

Hirsch, Barry T.

**Working Paper**

## Reconsidering Union Wage Effects: Surveying New Evidence on an Old Topic

IZA Discussion Papers, No. 795

**Provided in Cooperation with:**

IZA – Institute of Labor Economics

*Suggested Citation:* Hirsch, Barry T. (2003) : Reconsidering Union Wage Effects: Surveying New Evidence on an Old Topic, IZA Discussion Papers, No. 795, Institute for the Study of Labor (IZA), Bonn

This Version is available at:

<https://hdl.handle.net/10419/20040>

**Standard-Nutzungsbedingungen:**

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

**Terms of use:**

*Documents in EconStor may be saved and copied for your personal and scholarly purposes.*

*You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.*

*If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.*

IZA DP No. 795

## Reconsidering Union Wage Effects: Surveying New Evidence on an Old Topic

Barry T. Hirsch

June 2003

# Reconsidering Union Wage Effects: Surveying New Evidence on an Old Topic

**Barry T. Hirsch**

*Trinity University and IZA Bonn*

Discussion Paper No. 795  
June 2003

IZA

P.O. Box 7240  
D-53072 Bonn  
Germany

Tel.: +49-228-3894-0  
Fax: +49-228-3894-210  
Email: [iza@iza.org](mailto:iza@iza.org)

This Discussion Paper is issued within the framework of IZA's research area *Welfare State and Labor Market*. Any opinions expressed here are those of the author(s) and not those of the institute. Research disseminated by IZA may include views on policy, but the institute itself takes no institutional policy positions.

The Institute for the Study of Labor (IZA) in Bonn is a local and virtual international research center and a place of communication between science, politics and business. IZA is an independent, nonprofit limited liability company (Gesellschaft mit beschränkter Haftung) supported by the Deutsche Post AG. The center is associated with the University of Bonn and offers a stimulating research environment through its research networks, research support, and visitors and doctoral programs. IZA engages in (i) original and internationally competitive research in all fields of labor economics, (ii) development of policy concepts, and (iii) dissemination of research results and concepts to the interested public. The current research program deals with (1) mobility and flexibility of labor, (2) internationalization of labor markets, (3) welfare state and labor market, (4) labor markets in transition countries, (5) the future of labor, (6) evaluation of labor market policies and projects and (7) general labor economics.

IZA Discussion Papers often represent preliminary work and are circulated to encourage discussion. Citation of such a paper should account for its provisional character. A revised version may be available on the IZA website ([www.iza.org](http://www.iza.org)) or directly from the author.

## ABSTRACT

### **Reconsidering Union Wage Effects: Surveying New Evidence on an Old Topic**

I examine evidence on private sector union wage gaps in the U.S. The consensus opinion among labor economists of an average union premium of roughly 15 percent is called into question. Two forms of measurement error create a downward bias in standard wage gap estimates. Match bias results from Census earnings imputation procedures that do not include union status as a match criterion. Downward bias is roughly equal to the proportion of workers with imputed earnings, currently about 30 percent. Misclassification of union status causes additional attenuation in union gap measures. This bias has worsened as private sector density has declined, since an increasing proportion of workers designated as union are instead nonunion workers. Corrections for misclassification and match bias lead to estimated union gaps substantially higher than standard estimates, but with less of a downward trend since the mid-1980s. Private sector union gaps corrected for these biases are estimated from the CPS for 1973-2001. The uncorrected estimate for 2001 is .13 log points. Correction for match bias increases the gap to .18 log points; further correction for misclassification bias, based on an assumed 2 percent error rate, increases the gap to .24. Reexamination of the skill-upgrading hypothesis leads to the conclusion that higher union gap estimates are plausible. The conventional wisdom of a 15 percent union wage premium warrants reexamination.

JEL Classification: J3, J5, C81

Keywords: union wage differentials, measurement error, hot deck imputation, match bias

Barry T. Hirsch  
Department of Economics  
Trinity University  
San Antonio, TX 78212-7200  
USA  
Tel.: +1 210 999 8112  
Fax: +1 210 999 7255  
Email: [bhirsch@trinity.edu](mailto:bhirsch@trinity.edu)

## I. *Introduction*

The effect of labor unions on wages is one of the more heavily studied topics in empirical labor economics. There exists a reasonably strong consensus among economists that the average union-nonunion wage gap is about 15 percent, a consensus based in no small part on the influence of work by H. Gregg Lewis (1963, 1986). A survey of labor economists at leading universities by Fuchs, Krueger, and Poterba (1998) asked the question: “What is your best estimate of the percentage impact of unions on the earnings of their average member?” The median response was 15 percent and mean response 13.1 percent; dispersion was low. A meta-analysis by Jarrell and Stanley (1990), reexamining the studies surveyed by Lewis (1986), places the range of average union wage gaps at 9-12 percent, a bit lower than Lewis’s summary of these same studies, with union gaps rising with unemployment. Likewise, labor economics textbooks place the union wage gap in the range of 10-20 percent.

This paper interprets recent evidence on the union impact on wages. It makes no attempt to provide a comprehensive survey of union wage gap studies. Rather, the paper focuses in depth on recent strands in the literature, with particular attention given to work by the author. The evidence on union wage effects is found to be far less clear-cut than might be expected given the consensus among labor economists. In the end, one concludes that the “Lewis consensus” centering on 15 percent (or less) is too low. Current private sector union wage gaps from the CPS appear well in excess of 20 percent, even following a gradual decline since the mid-1990s. How one arrives at this conclusion makes for an interesting story, one involving arguments and evidence not so well known among labor economists.

Although standard estimates of the union gap in the U.S. are in the neighborhood of 15 percent, these estimates fail to account for two measurement problems in CPS data (as well as other data sets), each of which causes a downward bias to wage gap estimates. The first problem, misclassification error in the reporting of union status, is well understood by labor economists and a central focus of studies using *longitudinal* analysis. Misclassification (measurement error) bias has not been considered a serious concern for standard wage level studies. This assumption is no longer tenable. As true union density in the U.S. private sector has fallen to below 9 percent, the noise-signal ratio in union status has risen and downward bias from misclassification has become a serious problem. This paper provides corrected estimates of union wage gaps based on alternative assumptions about rates of misclassification (1, 2, and 3 percent).

The second source of downward bias (currently about 5 percentage points) is the result of imputation “match bias” (Hirsch and Schumacher, 2004). Workers who do not report earnings have them assigned (imputed) based on the earnings of a matched donor with a set of characteristics identical to the nonrespondent. CPS match criteria do *not* include union status, however, so most union workers who fail to report earnings are assigned the earnings of nonunion workers, while a few nonunion workers are assigned the earnings of union workers. Hirsch and Schumacher show that the proportional attrition in the union coefficient is roughly equal to the proportion of workers with imputed earnings, currently about 30 percent.

The combined effect of match bias and misclassification error is substantial – a downward bias in union wage gap estimates on the order of .10 log points, from roughly .25 to .15 in recent years. Match bias and misclassification error have roughly equal effects, although the magnitude of the latter bias depends critically on the assumed rate of misclassification, something about which we know far too little. Each of these sources of bias has become more severe over time, with rising imputation rates leading to increased match bias and falling private sector density leading to greater misclassification bias. Standard estimates of union wage gaps, therefore, substantially understate the level of the union wage gap, but overstate decline in the union wage gap since the mid-1980s. Independent of the biases addressed herein, the magnitude of the decline in the union wage gap is uncertain due to conflicting evidence from alternative data surveys (Hirsch et al., 2003).

My principal focus centers on the level and changes over time in the union-nonunion wage gap in the U.S. private sector, following accounting for measurement biases. The analysis is based on OLS wage level estimates, which raises the question of whether such methods are appropriate, given the possibility of unmeasured skills correlated with union status and skill upgrading by union employers who pay high wages. I discuss alternative estimation approaches, in particular selection models and longitudinal analyses accounting for worker fixed effects. Longitudinal evidence shows that selection is two-sided, with positive selection among workers with low measured skills and negative selection among workers with high measured skills (Card, 1996; Hirsch and Schumacher, 1998). As emphasized by Card, this form of two-sided selection is inconsistent with single-index selection models, calling into question gap estimates from such studies (Lee, 1978; Robinson, 1989).

The theoretical basis for the skill-upgrading hypothesis also warrants examination. Theory and

evidence suggest that unions may have little effect on *average* worker skills (Wessels, 1994). Union wage effects *following sorting* by employers and employees may differ little across worker skill groups, contrary to standard evidence. The average wage gap obtained across a broad sample of workers may be a better measure of the causal impact of unions on wages than are wage level estimates disaggregated by worker group, based on schooling, occupation, race, gender, etc. Absent good instruments for union status (i.e., predictors of union status that do not affect earnings), wage level analysis producing an average union wage effect may provide a reasonable approach for many (but not all) applications.

I first present an overview of union wage gap theory and estimation. This is followed by an analysis of two important measurement issues – match bias in wage gap estimates due to earnings imputation and bias from misclassification of union status. Estimates of union wage gaps that account for these biases are presented for the U.S. private sector for 1973-2001. That is followed by an appraisal of the skill-upgrading hypothesis and longitudinal union gap estimates. A final section examines briefly such topics as unions and earnings inequality, union wage effects in the public sector, differences in wage effects for union members and covered nonmembers, union effects on compensation (wages plus benefits), and the relationship between union status and working conditions.

## II. *Union Wage Gaps: Estimation Methods and Measurement Issues*

### A. *What Wage Effects Are Being Measured? Union Gaps and Gains*

Unions affect wages in a number of ways. Most important, union bargaining power leads to higher wages among covered workers. Collective bargaining leads to a more formalized governance structure and wage determination process, reducing management discretion. A result of this process is lower wage dispersion, evident through both a flattening of skill differentials and the standardizing of wages within job-seniority classifications (Freeman, 1980, 1982). As argued subsequently, the decrease in skill differentials is far less than suggested by standard wage regressions.

Unions also affect nonunion wages, albeit to a far lesser extent than union wages. On the one hand, “threat effects” lead nonunion employers to increase wages to deter union organizing. Threat effects are most apparent in sectors with highly unionized labor markets, defined by industry (e.g., airlines) or location. “Spillover” effects, on the other hand, produce downward pressure on nonunion wages. Such effects follow from a two-sector model in which the decrease in employment in the union sector increases labor supply in

the nonunion sector. A number of studies find evidence consistent with threat effects, based on a positive relationship between nonunion wages and industry union density. The weight of the evidence suggests that union wages grow with respect to industry density at a rate equivalent to or greater than nonunion wages, implying a constant or increasing union wage gap with respect to density. Evidence is less clear-cut with respect to metropolitan area density.<sup>1</sup>

Spillover effects are difficult to detect since they are widely dispersed. Were the negative effects on nonunion wages from spillovers sufficiently large to offset both the wage gains among union members and positive threat effects on nonunion wages, the total wage bill or labor's share of income should be unaffected by unionization. But isolating how union density affects labor's share is not easy. Aggregate time series data require that one distinguish the effect of changes in (highly trended) union density from that of other time-series variables that may influence labor's share. Considerable cross-sectional variation occurs in union density across industry, but this reflects differences among industries heavily affected by threat effects and those most affected by spillover effects, rather than their net effect economy-wide.<sup>2</sup>

Lewis makes a useful distinction between wage "gaps" and "gains." Gaps are wage differentials between union and nonunion workers. Gains are the effect of unionism on union wages, relative to what wages would be absent unionism. Were the net effect of unionism on nonunion wages zero, union gaps and gains would be equivalent. If union threat effects are dominant (thus raising nonunion wages), estimates of union gaps understate union gains. If spillover effects are dominant, gaps exceed gains. The focus in the literature (and in this paper) has been on estimates of the average union wage gap. Union gains are difficult to estimate given that we do not observe what wages would be in absence of unionization or the institutional structure (e.g., labor law) that enables unions. One can estimate union and nonunion wages including a control for union density (say at the industry level), thus providing estimates of union gaps and gains for each level of union density (e.g., Freeman and Medoff, 1981). Lewis (1986) is skeptical about the reliability of this approach. He is concerned that union density coefficients capture more than the effects of unionization

---

<sup>1</sup> One of the first micro-level wage equation papers to include measures of union density at the industry level was Freeman and Medoff (1981). Curme and Macpherson (1991) include union density in union and nonunion wage equations at the industry and metropolitan area levels.

<sup>2</sup> Macpherson (1990) examines union effects on labor's share based on cross-sectional variation in union density within manufacturing for the years 1973-1975 and 1983-1985. He obtains a positive relationship in each of the periods and concludes that unionization increases labor's share.



(for example, employer size), as seen by their sensitivity with respect to choice of control variables.

Finally, it is worth clarifying terminology. A term used frequently in the literature is union wage *premium*. The terms premium and gap are largely interchangeable, typically referring to the excess of union wages over wages for similar nonunion workers and jobs. As a matter of taste, I prefer the term gap; the term premium is used sparingly.

## B. *Specification*

My focus is on union gaps estimated with cross-sectional wage level regressions. A subsequent section discusses longitudinal gap estimates and variation in wage effects across worker groups. The obvious starting point is the standard semilogarithmic wage equation. This earnings function can be interpreted, rather heroically, as a structural human capital model (Mincer, 1974; Willis, 1986) or, more generally, as a reduced form price equation in which supply and demand determinants, including union status, are explanatory variables. The semilog form of the earnings function is justified by theory (both human capital and probabilistic descriptions of the earnings generations process) and statistical evidence on functional form (Heckman and Polachek, 1974). Given a semilog wage equation, a natural metric for measuring union gaps is the log wage differential, which is simply the difference in mean union log wages and mean nonunion log wages, conditional on worker and (ideally) job characteristics. Although log wage gaps can be readily converted to (approximate) percentage differentials, I present gaps in log point form.<sup>3</sup>

A logical starting point is a general model of both wage and union status determination.

$$U = Z\Psi + \gamma(W_u - W_n) + \mu_1 \tag{1}$$

$$W_u = X_u\beta_u + \mu_2 \tag{2}$$

$$W_n = X_n\beta_n + \mu_3. \tag{3}$$

Here  $U$  is union status; subscripts  $u$  and  $n$  designate union and nonunion;  $W$  is the individual log wage;  $X$  is a vector of wage determinants;  $\beta$  is the corresponding coefficient vector; and  $\mu$  is the error term. Interpreted as a structural union choice equation, individual workers select a  $U$  job conditional on a vector  $Z$  of individual characteristics and their relative wage opportunities in union and nonunion jobs. As emphasized by Abowd and Farber (1982), workers must select the union job queue *and* be selected from the queue by union

---

<sup>3</sup> A simple approximation of the percentage gap is  $[\exp(\beta-1)]100$ , where  $\beta$  is the log wage gap. Giles (1982) provides a comparison between this simple measure and more precise approximations.

employers. Thus, the union status equation might better be interpreted as a reduced form equation in which vector  $Z$  includes job characteristics (employer size, working conditions, etc.) that influence the benefits and costs of union organizing and provision of union services (Hirsch and Berger, 1984).

Although reliable estimation of the general model is not typically possible, it is useful to think about what such a model involves, if for no other reason than to help understand the limitations inherent in models estimated in the literature. In the Heckman-Lee selection approach (Lee, 1978; Duncan and Leigh, 1980; Hirsch and Berger, 1984; Robinson, 1989), a reduced form union status probit equation is first estimated (i.e., with  $W_u - W_n$ , the log wage gap, excluded), from which selectivity variables for the union and nonunion wage equations are constructed (so-called  $\lambda$ 's). In principle, these are intended to yield unbiased estimates of wage equation parameters and thus each worker's wage in the union and nonunion sectors. Once such information is available, one can calculate both "unbiased" union-nonunion wage gaps and "structural" estimates of the union status equation (i.e., with inclusion of  $W_u - W_n$ ). Such models can be identified and estimated in a statistical sense, based on (sometimes arbitrary) exclusion of a union status determinant from the wage equations or due to nonlinearity in the probit equation error term.

The reliability of results obtained from selection models has been questioned by prominent labor economists (Lewis, 1986; Freeman and Medoff, 1984). Skepticism initially stemmed from the enormous variance of estimates produced by such methods and because the high gap estimates sometimes obtained in such studies violated most economists' priors that skill upgrading should result in selectivity-adjusted estimates that are lower than standard estimates. Robinson (1989) provides a careful defense of such methods, but also obtains gap estimates exceeding those obtained from standard (nonselection) models.<sup>4</sup>

Selection model estimates of the union gap are unreliable for two principal reasons. First, economically meaningful identification of the selection model requires that there be at least one variable that determines union status but not the wage. Unfortunately, there are few obvious candidates among available measures. A less recognized but no less important shortcoming is that the selection model (as typically employed in the literature) is incompatible with the two-sided selection process that characterizes union status determination. Longitudinal evidence (Section V) implies that there is positive selection into unions

---

<sup>4</sup> Robinson (1989) offers the explanation that skills are heterogenous, with union workers having an absolute advantage in skills in union jobs and nonunion workers an absolute advantage in nonunion jobs.

(i.e., positive unmeasured attributes) among workers with low levels of measured skill and negative selection among workers with high levels of measured skill. As discussed by Card (1996), however, selection models used in the literature produce identical selectivity terms for those with equivalent probabilities of being a union member. Those in the bottom and top of the skill distribution have roughly equal union density rates (Card, 1996), implying that these workers are assigned similar selectivity adjustment terms, wholly inconsistent with there being two-sided selection. It is not surprising that selection models produce unreliable estimates of union wage effects.<sup>5</sup>

Absent the ability to estimate the full model shown above, a natural approach would be to “instrument” union status in order to avoid bias in the wage equations due to unmeasured wage determinants correlated with union status. But such an approach does not avoid the inherent problem of identifying appropriate instruments (i.e., good predictors of union status not correlated with unmeasured skills or other wage determinants). Apart from the issue of unmeasured skills, a union wage effect identified through IV methods would strictly apply only to the exogenous variation in union status associated with that particular variable(s) (Angrist and Krueger, 2001). It is uncertain whether such an estimate would be close to the average gap obtained (theoretically) with respect to all exogenous variation in union status. Despite these concerns, “discovery” of appropriate instruments, or other sources of exogenous variation in union status, would provide an important step toward estimation of unbiased union gaps.

In practice, union wage gaps are typically estimated with union status treated as exogenous. Union wage gaps can be and sometimes are estimated based on separate union and nonunion wage equations (Bloch and Kuskin, 1978). Decomposition of the union-nonunion wage differential into portions *explained* by differences in the  $X$ 's and a portion that is a union wage premium (i.e., due to different coefficients or payoffs in the union sector), begins with the following (nonunique) identity:

$$w_u - w_n = (x_u - x_n)\beta_n + (\beta_u - \beta_n)x_u \quad (4)$$

where lowercase  $w$  and  $x$  now represent *mean* values. Here the total union wage differential  $w_u - w_n$  is

---

<sup>5</sup> Closely related critiques of the selection model, as applied to union wage gap estimates, are provided in Farber (1983) and Lemieux (1998). It is not surprising that estimation of a structural model is difficult, given the complexity of the process by which union status is determined. There exists both worker demand for union services and a supply curve of union services based on the costs of organizing and satisfying the preferences of rank and file (Pencavel, 1971; Lazear, 1983; Hirsch and Addison, 1986, Ch. 2). Union employment requires both that workers be in the union queue and be

shown as the sum of the portion explained by differences in the  $X$ 's, evaluated using nonunion betas, and the union wage premium or gap, measured by union mean characteristics ( $X$ 's) times the difference in coefficients (note that this portion includes the difference in intercepts). Coefficients (apart from the intercept) tend to be smaller or “flatter” in the union than in the nonunion sector due to greater wage standardization in the union sector and because of greater homogeneity in the union work force for a given set of measured characteristics. In calculating the union gap (the right-hand-side term in (4)), “flatter” union betas are more than offset by a larger union intercept.<sup>6</sup>

The simplest and most common method for obtaining union wage gap estimates is the single equation approach:

$$W = X\beta + \Gamma U + \mu. \tag{5}$$

This is simply a restricted form of the general model, making the assumptions that union status is exogenous and slope coefficients equal, with the union gap  $\Gamma$  measured by the difference in the union and nonunion intercepts. For estimates from the private sector, this simple method produces an estimate not too different from that obtained from the right-hand-side term in equation (4). Estimates of the private sector wage gaps shown in subsequent sections will be based on a single equation (by year), as seen in equation (5).

### C. *Choice of Control Variables*

Estimates of the union gap can differ depending on the choice of control variables. That being said, most wage equations contain a fairly standard set of controls and, with a few exceptions noted below, estimates do not vary too greatly because of differences in controls (for an analysis, see Lewis, 1986; Blanchflower and Bryson, 2002). One of the more important choices is the use of occupation and industry controls (dummies). In general, control for occupation (even at a very broad level) is associated with higher union wage gap estimates, since union workers are concentrated in lower skill broad occupational categories. Detailed

---

selected by employers, with a wage premium creating excess demand for union jobs (Abowd and Farber, 1982). There is no presumption that there exists a price (dues) based equilibrium as in standard markets.

<sup>6</sup> The decomposition shown by (4) is not unique; for example, one can decompose the wage differential by:

$$w_u - w_n = (x_u - x_n)\beta_u + (\beta_u - \beta_n)x_n.$$

This provides a less intuitive decomposition, since the explained portion is based on the wage structure in the union sector and the union premium is measured by the additional wage that workers with nonunion characteristics would earn if paid according to the union wage structure. One could also take a weighted average of these two measures, letting the explained portion be the employment (or sample) weighted average of  $\beta_u$  and  $\beta_n$  times  $(x_u - x_n)$  and the remainder of the differential a measure of the gap. Owing to a large nonunion share, the weighted average approach would produce a premium estimate for the U.S. private sector close to the estimate obtained in (4).

controls for industry, by contrast, tend to lower union gap estimates, since industry controls capture the effects of unions on industry-wide wages and may control for compensating differentials associated with adverse working conditions.

Blanchflower and Bryson (2002), for example, omit occupation dummies and include relatively detailed industry (and state) dummies. Using otherwise similar data and estimation methods, they obtain lower CPS union wage gap estimates than do Hirsch and Schumacher (2004), who include broad occupation and industry dummies. My own preference is that broad occupation be included and interpreted as a crude control for unmeasured skills, since the schooling and potential experience variables available in the CPS are incomplete measures of human capital. Supporting this proposition is Hirsch and Schumacher's (1998) evidence that inclusion of occupation dummies sharply increases estimated union gaps in wage level equations but have little effect in wage change equations where unmeasured worker-specific skills (fixed effects) have been netted out.

Including industry dummies are appropriate to the extent that they measure differences in working conditions or skills not captured by occupation or other controls.<sup>7</sup> They also may be appropriate as controls if substantial industry rents exist independent of union coverage. The concern with inclusion of *detailed* industry dummies is that they entangle industry differentials and union wage effects. For example, union bargaining power within industries included in the Census transportation category (e.g., airlines, railroads, the U.S. Postal Service) increases wages for union and nonunion workers. With detailed industry controls, one roughly makes a comparison of wages between union members and nonmembers within each detailed industry (since all get "credit" for the industry dummy), whereas absent industry dummies one makes a comparison between union and nonunion workers with similar characteristics within broad industries.<sup>8</sup>

Two controls typically excluded from regression estimates of union wage gaps are employer size (company and establishment size) and years of tenure with a worker's current employer. Neither of these

---

<sup>7</sup> As seen subsequently, misclassification bias is worsened by inclusion of industry dummies, since these are correlated with observed union status. Bratsberg and Ragan (2002) examine differences across industries in union wage gaps. They find that industry differences in union gaps have narrowed over time.

<sup>8</sup> A similar issue arises if one wants to measure an industry wage effect in a highly unionized sector. For example, in measuring wage comparability between unionized U.S. Postal Service workers and similar levels of work in the private sector (as required by statute) control for union status would effectively compare union postal workers to unionized private workers (i.e., postal workers would be credited with a union "endowment" variable in the same way they are

wage correlates is available in the monthly CPS earnings files. Each of these exclusions leads to higher union gap estimates since employer size and tenure are positively correlated with earnings and union status. Whether it is better or worse to include these variables is not clear. The union wage gap is largest among small employers and relatively small among large employers. That is, unions do not increase wages by very much over and beyond what is obtained by nonunion employees in large firms and establishments. The wage advantage associated with employer size appears to be some combination of human capital (measured and unmeasured) matched with physical capital and economic rents (Brown and Medoff, 1989; Oi and Idson, 1999). If the size advantage were entirely due to skill, the question would then arise whether the matching on skill seen among nonunion employers applies equally to union employers. (Section V discusses unions and unmeasured skills.) If not, inclusion of employer size would bias downward union gap estimates. If matching on skill by large employers were the sole source of the size-wage advantage and occurred in both union and nonunion firms, control for employer size would be appropriate. My reading of the literature is that skill differences are insufficient to account for size-wage advantage in general and that selection on skill is less likely in union than in nonunion establishments. If my assessment is correct, exclusion of employer size may produce a less biased estimate of the union gap than would inclusion, albeit in opposite directions.

Control for tenure when estimating a union wage gap is problematic. The argument for inclusion is based on the presumption that tenure measures accumulated skills beyond those captured by an experience (or potential experience) variable. But length of tenure with an employer is simply an inverse function of quits (Farber, 1999). Quits are lower in union firms because of higher wages, benefits (e.g., defined benefit pension plans that penalize mid-career quits), and union voice (Freeman, 1980). Thus, causality runs not only from tenure to wages but also from wages to tenure (Abowd and Kang, 2001; Farber, 1999). Inclusion of tenure necessarily biases the union gap downward since the tenure coefficient reflects in part how wages affect quits. To the extent that tenure measures skills not reflected in general experience and other control variables, its exclusion biases the union gap upward.

### III. *Measurement Issues: Match Bias and Misclassification of Union Status*

Data quality is an important concern in most empirical endeavors. Estimation of union wage gaps is no

---

credited with schooling or experience). Exclusion of a union control would effectively compare postal workers to a weighted average of union and nonunion workers throughout the private sector (Hirsch, Wachter, and Gillula, 1999).

exception. This section provides discussion of two measurements issues. The first issue, “match bias,” refers to the downward bias in union gap estimates due to earnings imputation procedures that do not include union status as a match criterion. This problem is not well known, but has a major impact on union gap estimates. The second issue – misclassification in union status – is a well-known problem. Misclassification is generally thought to cause a substantial bias in longitudinal estimates of the union wage gap, but to have little effect on wage level estimates. Discussion here suggests that misclassification has a sizable effect on wage level estimates when true union density is low, as in the U.S. private sector. Taken together, match bias associated with earnings imputation and measurement error from misclassification of union status cause a large downward bias in estimates of union wage effects.

A. *Match Bias from Earnings Imputation*

A substantial proportion of workers surveyed in the CPS and in other household surveys do not report earnings, and these individuals have their earnings imputed by the Census. As analyzed by Hirsch and Schumacher (2004), earnings imputation in the CPS causes a substantial downward bias in union wage gap estimates, currently about 5 percentage points. Because of increases in nonresponse over time, match bias affects the trend as well as level of union gap estimates. CPS estimates that fail to account for match bias greatly understate the magnitude of the union gap and overstate a downward trend in the gap evident since the mid-1980s.

The monthly CPS-ORG (outgoing rotation group) files are a principal source of union gap estimates. Individuals who either refuse or are unable to report weekly earnings have earnings imputed using a “cell hot deck” method that matches each nonrespondent to an earnings “donor” with an identical set of match characteristics. There are 14,976 “cells” or combinations of match variables. Match criteria include categorical variables for schooling, age, occupation, hours worked, gender, and receipt of tips, commissions, or overtime. They do *not* include sectoral variables such as union status or industry.

Even if the Census matching provides an unbiased measure of average earnings, wage differential estimates are systematically understated when the attribute being studied is *not* a criterion used by the Census to match donors to nonrespondents. These include, among others, union-nonunion, industry, and public-

private wage gap estimates. Hirsch and Schumacher (2004) derive a general expression for match bias.<sup>9</sup>

Bias in the union wage gap, *absent covariates*, is shown to equal

$$B = [(1-\rho_u)\Omega_u + \rho_n\Omega_n]\Gamma, \quad (6)$$

where  $\Gamma$  is the unbiased log wage gap and the term in brackets is the attenuation in the gap estimate. The terms  $\Omega_u$  and  $\Omega_n$  measure the rate of imputation or nonresponse for union and nonunion workers, while  $\rho_u$  and  $\rho_n$  measure the proportion of union donors matched to union and nonunion workers. The attenuation term in brackets is equal to the sum of match error rates, the proportion of false negatives (nonunion donors matched to union nonrespondents) and false positives (union donors matched to nonunion nonrespondents).<sup>10</sup>

Note that if the nonresponse rate for union and nonunion workers is similar ( $\Omega_u = \Omega_n$ ) and the rate at which each is matched to a union donor is equivalent ( $\rho_u = \rho_n$ ), equation (6) reduces to:

$$B = \Omega\Gamma. \quad (7)$$

In this case, match bias can be approximated by  $\Omega\Gamma$ , where  $\Omega$  is the proportion of workers with imputed earnings and  $\Gamma$  is a log wage gap estimate free of match bias. Standard wage gap estimates from CPS samples including allocated earners must be multiplied by  $1/(1-\Omega)$  to approximate a wage gap free of match bias. For example, if 30 percent of the estimation sample has earnings imputed, the gap estimate is attenuated by about 30 percent. Intuitively, the 30 percent of the sample for whom earnings is imputed displays little or no union gap, since both union and nonunion nonrespondents are matched to similar proportions of union and nonunion donors.

Introduction of covariates *increases* the downward bias in gap estimates due to earnings imputation. Hirsch and Schumacher explain how this can be likened to right-hand-side measurement error bias and employ a bias expression developed by Card (1996).<sup>11</sup> Correlation between union status and other match criteria reduces bias via improvements in imputation match quality. For a given match quality, however, correlation between union status and right-hand-side covariates exacerbates bias in the union coefficient, since bias increases with the ratio of error variance to total variance *conditional on covariates*.

---

<sup>9</sup> Note that match bias is distinct from response bias. It would exist even if nonresponse were randomly determined.

<sup>10</sup> A similar but more complex match bias exists for longitudinal estimates of wage gaps, depending on whether a worker has earnings imputed in one or multiple years and whether the worker is a union joiner or leaver. Hirsch and Schumacher (2004) describe this bias. Unlike wage level studies, longitudinal studies often omit workers with allocated earnings in any year.



Hirsch and Schumacher compare the difference in union gap estimates in samples including and excluding nonrespondents with the theoretical match bias (from Card) based on union and nonunion imputation and donor match rates and the correlation between union status and all other control variables. They conclude that excluding imputed earners from the estimation sample is a reasonable approach, the degree of *observed* attenuation being highly similar to  $\Omega$ , the nonresponse or imputation rate. They show, however, that true bias exceeds the observed attenuation, with the nonrespondent population having a set of characteristics associated with a relatively high union gap. In short, the simplest approach for researchers is to exclude imputed earners from their estimation samples. This avoids the severe match bias that accompanies wage gap estimates associated with non-match criteria. In the case of union wage gaps, however, even this approach leaves gap estimates slightly biased.

Census imputation rates and the ability to identify imputed earners have varied over time, making it difficult to obtain time-consistent union gap estimates. For example, the 1973-1978 May CPS files do not provide imputed earnings for nonrespondents (earnings are recorded as missing). The apparent sharp decline in CPS union wage gap estimates after 1978, viewed as a puzzle in the literature (Freeman, 1984; Lewis, 1986), is in large part a result of the inclusion of imputed earners in CPS estimation samples beginning in 1979 following their exclusion in 1973-1978. During the period 1989-1993, only about a fifth of those with imputed earnings are identified as such by the Census allocation flag. For those years, Hirsch and Schumacher identify imputed earnings based on whether or not the “unedited” earnings field value is missing or contains an earnings value. Revisions in the earnings questions beginning in 1994 increased complexity and rates of imputation. There were no valid allocated earnings flags included with the CPS-ORG during 1994 and most of 1995. Hirsch and Schumacher address these issues and construct time-consistent union gap estimates for 1973-2001 that are purged of match bias.

#### B. *Misclassification Bias*

A second measurement issue concerns bias from the misclassification of union status. Random measurement error on a right-hand-side variable, such as union status, biases the coefficient toward zero, with attenuation

---

<sup>11</sup> When earnings are imputed, one is correctly measuring a *donor's* earnings and explanatory variables used in the match. The donor's union status and other nonmatch criteria are measured with error.

increasing with the ratio of noise to signal.<sup>12</sup> The principal source of information on CPS reporting error in union status comes from a January 1977 CPS validation study that compared employer reports of union status with worker-reported union status in the CPS (Mellow and Sider, 1983). Freeman (1984) and Card (1996) have used information from the validation study to adjust longitudinal union wage gaps for misclassification bias. Card concludes that both workers and employers misreport union status and that misreporting is roughly symmetrical, the probability of a nonunion worker misreporting union status is equal to the probability of a union worker misreporting. Card's best estimate is that the misclassification rate for union status is between 2.5 and 3 percent. Based on Card's analysis, Farber and Western (2002), for example, assume a symmetric misclassification rate of 2.7 percent in order to calculate a "corrected" union density series for the U.S.

I am not aware of any other data source with which to compare the 1977 CPS validation results. But there are reasons to be cautious in placing too much weight on estimates from this single survey. First, the sample size is small: Card's calculations from the validation survey were based on 1,718 men age 24-66. Second, as pointed out by Card, the union status question asked of employers and employees in the CPS validation survey is not identical to the union question(s) asked of employees in the CPS.<sup>13</sup> Misclassification rates could differ across the two sets of questions, although it is not obvious which set of questions should produce a higher error rate. Finally, the validation survey is now a quarter century old. Because of the small sample sizes, differences in the union question, and age of the survey, it is easy to imagine that the true union misclassification rate is below or above, say, 2.7 percent. In the discussion below, I use a "conservative" estimate of a 2 percent misclassification rate. Subsequent analysis presents union gap estimates adjusted for 1, 2, and 3 percent misclassification rates.

Misclassification bias is a first-order concern for longitudinal studies, which identify the union wage effect by the partial correlation of wage change with respect to changes in union status. Because true changes in union status over short time periods is a low-frequency event, much of the measured union status

---

<sup>12</sup> See Aigner (1973) and Card (1996), among others. Absent covariates, misclassification bias looks very much like the match bias formula presented earlier, with attenuation (proportionate bias) equal to the sum of match error rates. As shown by Card, bias is exacerbated if there are covariates, since the relevant "signal" from union status is that net of its correlation with other explanatory variables.

change reflects reporting error (Mincer, 1983; Freeman, 1984; Card, 1996).<sup>14</sup> Freeman and Card conclude that following adjustment for misclassification bias, longitudinal union wage gap estimates are similar to those obtained from OLS wage level regressions.

In contrast to the longitudinal studies, misclassification bias has not been regarded as serious for the estimation of union gaps from wage level regressions.<sup>15</sup> This assumption should be reconsidered. As Farber and Western (2002) show in their analysis of union density, were true union density 50 percent, symmetric misclassification would not bias density estimates, but bias becomes more serious as union density moves toward zero or one hundred percent. For example, if true union density were zero and there were a 2 percent misclassification rate, all 2 percent of the measured union density would be due to misclassification.

Letting  $u^*$  be the true but unobserved union status and  $u$  be observed union status, Farber and Western (2002) derive the following relationship between true union density  $\bar{u}^*$  and observed union density  $\bar{u}$  (likewise, see Card, 1996):

$$\bar{u}^* = (\bar{u} - \lambda) / (1 - 2\lambda), \tag{8}$$

where  $\lambda$  is the union misclassification rate, assumed in (8) to be the same for union and nonunion workers.

For example, given an observed union density in the private sector of 9 percent in 2001 and assuming that the misclassification rate is 2 percent, true union density is 7.3 percent (that is,  $(.09 - .02) / (1 - .04) = .0729$ ).

I am concerned with bias in the estimation of union wage gaps, and not the measurement of union density. Given high *nonunion* density, a 2 percent misreporting rate distorts by very little observed nonunion wages. However, observed private sector union density is below 10 percent, implying that the proportion misclassified in the union sample is high and that union wages are understated accordingly. Thus, a substantial downward bias exists for standard union wage gap estimates, at least those for the private sector.

To estimate the bias in union gap estimates resulting from misclassification error, I rely on Card's attenuation coefficients. An attenuation coefficient,  $\gamma$ , is bounded [0, 1], with  $\gamma = 1.0$  implying no attenuation

---

<sup>13</sup> The validation survey asked employers and employees whether their wage was determined by collective bargaining. The CPS question asks workers if they are members of a labor union or employee association like a union. Nonmembers are asked if they are covered by a collective bargaining agreement on their principal job.

<sup>14</sup> Here two wrongs make a right, with consistent misreporters not being classified as changing union status.

<sup>15</sup> A notable exception is Bollinger (2001), who emphasizes the wide bounds that both cross-section and fixed-effects union gap estimates can take over the full range of union misclassification rates. Bollinger does not provide "preferred" estimates of union wage gaps based on "best guess" misclassification rates. He calculates that the misclassification rate must be below 0.8 percent for the true coefficient in his fixed effects model to be bounded below the true coefficient in a wage level model. The 1977 CPS validation study suggests a far higher misclassification rate.

or bias, and  $\gamma = 0$  implying complete attenuation or a coefficient that goes to zero. Absent covariates, Card (1996, p. 959) derives the attenuation coefficient  $\gamma^0$  equal to:

$$\gamma^0 = \bar{u}^*/\bar{u} [(q_1 - \bar{u}) / (1 - \bar{u})] \quad (9)$$

Here  $q_1 = \text{Prob}(u = 1 \mid u^* = 1)$ , the probability of a true union member being classified as union. This is simply equal to  $(1 - \lambda)$  or .98 if we assume a misclassification rate  $\lambda$  equal to .02. As previously stated,  $\bar{u}^*$  and  $\bar{u}$  are the true and observed union densities, respectively, with  $\bar{u}^* = (\bar{u} - \lambda)/(1 - 2\lambda)$ .

Based on equation (9) above and assuming a misclassification rate  $\lambda$  equal to .02 and an observed private sector union density  $\bar{u} = .09$ , the implied attenuation is  $\gamma^0 = .792$ . Thus, private sector union wage gaps (unadjusted for covariates) are biased downward by a fifth (.208) given a misclassification rate of 2 percent, and by considerably more if misclassification exceeds 2 percent.

Attenuation is more severe given positive correlation between union status and other regression covariates. Letting  $R^2$  be the coefficient of determination from a regression of observed union status on other covariates, Card (1996, p. 960) derives an attenuation coefficient that for our application is equal to:

$$\gamma^1 = [\gamma^0 - (R^2/(1 - 2\lambda))]/(1 - R^2) \quad (10)$$

If  $R^2$  were equal zero,  $\gamma^1 = \gamma^0$ . Assuming that  $R^2 = .11$ , its average value across recent years (Hirsch and Schumacher, 2004), the attenuation coefficient is  $\gamma^1 = .761$ , which implies that private sector union gap estimates are biased downward by almost a quarter (.239). Stated alternatively, private sector union wage gap estimates (for recent years) must be multiplied by  $1/\gamma^1 = 1.31$  to eliminate misclassification bias.

Downward bias in the union coefficient is highly sensitive to the rate of misclassification, which I have assumed to be 2 percent (1 person in 50). A misclassification rate of  $\lambda = .03$  or 3 percent implies  $\gamma^0 = .686$  and  $\gamma^1 = .642$ . In short, downward bias in standard union gap estimates may be far more severe than previous figures suggest. Unfortunately, we simply know far too little about actual union misclassification rates. Given the importance of attenuation bias for low-frequency outcomes (e.g., union status in the private sector), future CPS validation surveys by the Census and BLS would have high value added.

The misclassification bias discussed in this section adds to what is already a substantial downward bias from imputation mismatch on union status. Obviously, attenuation from measurement error is less

severe if misclassification rates are low or union density is high.<sup>16</sup> The clear implication of this section is that measurement error bias should be carefully considered not only in longitudinal analyses of union wage gaps, but also in wage level analyses. Moreover, the severity of this bias has increased over time as union density has fallen in the private sector, at the same time that match bias from imputation has worsened as CPS nonresponse rates have increased. Both misclassification and match bias result in standard analyses understating the level of the union wage gap and overstating its decline. We turn to this evidence below.

#### *IV. Private Sector Union Wage Effects Over Time*

In this section, changes over time in the U.S. private sector union wage gap are examined, based on regression results from the CPS. Following the analysis in the previous section, results are presented based on standard CPS samples with imputed earners included, with imputed earners excluded, and with alternative assumptions about the rate of union misclassification (1, 2, and 3 percent). Union wage gap estimates increase substantially once imputed earners are excluded, consistent with a downward match bias from imputation described in Hirsch and Schumacher (2004). As described in the previous section, union gap estimates also increase substantially with the assumed rate of misclassification. The downward trend in the private sector union gap seen in standard estimates from the CPS is shown to result in part from attenuation due to match bias and misclassification bias.

##### *A. Union Wage Gaps with Match Bias and Misclassification Rate Adjustments*

Figure 1 shows estimates of private nonagricultural sector union wage gaps from the CPS for 1973-2001, taken from recent work by Hirsch and Schumacher (2004).<sup>17</sup> The base series in Figure 1 (shown in diamonds) provides estimates from the sample including workers with imputed as well as reported earnings; hence, these estimates suffer from the downward match bias described in the previous section.<sup>18</sup> The next series (in squares) provides correction for match bias by omitting workers whose earnings are allocated (for details, see Hirsch and Schumacher, 2004). As evident from the two series, inclusion of imputed earners biases union estimates gap downward roughly 3-4 percentage points prior to 1994 and by roughly 5

---

<sup>16</sup> An interesting question is how misclassification rates differ between the private and public sectors (see Section VI). Because the meaning of union membership and coverage is not well understood by workers in some public sector jobs, my expectation is that misclassification is less severe in the private sector. On the other hand, union density levels are lower in the private sector, making attenuation bias more severe for a given misclassification rate.

<sup>17</sup> Details regarding estimation and a listing of wage gap estimates by year are in Hirsch and Schumacher (2004).

percentage points in years since 1999. Standard estimates including workers with imputed earnings tend to understate the magnitude of the union wage gap, but exaggerate its decline.

Relying on estimates from the series corrected for match bias, the average private sector union gap rose during the 1970s, hit a high of .23 in 1984, and has slowly declined since that time, the exception being the early 1990s economic slowdown. By 2001, the private sector union gap is estimated to be .18 log points, not too far above the 15 percent “conventional wisdom” based on a much earlier period. The CPS sample including imputed earners, which makes up 30 percent of the CPS sample, produces a union wage gap estimate of .13 in 2001. The differences between series 1 and 2 illustrate clearly that union wage gap estimates from the CPS must account for what is a severe downward bias, either through removal of imputed earners or the use of alternative imputation methods (Hirsch and Schumacher, 2004).

Apart from the issue of match bias due to Census earnings imputation, are the estimates in Table 1 consistent with what other researchers would obtain? A recent paper by Blanchflower and Bryson (2002) helps answer this question. Blanchflower and Bryson also use the CPS for 1973-2001 and provide adjustments for imputation bias based on the methods described in Hirsch and Schumacher. However, they choose a specification that excludes broad occupation dummies but includes rather detailed industry dummies and state fixed effects. For reasons described in an earlier section, each of these adjustments drive down the union premium. Moreover, correlation between union status and the other covariates in the Blanchflower-Bryson specification will be higher than in Hirsch-Schumacher, since union status differs substantially by detailed industry and state. As seen in the previous section, such correlation exacerbates downward bias in the union gap estimate owing to misclassification error in union status. Whereas our log wage gap estimate for 2001, corrected for match bias, is .18, the corresponding Blanchflower-Bryson estimate is .14. Changes in the union wage gap over time are similar in the two studies. Blanchflower and Bryson provide further checks on specification sensitivity.

Shown next in Figure 1 is the estimated union wage gap series (in triangles) including not only the match bias adjustment, but also a correction for misclassification bias using the Card (1996) correction formula (equation 10). The diamond series is based on an assumed misclassification rate of 2 percent, which

---

<sup>18</sup> The May 1973-1981 CPS does not include imputed earnings for nonrespondents. The “diamond” wage gap estimates for 1973-1978 are approximations of what gaps would be for the full sample were earnings imputed for nonrespondents

earlier discussion characterized as “conservative.” Adjustment for misclassification has a substantial effect on union gap estimates. The attenuation coefficient  $\gamma^1$  is about .89 in 1973 and declines throughout the nearly thirty-year period, to a value of .77 in 2001 (downward bias of 11 and 23 percent, respectively). Comparing wage gap estimates from series 2 and 3, those adjusted for misclassification bias are about 2 percentage points higher during the mid-1970s, a bit over 4 percentage points in the mid-1980s, and currently in excess of 5 percentage points higher.

The bias from misclassification is sensitive to the assumed rate of misreporting. Figure 2 shows alternative union wage gap series, with corrections to the Hirsch-Schumacher series based on the assumption of misclassification rates of 1, 2, and 3 percent (the 2 percent series was shown previously in Figure 1). A misclassification rate of 1 percent has little effect on estimates, at least in early years when private sector density was well above current levels. In 2001, a 1 percent misclassification rate implies an attenuation coefficient  $\gamma^1$  of about .88 (downward bias of 12 percent) or about .02 log points off the union gap (.182 versus .206). A rate of 3 percent misclassification implies a large bias, with an attenuation coefficient  $\gamma^1$  of .65 (downward bias of 35 percent) or nearly .10 log points off the union gap (.182 versus .281).

These results show clearly that as private sector union density has fallen to low levels, attenuation in wage gap estimates due to misclassification of union status has increased. The bias is now serious, causing a substantial understatement of union wage gaps. Unfortunately, the exact degree of attenuation is highly sensitive to the rate of misclassification, a number about which we know little, apart from the error rate of 2.5-3 percent found in the 1977 CPS validation survey (Card, 1996).<sup>19</sup>

Although the precise degree of misclassification bias is uncertain, enough is known to cause us to reconsider the conventional wisdom of a 15 percent union wage gap (Fuchs et al., 1998). A misclassification rate of 3 percent (1 in 33 workers), in line with evidence from the 1977 CPS validation survey, yields estimates of union wage gaps since the mid-1980s of about .30 log points (this includes both the match bias and misclassification bias corrections). An error rate of 2 percent (1 in 50 workers) implies union wage gaps of roughly .25 since the mid-1980s. Union wage gaps at these levels are well beyond conventional beliefs.

---

(Hirsch and Schumacher, 2004).

<sup>19</sup> Note that adjusting for misclassification error requires knowledge of the error rate (assumed here), union density in each year, and the  $R^2$  obtained from a regression of union status on the other covariates. The annual  $R^2$ , calculated in

Based on the discussion in this section, I would argue that the profession should adjust upward its belief about the magnitude of union wage effects in the U.S. private sector. Just how large that upward adjustment ought to be, however, requires further study.

B. *Union and Nonunion Wage Trends in the CPS, ECI, and ECEC: Troubling Puzzles*

Adjustment for match bias and misclassification affect inferences regarding changes in the union wage gap over time. The evidence from standard estimates not accounting for these biases can be readily summarized. As recognized in the literature, union wage gaps tended to increase during the 1970s, a period of both a strong and weak recession and high unanticipated inflation.<sup>20</sup> Sample sizes were very small during 1979-1981 (there was no 1982 union question in the CPS), but it appears that union wage gaps peaked around 1984, consistent with evidence from the Employment Cost Index (ECI). Union wage gaps appeared to ease downward after around 1984, and then declined steadily after 1996, with a particularly sharp drop beginning in 1999. Inferences about the size of the decline must be moderated once one adjusts for imputation match bias and then again following a correction for misclassification bias.

Nothing in the analysis causes us to question the conventional wisdom that the union wage gap is countercyclical, noted long ago by Lewis (1963, 1986). Stated alternatively, one observes a lesser degree of wage cyclicality among union than among nonunion workers. An article by Grant (2001) surveys the literature and provides new evidence, concluding that nonunion wages tend to be somewhat procyclical and union wages largely acyclical.

As seen in Figures 1 and 2, all union gap series indicate a decline in the gap since 1984, but the extent of the decline is affected by adjustments for imputation and misclassification bias. Each of these adjustments lessen the decline in the union gap. In our preferred series with the 2 percent misclassification rate, much of the decline that occurs takes place in the years after 1997, a period when the labor market was relatively strong. This evidence supports the conclusion reached by Blanchflower and Bryson (2002) that there has been little underlying change in the union wage premium, the 1990s decline being the result of an unusually strong economy and likely to be temporary.

---

Hirsch and Schumacher (2004), declines over time from a high of about .25 in 1973 to .10 in 2001. Declining union density over time increases bias, while the declining correlation of union status with other covariates mitigates bias.



Does evidence on union and nonunion wages from other data surveys match the CPS? The most commonly used measure of wage growth is the Employer Cost Index (ECI), derived from the National Compensation Survey, published by BLS (USDOL, 2003). Its observational unit is occupation-by-establishment job cells observed in consecutive quarters. The ECI measures quarterly wage and compensation growth and has separate indices for private sector union and nonunion jobs beginning in 1975:3. Shown in Figure 3 is an updated version of a figure from Hirsch, Macpherson, and Schumacher (2003) (hereinafter HMS). It provides ECI union-nonunion differences, by quarter, in rates of wage growth during the previous four quarters (thus beginning with 1976:3). A positive value indicates faster union than nonunion wage growth (a rising union gap) and a negative value indicates slower growth (a falling gap).

The ECI evidence is qualitatively the same as that seen in the CPS. As evident in Figure 3, union wage growth exceeded nonunion growth during the late 1970s and early 1980s, briefly during the early 1990s recession, and once again in late 2001 and 2002 during the most recent recession, consistent with a countercyclical wage gap (BLS, 2003). Despite the recent pattern, for most of the years since the early to mid-1980s, union wages have risen more slowly than have nonunion wages. Figures on total compensation indicate less of a decline in the union compensation gap than in the wage gap, with greater growth in union than in nonunion benefits (Hirsch et al., 2003).

Despite the broad similarity in the ECI and CPS evidence on relative union-nonunion wage growth, HMS identify several troubling puzzles. The ECI indicates far more rapid decline in relative union wages than seen in CPS regression analysis of union wage gaps. Decline in the ECI is similar to that seen in the *unadjusted* (i.e., non-regression) CPS figures. HMS show that about half of the CPS decline in the unadjusted union gap is accounted for by changes in the relative composition of union to nonunion wage attributes, with the remaining half representing decline in the adjusted gap. The puzzle is that the ECI is intended to represent a fixed weight index, much akin to what is effected through regression controls. The ECI unit of observation holds the occupation, occupation skill level, industry, and establishment constant, as well as changes in employment (through the use of fixed weights). It cannot take into account (i.e., hold constant) worker characteristics such as schooling, age, or gender.

---

<sup>20</sup> The CPS did not include imputed earnings in May 1973-1978. So standard estimates in the literature are similar to the “squares” in Figure 1 for 1973-1978 and then show a large drop to the “diamonds” beginning in 1979, a puzzle

Even more puzzling is that the Employer Costs for Employee Compensation (ECEC), also drawn from the NCS shows a different pattern than the ECI. The ECEC uses current rather than fixed weights and allows job cells to move into and out of the sample. In principle, the ECEC should yield results more similar to the *unadjusted* CPS, rather than regression estimates. Surprisingly, the ECEC indicates highly similar union and nonunion wage growth, at odds with both the ECI and the CPS (unadjusted or adjusted for worker characteristics and sector). HMS are unable to resolve these puzzles in a satisfactory manner.

A reasonable conclusion seems to be that the weight of the evidence shows a decline in relative union to nonunion wages. The magnitude of the decline is in doubt, as is the question of whether the decline since the mid-1990s will be short-lived or permanent. Closing over time between union and nonunion *compensation* has been less than the closing in wages. Blanchflower and Bryson (2002) conclude that the 1990s decline in the union gap is likely to be temporary, primarily the result of the unusually strong economy. Focusing on the CPS evidence in Figure 1, it is clear that the series corrected for match bias and misclassification bias (the triangles) is more consistent with the “no trend thesis” than are uncorrected estimates (the diamonds). Given the increased competitiveness of the U.S. economy and the continuing decline in private sector union density, a rebound in the union gap to its *peak* level would be surprising. Evidence will tell (or may tell), but not for some time.

#### V. *Employer/Employee Selection, the Skill-Upgrading Hypothesis, and Longitudinal Wage Gaps*

Preceding sections have argued that union wage gaps are higher than standard estimates and conventional wisdom suggest. Whether union gap estimates well above 15 percent are plausible depends to no small degree on one’s priors regarding the relationship between union status and worker skills. This section addresses the issue, focusing on the skill-upgrading hypothesis.

Interpretation of union wage gap estimates depends in part on the correlation between union status and worker skills that are unmeasured in typical wage equations.<sup>21</sup> If unmeasured skills are positively correlated with union status, union gap estimates are likely to be biased upward, and vice-versa. A conventional argument, sometimes referred to as the “skill-upgrading” hypothesis, is that a union wage premium both allows and provides incentive for employers to upgrade the skill level of their work forces,

---

addressed without satisfaction by Freeman (1984) and Lewis (1986).

offsetting part of the higher wage. Despite a substantial empirical effort to account for unmeasured skills and facilitate estimation of unbiased union gap estimates, the literature does not point clearly toward the existence of substantial skill upgrading in union workplaces.

Absence of clear-cut evidence in favor of skill upgrading by union employers is not so surprising once one realizes that such behavior need not follow from theory. Wessels (1994) has provided a simple but persuasive challenge to the skill-upgrading hypothesis. If firms upgrade in response to a union wage increase, the union can then bargain in a future contract for an even higher wage in order to restore the premium.<sup>22</sup> Employers, anticipating this, may respond by not upgrading. Firms that upgrade will face higher future wage demands and will have distorted their factor mix, using a higher skill labor mix than is optimal given its technology. Wessels provides an explicit model in which it is assumed that labor quality augments capital productivity and the decision to hire higher (lower) skill workers results if the elasticity of substitution between labor and capital is greater (less) than unity. Wessels concludes that available evidence is not consistent with skill upgrading. In short, once one takes into account that unions acquire wage gains through bargaining power, it no longer follows that firms can offset, or will attempt to offset, cost increases through skill upgrading.

The selection mechanism within union companies may best be characterized as a form of two-sided selection (Abowd and Farber, 1982; Card, 1996). A queue of workers exists for union jobs. Hence, employers are able to avoid hiring workers in the lower tail of the ability distribution (ability is used broadly to include productivity-related attributes such as motivation and reliability). Yet few workers from the upper-tail of the ability distribution are either in or chosen from the union queue. Wage compression within union firms, both compression in skill differentials and contractually standardized wages, discourages applications by many of the most able workers. And, following the logic of Wessels, firms may not have incentive to screen for and select the most able workers in the union queue if such action may lead to an increase in future wage demands. In short, realistic models of worker and firm selection lead to the prediction that unionized work forces need not be systematically more or less skilled than a nonunion labor

---

<sup>21</sup> The same point holds true for omitted working conditions correlated with union status. This issue is discussed briefly in a final section.

force, but the skill distribution is more likely to be compressed. Not only are unions associated with compression in wages, but also compression in worker ability (Hirsch and Schumacher, 1998).

The estimation issue arising from the nonrandom determination (i.e., endogeneity) of union status is how to account for the possible bias in union gap estimates owing to correlation between unmeasured skill and union status. The literature has attempted to account for skill bias in two principal ways: (a) selection models that explicitly account for the process through which union status is determined, and (b) longitudinal models relating wage changes for individual workers to changes in union status, thus netting out unmeasured worker-specific skills fixed over time, i.e., transferable across jobs.<sup>23</sup>

Selection models were discussed in a previous section, where I argued that selection results on union wage gaps are unreliable for two reasons. First, these models typically employ a single index selection measure that is equivalent for workers with the same probability of union status. This is inconsistent with the two-sided selection results in Card (1996) in which workers with low and high measured skills have similar union probability rates, but opposite selectivity (e.g., positive selection for workers with low measured skills, and negative selection for those with high measured skills). Second, even were the single index selection model appropriate, meaningful identification requires use of one or more variables that are good predictors of union status but not earnings.

Longitudinal models identify a union wage effect from the wage changes of workers who switch union status (Mellow, 1981; Freeman, 1984; Card, 1996, Hirsch and Schumacher, 1998), thus controlling for unmeasured skills transferable across jobs. That is, if union workers possess high unmeasured skills, their wages would be higher than predicted in nonunion jobs and we should observe relatively small wage gains in

---

<sup>22</sup> This process is aptly described by Kalachek and Raines (1980, p. 68), who nonetheless stress the importance of skill upgrading by union employers. “Hence the stage is set for a neverending chase between the employer and the union. The faster the employer liquidates rents, the greater the pressure on the union to recreate them.”

<sup>23</sup> A distinct but related literature uses a production function approach to see if unionized firms or industries have higher technical efficiency, or measured output for given combinations of inputs (for surveys, see Addison and Hirsch, 1989; Booth, 1995). Studies attempt to measure whether union establishments are more productive owing to unionization, holding constant the capital-labor ratio and labor quality. If union status is a close proxy for unmeasured worker skills, it is likely that those same skills, unmeasured in a production function study, would show up as a positive union productivity effect. The most reliable economy-wide evidence, however, based on data at the line-of business or firm level (Clark, 1984; Hirsch, 1991) does not indicate positive union productivity effects, on average. A survey measuring the views of labor economists at top universities asked the question: “What is your best estimate of the percentage impact of unions on the productivity of unionized companies.” The median response was zero and mean response 3.1 percent (Fuchs et al., 1998, pp. 1392, 1418).

moving from nonunion to union jobs, as well as small wage losses in moving from union to nonunion jobs.<sup>24</sup>

Absent an accounting for misclassification error in union status, longitudinal analysis typically obtains far lower union gap estimates than those obtained from wage level equations. Some authors conclude that the lower estimates support the proposition that union workers have high unmeasured skills. As emphasized by Freeman (1984) and others, longitudinal estimates are seriously biased toward zero owing to measurement error. Although misclassification rates for union status may be relatively low, so is the incidence of true union switching over short time periods (one year in studies using the CPS). Thus a large proportion of those recorded as union switchers have not changed union status and display no correlation between wage change and recorded union status change. Freeman and Card correct their estimates based on union status information from the 1977 matched CPS-employer survey. Although using different methods and assumptions to correct bias, each concludes that corrected longitudinal estimates are similar to those from wage level equations. Hirsch and Schumacher (1998) make no explicit correction for misclassification, but instead rely on union gap estimates among workers who change detailed industry and occupation as well as union status, since such workers are most likely to be true job and union status changers. Their longitudinal estimates are lower than their wage level estimates. Taken together, longitudinal studies suggest that omitted ability bias has at most a modest effect on average union wage gap estimates.<sup>25</sup>

Although longitudinal studies do not greatly alter our assessment of *average* union wage effects, they very much affect assessment of union wage differences across worker skill groups. The conventional wisdom is that union wage premiums are largest for workers with low skills and smallest for workers with high skills. For example, standard union gap estimates are substantially larger for high school than for college graduates and for blue-collar than for white-collar workers. Such findings reinforce the thesis that unions decrease skill-based wage differentials.

As emphasized by Card (1996) and Hirsch and Schumacher (1998), longitudinal union gap estimates, which control for worker-specific skills, are far more similar across skill groups than are standard estimates. Among union workers with low measured skills (e.g., high school dropouts), there exists positive

---

<sup>24</sup> Ignored is the important issue of whether union status *change* is exogenous.

<sup>25</sup> This same conclusion need not apply to specific worker groups. For example, Hirsch (1993) concludes that low longitudinal estimates among truck drivers support the proposition that union drivers have high unmeasured skill (e.g., driving experience, reliability).

selection, union employers hiring and retaining those with high unmeasured skills (motivation, dependability, etc.). By contrast, among union workers with high measured credentials (e.g., college graduates), negative selection is more typical, with the most motivated and skilled workers not sorting into the union sector or being so highly valued as by nonunion employers. Union wage compression and rate standardization help produce this pattern of two-sided selection. But, following sorting and a more complete accounting for worker skills, union wage gap differences across skill groups are far less than commonly believed.

A vivid example is provided in Hirsch and Schumacher (1998). When they segment their sample by educational group, they obtain the following log point union gap estimates – dropouts: 0.24; high school grads: 0.20; some college: 0.17; and college graduates: 0.08. This is the standard pattern found in the literature, gap estimates falling sharply with measured skill (if occupational dummies are excluded, a far steeper gradient with respect to schooling is found). In sharp contrast, longitudinal union gap estimates for the same four groups of workers were as follows: 0.11, 0.12, 0.12, and 0.11 (longitudinal estimates, which control for worker-specific skills, are unaffected by occupational controls). That is, no discernable difference is found in union gap estimates across schooling groups, following sorting in the labor market. A similar result is seen using a multidimensional index of measured skills (i.e., predicted earnings from a sparse earnings function). Card obtains a similar qualitative pattern, but concludes that union wage effects do decline with skill level, although not by nearly so much as evident in standard estimates.<sup>26</sup>

In short, union-nonunion wage gaps for workers of equivalent ability do not vary nearly so much across worker groups as commonly believed. This result follows selection by individuals into the union queue and the selection of workers by firms from the queue. Union workplaces not only tend to have a compressed wage distribution, but also a compressed distribution of worker skills. The compressed skill distribution results in part from union success in compressing compensation. Moreover, work settings that best lend themselves to relative homogeneity in skills and wages are also the most likely to be successfully organized and maintain political support for union representation.

Besides aiding our understanding of what unions do, the conclusion that union wage effects vary relatively little across worker groups, following matching by workers and employers, has implications for

empirical studies. Because union gap estimates from wage level equations differ so markedly by worker group, the response by some is to move away from a single union wage gap estimate and instead estimate a family of union wage effects. Some applied researchers recommend the use of quantile regression methods or other semi-parametric (or nonparametric) methods that permit gap estimates to better fit the data. Discussion in this section suggests that flexible estimation methods that fail to account for two-sided selection need not produce reliable measures of union wage gaps. Union wage gaps estimated across a broad cross-section of private sector wage and salary employees may well produce more appropriate estimates. Attention should be aimed at controlling for positive selection at the low end of the measured skill distribution and negative selection at the high end of the measured skill distribution. Studies that model and estimate this process in a principled manner would provide a welcome contribution to the literature.

VI. *Further Topics: Inequality, Public Sector Unions, Membership vs. Coverage, Compensation, and Working Conditions*

I have focused on a narrow but important range of issues concerning unions and wages. In this section, brief mention is made of several additional issues. Some of these topics have been studied extensively in the literature (e.g., inequality), while others have not (e.g., working conditions).<sup>27</sup>

A. *Unions and Earnings Inequality*

Unionization is strongly associated with lower earnings dispersion or inequality. As carefully documented by Richard Freeman (1980, 1982), unions appear to lower dispersion in two major ways. First, union workplaces have compressed skill differentials – that is, smaller wage differences between workers in low and high skill positions. Second, wages are contractually determined or standardized, with pay often depending solely on a worker’s job position and seniority. Management has less discretion and is limited in its ability to vary pay across workers in similar positions. The decline in private sector unionization is

---

<sup>26</sup> Hirsch and Schumacher (1998) also provide a direct measure of two-sided selection, reporting AFQT percentile scores for union and nonunion workers with low and high measured attributes. Among workers with high measured attributes, union workers have AFQT scores 14-15 percentile points less than nonunion workers.

<sup>27</sup> Space prevents discussion of additional topics. Two that warrant exploration are an evaluation of matching techniques and the use of sample weights. Matching estimators would calculate a union gap as the mean difference in log wages between each union worker and a matched nonunion worker, where each match is based on a random selection from among all nonunion workers with the exact combination of attributes (i.e., hot-deck matching) or among nonunion workers with similar propensity scores (i.e., the probability of being a union member based on logit or probit estimates). In contrast to regression analysis, matching estimators provide an explicit weighting function based on union worker (“treatment group”) characteristics (Angrist and Krueger, 1999). A second issue concerns use of sample weights. The CPS and other household surveys are not fully representative, although they do contain weights intended

frequently identified as one of several important explanations for rising earnings inequality. In particular the sharp decline in union density during the 1980s corresponds to the period when earnings inequality rose most rapidly. A paper by Card et al. (2003) provides recent evidence and a careful documentation of the union-inequality relationship. Moreover, the strong empirical relationship between unions and inequality is consistent with the political economy of democratic unions that have incentive to foster support among average workers (the median voter) more than mobile marginal workers catered to by employers in a competitive labor market (Farber, 1986).

Despite the strong inverse correlation between unions and inequality, an important caveat is in order. It is difficult to know how much of this relationship is causal. As discussed in the previous section, longitudinal union gaps across measured skill groups vary little following sorting in the labor market (i.e., positive sorting on low-measured skill and vice-versa). Union wage policies affect sorting in the labor market between the union and nonunion sectors. Thus, differences in observed wage inequality between the union and nonunion sectors exaggerates unions' causal effects on economy-wide dispersion. Were unions to disappear, inequality among previously unionized *workers* would not rise to the level seen in nonunion sectors since there exists greater homogeneity or compression of worker skills and tastes within the union sector. Moreover, greater homogeneity is not only the result of sorting in response to union policies. Unions are more likely to organize successfully in sectors where worker preferences and skills are homogeneous (Hirsch, 1982). That is, causality runs from low dispersion to unionization as well as in the other direction.

#### B. *Public Sector Unionization*

Much of my discussion has concerned union wage effects in the private sector. Separate analyses of unionism in the private and public sectors are appropriate. First, the legal structure differs, the private sector being covered by the National Labor Relations Act (or the Railway Labor Act for workers in a few industries), while public sector workers are covered by a diverse mix of federal and state collective bargaining laws. Second, some public sector workers counted as members belong to employee associations that do not have collective bargaining rights or, more generally, the bargaining rights of public workers vary considerably across jurisdictions. Third, the impact of public sector workers may take place as much through

---

to provide population estimates. A weighted estimator should produce an average union gap estimate similar to what would be obtained based on a fully representative sample.



the political process as through collective bargaining and show up not only in higher wages, but also in public appropriations and employment (Freeman and Medoff, 1984). Fourth, there is likely to be more misclassification error in union status in the public sector since what constitutes union membership and coverage may not be clear to some workers.<sup>28</sup>

As discussed previously, downward bias in union wage estimates from misclassification is a more serious problem than generally believed. Bias is a function of the ratio of noise to signal in the union measure. Misclassification in the public sector is probably more severe than in the private sector. Yet true density is far higher in the public than private sector. We simply do not have enough information to know whether misclassification bias is more or less severe in the public than in the private sector.

For the above reasons, estimation of private and public sector union gaps are best kept separate. My focus on private sector gaps should not be taken to imply that study of the public sector is not important. It is important and increasingly so. Whereas 26 percent of all union members and 28 percent of covered workers were government employees in 1977 (the year the CPS adopted its current set of union status questions), corresponding figures were 44 and 45 percent in 2001 (Hirsch and Macpherson, 2003).

### C. *Membership versus Coverage*

Most of the literature on union wage gaps uses as its measure of union status a binary variable either for membership or coverage, comparing the selected union group to everyone else. Few studies distinguish between the three groups – union members, covered nonmembers, and non-covered nonmembers. Were there identical union wage effects from being a member and being a nonmember covered by a collective bargaining agreement, one would get a lower union gap estimate using membership, since covered nonmembers would be included in the control group (Jones, 1982). Just the opposite is found. As seen in Lewis (1986), the choice of union status variable typically has a small effect on union gap estimates, but membership gaps typically are a little larger than are coverage gaps.

In the work shown in this study, a union membership variable is used, since the coverage variable was not available until 1977. The number of private sector covered nonmembers is sufficiently small so that their inclusion in the nonunion control group produces union wage effect estimates nearly equivalent to what

---

<sup>28</sup> In the CPS, nonmembers are asked if they are covered by a collective bargaining agreement at their workplace. Misclassification of the coverage variable is believed to be particularly severe among federal workers, with workers in

would be obtained were covered nonmembers included as a distinct group (or removed from the analysis).

In U.S. studies that distinguish among the three groups, union members display a substantial wage advantage relative to covered nonmembers (Jones, 1982; Schumacher, 1999; Budd and Na, 2000). For example, focusing exclusively on right-to-work states, where union gaps tend to be higher than across all states, Schumacher (1999) finds a -.14 lower union gap for covered nonmembers than for members (and a -.10 log wage difference when the estimation sample is limited to covered workers, members and nonmembers). An interesting question then arises. Why is there a wage advantage to membership versus coverage? Schumacher investigates a number of sources. He finds little evidence to support the following explanations for the gap: (a) higher unmeasured skills among members than covered nonmembers; (b) lower job tenure among covered nonmembers than among members; (c) higher misclassification rates for covered nonmembers; or (d) industry and occupational wage effects.

Schumacher (1999) and Budd and Na (2000) conclude that union wage gaps are lowest in industry-occupation sectors in which there are large numbers of “free riders” (the ratio of nonmembers to total covered). What is not clear is the direction of causation. One possibility is that less free riding provides unions with greater bargaining power. Alternatively, covered workers may be less likely to free ride in jobs where they realize a substantial wage advantage. Both explanations are plausible and likely to matter. It will be up to future studies to sort out the relative importance of these two explanations.<sup>29</sup>

#### D. *Compensation*

Discussion to this point has been in terms of union *wage* effects. Ideally we would like to focus on *compensation* (wages plus benefits). The difficulty is that micro data sets providing worker characteristics (e.g., the CPS) at most provide indicator variables on the receipt of benefits (e.g., pension payments and health insurance) but not the dollar value of benefits. Published data on employer costs for benefits, for example, from the BLS’s Employer Costs for Employee Compensation surveys, do not permit the matching of dollar benefits to individual workers and their characteristics.

Although evidence is limited, that which exists indicates clearly that higher benefits are received by

---

the same work units providing different answers.

<sup>29</sup> Booth and Bryan (2001) use matched employer-employee data from Britain and instrumental variable estimation to account for membership endogeneity. Despite some variation across worker groups, the authors conclude that there is little systematic difference in wages between covered members and nonmembers.

union than by nonunion workers. Work by Freeman (1981) finds that the union benefit gaps greatly exceed wage gaps, leading to a union compensation gap that exceeds the wage gap by about 2 percentage points. Freeman finds the union benefit advantage particularly large for insurance and pensions.

Recent wage and benefit data from the *Employer Costs for Employee Compensation* indicates the large gap in benefits between union and nonunion workers. In 2001, the ECEC data (USDOL, 2002) show hourly compensation for nonunion workers of \$19.98, \$14.81 in wages and \$5.18 in benefits (included those government mandated). Compensation for union workers is \$27.80, \$18.36 in wages and \$9.45 in benefits. Current data confirm that union wage gaps understate union compensation gaps, perhaps substantially so.<sup>30</sup>

#### E. *Working Conditions*

A shortcoming of most union wage gap estimates, including those herein, is that they fail to control (at least directly) for working conditions. To the extent that union jobs have more adverse working conditions than do nonunion jobs (Duncan and Stafford, 1980), union gap estimates are biased upward. The extent of this bias is not known. Even if appropriate data were available, estimation of compensating differentials associated with working conditions is notoriously difficult (Rosen, 1986; Hwang et al., 1992). To some extent, working conditions are controlled for through the inclusion of occupation and industry dummies. As discussed previously, control for occupation drives union gap estimates upward while control for industry drives gap estimates downward. Use of detailed industry controls, while helping control for working conditions, also captures some of what we consider a union effect and exacerbates downward bias from misclassification error in union status.

The literature on compensating differentials finds that union workers are compensated more highly for risks than are nonunion workers (Viscusi and Aldy, 2003). One explanation is that this reflects a lack of information or some form of market failure in nonunion companies, one that is mitigated by union voice. Alternatively, if compensating differentials in the nonunion sector reflect competitive returns, higher differentials in union jobs is simply one way in which unions produce a premium for their members. It is difficult to distinguish between these explanations (for an informal effort to do so with respect to workers' compensation claims, see Hirsch et al., 1997).

---

<sup>30</sup> Before attaching too much weight to ECEC wage figures, note that nonunion (but not union) wage growth in the ECEC is substantially lower than that seen in the ECI or CPS (Hirsch et al., 2003).

## VII. *Summary*

Conventional wisdom among labor economists holds that the average union wage premium is approximately 15 percent. Accompanying this loose consensus have been various arguments that can lead one to argue that union gaps are either higher or lower than 15 percent. Whatever one's view, arguments presented herein show that there exists a substantial downward bias in standard union gap estimates. The bias emanates from two sources of measurement error. The first source, not well known among economists, is the "match bias" associated with Census earnings imputation procedures. The second is a well-known source of bias, but one thought to be unimportant for wage level estimates of union gaps – measurement error resulting from the misclassification of union status.

Based on a 2001 CPS sample of private sector nonagricultural workers, a .13 log wage gap is obtained using the full sample of workers with and without imputed earnings. The union gap rises to .18 when nonrespondents are excluded from the sample. Correcting for misclassification bias (with an assumed 2 percent rate) raises the private sector union wage gap from .18 to .24 log points in 2001. Based on misclassification rates of 1 or 3 percent, the union gap increases from .18 to .21 or .28, respectively.

The combined effect of misclassification and match bias is not only sizable, but has become larger over time as nonresponse rates have risen and private sector density has declined. I have presented time-consistent private sector union wage gaps for 1973-2001, with correction for match bias and misclassification bias. Absent these corrections, standard estimates substantially understate the true union wage gap and overstate decline in the union gap since the mid-1980s. The precise magnitude of the private sector gap and its decline depends very much on assumed rates of misclassification. Unfortunately, when one turns to data sources other than the CPS to clarify trends in union and nonunion wage growth, further puzzles arise.

How plausible is a union wage gap substantially greater than 15 percent? The answer depends in part on one's beliefs about employers' ability and incentives to offset union wage increases with skill upgrading of their work forces. It is argued herein that in a repeated bargaining framework, skill upgrading need not follow. Nor does empirical evidence reveal substantial skill upgrading among union employers. Longitudinal studies indicate that two-sided selection produces a skill distribution among union workers that is more compressed than among a nonunion work force with the same distribution of measured earnings attributes. Although wage level evidence produces union wage gap estimates that decline sharply with

respect to measured skill attributes (e.g., schooling), true union gaps may differ little across skill groups following labor market sorting and an accounting for two-sided selection.

Union wage gaps in the U.S., long thought to be about 15 percent, have been considered high as compared to union gaps in other countries (Blanchflower and Freeman, 1992; Blanchflower and Bryson, 2002). I argue that the average union wage gap in the U.S. private sector is substantially higher than previously believed. An even higher U.S. union wage gap makes the international comparison all the more notable.<sup>31</sup> One cannot avoid the conclusion that a relatively high wage premium in the U.S. explains *part* of the fierce management opposition to union organizing and the inability of unions to maintain private sector union density at past or current rates.

The study of union wage gaps has a long history in labor economics, although interest in this topic may be waning as density declines. Recent research suggests that there remains much to be learned. There never has been a consensus on the most appropriate methods by which union gaps should be estimated. The literature continues to await modeling and estimation approaches that can reliably identify the causal effects of unions on wages. Independent of the approach, the measurement issues addressed herein require one to revise one's assessment of both the level and trend in the average union wage gap. The conventional wisdom of a 15 percent union gap requires reconsideration. Whether a new consensus will emerge is uncertain.

---

<sup>31</sup> I do not know in which countries or data sources there exist match bias from earnings imputation. To the extent that union density is higher in other countries, measurement error bias will be less severe for given rates of misclassification. In countries where union status is less clearly defined than in the U.S., rates of misclassification may be high.

## REFERENCES

- Abowd, John M. and Changhui Kang. "Simultaneous Determination of Wage Rates and Tenure." Cornell University, March 2001.
- Abowd, John M. and Henry S. Farber. "Job Queues and the Union Status of Workers." *Industrial and Labor Relations Review* 35 (April 1982): 354-67.
- Addison, John T. and Barry T. Hirsch. "Union Effects on Productivity, Profits, and Growth: Has the Long Run Arrived?" *Journal of Labor Economics* 7 (January 1989): 72-105.
- Angrist, Joshua D. and Alan B. Krueger. "Empirical Strategies in Labor Economics." In O. Ashenfelter and D. Card, eds. *Handbook of Labor Economics, Vol. 3A*. Amsterdam: North-Holland, 1999, pp. 1277-1366.
- \_\_\_\_\_. "Instrumental Variables and the Search for Identification: From Supply and Demand to Natural Experiments." *Journal of Economic Perspectives* 15 (Fall 2001): 69-85.
- Blanchflower, David G. "Changes Over Time in Union Relative Wage Effects in Great Britain and the United States." In S. Daniel, P. Arestis and J. Grahl, eds. *The History and Practice of Economics. Essays in Honor of Bernard Corry and Maurice Peston, Vol. 2*. Northampton, Mass.: Edward Elgar, 1999, pp. 3-32.
- \_\_\_\_\_ and Alex Bryson. "Changes Over Time in Union Relative Wage Effects in the UK and the US Revisited." NBER Working Paper #9395, December 2002.
- Blanchflower, David G. and Richard B. Freeman. "Unionism in the United States and Other Advanced OECD Countries." In M. Bognanno and M. Kleiner, eds. *Labor Market Institutions and the Future Role of Unions*. Cambridge, Mass.: Blackwell Publishers, 1992, pp. 56-79.
- Bloch, Farrell E. and Mark S. Kuskin. "Wage Determination in the Union and Nonunion Sectors." *Industrial and Labor Relations Review* 31 (January 1978): 183-92.
- Bollinger, Christopher R. "Response Error and the Union Wage Differential." *Southern Economic Journal* 68 (July 2001): 60-76.
- Booth, Alison L. *The Economics of the Trade Union*. Cambridge: Cambridge University Press, 1995.
- \_\_\_\_\_ and Mark L. Bryan. "The Union Membership Wage-Premium Puzzle: Is there a Free-Rider Problem?" Center for Economic Policy Research, Discussion Paper No. 2879, July 2001.
- Bratsberg, Bernt and James F. Ragan, Jr. "Changes in the Union Wage Premium by Industry." *Industrial and Labor Relations Review* 56 (October 2002): 65-83.
- Brown, Charles and James Medoff. "The Employer Size-Wage Effect." *Journal of Political Economy* 97 (October 1989): 1027-59.
- Budd, John W. and In-Gang Na. "The Union Membership Wage Premium for Employees Covered by Collective Bargaining Agreements." *Journal of Labor Economics* 18 (October 2000): 783-807.
- Card, David. "The Effect of Unions on the Structure of Wages: A Longitudinal Analysis." *Econometrica* 64 (July 1996): 957-79.
- \_\_\_\_\_, Thomas Lemieux, and W. Craig Riddell. "Unionization and Wage Inequality: A Comparative Study of the U.S., the U.K., and Canada." NBER Working Paper #9473, January 2003.

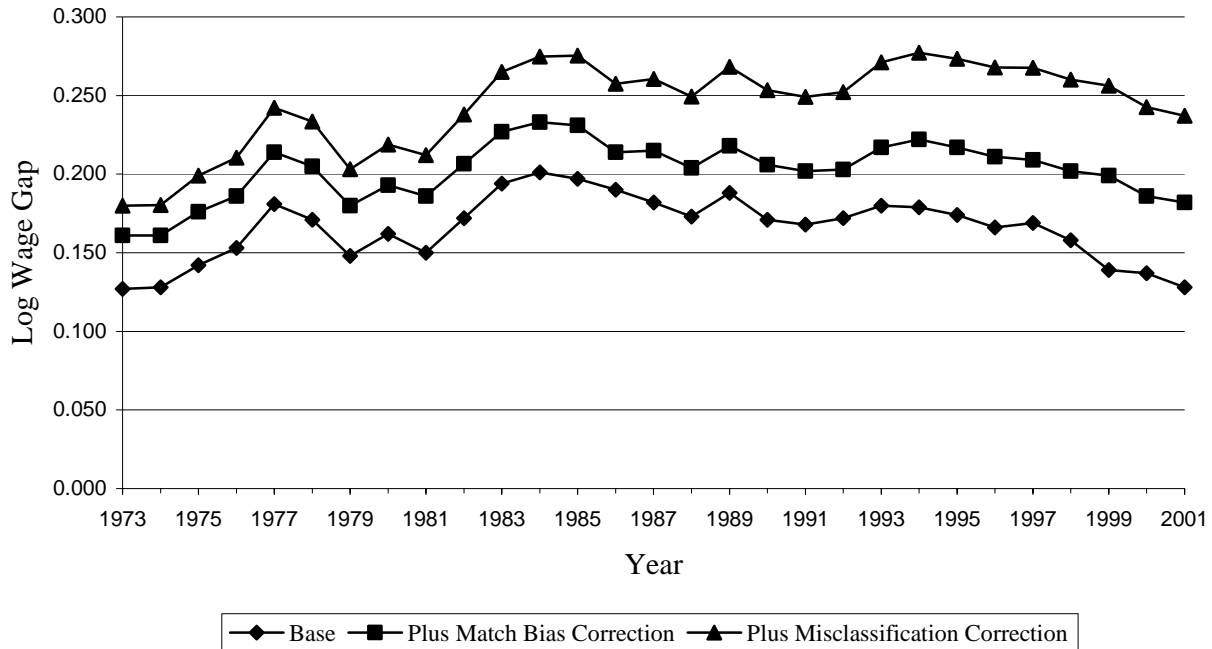
- Clark, Kim B. "Unionization and Firm Performance: The Impact on Profits, Growth, and Productivity." *American Economic Review* 74 (December 1984): 893-919.
- Curme, Michael and David A. Macpherson. "Union Wage Differentials and the Effects of Industry and Local Union Density: Evidence from the 1980s." *Journal of Labor Research* 12 (Fall 1991): 419-27.
- Duncan, Greg J. and Frank P. Stafford. "Do Union Members Receive Compensating Wage Differentials?" *American Economic Review* 70 (June 1980): 355-71.
- Duncan, Gregory M. and Duane E. Leigh. "Wage Determination in the Union and Nonunion Sectors: A Sample Selectivity Approach." *Industrial and Labor Relations Review* 34 (October 1980): 24-34.
- Farber, Henry S. "The Determination of the Union Status of Workers." *Econometrica* 51 (September 1983): 1417-38.
- \_\_\_\_\_. "The Analysis of Union Behavior." In O. Ashenfelter and R. Layard, eds. *Handbook of Labor Economics, Vol. II*. Amsterdam: North-Holland, 1986, pp. 1039-89.
- \_\_\_\_\_. "Mobility and Stability: The Dynamics of Job Change in Labor Markets." In O. Ashenfelter and D. Card, eds. *Handbook of Labor Economics, Vol. 3B*. Amsterdam: North-Holland, 1999, pp. 2439-83.
- \_\_\_\_\_ and Bruce Western. "Accounting for the Decline of Unions in the Private Sector, 1973-1988." In James T. Bennett and Bruce E. Kaufman, eds. *The Future of Private Sector Unionism in the United States*. Armonk, N.Y.: M.E. Sharpe, 2002, pp. 28-58.
- Freeman, Richard B. "The Exit-Voice Tradeoff in the Labor Market: Unionism, Job Tenure, Quits and Separations." *Quarterly Journal of Economics* 94 (June 1980): 643-74.
- \_\_\_\_\_. "Unionism and the Dispersion of Wages." *Industrial and Labor Relations Review* 34 (October 1980): 3-23.
- \_\_\_\_\_. "The Effect of Unionism on Fringe Benefits." *Industrial and Labor Relations Review* 34 (July 1981): 489-509.
- \_\_\_\_\_. "Union Wage Practices and Wage Dispersion Within Establishments." *Industrial and Labor Relations Review* 36 (October 1982): 3-21.
- \_\_\_\_\_. "Longitudinal Analyses of the Effects of Trade Unions." *Journal of Labor Economics* 2 (January 1984): 1-26.
- \_\_\_\_\_. "In Search of Union Wage Concessions in Standard Data Sets." *Industrial Relations* 25 (Spring 1986): 131-45.
- \_\_\_\_\_ and James L. Medoff. "The Impact of the Percent Organized on Union and Nonunion Wages." *Review of Economics and Statistics* 63 (November 1981): 561-72.
- \_\_\_\_\_. *What Do Unions Do?* New York: Basic Books, 1984.
- Fuchs, Victor R., Alan B. Krueger, and James M. Poterba. "Economists' Views about Parameters, Values, and Policies: Survey Results in Labor and Public Economics." *Journal of Economic Literature* 36 (September 1998): 1387-425.
- Giles, D.E.A. "The Interpretation of Dummy Variables in Semilogarithmic Equations: Unbiased Estimation." *Economics Letters* 10 (1982): 77-79.

- Grant, Darren. "A Comparison of the Cyclical Behavior of Union and Nonunion Wages in the United States." *Journal of Human Resources* 36 (Winter 2001): 31-57.
- Heckman, James J. and Solomon W. Polachek. "Empirical Evidence on the Functional Form of the Earnings-Schooling Relationship." *Journal of the American Statistical Association* 69 (June 1974): 350-54.
- Hirsch, Barry T. "The Interindustry Structure of Unionism, Earnings, and Earnings Dispersion." *Industrial and Labor Relations Review* 36 (October 1982): 22-39.
- \_\_\_\_\_. *Labor Unions and the Economic Performance of U.S. Firms*. Kalamazoo, Mich.: Upjohn Institute for Employment Research, 1991.
- \_\_\_\_\_ and John T. Addison. *The Economic Analysis of Unions: New Approaches and Evidence*. Boston: Allen & Unwin, 1986.
- Hirsch, Barry T. and Mark C. Berger. "Union Membership Determination and Industry Characteristics." *Southern Economic Journal* 50 (January 1984): 665-79.
- Hirsch, Barry T. and David A. Macpherson. "Union Membership and Coverage Database from the Current Population Survey: Note," *Industrial and Labor Relations Review* 56 (January 2003): 349-54, and accompanying data site <http://www.unionstats.com/>.
- Hirsch, Barry T., David A. Macpherson, and J. Michael DuMond. "Workers' Compensation Reciprocity in Union and Nonunion Workplaces." *Industrial and Labor Relations Review* 50 (January 1997): 213-36.
- Hirsch, Barry T., David A. Macpherson, and Edward J. Schumacher. "Measuring Union and Nonunion Wage Growth: Puzzles in Search of Solutions." In P. Wunnava, ed. *The Changing Role of Unions: New Forms of Representation*. Armonk, N.Y.: M.E. Sharpe, 2003.
- Hirsch, Barry T. and Edward J. Schumacher. "Match Bias in Wage Gap Estimates Due to Earnings Imputation." *Journal of Labor Economics* 22 (July 2004): forthcoming.
- \_\_\_\_\_. "Unions, Wages, and Skills." *Journal of Human Resources* 33 (Winter 1998): 201-19.
- Hirsch, Barry T., Michael L. Wachter, and James W. Gillula. "Postal Service Compensation and the Comparability Standard." *Research in Labor Economics* 18 (1999): 243-79.
- Hwang, Hae-shin, W. Robert Reed, and Carlton Hubbard. "Compensating Wage Differentials and Unobserved Productivity." *Journal of Political Economy* 100 (August 1992): 835-58.
- Jarrell, Stephen B. and T.D. Stanley. "A Meta-Analysis of the Union-Nonunion Wage Gap." *Industrial and Labor Relations Review* 44 (October 1990): 54-67.
- Jones, Ethel B. "Union/Nonunion Wage Differentials: Membership or Coverage?" *Journal of Human Resources* 17 (Spring 1982): 276-85.
- Kalachek, Edward and Frederic Raines. "Trade Unions and Hiring Standards." *Journal of Labor Research* 1 (Spring 1980): 63-75.
- Lazear, Edward P. "A Competitive Theory of Monopoly Unionism." *American Economic Review* 73 (September 1983): 631-43.
- Lee, Lung-Fei. "Unionism and Wage Rates: A Simultaneous Equations Model with Qualitative and Limited Dependent Variables." *International Economic Review* 19 (June 1978): 415-34.



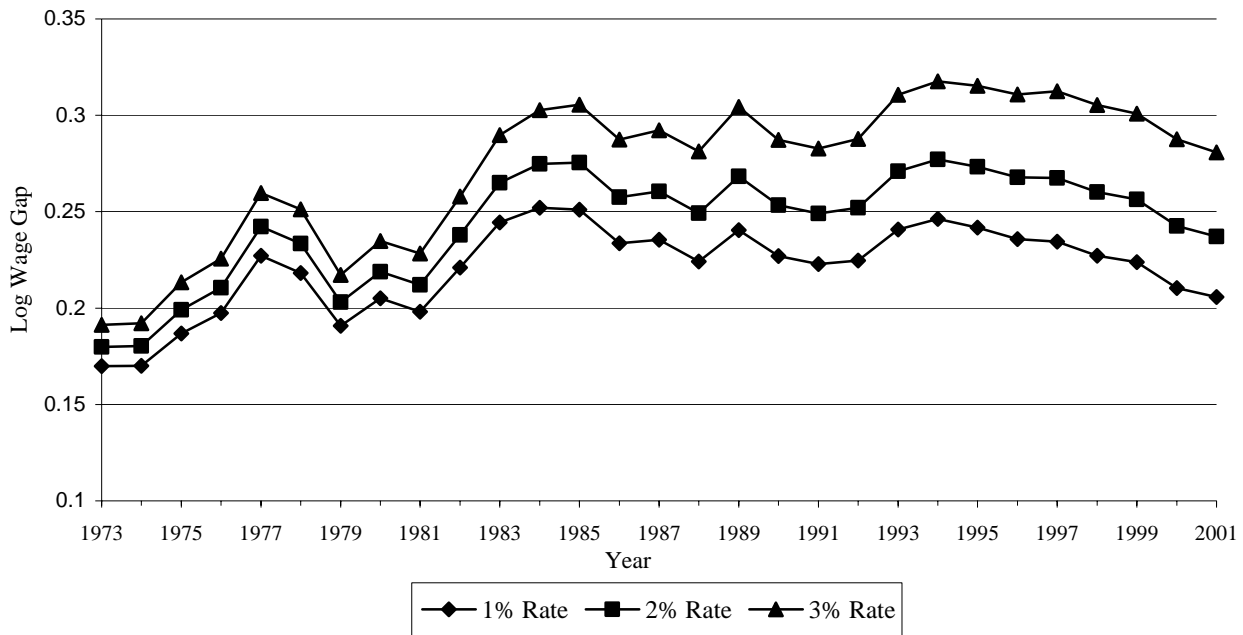
- Lemieux, Thomas. "Estimating the Effects of Unions on Wage Inequality in a Panel Data Model with Comparative Advantage and Non-Random Selection." *Journal of Labor Economics* 16 (June 1978): 261-91.
- Lettau, Michael K., Mark A. Loewenstein, and Aaron T. Cushner. "Explaining the Differential Growth Rates of the ECI and the ECEC." *Compensation and Working Conditions* 2 (Summer 1997): 15-23.
- Lewis, H. Gregg. *Unionism and Relative Wages in the United States: An Empirical Inquiry*. Chicago: University of Chicago Press, 1963.
- \_\_\_\_\_. *Union Relative Wage Effects: A Survey*. Chicago: University of Chicago Press, 1986.
- Macpherson, David A. "Trade Unions and Labor's Share in U.S. Manufacturing Industries." *International Journal of Industrial Organization* 8 (March 1990): 143-51.
- Mellow, Wesley. "Unionism and Wages: A Longitudinal Analysis." *Review of Economics and Statistics* 63 (February 1981): 43-52.
- \_\_\_\_\_ and Hal Sider. "Accuracy of Response in Labor Market Surveys: Evidence and Implications." *Journal of Labor Economics* 1 (October 1983): 331-44.
- Mincer, Jacob. *Schooling, Experience, and Earnings*. New York: Columbia University Press, 1974.
- \_\_\_\_\_. "Union Effects: Wages, Turnover, and Job Training." In J.D. Reid, Jr., ed. *New Approaches to Labor Unions*. Greenwich, Conn.: JAI Press, 1983, pp. 217-52.
- Oi, Walter Y. and Todd L. Idson. "Firm Size and Wages." In O. Ashenfelter and D. Card, eds. *Handbook of Labor Economics, Vol. 3B*. Amsterdam: North-Holland, 1999, pp. 2165-2214.
- Pencavel, John H. "The Demand for Union Services: An Exercise." *Industrial and Labor Relations Review* 24 (January 1971): 180-90.
- Robinson, Chris. "The Joint Determination of Union Status and Union Wage Effects: Some Tests of Alternative Models." *Journal of Political Economy* 97 (June 1989): 639-67.
- Rosen, Sherwin. "The Theory of Equalizing Differentials." In O.C. Ashenfelter and R. Layard, eds. *Handbook of Labor Economics, Vol. 1*. Amsterdam: North-Holland, 1986, pp. 641-92.
- Schumacher, Edward J. "What Explains Wage Differences Between Union Members and Covered Nonmembers?" *Southern Economic Journal* 65 (January 1999): 493-512.
- Viscusi, W. Kip and Joseph E. Aldy. "The Value of a Statistical Life: A Critical Review of Market Estimates throughout the World." NBER Working Paper #9487, February 2003.
- Wessels, Walter. "Do Unionized Firms Hire Better Workers?" *Economic Inquiry* 32 (October 1994): 616-29.
- Willis, Robert J. "Wage Determinants: A Survey and Reinterpretation of Human Capital Earnings Functions." In O.C. Ashenfelter and R. Layard, eds. *Handbook of Labor Economics, Vol. 1*. Amsterdam: North-Holland, 1986, pp. 525-602.
- U.S. Department of Labor, Bureau of Labor Statistics. *Employment Cost Index, Historical Listing*. January 30, 2003. Available at <http://www.bls.gov/ncs/ect/>.
- U.S. Department of Labor, Bureau of Labor Statistics. *Employer Costs for Employee Compensation, Historical Listing (Annual) 1986-2001*. June 19, 2002. Available at <http://www.bls.gov/ncs/ect/>.

Figure 1  
*Private Sector Union Wage Gaps, With and Without Match Bias and  
 Misclassification Bias Corrections, 1973-2001*



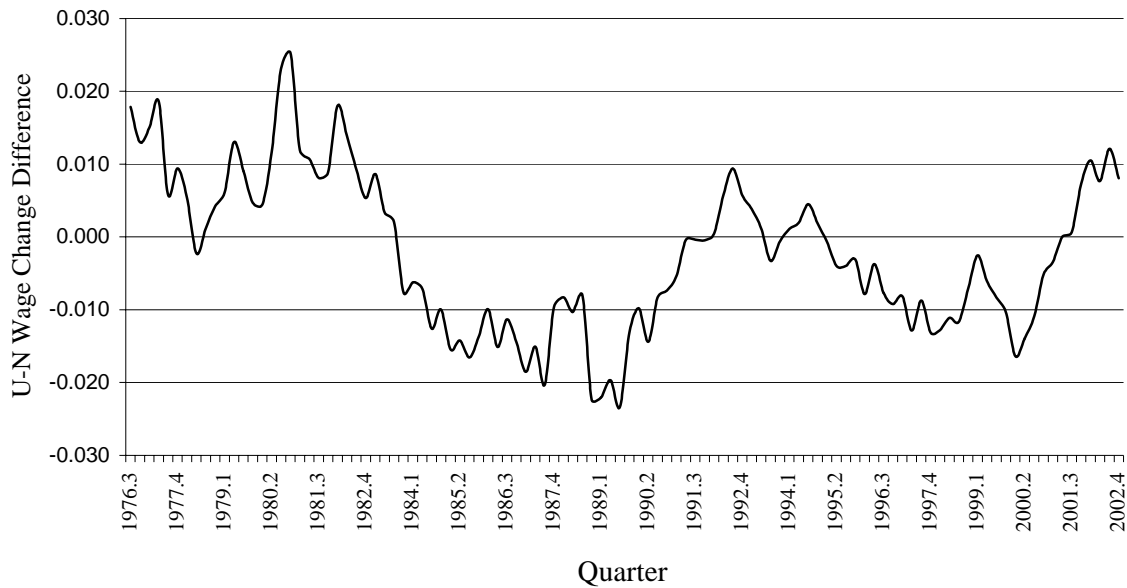
Source for the “Base” and “Base Plus Match Bias Correction” is Hirsch and Schumacher (2004). The series shown with “triangles” corrects for misclassification bias based on a 2 percent misclassification rate (see text). Data for 1973-1981 are from the May CPS Earnings Supplements and for 1983-2001 from the monthly CPS-ORG earnings files. There was no union status variable in 1982. The sample includes employed private sector nonagricultural wage and salary workers ages 16 and over with positive weekly earnings. The union gap for each year is the coefficient on a membership dummy variable in a regression where the log of hourly earnings is the dependent variable. Control variables included are years of schooling, experience and its square (allowed to vary by gender), and dummy variables for gender, race and ethnicity (3), marital status (2), part-time status, region (8), large metropolitan area, industry (8), and occupation (12). The series labeled “Base” include the full sample (workers with and without earnings imputed) for the years 1979-2001 and estimates of the full sample for 1973-1978. The “Corrected” series include only workers reporting earnings. Researchers who use all valid earnings records in CPS files would obtain gap estimates similar to the “squares” for 1973-1978 when CPS files do not include imputed earnings, and the “diamonds” beginning in 1979 when CPS files include imputed earnings values. See Hirsch and Schumacher (2004) and the text for details.

Figure 2  
*Private Sector Union Wage Gaps with Alternative Misclassification  
 Rate Assumptions, 1973-2001*



See the note to Figure 1. Each set of wage gap estimates starts with the series corrected for match bias, followed by corrections for misclassification bias based on assumed rates of 1 percent (diamonds), 2 percent (squares), and 3 percent (triangles). See the text for details.

Figure 3  
*Union and Nonunion Differences in ECI Annual Wage Change*



ECI Wage index values for 1975:3 through 2002:4 are from U.S. Bureau of Labor Statistics, Office of Compensation Levels and Trends, *Employment Cost Index: Historical Listing*, January 30, 2003, Table 7. U-N Wage Change Differences represent the 4-quarter change in log values for the union ECI index minus the nonunion change. A positive (negative) value represents an increase (decrease) in the union-nonunion wage gap.

## IZA Discussion Papers

No.	Author(s)	Title	Area	Date
780	A. Engellandt R. T. Riphahn	Temporary Contracts and Employee Effort	1	05/03
781	J. H. Abbring J. R. Campbell	A Structural Empirical Model of Firm Growth, Learning, and Survival	5	05/03
782	M. Güell B. Petrongolo	How Binding Are Legal Limits? Transitions from Temporary to Permanent Work in Spain	1	05/03
783	B. T. Hirsch E. J. Schumacher	Match Bias in Wage Gap Estimates Due to Earnings Imputation	5	05/03
784	O. Pierrard H. R. Sneessens	Low-Skilled Unemployment, Biased Technological Shocks and Job Competition	2	05/03
785	R. Almeida	The Effects of Foreign Owned Firms on the Labor Market	2	05/03
786	P. Mueser K. R. Troske A. Gorislavsky	Using State Administrative Data to Measure Program Performance	6	05/03
787	P. Garibaldi L. Pacelli A. Borgarello	Employment Protection Legislation and the Size of Firms	3	05/03
788	F. Büchel J. R. Frick	Immigrants in the UK and in West Germany – Relative Income Position, Income Portfolio, and Redistribution Effects	1	06/03
789	J. A. Cabral Vieira A. R. Cardoso M. Portela	Recruitment and Pay at the Establishment Level: Gender Segregation and the Wage Gap in Portugal	5	06/03
790	A. Cigno A. Luporini A. Pettini	Hidden Information Problems in the Design of Family Allowances	3	06/03
791	M. Güell	Fixed-Term Contracts and the Duration Distribution of Unemployment	1	06/03
792	L. Borghans B. ter Weel	What Happens When Agent T Gets a Computer? The Labor Market Impact of Cost Efficient Computer Adoption	5	06/03
793	T. Palokangas	Foreign Direct Investment, Labour Market Regulation and Self-Interested Governments	3	06/03
794	P. Frijters M. A. Shields S. Wheatley Price	Investigating the Quitting Decision of Nurses: Panel Data Evidence from the British National Health Service	1	06/03
795	B. T. Hirsch	Reconsidering Union Wage Effects: Surveying New Evidence on an Old Topic	3	06/03

An updated list of IZA Discussion Papers is available on the center's homepage [www.iza.org](http://www.iza.org).