

Response Error and the Union Wage Differential

Christopher R. Bollinger*

Broad variation in estimates of the union wage gap has perplexed labor economists. One specification error that is consistent with the observed variation is measurement error in reported union status. This article applies results of Bollinger (1996) to estimate a range for the union wage gap. Both a cross-sectional model and a fixed-effects model are estimated. In order for the true coefficient in the fixed-effects model to be bounded below the true coefficient in the cross-sectional estimates, measurement error would have to be less than 0.8%. The difference between the fixed-effects estimates and the cross-sectional estimates is primarily due to measurement error rather than to unobserved heterogeneity. An examination of differences in returns to union membership by industry, occupation, and educational level shows that these differences are largely robust to measurement error. Many of these differences would be found even if error rates were as high as 10% or more.

1. Introduction

An array of empirical estimates for the union wage differential has resulted from the variety of approaches to estimation. Lewis (1986) reviews the vast literature that attempts to estimate the union wage effect. Two extremes are represented by Mincer (1983), who estimates the wage differential to be 0.01, while Farber (1990) reports an estimate of 0.26. Clearly, specification error is the root cause of these differences. The literature has focused upon unobserved heterogeneity in worker status as the main specification error. Frequently, estimation approaches based on “within” estimators applied to fixed-effects models using panel data are used to account for the possibility of unobserved heterogeneity. As would be predicted by unobserved heterogeneity, Lewis (1986, p. 94) reports that “the panel wage gap estimates surveyed in the chapter on the average are roughly half as large as the corresponding cross-section estimates.” This is often taken as evidence for bias in the cross-section estimates due to the presence of fixed effects. However, these results are also consistent with measurement error in the report of the union status. Indeed, chapter 5 in Lewis (1986) focuses upon this possibility: “the difference might be the result of union status measurement error” (p. 94).

This article applies results of Bollinger (1996) to compare the effect of measurement error on cross-sectional estimates and panel estimates of the union wage differential. Bollinger (1996) establishes bounds for the slope coefficients of a linear regression when a binary regressor is thought to have measurement error. The results here do not identify a point estimate of the union wage differential; they relax many of the assumptions that are required to obtain the point estimates. These bounds serve the purpose of sensitivity analysis as called for by Leamer (1985).

* Department of Economics, University of Kentucky, Lexington, KY 40506, USA; E-mail erboll@pop.uky.edu.

I thank Chuck Manksi, Art Goldberger, Jim Walker, John Garen, Dan Black, and two anonymous referees for many helpful comments and suggestions.

Received May 1999; accepted September 2000.

Bounds for parameter estimates answer the question, “How sensitive to measurement error are the results we typically observe?” The results below establish that, in the presence of no prior information on the extent of measurement error in the union status, a remarkably wide range of estimates for the union wage differential are allowed: in cross section, from 15% to over 600%, and in fixed-effects panel model, from 5% to over 4000%. In spite of the wide range demonstrated by the bound, the analysis here reveals a number of important results. First, the bounds for the cross section are much tighter than those for the fixed-effects estimates, clearly demonstrating the extraordinary impact of measurement error on “within” estimators: The cross-sectional estimates are much less biased by measurement error. The upper bounds represent a case of maximal measurement error; additional information will substantially tighten these bounds. This allows the hypothetical question “How low does measurement error have to be for the panel estimates to be bounded below the cross sectional estimates?” to be answered. The results are striking: There must be less than 1% misclassification, a rate substantially lower than any estimates currently available. These two points suggest, as Lewis (1986) argues, that the cross-sectional estimates may be more reliable than the panel estimates. This implies that the differences between cross-sectional and fixed-effects estimates of the union wage differential are due to measurement error rather than unobserved heterogeneity.

This article also examines how robust differences in returns to union status across occupation, industry, and educational groups are to measurement error. It has typically been found in cross section that the union wage differential varies across these groups. One possibility for this finding is that the error rates differ across these groups. Here, many of the differences typically found in the union literature appear to be quite robust to measurement error: Rates even as high as 10% would support differences in some categories. In particular, it is found that workers in the construction and retail industry earn the highest union premium, but differences between construction and retail may be due to measurement error. Manufacturing and service industry workers earn the lowest (and a negative return for the financial industry), but measurement error may be the reason for differences between them. It is also found that service occupations and operators, fabricators, and laborers have the highest union premium, but measurement error may account for differences between service and operators, fabricators and laborers. Also, measurement error may account for differences between precision production craft and repair occupations and technical sales and administrative support. The educational findings are somewhat stronger. The return to union membership for those with no high school is clearly higher than any other category. The return to union membership is next highest for high school graduates and is clearly larger than for those with college. This relationship appears robust to measurement error.

This article differs from Bollinger (1996) in three important ways. Bollinger (1996) derives and proves the theorems upon which the analysis here is based. The theoretical results are the primary contribution of that paper. That paper uses a small subsample of the outgoing rotation groups of the May 1985 Current Population Survey (CPS) to illustrate the bounds and has a limited set of analysis examining the sensitivity of the bounds to additional information. This article extends the methodology in Bollinger (1996) to include “within” estimators applied to panel data. The methodology then allows a comparison between panel and cross-sectional bounds, which is a major focus of this article. This article also examines union differentials by six industry and six occupational categories and establishes that some of the differences in returns to union status may be due to differential response error across industry or occupational group, but some of the differences are robust. Finally, this article examines union differentials

by educational level. Here, in contrast to the occupation and industry category, the differences in return to union status are found to be quite robust to measurement error. The article also differs in a number of other dimensions: The sample includes all outgoing rotation groups from 1989, resulting in a much larger sample than in Bollinger (1996). The sample here is composed only of prime aged men, removing questions concerning selection of women into the labor force.

That measurement error exists is quite well documented. An excellent survey can be found in Bound, Brown, and Mathiowetz (2001). Freeman (1984), Peracchi and Welch (1995), Bound and Kreuger (1991), and Bollinger (1998) all explore measurement error in the CPS. Some papers that attempt to address measurement error in union status reports are Chowdhury and Nickell (1985), Mellow and Sider (1983), Freeman (1984), Jakubson (1991), Card (1996), Hirsch and Schumacher (1998), and Budd and Na (2000). In Mellow and Sider (1983), Freeman (1984), and Card (1996), auxiliary data from the 1977 CPS employer–employee match were used to estimate the misclassification rate in CPS reports of union status. This approach has considerable appeal. However, in order to use the matched data, one of two approaches is taken. Both Mellow and Sider (1983) and Freeman (1984) assume that the employer report of union status is without error. While certainly possible, Card (1996) argues convincingly that it is improbable. Card (1996) then assumes that the employer and employee can both make errors, but then must go on to assume that the error processes are independent yet have the same error rate. He further must assume that the rate of classifying union workers as nonunion is equal to the rate of classifying nonunion workers as union. Again, while possible, these are not trivial assumptions. Another concern when using the 1977 match data is the differences in the CPS questionnaire. Many of the reforms (both in 1988 and again in 1991) were designed to reduce measurement error. For a review of these reforms, see Polivka and Rothgeb (1993).

Chowdhury and Nickell (1985) do not arrive at estimates that are free from measurement error, but argue that an instrumental variable approach, using multiple years of union status data, reduces the bias. Similarly, Hirsch and Schumacher (1998) examine the effect of removing observations with allocated union status or proxy interviews. The data used here remove allocated observations, but Bollinger and David (1997) find that proxy interviews are at least as accurate as actual interviews. Hirsch and Schumacher (1998) further explore using changes in occupation or industry coincident with changes in union status to reduce measurement error. Budd and Na (2000) argue that agreement in reports across years indicates more likelihood that the reports are accurate and use this information in cross-sectional estimates. Jakubson (1991) uses a method of moments approach, which is quite general but, among other problems, requires a minimum of three annual observations on each individual. While Jakubson's approach is useful, it cannot be applied to the main data set of choice for estimation of the determinants of wages: the Current Population Survey. Other data sets are less reliable due to sampling frame and sample size. In each of these studies, however, the authors find that the differences between the cross-sectional estimates and estimates based on a fixed-effects model are due largely to measurement error.

Section 2 of this paper briefly describes the effect of misclassification in the union status variable on estimates of the wage differential and summarizes the results from Bollinger (1996) for the linear model. In section 3, general descriptive statistics and traditional regression results are reported. Section 4 compares and contrasts the general bounds for the cross-sectional and fixed effects models described in section 2. In section 5, the union wage premium is examined by industry, occupation, and educational attainment. Section 6 contains concluding remarks.

2. Methodology

It has long been understood that measurement error in explanatory variables causes least-squares estimates of parameters in a linear model to be inconsistent. This fact was pointed out by econometricians as early as Frisch (1934), Koopmans (1937), and Reiersøl (1950). A number of identifying assumptions have been suggested to remedy this situation; see Fuller (1987) or Aigner et al. (1984) for excellent surveys. The literature on measurement error focuses on the classical errors-in-variables (CEIV) model. This model assumes that the observed variable differs from the true variable by a random component that is uncorrelated with all other variables in the model.

The CEIV model cannot be used as a framework for the problem of response error in the union status variable. If Z is the union status variable ($Z = 1$ if the worker is a union member, zero otherwise), the difference between the true union status variable and the reported union status variable, $X - Z$, cannot be uncorrelated with the true union status variable. The response error can be either 0 or 1 if $Z = 0$, and it can be either 0 or -1 if $Z = 1$. Model I is a simple binary misclassification (BEIV) model:

$$Y = \alpha + \beta Z + u, \quad E[u|Z] = 0 \quad (1)$$

$$\Pr[X = 1|Z, Y] = (1 - q)Z + p(1 - Z) \quad (2)$$

$$\Pr[X = 0|Z, Y] = qZ + (1 - p)(1 - Z), \quad (3)$$

$$p + q < 1. \quad (A1)$$

The researcher is only able to observe the pair $\{Y_i, X_i\}$ and wishes to estimate β . The variable X is the reported union status variable. The variable Z is the true union status. Hence, p is the probability of being classified as a union member when the individual is not in a union; q is the probability of being classified as not in a union when the individual is actually a union member. The assumption that $p + q < 1$ insures that covariance of the true union status and the observed union status is positive. In other words, the measurement error is not so severe that the definition of the variable has been reversed.

Aigner (1973) studied this type of model and showed that, as in the CEIV model, the ordinary least squares (OLS) estimate of β is biased toward zero, and in general, β was not identified. Rather than impose further possibly erroneous restrictions on the model to gain identification, the approach taken here establishes bounds on the model parameters. The idea of bounding parameters in the classic errors-in-variables model was first suggested by Frisch (1934) and extended by Koopmans (1937). Further extensions by Klepper and Leamer (1984) and Klepper (1988) have been considered.

Bollinger (1996) establishes bounds for the parameters of the BEIV model, which are applied here. In the simple model above, the lower bound is the slope coefficient (b) from the least squares regression of Y on X . This simply uses the well-known result that the OLS slope is attenuated due to measurement error. In classical measurement error models, Frisch (1934) first established that the inverse of the reverse regression (d) provides an upper bound; that is, by regressing X on Y and taking the inverse of the resulting slope (d), an upper bound for the true parameter β can be found. In the BEIV model, this upper bound can be improved upon. The additional information implied by the special distribution of the response error leads to an upper bound that is a linear combination of b and d . Specifically, Bollinger showed that

$$\beta \leq \max \begin{cases} P_x b + (1 - P_x) d \\ (1 - P_x) b + P_x d, \end{cases} \quad (4)$$

where P_x is the mean of the observed X (sample proportion of ones). In addition to the bounds on β , Bollinger (1996) provides bounds on p and q . Bollinger (1996) also extends the results to include other regressors. The modified bounds are then

$$\beta \leq \max \begin{cases} P_x(1 - R_{xw}^2)b + ((1 - P_x) + P_x R_{xw}^2)d \\ (1 - P_x)(1 - R_{xw}^2)b + (P_x + (1 - P_x)R_{xw}^2)d, \end{cases} \quad (5)$$

where b is the slope coefficient on X from the regression of Y on the mismeasured X and any other regressors W , d is the inverse of the slope coefficient on Y from the reverse regression of X on Y and other regressors W , and R_{xw}^2 is the R^2 from the regression of X on the other regressors W . Throughout this article, these regressors W are assumed to be measured without error. An additional key assumption is that the measurement error process is independent of these regressors W . There are three important aspects to interpretation of the bounds that need to be discussed. The first are the conditions associated with β being equal to the upper or the lower bound. The second is the interpretation of estimated rather than population bounds. The third is how the approach can be extended to apply to models with other regressors and fixed effects models.

The lower bound is achieved when no measurement error is present. As is the case with all measurement error bounds, the lower bound is consistent with no measurement error: The only stochastic aspect to the relationship between Y and X is due to the residual u . The upper bound is consistent with maximal measurement error: The only stochastic aspect to the relationship between Y and X is due to measurement error. In contrast with the CEIV model, at the upper bound, the measurement error must be allocated to errors of omission ($X = 0$ when $Z = 1$, represented by q) or errors of commission ($X = 1$ when $Z = 0$, represented by p). The upper bound is then associated with a particular allocation of errors of omission and commission. The upper bound is associated with a lopsided error process. Either p will be zero and q will be large or q will be zero and p will be large. This allows the researcher to determine if the particular process necessary to achieve the upper bound is sensible.

Estimated bounds can be interpreted one of two ways. The first parallels the classical interpretation of point estimates: The estimated bounds are an estimate of the true parameters of interest, the upper and lower bounds. In population, these bounds would be a 100% confidence interval. In sample, one can construct a 95% confidence interval around each bound's estimate.

However, another very interesting interpretation is available. The estimated bounds must, by definition, contain the estimate of β that would be obtained if this data set did not have any measurement error; that is, if the researcher were suddenly given the correct data for Z in this particular sample and were to utilize OLS to estimate Equation 1, the estimate $\hat{\beta}$ would lie within these bounds.

In the empirical section, a model including additional control variables is estimated. Bounds for slope coefficients on all variables are presented. These bounds are referred to as the left and right bounds. The lower bound on the union coefficient is associated with a bound on each of the other coefficients in the model. These terms are collectively called the left bound. The upper bound on the union coefficient is associated with a bound on each of the other coefficients in the model. These terms are collectively called the right bound. For some variables, the left

bound is the lower bound (e.g., the union coefficient and education coefficient). For other cases, the left bound is the upper bound (e.g., the experience coefficient and the race coefficient). A particular feasible value of β_1 implies a particular set of values for the other coefficients β_2 .

The results in Bollinger (1996) do not specifically consider fixed-effects models but can be extended to cover this case. Equation 1 would be modified to

$$Y_{it} = \alpha + \beta Z_{it} + \theta_i + u_{it}. \quad (6)$$

The error terms u_{it} are assumed to be mean zero and mutually independent both across individuals and over time. One approach to estimation is to include an $N - 1$ vector of dummy variables (δ_j) for each individual, thus making the empirical specification

$$Y_{it} = \alpha + \beta Z_{it} + \sum_{j=1}^{N-1} \delta_{ij} \theta_j + u_{it}, \quad (7)$$

where $\delta_{ij} = 1$ if $i = j$ and is zero otherwise (an excellent discussion of estimating fixed effects models can be found in Wooldridge 2000). This specification is algebraically equivalent to the more popular differences-from-means approach. In data where only two periods are observed, the estimates are algebraically equivalent to first difference estimates as well. The dummy variable specification, though, is easily analyzed using the results of Bollinger (1996). The vector of dummy variables is now just a part of the "other regressors" W . The assumption that the measurement error process be independent of W now implies that the error propensity is the same across all individuals. It rules out fixed effects in the error process. This is equivalent to the assumptions made by Freeman (1984), Jakubson (1991), Card (1996), and Budd and Na (2000). Additionally, it tightens the bounds on the response error rates p and q . In Bollinger (1996), it is shown that $p < P_x(1 - R_{xw}^2)(1 - b/d)$ and $q < (1 - P_x)(1 - R_{xw}^2)(1 - b/d)$. The R^2 from the regression of X on all other regressors, including the $N - 1$ vector of dummy variables, is, of course, very large. But it does not result in tighter bounds on β . The additional information provided by the presence of other regressors is useful in reducing the potential total amount of response error but, because these regressors also significantly reduce the amount of other error (the variance of u), measurement error potentially increases as a percentage of total error in the system.

The methodology in Bollinger (1996) also allows external information about p and q to be applied to result in tighter bounds. This allows two types of approach: First, information from validation studies such as Freeman (1984) can be used to tighten the bounds. Second, one can turn the question around and find levels of p and q that would support a particular range for the union coefficient. While Bollinger (1996) uses the former approach, this article makes extensive use of the later approach in examining the robustness of standard results to the possibility of measurement error. It should be noted that there are many ways to bring in additional information to tighten the bounds. As noted above, the width of the bounds is due to randomness in the structural model (as represented by the variance of the error term u) plus randomness in the relationship between observed union status and the reported union status. The bounds above are derived by allocating the total amount of randomness to one category or the other. Placing limits on the measurement error rates (p and q) forces more of the total randomness to be allocated to the structural equation (in terms of the variance of u). Klepper (1988) uses another approach; by placing lower limits on the amount of randomness attributed to the structural equation, he is able to achieve tighter bounds. Another approach may be to

bound the range of the dependent variables (which would result in a bound on the variance of u). However, each of these approaches will have a corresponding bound achieved by limiting p and q ; that is, any approach that uses other information to achieve tighter bounds can be recast in terms of bounds on p and q .

3. Data

The data used in this study are from the 1989 and 1990 Current Population Survey (CPS) outgoing rotations (rotation groups 4 and 8). All respondents in the outgoing rotation groups of the survey are asked questions about union status and earnings. The sample analyzed here consists of adult males aged 18 to 65 in rotation group 4 during 1989 and rotation group 8 during 1990 who worked full time in nonagricultural private wage and salary positions in the week prior to the interview. The 1989 and 1990 samples are matched based first upon household identification number and line number. The resulting matched pairs were then checked for agreement on age, race, and education. Individuals whose age decreased or increased by more than two years were eliminated. Approximately 5% of the sample either reported the same age as the previous year or an age two years larger. This variation was allowed due to differences in timing of the interviews from year to year. By definition, one would expect the outgoing rotation interview to fall in the same month as the birthday of 8% of the sample. Therefore, a possibility exists that the interview in 1989 occurred prior to an individual's birthday but the interview in 1990 occurred after or vice versa (for a more complete discussion, see Peracchi and Welch 1995). No changes in race were allowed. Since only men were matched, this also ensures no changes in gender. Individuals whose education decreased or increased by more than one year were discarded. This results in a sample of 14,347 individuals for two years, a total of 28,694 observations. Table 1 presents descriptive statistics of the matched sample for the variables used in the analysis. The reported means are based on 1989 values but are consistent (given the matching procedure) with the corresponding 1990 values (e.g., experience is one year larger but education remains unchanged).

The CPS reports the responses to two union status questions. In this study, the union status variable is one if the person is a union member. Table 1 shows 20% of the respondents report membership in a union. Additionally, the CPS asks nonunion members if they are covered by a collective bargaining agreement. Table 1 shows that 22% of the respondents are either union members or covered by such an agreement (2% are covered nonmembers). The results below are not appreciably different if the broader member or covered nonmember variable is used. Budd and Na (2000) and Jones (1982) both find differences in return between union members and covered nonmembers, so this article focuses on the union member return.

In addition to log wages and union status, measures of education in years, experience (age - education - 6) in years, and race (one if black) are also used. The data are also subdivided into broad industry and occupation categories. Descriptive statistics for these variables are reported in Table 1. The eighth row of Table 1, titled Raw union differential, gives the difference in average log weekly wage between union and nonunion workers. The standard error of that estimate is reported in the standard deviation column.

Table 2 provides traditional OLS estimates of a standard cross-sectional wage equation and estimates of the fixed-effects model. As is typically found, the union coefficient is positive and statistically significant, with the cross-sectional estimate substantially larger than the fixed-ef-

Table 1. Sample Means (Based on 1998 Observations of Matched Panel, $N = 14,347$)

Variable	Mean	Standard Deviation
Hourly wage	13.36	7.48
Log wage	2.46	0.52
Union member	0.20	0.40
Union coverage	0.22	0.41
Education	13.12	2.63
Experience	19.61	11.35
Black	0.07	0.26
Raw union differential	0.08	0.007
Industries		
Construction	0.096	0.29
Retail	0.123	0.328
Transportation, communications, and utilities	0.105	0.307
Wholesale	0.072	0.258
Manufacturing	0.377	0.485
Services industry	0.151	0.358
Financial, insurance, and real estate	0.058	0.235
Occupations		
Service	0.049	0.216
Operators, fabricators, and laborers	0.259	0.428
Precision production, craft, and repair	0.245	0.430
Technical, sales, and administrative support	0.201	0.401
Managerial and professional	0.247	0.431
Education		
No high school	0.044	0.021
Some high school	0.083	0.28
High school graduate	0.426	0.49
Some post high school	0.198	0.40
College plus	0.246	0.44

facts estimate. The 15% wage gap implied by the estimate is in the range typically reported (Lewis 1986, Tables 4.1, 4.2, 5.1, and 5.3), while the 5% estimate in the fixed-effects model is typical also. Other coefficients are not remarkable. The return to education is 9% per year. Experience has the usual diminishing marginal returns shape. Being black reduces wages by approximately 21.5%. This is consistent with other findings.

4. Comparison of Cross-Sectional and Fixed-Effect Bounds

The most general set of bounds estimable is for the case where the only assumption about the error rates is A1, which requires $p + q < 1$ and the independence from other regressors. Table 3 presents these bounds for both the cross-sectional model and a model with fixed effects. Since the left bounds are consistent with no measurement error, they are simply the slope coefficients from the OLS regressions presented in Table 2. Heteroskedastic consistent standard errors are reported in parentheses below each parameter estimate.

As can be seen in Table 3, the right (or upper) bounds on the union status variable are very large. The right bounds on other coefficients are also remarkable. The cross-sectional

Table 2. Base Regression Results

	Cross Section	Fixed Effects
Constant	0.732 (0.018)	—
Union	0.154 (0.006)	0.052 (0.010)
Education	0.096 (0.001)	0.107 (0.012)
Experience	0.038 (0.0008)	0.084 (0.004)
Experience ²	-0.0006 (0.00001)	-0.0006 (0.0001)
Black	-0.215 (0.010)	—
Sample size	14,437	28,694

Cross-sectional estimates based on 1998 observations of the matched 1998/1999 outgoing rotation group panel. Fixed-effects estimates use differences from means.

results place the union wage differential in the interval [0.15, 6.73]. The fixed-effect bounds are even wider, placing the union wage differential in the interval [0.05, 48.45]. Most would argue that a union wage differential of over 600% is not possible, and a differential in the 4800% range might be called ridiculous. Recalling that the upper bounds in each case are consistent with maximal error, the upper bound is not likely to be achieved. However, in the absence of additional information, this is the range supported by the data.

Prior to further examining the relationship between the cross-sectional and fixed-effects models, it is instructive to consider the bounds on other parameters. In the cross-sectional results, the intervals for black, [-0.22, -0.67], and education, [0.10, 0.34], do not cross the origin, suggesting that current estimates of the return to education and the effect of race are too low in magnitude due to response error in union status. However, the general conclusions that being

Table 3. Bounds for Both Models, Comparing Cross-Sectional to Fixed-Effects Model

	Cross-Sectional Model		Fixed-Effects Model	
	Left (Lower)	Right (Upper)	Left (Lower)	Right (Upper)
Constant	0.732 (0.018)	-2.886 (0.11)	—	—
Union	0.154 (0.006)	6.73 (0.258)	0.052 (0.010)	48.22 (9.63)
Education	0.096 (0.001)	0.336 (0.009)	0.107 (0.012)	-0.008 (0.46)
Experience	0.037 (0.0009)	-0.062 (0.006)	0.084 (0.004)	-0.144 (0.134)
Experience ²	-0.0006 (0.0001)	0.001 (0.0001)	-0.0006 (0.0001)	0.007 (0.003)
Black	-0.225 (0.024)	-0.6659 (0.087)	—	—
Sample size	14,347	14,347	28,694	28,694

Cross-sectional estimates based on 1998 observations from matched panel. Fixed-effects estimates based on differences from means estimator. Lower bounds based on no measurement error; upper bound represents maximal measurement error.

black is associated with lower wages and education is associated with higher wages are supported. The story is not so clear in the fixed-effects model estimates: The range for education is $[-0.08, 0.107]$. This suggests that measurement error in the union status variable may mask an education coefficient that is lower in panel estimates than in cross-sectional estimates. The coefficients on experience and experience squared both cross the original in each model. Thus, the measurement error may lead to an overstatement of the magnitude of the return and the amount of curvature.

Comparison of the intervals for the union coefficient reveals much about the impact of measurement error on estimates of the union wage differential. The fact that the interval for the cross-sectional results is completely contained in the interval for the fixed-effects model is important. This clearly quantifies the suggestion of Lewis (1986) that the cross-sectional estimates are more reliable. Further, this range implies that, for some measurement error levels, the fixed-effects coefficient may be greater than the cross-sectional coefficient. While this may seem an incongruous finding, it supports a wide body of literature that uses control functions or instrumental variables to account for unobserved heterogeneity (see Robinson 1989).

In addition to the bounds on the model slope coefficients, bounds on the error probabilities and the true probability of being in a union are implied by the analysis. In the cross-sectional analysis, the maximum value that p (the probability of misclassifying a nonunion worker as a union member) can attain is 0.189 while the maximum value that q (the probability of misclassifying a union worker as nonunion worker) can attain is 0.730. Recall that p and q cannot simultaneously achieve their maxima. In this case, the right bounds are associated with the maximum value of p and with $q = 0$. This also implies that the right bounds are associated with P_z (true proportion of union members) equal to 0.020. This is a lower bound for P_z . This implies that, to achieve the right bounds, only 2% of the population would actually be unionized but 18.9% of the nonunion workers would report being unionized. Clearly, this is unlikely to be the true structure.

The bounds for p and q in the fixed-effects analysis are much tighter: $p < 0.015$, while $q < 0.059$. Although the upper bound on p is slightly lower than the amounts found using the 1977 employer-employee match data, it is not strikingly lower and could easily be the result of changes in the CPS questionnaire or the awareness of union members. The results also place a much tighter bound on union membership: 19 to 22%.

Although the ranges presented are very large, it should be noted that these bounds do answer the question, "What values of β can be supported by the data if measurement error in union status is present?" This answer demonstrates that a wide range of values is feasible. It clearly establishes that measurement error may very well be the source of much of the disagreement among estimates. The difference between cross-sectional and fixed-effects models is small compared with the range of values expressed by these bounds, supporting Lewis' (1986) statement that measurement error may cause substantial bias. Further, it cautions us that rules of thumb and guesses concerning the measurement error could be far from the mark.

Clearly, additional information can be brought to bear to obtain sharper bounds. A logical first step is to ask what the upper bound for the cross-sectional estimates would be if the upper bounds on p and q from the fixed-effects model are imposed. Table 4 presents these new, much sharper bounds: The union wage differential (in cross section) is now between $[0.154, 0.168]$. What is particularly interesting here is to note that while the new bounds in cross section are much tighter, this information does not yield tighter bounds for the fixed effects estimates. This

Table 4. Bounds on Cross-Section Estimates Using p, q Information from Fixed Effects ($N = 14,347$), $p \leq 0.0153$, $q < 0.0591$

	New Right Bound (Upper)
Constant	0.719 (0.018)
Union	0.168 (0.007)
Education	0.097 (0.001)
Experience	0.004 (0.001)
Experience ²	-0.0006 (0.0001)
Black	-0.217 (0.010)

Cross-sectional estimates from 1998 observations of matched panel. Right bound represents maximal measurement error allowed under constrained error rates.

again underscores the fact that the fixed effect estimates are far more biased by measurement error than the cross sectional estimates.

An interesting hypothetical question is how low measurement error needs to be for the cross-sectional estimates to be the same or more than the fixed-effects estimates. This question is answered in three ways. First, what levels for *a priori* bounds on p and q would result in the upper bound for the fixed-effects model being equal to the upper bound for the cross-sectional model. This would place the cross-sectional estimates clearly in the upper range of the fixed-effects estimates. The first two columns in Table 5 present these bounds for the cross-sectional and fixed-effects estimates when p and q are both bounded below 0.00835, that is, response error would need to be lower than 0.835%. It should be noted that there is actually a continuum of bounds for p and q that would be available (higher p would require lower q and vice versa); requiring the bounds to be equal serves as an intuitive index. Note that this results in a very sharp bound for the cross-sectional estimates. The cross-sectional coefficient on the union variable would be between 0.154 and 0.160.

A second approach is to find *a priori* bounds on p and q so that the upper bound on the fixed-effects region was equal to the lower bound on the cross-sectional region. The third column in Table 5 presents the bounds for the fixed-effects estimates when p and q are bounded below 0.0819. The effect on the cross-sectional upper bounds (relative to the first column) is minimal and so is not presented. Finally, the fourth column presents the case where *a priori* bounds on p and q are found so that the fixed-effects union coefficient must lie below 0.10. This value was chosen as being near the top of the range of the typical fixed-effects estimates. Here, p and q must be below 0.00593.

In order for the cross-sectional estimates to be an overstatement of the true union wage differential, measurement error must be below 1%. While this is possible and should not be ruled out, measurement error in the range of 1% or larger would result in the cross-sectional estimates of the union wage differential to be understatements of the true union wage effect, understated both due to measurement error and due to unobserved heterogeneity. The conclusion that most or all of the difference between the cross-sectional and fixed-effects estimates is due to measurement error (as suggested by Lewis 1986; Freeman 1984; and Card 1996) is clearly

Table 5. Restrictions on p , q to Support Cross-Section Union Coefficient Larger Than Fixed Effects

	Upper Bounds Equal ($p, q < 0.0084$)		Intervals Not Overlapped ($p, q < 0.0082$)	Fixed Effect < 0.10 ($p, q < 0.0059$)
	(Right (1))	(Right (2))	Fixed Effect (Right (3))	Fixed Effect (Right (4))
Constant	0.723 (0.018)	—	—	—
Union	0.1601 (0.006)	0.1601 (0.173)	0.154 (0.1578)	0.10 (0.051)
Education	0.096 (0.001)	0.107 (0.012)	0.107 (0.012)	0.107 (0.012)
Experience	0.038 (0.001)	0.086 (0.004)	0.083 (0.004)	0.083 (0.004)
Experience ²	-0.0006 (0.0001)	-0.0006 (0.0001)	-0.0006 (0.0001)	-0.0006 (0.0001)
Black	-0.216 (0.010)	—	—	—
Sample size	14,347	26,694	26,694	26,694

Cross-sectional estimates from 1998 observations of matched panel. Fixed-effects estimates based on differences from means. New right bounds represent constraints on error rates, which support conclusion that fixed-effects coefficients are below cross-sectional coefficient.

supported by this analysis. Only modest amounts of measurement error are required for this to occur.

In fact, Robinson (1989) argues that unobserved heterogeneity is present but is biasing cross-sectional estimates downward. A full analysis of measurement error in control function and IV estimates is beyond the scope of this article, but results by Black, Berger, and Scott (2000) and Frazis and Loewenstein (1999) suggest that a suitable modification of Robinson's IV estimates have the potential to correct for measurement error. Robinson's (1989) tests include (as a part of the null hypothesis) that measurement error is not present; therefore, rejection (as he finds) may simply indicate measurement error.

For these reasons, it is plausible to conclude that the cross sectional estimates are as good or better than any other estimates. And further, that differences between those estimates and the fixed effects results are due, primarily, to measurement error. Peracchi and Welch (1995) also argue that the matching process in the CPS may lead to sample bias. For these reasons, the remaining analysis focuses upon cross sectional estimates.

5. Union Wage Differentials by Subpopulations

An important empirical regularity is that the benefit of unionization differs across industries, occupations, and educational level. Both Card (1996) and Hirsch and Schumacher (1998) examine this issue. Typically, the differences are interpreted to have economic content. For example, workers with low education benefit from unions more than those with higher education. The relationships implied by the cross-sectional regressions of wage on union status (and other control variables) by each of these subcategories are potentially biased by measurement error if measurement error is different across categories. For example, if two industries have

Table 6. Bounds on Union Coefficient by Industry ($N = 14,347$)

	Lower	Upper	(p, q) Bounds Preserving Order
Construction	0.330	1.922	—
Retail	0.226	10.345	0.031
Transportation, communication, and utilities	0.156	3.075	0.120
Wholesale	0.1264	16.77	0.018
Manufacturing	0.039	14.62	0.178
Service	0.15	252.3	0.039
Financial, insurance, and real estate	-0.141	-51.2	Any

Cross-sectional estimates from 1998 observations of matched panel by industry group. Third column represents restrictions on error rates supporting nonoverlapping intervals.

the same return to unions but one industry has more response error in the report of unionization, it may appear as though that group has a lower return to union membership. This section, similar to the analysis in the previous section, asks the question, "How low does measurement error have to be for the ordering implied by cross-sectional estimates to be reliable?" However, it allows the misclassification rates to vary systematically with the subpopulation (industry, occupation, or educational level). It seems plausible that the misclassification rates may differ across these groups (only the variation by education would violate the assumptions in the previous section). Further, it allows the return to unionization to vary also. It also demonstrates that the bounds are generally substantially tighter in these subgroups than in the cross section as a whole. This implies that much of the range of the bounds in the cross section is due to error in the model (the variance of u) rather than to measurement error.

Industry

The data are divided into seven broad industry categories based on the self-reported industry of the individual. The seven categories are construction (9.6% of total sample; see Table 1); retail (12.3%); transportation, communications, and utilities (TCU, 10.5%); wholesale (7.2%); manufacturing (37.6%); services (18.3%); and financial, insurance, and real estate (FIRE, 5.8%).

Table 6 presents the lower and upper bounds on the union coefficient from the cross-sectional regressions by industry. Education, experience, and race were again used as control variables. The coefficients on these variables are similar to those found in the whole cross section and so are not reported here. Recall that the lower bound, presented in the first column, is the coefficient from the regression of log wage on union and the other control variables and hence represents the value that would be obtained by researchers who do not control for measurement error. The second column presents the upper bound derived from the inverse of the reverse regression. The third column represents a set of *a priori* bounds on p, q such that the upper bound on the union coefficient would be equal to the lower bound in the preceding row of the table. For example, in order for the upper bound on the union coefficient for retail workers to be equal to 0.3297 (the lower bound from the construction workers), the misclassification rates would need to be no larger than 0.0308.

Considering the lower bounds (which are the results from the usual linear regression), the construction trades have the highest return to unionization while the typically white-collar FIRE

Table 7. Bounds on Union Coefficient by Occupation ($N = 14,347$)

	Lower	Upper	(p, q) Bounds Preserving Order
Service	0.325	3.58	—
Operators, fabricators, and laborers	0.290	1.54	0.0417
Precision production, craft, and repair	0.201	2.3425	0.102
Technical, sales, and administrative support	0.114	19.2	0.044
Managerial and professional	-0.002	-354	any

Cross-sectional estimates from 1998 observations of matched panel by industry group. Third column represents restrictions on error rates supporting nonoverlapping intervals.

industry actually has a negative (but insignificant) effect. The lowest positive returns are for services and manufacturing. These results are comparable with other studies (see Lewis 1986 for examples). It is striking to note how much variation exists in the upper bounds. The upper bound for construction workers is less than 2, while the upper bound for service workers is over 200. The upper bound varies with two underlying terms: the noise in the relationship and the measurement error. A large upper bound could be due to lots of noise or to lots of measurement error.

The third column presents the *a priori* bounds on p and q necessary for the usual ordering to be robust to measurement error. For example, for the upper bound on the TCU union coefficient to be 0.2256 (equal to the lower bound on retail), measurement error in union status in TCU would only have to be less than 12%. This seems quite plausible based on the estimates from the employer–employee match (Freeman 1984). It should be noted that the employer–employee match estimates have not been examined by these categories, and the small sample may prevent reliable estimation. The most restrictive case is for the comparison between TCU and wholesale. The error rates in wholesale would have to be below 1.8% in order to bound the wholesale union coefficient below the TCU coefficient. While plausible (meeting some of the estimates from Freeman 1984, for example), the conclusion that the union coefficient for wholesale is lower than for TCU is tenuous. The comparisons between retail and construction and services and manufacturing are similarly tenuous but still only require error rates below 3 and 4%, respectively. It seems clear, though, that construction and retail workers have a larger return to unionization than any other category. Next are TCU and wholesale, with manufacturing and service at the bottom of the positive returns. As noted above, FIRE has a negative estimate and so is the lowest.

Occupation

The data were divided into five occupational categories: services (4.9%); operators, fabricators, and laborers (25.9%); precision production, craft, and repair (24.5%); technical, sales, and administrative support (20.1%); and managerial and professions (24.7%). Table 7 presents an analysis similar to that done for industries. The first and second columns represent the bounds on the union coefficient from the cross-sectional regressions by occupation. Again, the coefficients on the other control variables (education, experience, and race) are similar to the population as a whole and are not reported. The third column presents the bounds for p and q that place the upper bound for that row equal to the lower bound for the previous row.

Table 8. Bounds on Union Coefficient by Education ($N = 14,347$)

	Lower	Upper	(p, q) Bounds Preserving Order
No high school	0.325	2.32	—
Some high school	0.284	2.13	0.037
High school grad	0.186	3.26	0.108
Some post high school	0.105	10.13	0.087
At least college	-0.10	-61.5	Any

Cross-sectional estimates from 1998 observations of matched panel by industry group. Third column represents restrictions on error rates supporting nonoverlapping intervals.

As is typically found (again, see Lewis 1986), the return to unionization for managers and professionals is negative (although not significant). Service workers have the largest return. This may seem at odds with the findings for industry in the previous section, but the definitions of the occupations listed as services and the industries listed as services provide explanation. Service occupations are often found in nonservice industries. For example, cleaning and building service occupations (codes 448–455) include janitors who may work in a variety of industries. So janitors working in retail industries may dominate the results for occupations. Operators, fabricators, and laborers are second, and production craft and repair workers are third. The lowest positive return is the return for technical, sales, and administrative support.

As with industries, the amount of error that would support the ordering found in cross section is not unreasonable. If error rates for operators, fabricators, and laborers were less than 4%, that would preserved their order below services. If error in production, craft, and repair occupations was below 10%, their return is bounded below the return for operators, fabricators, and laborers. Finally, error rates in technical, sales, and administrative support occupations need only be less than 4.3% to bound them below precision production, craft, and repair workers. It seems quite clear, at any rate, that service occupations and operators, fabricators, and laborers have a higher return than for production craft and repair or technical, sales, and administrative support occupations.

Education

One of the most important overall determinants of earning is education. As would be predicted, the impact of unionization appears highest for those workers with the least education. Table 7 presents the bounds for the union differential by five educational categories: highest grade attained less than 9 (no high school [HS], 4.4%), highest grade completed between 9 and 11 (some HS, 8.3%), high school graduate (HS, 42.6%), highest grade attained between 13 and 15 (some post-HS, 19.8%), highest grade attained at least 16 (at least college, 24.6%). Similar to the previous two tables, the first column reports the lower bound on the union wage differential, while the second column presents the upper bound. The third column presents the error rates that would support the ordering from the OLS results (which coincide with the lower bound). Again, the coefficients on education, experience, and race are not reported but are similar to the coefficients in the full cross section. These results are similar to those of Hirsch and Schumacher (1998).

As with the previous two sections, the upper bound tends to be much smaller than in the entire cross section, with only the some post-HS category larger. The union effect for college graduates is negative and significant. Similar to the previous section, the cross-sectional ordering

seems unlikely to be affected by measurement error. Except for the case of some high school to no high school, the comparisons only require that error rates be less than the 8–10% level. Even the comparison between some high school and no high school would only require error rates less than 3.7%. It seems quite likely that, even with relatively severe measurement error, we could conclude that individuals without a high school diploma have a higher union wage differential than those with a high school diploma. Further, we could conclude that individuals with more than a high school education have a lower rate than those with just high school.

6. Conclusions

Measurement error in the report of the union status has long been suspected to account for some or all of the difference between estimates of the union wage differential based on fixed-effects models and estimates from cross-sectional models. The bounds presented here demonstrate that measurement error would have to be very low for unobserved heterogeneity to be the main cause of the much lower estimates of the union wage differential found based on fixed-effects models. Indeed, it seems far more plausible that the cross-sectional estimates are much less biased by unobserved heterogeneity than by measurement error. This supports conclusions drawn by Lewis (1986) and others.

While the difference between cross-sectional and panel estimates is not supported when measurement error is taken into account, differences in the union wage differential between industrial, occupational, and educational groups seem very likely to be supported even in the presence of measurement error.

The fact that a wide range of values for the union wage gap is allowed for by measurement error suggests that further research addressing this problem is warranted. In particular, the demonstrated value of information concerning the rates of response error suggests that the research might focus on obtaining estimates for these results. The work of Freeman (1984) is based on data prior to the 1980s. Clearly, a method of updating these results would be fruitful. Validation data of the type used by Freeman (1984) and others serves a useful purpose and is highly valuable. This article has quantified the value of such additional information.

References

- Aigner, Dennis J. 1973. Regression with a binary independent variable subject to errors of observation. *Journal of Econometrics* 1:49–59.
- Aigner, Dennis J., Cheng Hsiao, Arie Kapteyn, and Tom Wansbeck. 1984. Latent variable models in econometrics. In *Handbook of econometrics*, volume II, edited by Z. Griliches and M. D. Intriligator. New York: Elsevier Science Publishers BV, pp. 1323–93.
- Black, Dan A., Mark C. Berger, and Frank A. Scott. 2000. Bounding parameter estimates with nonclassical error. *Journal of the American Statistical Association*. 95:739–48.
- Bollinger, Christopher R. 1996. Bounding mean regressions when a binary regressor is mismeasured. *Journal of Econometrics* 73:387–99.
- Bollinger, Christopher R. 1998. Measurement error in the current population survey: A nonparametric look. *Journal of Labor Economics* 16:576–94.
- Bollinger, Christopher R., and Martin H. David. 1997. Modeling discrete choice with response error: Food stamp participation. *Journal of the American Statistical Association* 92:827–35.
- Bound, John, Charles Brown, and Nancy Mathiowetz. 2001. Measurement error in survey data. *Handbook of Econometrics* In press.

- Bound, John, and Alan B. Krueger. 1991. The extent of measurement error in longitudinal earnings data: Do two wrongs make a right? *Journal of Labor Economics* 9:1–24.
- Budd, John W., and In-Gang Na. 2000. The union membership wage premium for employees covered by collective bargaining agreements. *Journal of Labor Economics* 18:783–807.
- Card, David. 1996. The effect of unions on the structure of wages: A longitudinal analysis. *Econometrica* 64:957–80.
- Chowdhury, Gopa, and Stephen Nickell. 1985. Hourly earnings in the United States: Another look at unionization, sickness and unemployment using PSID data. *Journal of Labor Economics* 3:38–69.
- Farber, Henry S. 1990. The decline of unionization in the United States: What can be learned from recent experience? *Journal of Labor Economics* 8:75–105.
- Frazis, Harley, and Mark A. Loewenstein. 1999. Instrumental-variable bounds for a mismeasured binary independent variable in a linear regression. Unpublished paper, Bureau of Labor Statistics.
- Freeman, Richard B. 1984. Longitudinal analysis of trade unions. *Journal of Labor Economics* 2:1–26.
- Frisch, R. 1934. *Statistical confluence analysis by means of complete regression systems*. Oslo: University Institute for Economics.
- Fuller, Wayne A. 1987. *Measurement error models*. New York: Wiley and Sons.
- Hirsch, Barry T., and Edward J. Schumacher. 1998. Unions, wages and skills. *Journal of Human Resources* 33:201–19.
- Jakubson, George. 1991. Distinguishing unobserved heterogeneity and measurement error in panel estimates of the union wage effect. ILR–Cornell Working Paper no. 206.
- Jones, Ethel B. 1982. Union/nonunion differentials: Membership or coverage? *Journal of Human Resources* 17:276–85.
- Klepper, Steven, and Edward Leamer. 1984. Consistent sets of estimates for regressions with errors in all variables. *Econometrica* 52:163–83.
- Klepper, Steven. 1988. Bounding the effects of measurement error in regressions involving dichotomous variables. *Journal of Econometrics* 37:343–59.
- Koopmans, T. 1937. *Linear regression analysis of economic time series*. Amsterdam: Netherlands Econometric Institute, Harrlem-de Erwen F. Bohn N.V.
- Leamer, Edward. 1985. Sensitivity analysis would help. *American Economic Review* 75:308–13.
- Lewis, H. Gregg. 1986. *Union relative wage effects: A survey*. Chicago: University of Chicago Press.
- Mellow, Wesley, and Hal Sider. 1983. Accuracy of response in labor market surveys: Evidence and implications. *Journal of Labor Economics* 1:331–44.
- Mincer, Jacob. 1983. Union effects: Wages, turnover, and job training. *Research in Labor Economics* Supplement 2:217–52.
- Peracchi, Franko, and Finis Welch. 1995. How representative are matched cross sections? Evidence from the Current Population Survey. *Journal of Econometrics* 68:153–79.
- Polivka, Anne E., and Jennifer M. Rothgeb. 1993. Overhauling the Current Population Survey: Redesigning the questionnaire. *Monthly Labor Review* 116:10–28.
- Reiersøl, Olav. 1950. Identifiability of a linear relation between variables which are subject to error. *Econometrica* 18:375–89.
- Robinson, Chris. 1989. The joint determination of union status and union wage effects: Some tests of alternative models. *Journal of Political Economy* 97:639–67.
- Wooldridge, Jeffrey M. 2000. *Introductory econometrics*. New York: South-Western College Publishing, Thomson Learning.