48	-.35	.12	-.61	.16	.29	.60	-1.04	1.24	-.04	.39	.37

Notes. (1) PSM estimates of log gross hourly wage differential between members and matched non-members. % differential is exp(mean log wage of members – mean log wage of non-members). Sample size is weighted with the combined match weight and sample weight. (2) Figures in parentheses are equivalent % differentials derived from WLS estimates run on the identical matched sample weighted by the matching weight described in the text. (3) F-stat and P-value relate to significance of PSM estimate of mean difference in log wages between members and matched non-members. (4) Propensity scores are derived from probit estimation of union membership (0,1) accounting for complex survey design, namely sampling weights, clustering and stratification). (5) Propensity scores based on probits incorporating individual-level data only contain the following controls, all of which are dummies unless otherwise stated: gender, age (6 dummies), highest academic qualification (5 dummies), health problem, non-white, married or living as married, occupation (9 dummies), permanent contract, hours and hours squared (continuous), gender segregation for employee’s job at the workplace (5 dummies), region (11 dummies), unemployment in local labour market below 5%, respondent says union on-site, respondent doesn’t know if union is on-site. (6) Propensity scores based on probits incorporating individual and workplace-level data contain all variables in (4) except the two union dummies plus: employer recognition of union for pay bargaining, union density at workplace (continuous), % working part-time, % employees who are women, % manual employees, % aged under 20 years, workplace size (4 dummies), single-establishment organisation, workplace supplies goods/services to other companies, administrative or head office, foreign ownership, industry (9 dummies). (7) In the sub-group analysis of covered occupations (model (2)) union recognition is dropped because it is another coverage variable. (8) Bootstrap shows the 95% confidence interval for the bootstrapped estimates. (9) Support loss is % members lost through enforcement of common support using 0.002 caliper. (10) Mean dif. is the mean difference in propensity scores for the members and matched non-members. (11) Mean NM wt. is the mean match weight for non-members. (12) Max NM wt. is the maximum match weight for non-members. (12) Bias is the mean absolute standardised bias after matching. (13) Bias is the difference of the sample means in the treated and matched non-treated subsamples as a percentage of the square root of the average of the sample variances in the treated and non-treated groups.

Table 2 presents the PSM analyses. These are run on identical samples to those used in the OLS estimates in Table 1. The sample sizes shown in the third row of Table 2 are smaller than those appearing in Table 1 because, in the process of matching members to their nearest neighbours, PSM leaves many non-members out of the estimation sample (see Section IV and the technical appendix for details). In addition, a small number of members have no support in the non-member population, so it is not possible to estimate membership effects for this subset. Fortunately this group tends to be small, ranging between 3%-6% in most cases (see row 7 of Table 2). This means common support is not a problem, so PSM can estimate the effect of membership on members for nearly all of the member population. To aid comparison with OLS, the figures in parentheses in row 2 are the wage premia estimated with OLS on the post-match samples that are identical to the samples on which the PSM effects are estimated. As in the case of the OLS estimates in Table 1, PSM estimates are run with individual controls only, and with individual plus workplace controls.

The results are striking. Column 1 shows the membership premium based on matching with individual data alone is estimated to be 8.9%. When workplace data are used in the matching this premium disappears and is even negatively signed (column 2). As in the case of the OLS estimates in Table 1, the introduction of workplace-level data always reduces the membership premium, confirming the potential for upward bias in estimates based on individual-level data – whether estimated via OLS or PSM. But, in stark contrast to the OLS estimates in Table 1 and those using the matched data in row 2 of Table 2, there is no significant membership premium for any type of employee where matching is based on individual and workplace-level data. The OLS premia in parentheses in column 2 are always statistically significant, so one can discount sample size differences and common support enforcement as reasons for differences in the OLS and PSM results. Rather, the OLS estimates are upwardly biased due to the linear functional form assumption. Of course, it is arguable that the OLS models are simply misspecified and that results could be reconciled through the addition of appropriate interaction terms. In practice, this requires a great deal of trial and effort. This is forcefully illustrated by the fact that the OLS-generated premium remains large and significant across coverage, gender and broad occupation, three dimensions where one is most likely to find heterogeneous membership effects. Yet, in each case, the PSM-generated premia are not significant.[15] It is true that, at 6.9%, the PSM estimate of the premium for manual workers comes close to statistical significance, with the bootstrapped 95% confidence interval only just straying into negative territory. So, if there is a membership premium for anyone, it is for manual workers.

Sensitivity analyses. Table 3 presents four sensitivity analyses. The first row reproduces results from Tables 1 and 2 for the 'baseline' estimates. The first two sensitivity analyses involve alterations to the X vector used in the OLS estimation and estimation of the propensity score. The third sensitivity analysis estimates effects on weekly wages, as opposed to hourly wages. The fourth involves splitting the analyses according to union strength at the workplace employing the worker. In each case, the first column presents results from the OLS using individual-level and workplace-level controls. The second column presents the PSM results, and the third column presents the OLS results run on the matched data.

[15] The membership premium for uncovered workers where matching is based on individual and workplace-level data is –2.4%.

Table 3
Sensitivity analyses

	OLS, unmatched data	PSM	OLS, matched data
1. Original estimates for whole private sector	6.1% (4.58)	-1.5% (.7018)	5.1% (2.87)
Sample size	14,531	5,593	5,593
2. Exclude union recognition and union density from workplace variables	11.7% (7.45)	1.8% (.6498)	8.6% (5.03)
Sample size	14,867	6,383	6,383
3. Add workplace training and workplace tenure to individual variables	4.8% (3.56)	-1.5% (.6810)	5.7% (3.21)
Sample size	14,450	5,550	5,550
4. Change dependent variable to log gross weekly wages	7.8% (5.50)	-2.5% (.5977)	5.4% (3.06)
Sample size	14,531	5,593	5,593
5a. Employees in workplaces with 50%+ union density	8.9% (4.03)	2.6% (.5968)	4.7% (1.85)
Sample size	3,949	3,582	3,582
5b. Employees in workplaces with <50% union density	5.0% (3.23)	-1.9% (.6301)	4.9% (2.51)
Sample size	10,582	2,263	2,263

Notes: (1) Percentages are the membership differentials based on exponentiated differences in log wages between members and non-members. Wages are gross hourly, apart from in 3 where weekly gross wages are estimated. (2) For OLS estimates, figures in parentheses are asymptotically robust t-ratios; for PSM they are p-values for f-statistics. (3) All estimates incorporate individual-level and workplace-level data. (4) All estimates are for the whole private sector, except for analyses run on high and low union density workplaces.

Although there is a sizeable union wage premium literature conditioning on bargaining coverage and union density, it is at least arguable that density and union recognition are endogenous with respect to membership in that these workplace features are, in part, a function of individuals' decisions to unionise. A comparison of results in row 2 Table 3 with the whole private sector estimates in row 1 shows the premium estimates rise in the absence of density and recognition controls. However, the pattern of results remains the same, with OLS producing sizeable and statistically significant premia, whereas the PSM estimate is small and statistically non-significant.

For reasons set out in Section IV, workplace training and workplace tenure were omitted from the earlier estimates. Their inclusion in row 3 of Table 3 makes no difference at all to the PSM estimates, and very little difference to the OLS using matched data. The premium estimated with OLS on unmatched data falls a little.

As noted in Section III, there is the potential for measurement error in the hourly earnings measure and estimates of the wage premium can differ across hourly and weekly earnings measures due to different working patterns of members and non-members. Row 4 in Table 3 therefore shows the sensitivity of results to the use of a weekly earnings measure. Again, the pattern of results is largely unchanged.

Empirical evidence for Britain (Stewart, 1987) and the United States (Schumacher, 1999) indicates that the union premium is higher where union density is higher. This may be because a higher incidence of 'free-riding' can weaken union bargaining strength, or else causation may work the other way if the incentive to join a union is higher where the union commands a larger premium. Splitting the analysis into employees working in lower and higher density workplaces offers limited support for the proposition that the membership premium is higher where the union is stronger. Using OLS to estimate the membership effect on unmatched data, the premium is much larger among employees in workplaces with 50%+ density than it is among those located in workplaces with less than 50% density. There is also a differential using PSM although the premium is not significant in either case.

Across all these sensitivity analyses, OLS identifies a sizeable and statistically significant membership premium whereas PSM finds no significant premium, supporting the main conclusion from the baseline analyses.

VI. Conclusions

The paper estimates the union wage premium in Britain's private sector at the end of the 20th Century after nearly two decades of union decline. It examines the performance of the linear estimator, which continues to dominate the literature, alongside a semi-parametric technique known as Propensity Score Matching. The techniques share the same key identifying assumption, namely that selection into membership is captured with observable data, an assumption which means both techniques demand very good employee and employer data to capture the selection process. These data are not usually available in individual or household surveys, raising questions about the possible size and direction of biases in the estimated union membership premium. Capitalising on rich linked employer-employee data, the paper investigates the performance of the two estimators and, by altering the information set entering estimation, shows the sensitivity of OLS and PSM results to data quality.

Using OLS and the full set of individual and workplace-level data, the membership premium varies between 5% and 12%, depending on worker type and model specification. However, using identical data, PSM estimates indicate no significant membership premium in the private sector, or across sub-groups of workers (covered, uncovered, men, women, manuals, non-manuals, those in highly unionised and less unionised workplaces). The results suggest the linear functional form assumption results in an upward bias in the OLS estimates with these data. In both OLS and PSM estimates, the addition of workplace-level data to the individual data substantially reduces the size of any membership effect, suggesting that some of the union effect attributed to membership in analyses based on individual or household data is actually related to the 'better' paying workplaces that members enter. Failure to take account of this results in upwardly biased estimates of the membership effect.

In a recent paper which also uses the linked employer-employee data in WERS, Booth and Bryan (2001) estimate the membership premium among covered workers in Britain's private sector. They also find significant premia using OLS but no significant premia when estimating using IV techniques. Although their paper differs in a number of ways from this paper, it lends further support to the suggestion that OLS estimates may be upwardly biased. Taken together, the two papers support the contention that there was no union membership wage premium in the late 1990s for Britain's private sector workers.[16]

If this situation persists, the question raised is: why don't union members leave the union? Well, they have been leaving: union density is in decline, even within unionised workplaces (Millward et al, 2000, chapter 5). However, evidence for the period 1983-2001 indicates that the rate at which employees have left membership has not risen and that the decline in membership is due to an increase in employees who have never been members (Bryson and Gomez, 2002b). It may well be that the returns to membership have declined.

---

[16]There are a number of ways in which Booth and Bryan's (2001) paper differ from this paper. First, they use a more restricted data set, namely workplaces with 25 or more employees including an interview with a worker representative. These workplaces are not representative of the whole private sector. Secondly, they use interval regression techniques to estimate log hourly wages with the banded pay data. Interval regression analyses were run to test the sensitivity of the results in this paper, but the results are very similar. Third, the IV technique recovers a different parameter from the matching approach: IV recovers the local average treatment effect, whereas the matching estimator and OLS are used here to recover the effect of treatment on the treated. Fourth, the IV technique explicitly accounts for unobserved correlation between wages and membership status, whereas matching only tackles selection on observables.

However, there are at least three reasons why it is not possible to read off employees' likely membership intentions from the wage premium they face. The first is that, as some suggest, the trend may not be secular. Instead, the premium may counter-cyclical, rising when economic conditions deteriorate (Blanchflower and Bryson, 2003). In any event, one can make the case for remaining in a union if – as the sensitivity analyses suggest - union density has a role to play in the size of the wage premium unions can extract from the employer. If members were to leave, the prospects of bargaining for better wages will deteriorate. Second, members are unlikely to value membership purely in terms of the wage mark up unions command. Members also benefit directly by unions' efforts to improve non-pecuniary benefits, job security, the handling of grievance and disciplinary matters, and by encouraging management to treat all employees more fairly. Third, many join and remain members because they are ideologically committed to doing so.

References

Abowd, John M., Francis Kramarz and David N. Margolis. “High Wage Workers and High Wage Firms.” *Econometrica* 67 (March 1999): 251-333

Airey, Colin, Jon Hales, Rosemary Hamilton, Christos Korvovessis, Anthony McKernan and Susan Purdon. *The Workplace Employee Relations Survey (WERS) 1997-8: Technical Report*. National Centre for Social Research, 1999.

Andrews, Martyn J., Mark B. Stewart, Joanna K. Swaffield, and Richard Upward. “The estimation of union wage differentials and the impact of methodological choices”. *Labour Economics*, 5 (1998): 449-474

Blackaby, David, P. Murphy and Peter Sloane. “Union membership, collective bargaining coverage and the trade union mark-up for Britain”. *Economic Letters* 36 (1991):203-208

Blakemore, Arthur E., Janet C. Hunt and B.F. Kiker. “Collective bargaining and union membership effects on the Wages of Male Youths”. *Journal of Labor Economics*, 4, (April 1986): 193-211

Blanchflower, David and Alex Bryson. “Changes over time in union relative wage effects in the UK and the US Revisited”, forthcoming as chapter 7 in John T. Addison and

Claus Schnabel (eds.), *International Handbook of Trade Unions,* Edward Elgar, Cheltenham England and Northampton Mass., USA, 2003

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

______________________ and Mark Bryan. "The Union Membership Wage-Premium Puzzle: Is There a Free Rider Problem", CEPR Discussion Paper No. 2879, 2001

Blundell, Richard and Monica Costa Dias. *Alternative Approaches to Evaluation in Empirical Microeconomics* Centre for Microdata Methods and Practice, Working Paper CWP10/02, available at http://cemmap.ifs.org.uk/working_papers.shtml, 2002

Bryson, Alex. *The Union Membership Wage Premium: An Analysis Using Propensity Score Matching*, Centre for Economic Performance Discussion Paper No. 530, London School of Economics, 2002

______________________, Lorenzo Cappellari, and Claudio Lucifora. "Why So Unhappy? The Effect of Union Membership on Job Satisfaction", Centre for Economic Performance Working Paper No. 1183, London School of Economics, 2002

______________________ and Raphael Gomez. "You Can't Always Get What You Want: Frustrated Demand for Unionisation in Britain", Centre for Economic Performance Working Paper #1182, London School of Economics, 2002a.

______________________ "Marching on together? Recent trends in union membership". In Alison Park, John Curtice, Katarina Thomson, Lindsay Jarvis, and Catherine Bromley (eds.) *British Social Attitudes: the 19th Report*, London: Sage, 2002b.

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

Dehejia, Rajeev H. and Sadek Wahba. "Causal Effects in Nonexperimental Studies: Reevaluating the Evaluation of Training Programs'. *Journal of the American Statistical Association*. 94(1999): 1053-1062

Farber, Henry. "Notes on the Economics of Labor Unions". Working Paper No.452. Princeton University, Industrial Relations Section, 2001.

Forth, John and Neil Millward. "Union effects on pay levels in Britain". *Labour Economics* 9 (2002): 547-561

Freeman, Richard B. "Longitudinal Analyses of the Effects of Trade Unions". *Journal of Labor Economics* 2 (January 1984): 1-26

______________________and James L. Medoff. *What Do Unions Do?*, New York, Basic Books, 1984

Frölich, Markus, Almas Heshmati and Michael Lechner. “A Microeconomtric Evaluation of Rehabilitaion of Long-term Sickness in Sweden”, St. Gallen Working Paper, 2001

Green, Francis. “The trade union wage gap in Britain: some new estimates”. *Economic Letters* 27 (1988): 183-187.

Heckman, James, Hidehiko Ichimura, Jeffrey Smith and Petra Todd. “Characterizing selection bias using experimental data”. *Econometrica* 66(1998): 1017-1098.

______________________, Hidehiko Ichimura, and Petra Todd. “Matching as an econometric evaluation estimator: evidence from evaluating a job training programme”. *Review of Economic Studies* 64 (1997): 605-654.

______________________, Robert LaLonde and Jeffrey Smith. “The Economics and Econometrics of Active Labor Market Programs” in Orley Ashenfelter and David Card (eds.) *The Handbook of Labour Economics*, Vol. III, Amsterdam: North Holland, 1999

Hildreth, Andrew K. G. “What has happened to the union wage differential in Britain in the 1990s”. *Oxford Bulletin of Economics and Statistics*, 61 (1999): 5-31

______________________ “Union Wage Differentials for Covered Members and Nonmembers in Great Britain”. *Journal of Labor Research* 21, (Winter 2000): 133-147.

Hirsch, Barry T. and David A. Macpherson *Union Membership and Earnings Data Book: Compilations from the Current Population Survey (2000 Edition).* Washington, D.C.: Bureau of National Affairs, 2000.

______________________ and Edward J. Schumacher. "Match Bias In Wage Gap Estimates Due To Earnings Imputation", Mimeograph, Trinity University, available at www.trinity.edu/bhirsch/ or www.ssrn.com, 2002.

______________________. "Private Sector Union Density and the Wage Premium: Past, Present and Future". *Journal of Labor Research* 22 (2001): 487-518.

Holland, Paul W. "Statistics and Causal Inference". *Journal of the American Statistical Association*, December, 81 (1986): 945-960

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

Lewis, H. Gregg. *Union Relative Wage Effects: A Survey*. University of Chicago Press, Chicago, Illinois, USA, 1986

Machin, Steve. "Does it still pay to be in or to join a union?" LSE Working Paper, 2001.

Millward, Neil, Alex Bryson and John Forth. *All Change at Work? British Employment Relations, 1980-98, as portrayed by the Workplace Industrial Relations Survey Series*, Routledge, London, 2000

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

Rosenbaum, Paul R. and Donald B. Rubin. “The central role of the propensity score in observational studies for causal effects”. *Biometrica* 70 (1983): 41-50.

______________________ “Constructing a Control Group Using Multivariate Matched Sampling Methods that Incorporate the Propensity Score”. *The American Statistician* 39 (1985): 33-38

Schumacher, Edward J. “What Explains Wage Differences between Union Members and Covered Nonmembers?”. *Southern Economic Journal*, 65(1999): 493-512

Sianesi, Barbara. “An Evaluation of the Active Labour Market Programmes in Sweden”. IFAU Working Paper #2001: 5

Stewart, Mark. “Collective bargaining arrangements, closed shops and relative pay”, *Economic Journal*, 97 (1987): 140-155

_______________ "Union wage differentials in an era of declining unionisation", *Oxford Bulletin of Economics and Statistics*, 57 (1995): 143-66

Swaffield, Joanna K. "Does measurement error bias fixed-effects estimates of the union wage effect?" *Oxford Bulletin of Economics and Statistics*, 63 (4, 2001): 437-457

Wessels, W. (1994), 'Do Unionized Firms Hire Better Workers?', Economic Inquiry, 32 (4), 616-629.

Technical Appendix: Causal inference through statistical matching

To establish whether the union membership wage premium is due to membership, or is due to systematic differences in personal, job and workplace characteristics across members and non-members, we need to isolate the causal effect of union membership on wages. Let us conceive of union membership as if it were a 'treatment' that the individual receives. We wish to evaluate the causal effect of this treatment (treatment *1*) relative to non-membership (treatment *0*) on an outcome variable, *Y*, gross earnings. Let Y_1 be earnings if the individual received treatment 1 (that is, where the individual is a union member) and Y_0 be the earnings that would result if the same individual received treatment 0 (non-membership). Let us denote the binary indicator of the treatment actually received as $D \in \{0,1\}$, while *X* is a set of attributes which are not affected by the treatment (demographic, job and workplace-related).

The effect of treatment 1 on individual *i* as measured by *Y* and relative to treatment 0 is:

$$? = Y - Y_{0I} \qquad \text{(A1)}$$

which is simply the difference between the individual's potential outcome if 'exposed' to membership and the individual's potential outcome from non-membership. To estimate the impact of membership on members' earnings, it is necessary to know what the outcome would have been if the individual had *not* been a member. The problem is that we can not observe the counterfactual, namely the outcome which would have resulted if an individual had made an alternative choice (that is, if members had chosen non-membership, and *vice versa*). Either Y_{1i} or Y_{0i} is missing for each i . Thus, our problem is one of estimating missing data. This counterfactual cannot be inferred directly from the outcomes of non-members since they are likely to differ substantially in their characteristics from members. To overcome this selection problem, researchers must choose from a range of evaluation methods, the choice being determined by a number of factors including the richness of the data and the nature of the treatment. Because it is impossible to observe the individual treatment effect, each method relies on generally untestable assumptions to make causal inferences (Holland, 1986). In order to identify individual treatment effects, it is necessary to make very strong assumptions about the joint distribution of Y_{1i} and Y_{0i}. However, the *average* treatment effect at the population or sub-population level can be identified under generally less stringent assumptions, some of which are set out below. Among the parameters that only depend on the marginal distributions of Y_{1i} and Y_{0i} is the parameter most commonly estimated and the one estimated in this paper, namely the mean impact of treatment on the treated:

$$q = E(Y_1 - Y_0 \mid D = 1, X) \quad = E(Y_1 \mid D = 1, X) - E(Y_0 \mid D = 1, X) \qquad \text{(A2)}$$

where D=*1* denotes treatment (membership), D=*0* denotes non-treatment (non-membership) and X is a set of conditioning variables. In assessing the expected treatment effect for individuals who are union members, we are addressing the question of how members' earnings compare with what they would have received had they not been members, on average.[17]

For members we observe Y_1 so that the average observed outcome for participants is an unbiased estimate of the first component of the effect of treatment on the treated $E(Y_1 \mid D = 1, X)$. The evaluation problem arises from the term $E(Y_0 \mid D = 1, X)$. This is the mean of the counterfactual which, since it is unobservable, must be identified and estimated on the basis of some usually untestable identifying assumptions justifying the use of the observable pairs $(Y_1, D = 1), (Y_0, D = 0)$.

Members may not be a random sample of all employees. If there are systematic differences in characteristics across members and non-members that are likely to influence earnings, failure to take account of these will bias any estimate of the union membership effect on earnings. Thus, $E(Y_1 \mid D = 1) - E(Y_0 \mid D = 0)$ would in general be biased for the effect of treatment on the treated. An exception is when the independence assumption $Y_0 \perp D$ can be invoked. This is credible where the random assignment of individuals to treatment ensures that potential outcomes are independent of treatment status. In this situation, $E(Y_0 \mid D = 1) = E(Y_0 \mid D = 0) = E(Y \mid D = 0)$ so that the treatment effect can be consistently estimated by the difference between the observed mean of the outcome variable for the treatment group and the observed mean for the non-treatment group.

In the absence of random assignment, one option is to construct a comparison group based on statistical matching. Matching estimators try to resemble an experiment by choosing a comparison group from all non-participants such that the selected group is as similar as possible to the treatment group in observable characteristics. Matching can yield

[17] To obtain the average treatment effect on the non-treated $E(Y_1 - Y_0 \mid D = 0)$ the procedure is applied symmetrically. The average treatment effect $E(Y_1 - Y_0)$ is a weighted average of the treatment effects for the treated and non-treated.

unbiased estimates of the treatment impact where differences between individuals affecting the outcome of interest are captured in their observed attributes. This assumption, which is often referred to as the Conditional Independence Assumption (CIA), is the key identifying assumption underpinning the matching methodology. The precise form of the CIA depends on the parameter being estimated. For the treatment on the treated parameter, the CIA requires that, conditional on observable characteristics, potential non-treatment outcomes are independent of treatment participation. Formally,

$$E(Y_0 \mid X, D = 1) = E(Y_0 \mid X, D = 0) \tag{A3}$$

Thus, CIA requires that the chosen group of matched controls does not differ from the group of treated by any variable which is systematically linked to the non-participation outcome Y_0, other than on those variables that are used to match them. This permits the use of the matched non-participants to measure how participants would have fared, on average, had they not participated.

The plausibility of the CIA depends on the informational richness of the data since the set of X's should contain all the variables thought to influence *both* participation (that is, membership) *and* the outcome (earnings) in the absence of participation. We discuss how likely it is that the CIA is met in this analysis in Sections 3 and 4.

Under CIA,

$$E(Y_1 \mid D = 1) - E(Y_0 \mid D = 1) = E_{x|D=1}\{E(Y|X, D = 1) - E(Y|X,D = 0)\} \tag{A4}$$

Hence, after adjusting for observable differences, the mean of the no-treatment (potential) outcome is the same for those receiving treatment as for those not receiving treatment. This allows non-participants' outcomes to be used to infer participants' counterfactual outcomes. However, this is only valid if there are non-participants for all participants' values of X (this is known as the support condition):

$$Pr\,(D = 1 \mid X) < 1 \tag{A5}$$

This ensures that all treated individuals have a counterpart in the non-treated population for each X for which we seek to make a comparison. If there are regions where the support of X does not overlap for the treated and non-treated groups, matching can only be performed, and the treatment parameter, q, retrieved, over the common support region. If treated individuals have no support in the non-treated population, they are dropped from analysis and the estimated treatment effect is redefined as the mean treatment effect for those treated falling within the common support.

Matching operates by constructing, for those participants with support, a counterfactual from the non-participants. There are a number of ways of defining this counterfactual (Heckman et al., 1997). Once the counterfactuals are identified, the mean impact of union membership can be estimated as the mean difference in the outcomes of the matched pairs.

A refinement to the matching approach was introduced by Rosenbaum and Rubin (1983). If the CIA is met and there is common support then:

$$Y_0 \perp D \mid P(X) \text{ for } X \text{ in } \mathbf{X} \tag{A6}$$

where $P(X)$ is the propensity score, the conditional probability of participating in the treatment – in our case, the probability of being a union member – given a vector of observed characteristics X.[18] Formally,

$$P(X_i) = Pr(D_i = 1 \mid X_i) \qquad \text{(A7)}$$

Rosenbaum and Rubin show treatment and the observed covariates are conditionally independent given the propensity score, that is:

$$D_i \perp X_i \mid P(X_i) \qquad \text{(A8)}$$

The advantage of Rosenbaum and Rubin's innovation is that the dimensionality of the match can be reduced to one. Rather than matching on a vector of characteristics, it is possible to match on just the propensity score. This is because, as Rosenbaum and Rubin show, by definition treatment and non-treatment observations with the same value of the propensity score have the same distribution of the full vector of regressors X. Having matched on the propensity score, the mean impact of union membership is estimated as the mean difference in the outcomes of the matched pairs.

[18] $P(X)$ is shorthand notation for $P(D=1|X)$.

Appendix Table 1: Individual union membership status in the private sector

Demographics:	
Age (ref.: under 20)	
20-24 years	.260 (2.12)
25-29 years	.463 (4.11)
30-39 years	.736 (6.54)
40-49 years	.834 (7.03)
50+ years	.772 (6.39)
Highest academic qualification (ref: none)	
CSE	.069 (1.04)
GCSE	-.049 (0.76)
A-level or equivalent	-.095 (1.27)
Degree or post-graduate	-.255 (2.71)
Female	.008 (0.14)
Married or living as married	.074 (1.61)
Health problem	.065 (0.84)
Member of non-white ethnic group	.155 (1.70)
Job-related:	
Occupational classification (ref.: operative)	
Manager/senior administrator	-1.084 (8.27)
Professional	-.588 (5.39)
Associate professional and technical	-.500 (4.30)
Clerical and secretarial	-.994 (9.01)
Craft and skilled service	-.139 (1.87)
Personal and protective service	-.952 (6.43)
Sales	-.472 (5.13)
Other unskilled occupations	-.649 (7.27)

Permanent contract	.209 (1.92)
Hours worked (continuous)	.017 (2.50)
Hours worked squared	-.000 (1.48)
Occupation performed solely by men	.115 (1.98)
Workforce composition:	
Percentage female is <25%	.064 (0.97)
Percentage part-time is <10%	-.165 (2.34)
No workers aged under 20 years	-.055 (0.86)
No manual workers	.249 (3.33)
Workplace:	
Union recognised for pay bargaining	.450 (5.33)
Union density	.026 (19.63)
Size (ref: 10-99 employees)	
100-199 employees	.076 (1.10)
200-499 employees	.279 (3.78)
500+ employees	.326 (4.18)
Foreign-owned	-.042 (0.65)
Single independent establishment	-.153 (2.09)
Workplace activity (ref: producers of goods/services for consumers, producers for other parts of organisation, non-producers)	
Administrative office only	.043 (0.36)
Supplier to other companies	-.168 (2.82)
Industrial classification (ref.: manufacturing, utilities, construction)	
Wholesale and retail distribution	.040 (0.49)
Hotels and Restaurants	.547 (3.46)
Transport and communication	.084 (0.81)
Financial Services	.275 (2.57)
Other business services	.197 (1.67)
Other	.528 (4.48)

Location (ref: East, East Midlands, London, South East, Yorkshire and Humberside, North East)	
North	-.184 (1.74)
North West	.076 (0.88)
Scotland	-.027 (0.35)
South West	-.093 (1.03)
Wales	.082 (0.51)
West Midlands	-.017 (0.22)
Local labour market conditions:	
Unemployment rate of 5%+	.155 (3.21)
Constant	-2.689 (11.46)
Observations	14,531
F-stat	51, 972 = 51.41 Prob >f 0.0000

Note: Absolute value of t-statistics in parentheses

Appendix Table 2: Imbalance in means between treated and matched comparators, plus standardised differences (%)

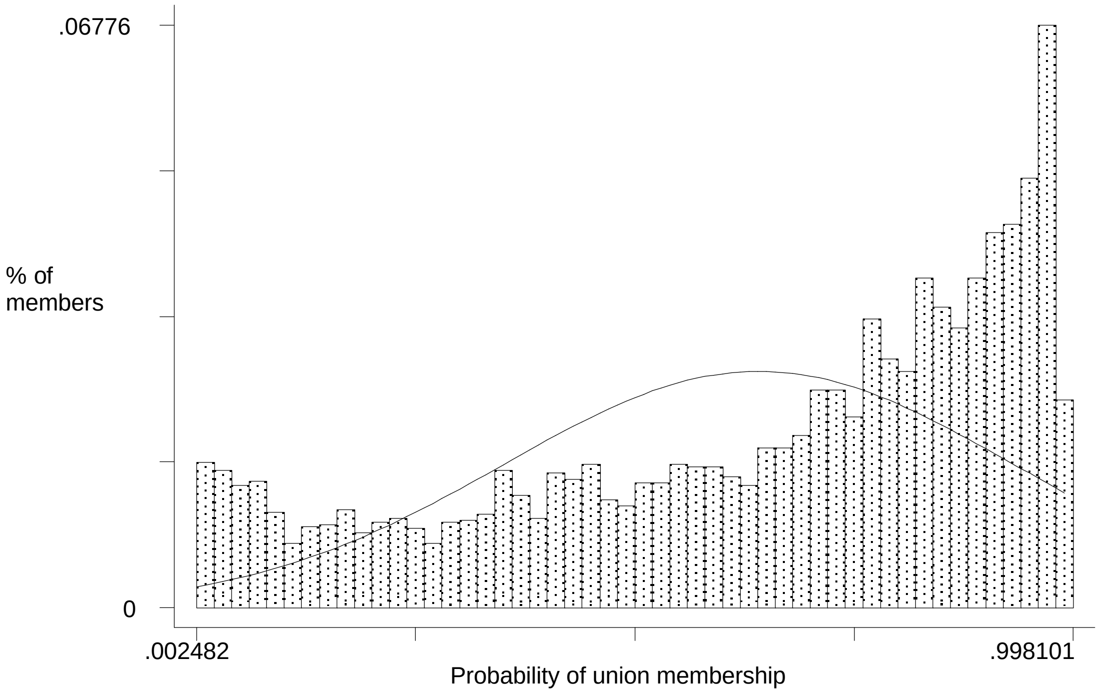
	Non-members pre-match	Non-members matched	Members	% bias before match	% bias after match
Age					
20-24 years	.11	.04	.04	-30.35	-2.23
25-29 years	.16	.12	.11	-12.96	-1.03
30-39 years	.27	.32	.31	9.86	-1.32
40-49 years	.20	.28	.30	23.08	4.56
50+ years	.20	.24	.23	8.03	-1.47
Highest academic qualification					
CSE	.12	.13	.13	4.45	1.20
GCSE	.28	.27	.28	-.76	.95
A-level or equivalent	.17	.16	.15	-4.52	-3.01
Degree or post-graduate	.22	.15	.15	-18.64	.32
Female	.50	.38	.35	-31.41	-5.73
Married or living as married	.64	.75	.76	28.84	3.81
Health problem	.05	.07	.06	7.60	-2.09
Member of non-white ethnic group	.04	.04	.03	-1.28	-4.89
Occupational classification					
Manager/senior administrator	.15	.09	.08	-21.80	-1.03
Professional	.09	.09	.09	-.66	2.09
Associate professional and technical	.08	.13	.10	6.89	-13.27
Clerical and secretarial	.23	.19	.19	-10.27	2.21
Craft and skilled service	.08	.18	.18	31.73	.75

Personal and protective service	.08	.02	.02	-25.75	-.34
Sales	.13	.07	.07	-19.12	-1.01
Other unskilled occupations	.09	.06	.06	-12.91	.58
Permanent contract	.94	.95	.97	15.29	8.43
Hours worked (continuous)	37.54	39.14	39.69	19.94	4.83
Hours worked squared	1578.35	1635.27	1671.84	12.20	4.24
Occupation performed solely by men	.15	.25	.27	31.56	4.34
Percentage female is <25%	.26	.45	.47	47.71	5.68
Percentage part-time is <10%	.53	.61	.65	25.79	7.16
No workers aged under 20 years	.22	.29	.30	17.85	3.02
No manual workers	.21	.22	.22	2.22	.67
Union density	11.81	59.97	60.14	192.45	.68
Union recognition	.26	.87	.87	159.89	2.00
Workplace size					
100-199 employees	.18	.22	.21	6.02	-3.17
200-499 employees	.17	.35	.33	38.91	-4.95
500+ employees	.08	.19	.18	30.59	-3.69
Foreign-owned	.18	.21	.21	6.59	-1.58
Single independent establishment	.29	.13	.10	-49.32	-6.89
Workplace activity					
Administrative office only	.07	.03	.04	-11.75	6.33
Supplier to other companies	.34	.25	.24	-22.70	-1.44
Industrial classification					
Wholesale and retail distribution	.23	.11	.11	-33.94	-.34

Hotels and Restaurants	.07	.02	.01	-29.16	-1.01
Transport and communication	.05	.13	.12	29.41	-5.56
Financial Services	.07	.13	.13	20.21	0
Other business services	.16	.04	.03	-44.07	-3.74
Other	.15	.11	.09	-17.32	-5.64
Location					
North	.04	.07	.08	16.28	7.07
North West	.08	.13	.13	18.85	-.49
Scotland	.09	.12	.11	8.65	-2.10
South West	.09	.10	.11	5.53	1.84
Wales	.04	.05	.05	4.52	.13
West Midlands	.08	.10	.09	4.86	-4.07
Unemployment rate of 5%+	.48	.60	.56	16.82	-8.47
Average absolute standardised bias pre-match, whole sample					24.54
Average absolute standardised bias post-match, whole sample					3.21
Average absolute standardised bias pre-match, matched sample					8.90
Average absolute standardised bias post-match, matched sample					-.35
Absolute bias reduction					61.29

Graph A1: Predicted union membership probability for members and non-members, whole private sector

Members



Non-members