Implications of Mean-Reverting Measurement Error for Longitudinal Studies of Wages and Employment

Bonggeun Kim, University of Canterbury
Gary Solon, University of Michigan

December 2003

Abstract: This paper examines the implications of mean-reverting measurement error for two influential literatures based on longitudinal survey data: (1) the literature on real wage variation over the business cycle and (2) the literature on intertemporal substitution in labor supply. Accounting for mean-reverting measurement error suggests that real wages may be even more procyclical than indicated by recent longitudinal studies. We also find that the instrumental variables estimator commonly used in intertemporal substitution studies is inconsistent if changes in earnings and hours of work are measured with different degrees of mean-reversion, but the magnitude of the resulting inconsistency appears to be small.

The authors are grateful for comments from the referees, John Bound, Charlie Brown, Chul-In Lee, Steve Pischke, and seminar participants at the University of Michigan, Michigan State University, Purdue University, the University of Waikato, and the 2003 Australasian meetings of the Econometric Society.
Implications of Mean-Reverting Measurement Error for Longitudinal Studies of Employment and Wages

I. Introduction

The textbook errors-in-variables model assumes that measurement error is independent of true values. As Bound, Brown, and Mathiowetz (2001) emphasize, however, this assumption is often violated by real-world microdata. Instead, Siegel and Hodge (1968), Bound and Krueger (1991), Bound et al. (1994), Kane, Rouse, and Staiger (1999), Black, Berger, and Scott (2000), and others have documented negative correlations between measurement error and true values in survey data on earnings, hours of work, and years of schooling. Although the implications of such “mean-reverting” measurement error for estimating wage returns to schooling have been investigated, the implications for many other important literatures remain unexplored.

This paper examines the implications of mean-reverting measurement error for two influential literatures based on longitudinal survey data on earnings and hours of work. In Section II, we reconsider the evidence on cyclical variation in real wages in light of mean-reverting measurement error. In Section III, we reconsider the properties of the instrumental variables estimator commonly used in studies of intertemporal substitution in labor supply.

II. Real Wages over the Business Cycle

Until a few years ago, macroeconomists firmly believed that real wages in the United States (and elsewhere) are nearly noncyclical. This conventional wisdom – discussed at length in Solon, Barsky, and Parker (1994) – was based on a long history of
aggregate time series evidence. The stylized fact of weak wage cyclicality spawned
umerous macroeconomic theories designed to explain how large cyclical swings in
employment could be accompanied by relatively little wage variation. These theories
included efficiency wage models, implicit contract models in which employers provide
real wage insurance to workers, and insider-outsider models.

More recently, a series of studies based on longitudinal microdata has
demonstrated that real wages in the United States actually are quite procyclical.1 As
shown in Solon, Barsky, and Parker (1994), the true procyclicality of real wages is
obscured in aggregate time series data because of a composition bias: the aggregate
statistics are constructed in a way that gives more weight to low-skill workers during
expansions than during recessions. In contrast, most of the studies based on longitudinal
microdata have been able to avoid composition bias by controlling for worker “fixed
effects,” usually by differencing them out.

For example, the fixed-effects model specified by Solon, Barsky, and Parker is

\[ \ln w_{it} = \alpha_i + \gamma_1 X_{it} + \gamma_2 X_{it}^2 + \gamma_3 t + \gamma_4 t^2 + \beta U_t + \epsilon_{it} \]

where \( \ln w_{it} \) is the \( i^{th} \) worker’s log real wage rate in year \( t \), the fixed effect \( \alpha_i \)
represents the combined effect of time-invariant characteristics of worker \( i \), \( X_{it} \) is
worker \( i \)’s years of work experience as of year \( t \), \( U_t \) is a business-cycle indicator such
as the civilian unemployment rate, and \( \epsilon_{it} \) is an error term orthogonal to the explanatory
variables. A convenient way to control for the fixed effect is to first-difference equation
(1) to get

---

1 See Section IV of Solon, Barsky, and Parker (1992) for a review of this literature.
where $\delta_0 = \gamma_1 - \gamma_2 + \gamma_3 - \gamma_4$, $\delta_1 = 2\gamma_2$, and $\delta_2 = 2\gamma_4$. With the unemployment rate as the cycle indicator, $\beta > 0$ as real wages are countercyclical, noncyclical, or procyclical.

Using a sample of male heads of household from the Panel Study of Income Dynamics (PSID), Solon, Barsky, and Parker estimate $\hat{\beta} = -0.014$ (with estimated standard error 0.002), indicating that, when the unemployment rate rises by an additional percentage point, real wage growth tends to decline by about 1.4 percentage points. This estimate is more than twice as procyclical as the estimate based on aggregate time series data for the same period, and it also is quite typical of other researchers’ estimates based on longitudinal microdata from both the PSID and the National Longitudinal Surveys of labor market experience. In light of the robust finding that real wages appear quite procyclical once composition effects are accounted for, macroeconomic theories designed to explain the supposed weakness of wage cyclicality may be unnecessary, and theories that predict substantially procyclical wages become more credible.

In most longitudinal studies of wage cyclicality, the hourly wage has been measured as the ratio of annual earnings to annual hours of work. Notwithstanding considerable evidence of substantial reporting error for both the numerator and denominator, the literature has given little attention to errors-in-variables bias. The implicit assumption has been that the error-ridden measure of the dependent variable in equation (2) follows the textbook errors-in-variables model

\[
\Delta \ln w_{it} = \delta_0 + \delta_1 X_{it} + \delta_2 t + \beta \Delta U_t + \Delta \varepsilon_{it}
\]
where the measurement error $\Delta v$ has zero mean and is orthogonal to the true wage growth $\Delta \ln w$ and to each of its determinants in equation (2). If the textbook model were appropriate, the measurement error in the dependent variable would merely add a well-behaved measurement error component to the regression equation’s error term, which would not cause any inconsistency in least squares estimation of the parameters of the equation (2).

The textbook errors-in-variables model in equation (3) can be viewed as a restricted version of the more general model

$$\ln w^* - \ln w^* = \theta + \lambda(\ln w - \ln w^* - s) + v - v^* - s.$$  

Obviously, the textbook model is the special case that assumes $\theta = 0$ and $\lambda = 1$. The assumption that $\theta = 0$ is inconsequential for purposes of estimating the slope coefficients in equation (2), but, as will be shown below, the assumption that $\lambda = 1$ is very consequential. The evidence reported in Bound et al. (1994), however, suggests that $0 < \lambda < 1$, a pattern that Bound and Krueger (1991) have dubbed “mean-reverting” measurement error.

The analysis in Bound et al. is based on the PSID Validation Study, which combined accurate information on earnings and hours from a company’s records with information from a PSID-like survey of the company’s workers. Bound et al.’s comparison of the survey reports of earnings and hours with the company records provides a wealth of information on properties of the measurement error in worker-reported earnings and hours. Furthermore, because the validation study was conducted in two waves four years apart, it also provides information on measurement error in changes in earnings, hours, and their ratio.
The results in Bound et al. include estimates of $\lambda$ in regressions like equation (4) for the changes in log annual earnings, log annual hours, and log hourly wage.\(^2\) For earnings growth, Bound et al.’s initial least squares estimate of $\lambda$ is 0.779 (with estimated standard error 0.041). With twelve outliers excluded, they reestimate $\lambda$ at 0.782 (0.045) and, when they reestimate with the full sample by least absolute deviations instead of least squares, $\hat{\lambda}$ becomes 0.853 (no standard error estimate reported). The corresponding estimates for hours growth are 0.834 (0.065), 0.862 (0.070), and 0.830, and the estimates for wage growth are 0.671 (0.212), 0.743 (0.193), and 0.790. All of these estimates of $\lambda$ are substantially less than the textbook model’s value of 1 and, in most cases for which estimated standard errors are reported, the deviation from 1 is statistically significant. Pischke’s (1995) reanalysis of the validation study’s earnings data replicates Bound et al.’s main findings and also suggests that the mean-reverting measurement error stems mainly from underreporting of the magnitude of transitory fluctuations. In effect, some workers asked to report their earnings for the preceding year tend to shade their answers towards their usual earnings.\(^3\)

Before analyzing the implications of mean-reverting measurement error for estimation of equation (2), it is worth asking what models of measurement error in levels

\(^2\)What Bound et al. actually report is the estimated slope coefficient in the regression of the measurement error on the true value, but this estimate is readily translated into our $\lambda$ simply by adding 1. For example, their estimate of –0.221 for the slope coefficient in the regression of the earnings growth measurement error on the true earnings growth implies an estimated $\hat{\lambda}$ of $1 - 0.221 = 0.779$.

\(^3\)As summarized in Bound, Brown, and Mathiowetz (2001), several other studies also have documented mean-reverting measurement error in the levels of earnings and hours. Using a match of Current Population Survey responses to Social Security earnings records, Bound and Krueger (1991) provide evidence on mean-reversion in measured changes in log annual earnings, estimating $\hat{\lambda}$ at 0.646. Only the PSID Validation Study, however, contains information on measurement error in changes of both earnings and hours.
are consistent with the differenced model in equation (4). One obvious candidate is
simply
\begin{equation}
\ln w_{it}^* = \eta_i + \lambda \ln w_{it} + v_{it}
\end{equation}
where \( \eta_i \) is an individual-specific fixed effect for reporting error. Differencing equation (5) leads directly to equation (4) with \( \theta = 0 \). Another candidate, which reflects Pischke’s suggestion that the mean-reversion is concentrated on the transitory variation, is
\begin{equation}
\ln w_{it}^* = \eta_i + \alpha_i + \gamma_1 X_{it} + \gamma_2 X_{it}^2 + \gamma_3 t + \gamma_4 t^2 + \lambda (\beta U_{it} + \varepsilon_{it}) + v_{it}.
\end{equation}
As seen by comparing this equation to equation (2) for the true wage variable, this model assumes that the mean-reversion does not apply to the systematic part of wage evolution, but only to the transitory variation, including the part associated with cyclical fluctuations. Differencing equation (6) leads to
\begin{equation}
\ln w_{it}^* - \ln w_{i,t-3}^* = \theta + \lambda (\ln w_{it}^* - \ln w_{i,t-3}^*) + 2s(1-\lambda)\gamma_3 X_{it} + v_{it} - v_{i,t-3},
\end{equation}
where \( \theta = (1-\lambda)(s\gamma_1 - s^2\gamma_2 + s\gamma_3 - s^2\gamma_4 + 2s\gamma_4 t) \). If this model is correct and \( \lambda \neq 1 \), then Bound et al.’s estimation of \( \lambda \) in equation (4) is subject to an omitted-variables bias from the failure to control for \( X_{it} \). If \( 0 < \lambda < 1 \), the bias is upward, so that Bound et al.’s analysis is biased against finding mean-reversion.\(^4\) A variant of equation (6) that would allow for mean-reversion in the systematic part with a \( \lambda \) greater than the \( \lambda \) for the transitory part would still lead to a differenced equation similar to equation (7).

What do these models of mean-reverting measurement error imply for the estimation of cyclicality in real wages? Substituting equation (4) or (7) into equation (2)\(^4\)

\(^4\) We impound the term involving \( t \) in the intercept \( \theta \) because, with only one difference available in the PSID Validation Study data, \( t \) is a constant.
yields

\[ \Delta \ln w_{it} = \tilde{\delta}_0 + \tilde{\delta}_1 X_{it} + \tilde{\delta}_2 t + \lambda \beta \Delta U_t + (\lambda \Delta \epsilon_{it} + \Delta \nu_{it}) \]

where \( \tilde{\delta} = \delta \) if equation (7) pertains and \( \tilde{\delta} = \lambda \delta \) if equation (4) pertains. In either case, equation (8) is the regression equation for measured wage growth that the empirical researcher actually is able to estimate. And in either case, since both components of the composite error term \( \lambda \Delta \epsilon_{it} + \Delta \nu_{it} \) are orthogonal to the regressors, least squares estimation of equation (8) identifies the coefficients of the equation. Notice, however, that the coefficient of \( \Delta U_t \) is not the original wage cyclicality parameter \( \beta \), but rather is \( \beta \) rescaled by the measurement error parameter \( \lambda \). If the standard textbook assumption that \( \lambda = 1 \) were true, then least squares estimation of equation (8) would consistently estimate \( \beta \). But once we recognize the likelihood that \( \lambda < 1 \), then \( \beta \) is not separately identified without additional information.

With the caveat that the PSID Validation Study evidence comes from the employees of only one company, we will proceed on the assumption that mean-reversion in those employees’ survey reports is similar to the unobserved mean-reversion for the national PSID sample.\(^5\) If that assumption is approximately correct, the validation study’s estimates of \( \lambda \) suggest that the longitudinal estimates of the response of real wages to unemployment fluctuations have been substantially attenuated. For example, recall Solon, Barsky, and Parker’s (1994) estimate that, for male heads of household, the

---

\(^5\) This would follow from an assumption that the distribution of the measured variable conditional on the true value is the same in the validation study as in the full PSID. This reverses Chen, Hong, and Tamer’s (2002) “main assumption” that the distribution of the true variable conditional on the measured variable is the same between the “auxiliary data set” and the “primary data set.” Although their assumption is applicable in some contexts, it is clearly inappropriate for the PSID Validation Study. See Chen, Hong, and Tamer’s literature review for a summary of methodological research on econometric estimation in the presence of measurement error.
coefficient of $\Delta U_t$ in equation (8) is -0.014. If we convert that estimate of $\lambda \beta$ into an estimate of the true wage cyclicality parameter $\beta$ by dividing through with Bound et al.’s $\hat{\lambda} = 0.67$ for measured wage growth, we obtain $\hat{\beta} = 0.021$ (with standard error estimated at 0.007 by the delta method). Thus, correcting for mean-reverting measurement error may lead to as much as a 50% proportional increase in the estimated procyclicality of real wages.\(^6\)

Relative to the previous literature based on aggregate time series evidence, the new longitudinal literature delivered the message that real wages are much more procyclical than macroeconomists had thought. Accounting for the impact of mean-reverting measurement error, however, indicates that even the new longitudinal studies may well have underestimated the procyclicality of real wages to a considerable degree. This finding amplifies the main new learning from the longitudinal literature – that theories designed to explain the supposed fact of noncyclical wages are misguided, and that theories consistent with substantially procyclical wages now appear more credible.

### III. Intertemporal Substitution in Labor Supply

One of the theories potentially consistent with procyclical real wages is Lucas and Rapping’s (1969) model of intertemporal substitution in labor supply. According to Lucas and Rapping, cyclical fluctuations in the labor market are comprised mainly of labor demand shifts along a short-run labor supply curve that is positively sloped because of intertemporal substitution behavior. If the short-run labor supply curve is sufficiently

\(^6\) If Bound et al.’s estimation of $\lambda$ is biased upward by their failure to control for $X_{it}$ in equation (7), then even this correction might be too small.
elastic, the demand shifts could induce large employment fluctuations along with procyclical real wages.

The question of how elastic the short-run labor supply curve is has motivated numerous empirical studies. MaCurdy (1981) and many others have used PSID data to estimate the model

\[ \Delta \ln h_{it} = \gamma + \beta \Delta \ln w_{it} + \varepsilon_{it} \]

where as before \( \Delta \ln w_{it} \) is the \( i^{th} \) worker’s change in log real wage from year \( t - 1 \) to year \( t \), \( \Delta \ln h_{it} \) is his change in log annual hours of work, \( \beta \) is now the intertemporal elasticity of substitution in labor supply, and \( \varepsilon_{it} \) is an error term orthogonal to the explanatory variables.\(^7\)

The researchers in this literature have been highly cognizant of the substantial measurement error in survey reports of year-to-year growth in real wages and hours of work. They also have been sophisticated in their recognition that, because wage growth is measured as the difference between growth in annual earnings and growth in annual hours, the measurement error in hours growth as the dependent variable is negatively correlated with the measurement error in the key explanatory variable, wage growth. Accordingly these researchers have estimated the model in equation (9) by instrumental variables (IV), using age and other variables as instruments for measured wage growth. Most of the resulting estimates of the intertemporal substitution elasticity \( \beta \) have been small, often no larger than 0.2, and many critics of the Lucas-Rapping model have observed that such elasticity estimates are too small to generate the observed aggregate

\(^7\) If \( \gamma \) is replaced by \( \gamma_i \), this section’s analysis is the same with all variables measured as deviations from period-specific means.
fluctuations in employment as labor-supply responses to cyclical variation in real wages. More recently, Lee (2001) has demonstrated that most of the IV estimates have been subject to a substantial downward finite-sample bias. Interestingly, his corrected estimates, which run around 0.5, are quite similar to the estimates in Oettinger’s (1999) clever study of intertemporal substitution in the labor supply of baseball stadium vendors.

Despite the careful attention to measurement error in the intertemporal substitution literature, most researchers have not taken account of the mean-reverting nature of the measurement error in earnings, hours, and hence hourly wages. Our purpose, therefore, is to check what mean-reverting measurement error implies for the properties of the IV estimators typically used in this literature.

We model the measurement error processes for earnings and hours growth as

\[ \Delta \ln {E^*_{it}} = \theta_E + \lambda_E \Delta \ln E_{it} + \Delta v_{Eit} \]

and

\[ \Delta \ln {h^*_{it}} = \theta_h + \lambda_h \Delta \ln h_{it} + \Delta v_{hit} \]

where \( \Delta \ln E_{it} \) and \( \Delta \ln h_{it} \) are the true values of earnings and hours growth, \( \Delta \ln E^*_{it} \) and \( \Delta \ln h^*_{it} \) are the error-ridden measures, \( 0 < \lambda_E \leq 1 \) and \( 0 < \lambda_h \leq 1 \), and \( \Delta v_{Eit} \) and \( \Delta v_{hit} \) are zero-mean random errors uncorrelated with \( \Delta \ln E_{it} \), \( \Delta \ln h_{it} \), or \( \varepsilon_{it} \). Since wage growth is measured as \( \Delta \ln w^*_{it} = \Delta \ln E^*_{it} - \Delta \ln h^*_{it} \), it follows that

---

8 One exception is French (2000), who considers the special case in which lagged wage growth is used as the instrument for current wage growth in IV estimation of equation (9). Consistent with our analysis, he finds that correcting for mean-reversion in measurement error makes only a small difference for his estimation of \( \beta \). Another exception is Mulligan (1999), who inflates his estimates of \( \beta \) on the assumption that, although hours of work are subject to mean-reverting measurement error, there is no offsetting mean-reversion in the measurement of earnings. Our analysis would lead to corrections similar to his if we also assumed no mean-reversion in earnings measurement. That assumption, however, is strongly rejected by the evidence from both the PSID Validation Study and Bound and Krueger’s (1991) analysis of matched data between the Current Population Survey and Social Security earnings records.
(12) \[ \Delta \ln w_{it}^* = (\theta_E - \theta_h) + \lambda_E \Delta \ln w_{it} + (\lambda_E - \lambda_h) \Delta \ln h_{it} + \Delta \nu_{pit} - \Delta \nu_{hit}. \]

Now suppose a valid instrumental variable \( z_{it} \) (which is correlated with true wage growth \( \Delta \ln w_{it} \), but not with \( \epsilon_{it}, \Delta \nu_{pit} \), or \( \Delta \nu_{hit} \)) is used as an instrument for measured wage growth \( \Delta \ln w_{it}^* \) in the regression of measured hours growth on measured wage growth. Using equations (9), (11), and (12) to derive the probability limit of the IV estimator of \( \beta \) shows that

\[
\text{plim } \hat{\beta} = \frac{\text{Cov}(z_{it}, \Delta \ln h_{it}^*)}{\text{Cov}(z_{it}, \Delta \ln w_{it}^*)} = \frac{\lambda_h \beta}{\lambda_h + (\lambda_E - \lambda_h)(1 + \beta)}. 
\]

Unless, contrary to the relevant economic theory, \( \beta \) equals either 0 or \(-1\), the IV estimator of \( \beta \) is consistent only if \( \lambda_E = \lambda_h \), that is, only if the measurement errors in earnings growth and hours growth happen to have just the same degree of mean-reversion. Of course, a special case of this is the classical case \( \lambda_E = \lambda_h = 1 \). More generally, if \( \beta > 0 \) as predicted by theory,

\[
\text{plim } \hat{\beta} \leq \beta \text{ as } \lambda_E \leq \lambda_h. 
\]

Again we can gauge the empirical importance of the estimator’s inconsistency by referring to the PSID Validation Study findings reported by Bound et al. (1994). Recall that their least squares estimates of \( \lambda_E \) and \( \lambda_h \) are respectively 0.779 and 0.834. As these estimates of \( \lambda_E \) and \( \lambda_h \) are not greatly different from each other, one might expect the resulting inconsistency in \( \hat{\beta} \) to be small. Indeed, equation (13) implies that if these
were the true values of $\lambda_E$ and $\lambda_H$ and if the true $\beta$ were 0.5, the probability limit of $\hat{\beta}$ would be 0.55. Bound et al.’s alternative estimates of the $\lambda$’s based on least squares with outliers excluded and on least absolute deviations estimation for the full sample imply respective $\hat{\beta}$ probability limits of 0.58 and 0.48.

Thus, in this case, unlike the case of estimating the cyclicality of real wages, accounting for mean-reverting measurement error does not lead to a large change in the implied value of the key parameter. While there may be other reasons to worry about the accuracy of IV estimates of the intertemporal substitution elasticity – such as the finite-sample bias highlighted by Lee (2001) or the question of whether the instruments are really uncorrelated with the labor supply equation’s error term – mean-reversion in the measurement errors for earnings and hours does not appear to be a major source of inconsistency.

Beyond our specific results on real wage cyclicality and intertemporal substitution in labor supply, our analyses also suggest two more general lessons. First, information on the nature of measurement error – such as that provided by the PSID Validation Study – is extremely valuable. It is unfortunate that, for some key variables such as longitudinal change in hours of work, the PSID Validation Study is presently the only source of such information. Second, our analyses of wage cyclicality and labor supply illustrate the tractability and usefulness of accounting for realistic departures from the textbook errors-in-variables model. We agree with Bound, Brown, and Mathiowetz (2001) that “the possibility of non-classical measurement error should be taken much more seriously by those who analyze survey data, both in assessing the likely biases in
analyses that take no account of measurement error and in devising procedures that
‘correct’ for such error.”
References


Lee, Chul-In, “Finite Sample Bias in IV Estimation of Intertemporal Labor Supply


