MASSACHUSETTS INSTITUTE of TECHNOLOGY
LIBRARIES

M.I.T. LIBRARIES - DEWEY

Digitized by the Internet Archive
in 2011 with funding from
Boston Library Consortium Member Libraries

http://www.archive.org/details/measurementerror00pisc

HB31
.M415
no.
94-1

Dewey

working paper department of economics

MEASUREMENT ERROR AND EARNINGS DYNAMICS:
Some Estimates from the PSID Validation Study

Jorn-Steffen Pischke

Dec. 93 94-1

massachusetts institute of technology

50 memorial drive
cambridge, mass. 02139

MEASUREMENT ERROR AND EARNINGS DYNAMICS:
Some Estimates from the PSID Validation Study

Jorn-Steffen Pischke

Dec. 93 94-1

APR 5 1994

Measurement Error and Earnings Dynamics: Some Estimates From the PSID Validation Study

Jörn-Steffen Pischke
Massachusetts Institute of Technology
Department of Economics
Cambridge, MA 02139-4307

November 1992
Revised: December 1993

Measurement Error and Earnings Dynamics:
Some Estimates From the PSID Validation Study

Jörn-Steffen Pischke

ABSTRACT

Previous empirical work on measurement error in survey earnings data has shown that the variance of the measurement error is roughly constant over time, it is negatively correlated with true earnings, and it is autocorrelated with previous measurement errors. This paper proposes a simple model where the measurement error stems from underreporting of transitory earnings fluctuations and a white noise component. The model fits well to data from the PSID Validation Study. While it implies that there will be some bias in estimates of variances and autocovariances these biases are of the same magnitude so that autocorrelations will be estimated roughly correctly.

Keywords: Panel data, reporting error, income, covariance structure model, minimum distance estimator.

1. INTRODUCTION

It is widely recognized that measurement error is a pervasive phenomenon in microeconomic survey data and that errors-in-variables can lead to inconsistent estimates in many econometric models. In a linear regression context, attenuation bias can be cured if suitable instruments are available. Empirical studies trying to identify, for example, earnings or consumption processes typically have to ignore measurement error altogether or assume that it follows a particular parametric structure (e.g. uncorrelated white noise). Absent any direct knowledge about the properties of the measurement error process such assumptions are invariably necessary for identification. This paper tries to establish some simple relationships between measurement error and the dynamics of true earnings.

Systematic research on response error plays an important role in other social sciences (see Sudman and Bradburn 1974) but little work has been undertaken to assess the magnitude and properties of reporting error in earnings data. The few studies that have tried to obtain validation variables to assess the extent of measurement error directly make the common identifying assumptions highly questionable. The earlier studies in this area have usually obtained two independent observations on the same variable. Both these records were potentially error ridden. The study by Mellow and Sider (1983) is an example of this approach. They used matched employee-employer responses to Current Population Survey (CPS) questions and compared the results of separate wage regressions for the two variables. Such studies can be instructive but yield little insight on the direct nature of the measurement error.

The first study trying to obtain accurate validation records for typical labor market variables was undertaken by researchers at the Institute for Social Research at the University of Michigan. They obtained payroll records at a cooperating firm and simultaneously administered the Panel Study of Income Dynamics (PSID) questionnaire to the covered employees (the study is therefore known as the PSID Validation Study). Due to careful checking of the firm records the research team claims to have obtained very reliable validation information. Their results are reported in Duncan and Hill (1985) and Duncan and Mathiowetz (1985). This study has an optimistic message:

more than 80 % of the variability of earnings is due to actual signal. The unfortunate part of their results was that the measurement error in earnings seems correlated with many typical regressors in earnings equations, thus violating the assumptions of the classical errors-in-variables-model.

Census data sets, like the CPS have been matched to administrative records at various occasions for the purpose to assess survey data quality. David, Little, Samuhel, and Triest (1986) used a match between the CPS and Internal Revenue Service records to assess the error introduced by the CPS hot deck procedure to impute missing earnings values. Bound and Krueger (1991) analyzed a similar match between the CPS and Social Security Administration records to study the properties of measurement error. The CPS-SSA match has various advantages compared to the PSIDVS. It is more representative of the whole population and contains many more observations. Moreover, it is a two period panel referring to 1976 and 1977. The disadvantage is that Social Security earnings may be erroneous as well and these earnings are top coded at the Social Security maximum. Bound and Krueger pointed out two new findings with the CPS-SSA match. Measurement error, constructed by subtracting social security earnings from earnings reported in the CPS, is negatively correlated with true earnings, a phenomenon they refer to as mean reverting measurement error. Secondly, measurement error is significantly positively autocorrelated in two adjacent periods. Both features defy the classical measurement error model and seriously complicate the estimation of earnings dynamics.

These findings were corroborated on a second wave of the PSID Validation Study by Bound, Brown, Duncan and Rogers (1989, in press). The second wave was collected four years later through a similar survey at the same firm. Bound et al. also find mean reversion and a positive autocorrelation in measurement error extending over a four year period. In addition, they develop a nice theoretical framework to summarize the impact of measurement error in a regression context and conclude that the impact of measurement error in the level of earnings is likely to be small.

None of these studies shed direct light on the question of how to correct estimates of earnings dynamics as they are often used in analyses of the intertemporal allocation of resources by households or individuals over time. Generally, very restrictive assumptions at least on the dynamic structure of the measurement error have to be made in studies of life cycle labor supply (e.g. Abowd and Card 1989) or consumption (e.g. Hall and Mishkin 1982). To improve on this we need to know how measurement error relates to earnings over time. In this paper I use the PSID Validation Study to address this issue. The fact that survey earnings are only available for two years imposes some limits. Fortunately, the dataset contains validation earnings for up to six years thus providing a lot of information on the true earnings process.

I begin in section 2 with a description of the data set and some summary statistics about the sample. I also revisit the basic findings of Bound and Krueger (1991) and Bound et al. (in press) and show that they largely hold up in my sample. Section 3 presents a dynamic model for both measurement error and earnings. The estimation results provides some evidence that individuals underreport transitory earnings fluctuations. Quantitatively, this is only a small fraction of the total measurement error, the bulk is explained by a classical white noise component. The following section discusses the impact of these results for the study of earnings dynamics, in particular the estimation of permanent and transitory components of earnings. Section 5 concludes.

2. DATA AND BASIC FACTS ABOUT MEASUREMENT ERROR

For the PSID Validation Study a survey questionnaire analogous to the one used for the PSID but shorter was administered to a number of employees at an unnamed Detroit area manufacturing firm. At the same time, payroll records for these employees were obtained from the firm. The producers stress the cooperation of the firm involved which allowed an error free transcription of the company records.

The PSIDVS consists of two waves. In the first wave, 387 employees were interviewed in 1983. This survey included questions about annual earnings in 1981 and 1982. A second survey was conducted in 1987 which included questions about earnings for every year from 1981 to 1986. Validation earnings for all these years are provided as well. For the 1987 wave an attempt was made to reinterview as many of the original participants as possible. 275 of the original 1983 participants became eventually part of the panel. Additional hourly workers were interviewed in the 1987 survey so that a total of 492 interviews could be coded for that year. Thus, both survey years consist of a cross section larger than the overlap that creates the panel. Since I will concentrate on measurement error for earnings with a one year recall I will refer to them as the 1986 and 1982 cross sections and the panel.

Before proceeding I eliminated observations that contained assigned values or zeros in the reported or validation earnings in the relevant years. I also eliminated four more observations because the coded validation earnings seem questionable and they deviate significantly from the rest of the sample. Bound et al. (in press) also work with a sample where they remove some outliers and question the validity of two of the 1986 validation values. Details on the omitted observations are given in the Appendix. In line with the literature, I take logs of the income variables thus asserting that measurement error is proportional. All of the qualitative results hold up for earnings in levels and additive measurement error.

Tables 1 and 2 provide some basic sample statistics for the panel dataset and the 1986 cross section, respectively. Most striking about the sample is the high tenure, 19 years on average. This is of course due to the sampling design, only employees continuously with the firm between 1981 and 1987 became part of the sample. The sample is therefore also older than the labor force in general. Sample members are relatively well educated, the proportion of women is very low. About half of the panel are salaried employees. Earnings in this sample are higher than in representative cross sections. Remarkably, validation and reported earnings do not differ much in either mean or variance. This implies that "measurement error," calculated as the difference between reported and validation earnings, has a mean close zero. The standard deviation of the measurement error is quite large, on the other hand. Its skewness ranges from -1.3 to 1.4 and the kurtosis from 8 to 11 indicating only small asymmetries but fat tails in the distribution of the measurement error. Plots of the distribution can be found in Bound and Krueger (1991) and Bound et al. (in press).

I will use this definition of measurement error in the analysis that follows. This implies that validation earnings are measured correctly, an assumption also made by Duncan and Hill (1985) and Bound et al. (in press). Identification of the moments of the survey measurement error necessitates some within period restrictions on the joint distribution of the measurement error in the survey and validation variables. Given that some gross outliers in the validation variables have been eliminated the assumption of correct validation earnings seems as good as any other.

Simple cross section regressions of the log measurement error so defined on demographics did not reveal any particular pattern and the explanatory power of these regressions is low. These results are therefore not reported here.

Bound and Krueger (1991) established two important findings using the CPS-SSA match. They found that measurement error is positively correlated over time. The autocorrelation between the two observations one year apart is about 0.40. Secondly, there is a negative correlation between measurement error and true earnings in their data set. This latter finding has been confirmed by Bound et al. (1989) on the PSIDVS. Bound et al. also found a positive but smaller autocorrelation for the error four years apart. I will show that these findings largely hold up on the samples I drew from the PSIDVS data. In revisiting the earlier results I will pay particular attention to the distinction between a recession year (1982) and a boom year (1986).

Table 3 reports the basic results. The first columns report the variances of measurement error (m), true earnings (y), and reported earnings (x). The variance of measurement error is between 15 and 25 % of the variance of earnings. The variance of the measurement error seems much more stable between 1982 and 1986 than the variance of earnings. This implies that the reliability ratio σ_y^2/σ_x^2, which plays a crucial role for the classical measurement error model in a regression context, will be higher in recessions. However, this is entirely related to the variance of the signal, a finding also stressed by Bound et al. (1989).

The last two columns report the regression coefficients of measurement error on true earnings (b_{my}) and on reported earnings (b_{mx}). Bound et al. (in press) show that these regression coefficients take the role of the reliability ratio for calculating the bias introduced by measurement error in more general models; the reader is referred to their paper for the algebraic details. My results are roughly similar to theirs. The negative correlation between measurement error and true earnings is stronger for the recession year 1982 than for the boom year 1986. For the panel this coefficient is actually estimated to be positive in 1986. However, a negative and significant coefficient is found for the 1986 cross section. This difference stems partly from the fact that the cross section contains more hourly workers and the correlation is more strongly negative for hourly workers. For the 1986 cross section the coefficients are -0.170 (0.056) for hourly workers and 0.019 (0.035) for salaried employees. For the panel the 1986 coefficients are -0.064 (0.067) for hourly workers and 0.020 (0.043) for the salaried group. The coefficient for the hourly workers differs to quite a substantial degree between the cross section and the panel sample. For 1982 the panel coefficients are -0.105 (0.056) for hourly workers and -0.029 (0.052) for salaried employees. Thus the general pattern holds up within the subgroups: the correlation is more negative in 1982 for both of them. Hourly workers have always a lower correlation.

The correlation between the measurement error and survey earnings is positive but shows the same pattern: it is higher in the boom. This has to follow by definition since $cov(m,x) = cov(m,y) + var(m)$. Since the variance of m is rather stable the two covariance terms have to follow the same relationship over time.

Table 3 reports results for the basic samples as well as for subsamples that restrict the log of true earnings to be above 9 (about $8,100) in each year. For the 1986 cross section all observations fulfilled that restriction. There are a few observations with very low earnings in the earlier years of the sample that may strongly influence the estimates involving y. As can be seen from the table, this mainly reduces the estimate of the variance of y and x in 1982 but changes none of the other estimates very much. All the results hold up in the restricted samples as well.

The correlation between the 1986 and 1982 measurement error is 0.094 (0.065) in the panel data set. This is slightly lower than what Bound et al. (in press) found. They pointed out already that this correlation is too high to be compatible with an AR(1) structure of the measurement error. Taking the Bound and Krueger (1991) estimate of 0.40 for the first order autocorrelation as a benchmark, the fourth order autocorrelation would be 0.026. One alternative possibility is that there exists a person specific fixed effect in misreporting combined with an AR(1).

Let me summarize the three main findings about measurement error at this stage: 1) The variance of the measurement error is roughly constant over time, it may be slightly higher in recessions. 2) Measurement error is weakly negatively correlated with true earnings, the correlation is stronger in recessions and is stronger for hourly workers. 3) Measurement error is positively autocorrelated over time, the correlation is declining slowly over time.

Since the PSIDVS only contains two observations on measurement error I cannot distinguish between different hypothesis on the correlation of the measurement error over time. However, the dataset contains relatively rich information on true earnings. I will therefore focus on the implications 1) and 2) reported above. In the next section, I use a simple model to capture the relationship between the dynamic structure of true earnings and measurement error.

3. A DYNAMIC MODEL FOR EARNINGS AND MEASUREMENT ERROR

In many applications it is useful to decompose earnings into a permanent and a transitory component. Different events are associated with these components. Promotions or job loss in a high paying industry lead to permanent changes in earnings while overtime or temporary layoffs induce temporary variations. Obviously, it is the changes in permanent earnings that are related to the more important events in people's lives. According to Sudman and Bradburn (1974) events of greater importance ("salience") are recalled better by survey respondents. This behavioral model suggests that respondents underreport transitory earnings changes. Such a model is also consistent with the findings reported above since transitory income fluctuations tend to be larger in recessions.

A simple statistical formulation of such a model is as follows. Let annual earnings be composed of a permanent (random walk) component and a transitory (white noise) component, i.e.

$$\begin{aligned} y_{it} &= z_{it} + \eta_{it} \\ z_{it} &= z_{it-1} + \varepsilon_{it} \end{aligned} \tag{1}$$

In the estimation it is not possible to identify time varying variances for both shocks for each period. Some restriction is necessary; I chose to let the transitory shock have time varying variances while the variance of the random walk component is constant over time.

$$var(\varepsilon_{it}) = \sigma_\varepsilon^2, \ \ var(\eta_{it}) = \sigma_{\eta t}^2, \ cov(\varepsilon_{it}, \eta_{is}) = 0 \tag{2}$$

The process for measurement error is given by

$$\begin{aligned} m_{it} &= \alpha\eta_{it} + \mu_i + \xi_{it} \\ var(\xi_{it}) &= \sigma_{\xi t}^2 \\ var(\mu_i) &= \sigma_\mu^2 \end{aligned} \tag{3}$$

Measurement error consists of three components. The first is correlated with the transitory income shock, we expect $\alpha < 0$. The second component is a person fixed effect in the measurement error unrelated to earnings. Recall that this will yield results that are consistent with the autocorrelation in measurement error found by Bound and Krueger (1991). Given the limitations of the dataset at hand no richer structure of the dynamics of the measurement error can be identified; the fixed effect is just an arbitrarily chosen alternative to model the autocorrelation. The last component of the measurement error is white noise. In this simple model measurement error is not related to permanent earnings at all.

For the estimation it is useful to difference the earnings equation (1) to eliminate heterogeneity introduced by varying initial conditions (i.e. person effects in the earnings process).

$$\Delta y_{it} = \varepsilon_{it} + \eta_{it} - \eta_{it-1} \tag{4}$$

The resulting model for earnings and measurement error defined in equations (2), (3), and (4) implies the following moment conditions

$$\begin{aligned}
var(\Delta y_{it}) &= \sigma_{\varepsilon}^2 + \sigma_{\eta t}^2 + \sigma_{\eta t-1}^2 \\
cov(\Delta y_{it}, \Delta y_{it-1}) &= -\sigma_{\eta t-1}^2 \\
var(m_{it}) &= \alpha^2\sigma_{\eta t}^2 + \sigma_{\xi t}^2 + \sigma_{\mu}^2 \\
cov(\Delta y_{it}, m_{it}) &= \alpha\sigma_{\eta t}^2 \\
cov(\Delta y_{it}, m_{it-1}) &= -\alpha\sigma_{\eta t-1}^2 \\
cov(m_{it}, m_{it-j}) &= \sigma_{\mu}^2
\end{aligned} \tag{5}$$

All other sample covariances are restricted to be zero. This is admittedly a very simple minded model. However, it helps to put some structure on the data and it fits amazingly well. But recall that I restrict the sample so that the log of true earnings is above 9. Including observations with smaller earnings often led to results that were strongly influenced by a single or a few observations, especially in the estimates of the earnings process.

I fit the moment conditions in (5) to the data by first estimating the unrestricted covariance matrix of earnings changes (for 1982 to 1986) and measurement errors (for 1982 and 1986) and then imposing the restrictions using a minimum distance estimator (see Chamberlain 1984 and Abowd and Card 1989). This strategy has a number of advantages compared to, say, Maximum Likelihood estimation. No distributional assumptions are necessary. The estimation is therefore robust to the departures from normality typically observed in actual data. Simple specification tests are available, allowing a general assessment of the fit of the model (Chamberlain 1984 and Newey 1985). Given that the first stage unrestricted covariance matrix is available, failures of the model can be traced to their sources relatively easily. If optimal weights are chosen the second stage estimates are asymptotically efficient. Finally, in large samples the two stage procedure is computationally simpler than MLE because the nonlinear maximization has to be performed only on a small set of moments rather than the entire dataset.

The minimum distance estimator for a sample of size T is given by

$$\hat{b}_T = argmin(c_T - f(b))'W_T(c_T - f(b)) \tag{6}$$

where c_T is a vector of r sample covariances, $f(b)$ is a function of the parameters of interest, and W_T is a positive semi-definite weighting matrix that possibly depends on the data. Define $F = \partial f(b)/\partial b$, the jacobian of the moment conditions. If the weighting matrix is the inverse of the matrix of sampling covariances of the estimates of the sample moments then the estimator has certain optimality characteristics and T times the quadratic form being minimized has a χ^2 -distribution with degrees of freedom equal to r - $rank(F)$. In this case the optimal weighting matrix is the inverse of the fourth moment matrix of earnings changes and measurement errors.

In empirical applications the optimally weighted minimum distance estimator tends to underfit the variances on the main diagonal of a covariance model like this. This results from the correlation between the second moments and their sampling variances, which are calculated from the same data The problem tends to be worse the fatter the tails of the empirical distribution (see Altonji and Segal 1993). In the sample the kurtosis of log wage changes ranges from 10 to 27. I experimented with the independently weighted estimator suggested by Altonji and Segal where the sample is split and the moments and their sampling variances are calculated from different portions of the data. The results were extremely imprecise and highly dependent on the particular sample split used. Instead, I use a weighting matrix with the optimal weights on the main diagonal and zeros elsewhere. This corresponds to weighted nonlinear least squares on the sample moments. Each moment condition receives a weight that is inversely proportional to its accuracy but no dependence between moments is allowed for. This estimator seems to yield very reasonable estimates.

The asymptotic distribution of the estimator in (6) is

$$\sqrt{T}(\hat{b}_T - b) \xrightarrow{D} N(0,(F'WF)^{-1}F'WV^{-1}WF'(F'WF)) \tag{7}$$

$$V = E(cc')$$

Newey (1985) has shown that an asymptotic χ^2 -statistic can be formed for this model as

$$T(c_T - f(\hat{b}_T))'Q_T^-(c_T - f(\hat{b}_T)) \sim \chi^2_{r-rank(F)}$$

$$Q_T = P_T V_T P_T'$$

$$P_T = I - F_T(F_T'W_TF_T)^{-1}F_T'W_T \tag{8}$$

where Q_T^- is a generalized inverse of Q_T .

The estimation results are given in Table 4. The specification test reveals that in the small sample at hand there is no evidence that the model does not fit the data. This may not mean that the model is correct, it may just reflect on the low power of the test with a sample of only 231 observations purged of gross outliers. The estimate of α is about -0.25, meaning that individuals underreport transitory earnings by 25 %. Unlike most of the variances in the model, this coefficient is not very well determined. The standard deviation of the white noise component in measurement error is quite large, around 0.10 in both years. There is no evidence that this variance differs across periods. This means that the underreporting of transitory earnings accounts only for a small part of the total variance in measurement error, 9 % in 1982 and 5 % in 1986. More than 80 % of the variance is due to the white noise component.

The estimates for the earnings process also seem quite sensible. They imply that between 75 and 90 % of the variation in earnings changes is due to the transitory component. Transitory earnings shocks are much more important in the recession years 1982-83 than in the boom years 1985-86, the standard deviation is about twice as large in the recession. This explains part of the stronger negative correlation between earnings and measurement error in a recession year like 1982 compared to a boom year like 1986. This is the second major fact about measurement error from the previous section.

Table 5 reports the actual covariance matrix used in the estimation and table 6 displays the fitted covariance matrix from the model. These tables provide useful information on the short-comings of the model. The model explains most of the variances and autocovariances of the earnings process well; it overpredicts the autocovariances between 1983/84 and 1984/85. This stems from the fact that a large value for the variance of the transitory earnings component is necessary to fit the earnings variances in 1983 and 1984. This is the only parameter entering the autocovariance. This deviation may be due to the fact that the variance of the permanent component is nonstationary as well and high during the early years of a boom because not everybody benefits from an upswing at the same time.

The model predicts three non-zero covariances between earnings and measurement error ($cov(\Delta y_{i82}, m_{i82})$, $cov(\Delta y_{i83}, m_{i82})$, $cov(\Delta y_{i86}, m_{i86})$). It fits all three of them very well. However, there is one other non-zero covariance that the model misses, between earnings changes in 1982 and measurement error in 1986. This correlation of -0.21 is rather peculiar since no similar correlation exists with earnings in any of the later years up to 1986. While this single correlation is significant, a χ^2-test for the joint hypothesis that all seven covariances $cov(\Delta y_{it-j}, m_{it})$ which should be zero are indeed zero has a p-value of 0.43. Thus, the correlation between 1982 earnings changes and the 1986 measurement error may just be due to sampling variation.

The model has a number of other implications. Many sample members had negative transitory earnings shocks in 1982 due to the recession. Thus, the measurement error should be positive on average for that year. According to table 1 this is indeed the case but the mean of the measurement error is very small (and not statistically different from zero).

An alternative test for the model can be designed by considering equation (3) as the basis for a linear regression where η_{it} is replaced by current earnings. Obviously, if the model was true this would result in a downward inconsistent estimate of α since earnings contains both a temporary and a permanent component. However, a consistent estimate could be obtained by instrumenting earnings with a variable that is just related with transitory earnings changes but not with permanent movements in earnings.

For the sample at hand, relatively stable manufacturing workers, changes in annual hours will satisfy this criterion relatively well. Between 1981 and 1982 much of these changes are due to changes in temporary layoffs and unemployment while between 1985 and 1986 most will be due to variation in overtime. However, hours are only available for hourly workers, cutting the sample size significantly. Table 7 displays the results of regressing the measurement error on earnings using hours as an instrument for the panel sample as well as the larger cross sectional samples. The coefficient estimates are similar to the ones obtained from the structural model above except for the estimate for the 1986 cross section. This yields a point estimate of -0.54, indicating that for this group one half of transitory earnings movements may go unreported. It should be kept in mind that some of the difference in these estimates will be due to sampling variance.

We would like to assess again how much of the variance of the measurement error is due to the underreporting of transitory income, i.e. we like an estimate for the quantity $\alpha^2\sigma_\eta^2/\sigma_m^2$. Obviously, the standard R^2 defined as explained sum of squares divided by total sum of squares will not estimate this quantity since the variance of the regressor, earnings, exceeds σ_η^2. A consistent estimator for the explained variance is available under the null that the model is true. It is given by

$$\hat{R}^2 = \hat{\alpha}_{IV} \frac{cov(m_{it}, y_{it})}{var(m_{it})} \tag{9}$$

This quantity is reported at the bottom of table 7. For the panel sample only 7 % of the variation in measurement error is due to underreporting. This is again similar to the finding above using the structural estimates. In the cross section samples the explained variance is two to three times as large but the major contribution comes still from other factors like the white noise measurement error component. In a more representative sample the model might actually perform better. This will be true if the variance of the white noise component of the measurement error is the same in the population as in the sample while the variance of the transitory component in earnings is higher. Since many movements in earnings are due to mobility between firms and separations this is likely to be true.

4. IMPLICATIONS FOR THE ESTIMATION OF EARNINGS DYNAMICS

One of the motivations for this paper is to assess to what degree measurement error biases estimates of the parameters in a dynamic model of earnings. Consider the simple model in (1) again. We want to estimate the variances of the innovation to the permanent component, ε_{it}, and to the transitory component, η_{it}. The only data available are contaminated by measurement error, i.e. data are only available on x_{it}. Instead of the true model in (4) we have

$$\Delta x_{it} = \varepsilon_{it} + (1+\alpha)(\eta_{it} - \eta_{it-1}) + \xi_{it} - \xi_{it-1} \tag{10}$$

The moment conditions are

$$var(\Delta x_{it}) = \sigma_\varepsilon^2 + (1+\alpha)^2(\sigma_{\eta t}^2 + \sigma_{\eta t-1}^2) + \sigma_{\xi t}^2 + \sigma_{\xi t-1}^2$$

$$cov(\Delta x_{it}, \Delta x_{it-1}) = -(1+\alpha)\sigma_{\eta t-1}^2 - \sigma_{\xi t-1}^2 \qquad (11)$$

The estimates of the variance of true earnings will be confounded by two bias terms of opposite sign. First, the variance will be underestimated because $\alpha < 0$. On the other hand, the white noise term in measurement error ξ_t introduces an additional positive variance component. Similar biases arise for the first autocovariance. Tables 8 and 9 summarize these biases for 1982 and 1986 using the estimates from Table 4.

The first column reports the bias due to underreporting of transitory income, it is given by $((1+\alpha)^2 - 1)(\sigma_{\eta t}^2 + \sigma_{\eta t-1}^2)$. Column (2) gives the bias due to the white noise component in the measurement error. I use $2\sigma_{\xi t}^2$, assuming that the variance of the noise is the same in two adjacent periods. Recall that this variance is only estimated for the survey years 1982 and 1986. Notice that this component has substantial variance, in the same order of magnitude as the variance of earnings changes (column 4). However, the total bias is much lower because of the underreporting of the transitory component. This is especially true for the recession year 1982 because of the large variance of the transitory component.

The first panel uses actual period *t* and period *t-1* estimates of the variances in the transitory earnings shock. This creates an asymmetry with the white noise error component. The second panel in the table uses only period *t* estimates for both variables and assumes the variances were the same in *t-1*. This yields more extreme values in the bias due to underreporting of the transitory component especially in 1982 because this was the highest variance year.

The reliability ratio for 1982, i.e. the ratio of variance of true earnings changes to the variance of survey earnings changes $\sigma_{\Delta y}^2/\sigma_{\Delta x}^2$, is around 0.8. This is somewhat higher than the 0.65 that Bound and Krueger (1991) found for 1976-77. The picture is much bleaker for the boom year 1986 where the reliability ratio is around one half. One should keep in mind that this comparison is not very accurate as the total variance in Bound and Krueger is much larger since their data come from the entire economy while my data just cover one plant. The variance in within plant earnings changes is of course much lower. Bound et al. (in press) report the variance of reported earnings changes to be about three times as high in the PSID than in the PSIDVS. The white noise and individual effects of the measurement error are presumably comparable in the population and the single plant use here. Hence, there will be two sources of reduction in the biases from measurement error in a representative sample. First, the variance of η_{it} will rise, so if α is the same in the population, mean reversion will be stronger in a representative sample. Secondly, the reliability ratio will be higher because the variance of the signal is larger. Furthermore, it is worth pointing out that the strong business cycle pattern present in the plant level PSIDVS data is not present in the PSID. Thus, the differences between 1982 and 1986 should be much lower in the population.

Return to the analysis of the PSIDVS data. For the first autocovariance biases can only be computed by using period *t* innovation variances and assuming that the variances were the same during the two previous years. This yields the results given in table 9. The covariance in the first line of column 4 is the one between 1983 and 1982. The biases in the first autocovariance show a very similar business cycle pattern as the variance.

Table 10 presents some results for the decomposition of the variance of earnings changes into a permanent effect (σ_ε and a transitory effect σ_η . These are based on the second panel in table 9 which is consistent with the assumptions for the autocovariances in table 8. Columns (2) and (3) present the estimates for the standard deviations of the innovations. The estimates based on the validation earnings from table 4 are given below in parentheses. The transitory component is identified from the autocovariance. Given its estimate, the permanent component can be identified from the variance. For 1982 the results are not compatible with a random walk plus noise model for earnings: the first order autocorrelation is less than -0.5. For 1986 both components are overestimated. As a measure of the relative importance of the two components the contribution of permanent innovations to the total variance of earnings changes is given in column (4). It can be seen that the measurement error biases both components in a similar way. This would not be true if measurement error was purely white noise. Then only the variance of the transitory component would be overstated while the variance of the permanent component would be estimated accurately. The relative contributions would therefore be affected.

This is relatively good news for the estimation of earnings dynamics: the negative correlation of measurement error with transitory earnings attenuates the role of white noise measurement error. The low reliability ratios for this sample should not be seen too pessimistically. The sample exhibits much less variation in the true variables than typical labor market datasets that include people in many different jobs, industries, and regions. Furthermore, the estimated permanent/transitory decomposition is roughly correct since the biases in the estimated variances and autocovariances are of about equal magnitude. This result will only carry over to a more representative sample as long as the transitory earnings variance does not rise too much. A small rise will actually be helpful in further reducing the impact of measurement error. If the transitory variance becomes sufficiently large the bias due to the underreporting feature will dominate the bias due to the white noise error component. This would imply that the importance of the transitory component might even be understated in a permanent/transitory decomposition.

5. CONCLUDING COMMENTS

This paper analyzes the structure of measurement error in earnings in relation to the dynamics of the true earnings process. The PSID Validation Study is a fairly unique data set that allows such an investigation. But the data also have important limitations. The small number of observations in the data set make it hard to statistically discriminate between opposing hypotheses. Furthermore, only two observations on survey earnings are available and these are not from adjacent years. Finally, the data come from a single plant and may therefore not be very representative. Most importantly, the variance in the responses is also limited by this fact.

The results obtained here are qualitatively similar to the findings of Bound and Krueger (1991) on the CPS-SSA match file. The PSIDVS data exhibit mean reverting measurement error in earnings, i.e. the measurement error is negatively correlated with validation earnings. However, this correlation is weaker than in Bound and Krueger. The results differ between a boom and a recession year mainly because the variance of the earnings process is different in these states while the properties of the measurement error changes little. To the extend that the economy as a whole is less cyclical than manufacturing this finding will not easily carry over to a larger population. I also find a positive autocorrelation in the measurement error over four years.

I characterize the data using a model where individuals underreport the transitory component of earnings. This seems like a sensible behavioral hypothesis since permanent income changes are often associated with more important events and should therefore be remembered better. This part of the measurement error implies that the transitory component in earnings will be underestimated in a model of earnings dynamics. However, the white noise component in measurement error more than offsets this effect. This implies that the permanence in earnings changes may be understated somewhat if inferences are made on the basis of data with measurement error. It would be wrong, however, to attribute the entire transitory component of earnings changes to measurement error. A better rule of thumb is probably to treat the estimated transitory/permanent decompositions of earnings as correct.

Overall, the model fits the data well when estimated as a covariance structure model or when identified using an instrumental variables strategy. Some features of the model are not fully consistent with the stylized facts about measurement error. The correlation of measurement errors over time cannot be fully accounted for by a person fixed effect in misreporting. A more extensive panel of survey and validation records would be necessary to investigate this aspect in the data. In more representative samples, earnings dynamics also do not fit the random walk plus white noise model as well. If the transitory component of earnings is itself serially correlated measurement error will also exhibit richer dynamics. On the other hand, this may also call for a more elaborate model of misreporting.

ACKNOWLEDGEMENTS

I thank David Card, François Laisney, and participants at the Princeton and University of Mannheim Labor Lunches for helpful comments as well as an Associate Editor and two referees for valuable suggestions.

APPENDIX: OMITTED OBSERVATIONS

Here I provide some information on the four observations which I omitted from the sample. Each observation was omitted because the validation record is questionable. This does not mean that the record is actually erroneous but on the basis of the available information there is a possibility that these records may be coded incorrectly. Since the four observations have a large influence on any (least squares) statistics calculated with the sample I find it saver to exclude them from the analysis. I should stress, however, that the decision to exclude these particular observations is taken rather arbitrarily, no particular cutoffs were used.

Each individual in the PSIDVS that is part of the 1983 wave has a 1983-ID, similarly, each individual in the 1987 wave has a 1987-ID. All four outliers are part of the panel. I identify them by their 1987-ID. The four omitted observations have the IDs 20, 127, 191, and 307. Information about validation and reported earnings is displayed in table A1 for these individuals. The questionable records are highlighted.

Observation 20 is an hourly worker with 26 years of tenure earning an hourly wage of about $10 in each year from 1982 to 1986. In 1981 he only worked 64 hours. It seems inconceivable that someone keeps working for less than a dollar an hour even for such a short time.

The other three observations are salaried employees, no hours information is available for them. They all have 16 or more years of education and between 12 and 18 years of tenure. In each case the highlighted earnings reflect a significant drop in normal earnings levels. The respondents report earnings much in line with earnings in the other years. It seems peculiar that the respondents would not remember earnings changes of such a magnitude after one or two years. There are no other college educated employees in the sample that had earnings changes even approaching the same magnitude.

REFERENCES

Abowd, J. and Card, D. (1989), "On the Covariance Structure of Earnings and Hours Changes," *Econometrica,* 57, 411-445.

Altonji, J. G. and Segal, L. (1993), "Small Sample Bias in GMM Estimation of Covariance Structures," working paper, Department of Economics, Northwestern University.

Bound, J., Brown, C., Duncan, G. J., and Rodgers, W. L. (1989), "Measurement Error in Cross-sectional and Longitudinal Labor Market Surveys: Results from Two Validation Studies," Working Paper No. 2885, National Bureau of Economic Research

_____ (in press), "Evidence on the Validity of Cross-Sectional and Longitudinal Labor Market Data," *Journal of Labor Economics.*

Bound, J. and Krueger, A. B. (1991), "The Extent of Measurement Error in Longitudinal Earnings Data: Do Two Wrongs Make a Right?" *Journal of Labor Economics,* 9, 1-24.

Chamberlain, G. (1984), "Panel Data," in *Handbook of Econometrics* (Vol. 2), eds. Z. Griliches and M. D. Intriligator, Amsterdam: North Holland, 1247-1318.

David, M., Little, R. J. A., Samuhel, M. E., and Triest, R. E. (1986), "Alternative Methods for CPS Income Imputation," *Journal of the American Statistical Association,* 81, 29-41.

Duncan, G. J. and Hill, D. H. (1985), "An Investigation of the Extent and Consequences of Measurement Error in Labor-economic Survey Data," *Journal of Labor Economics,* 3, 508-532.

Duncan, G. J. and Mathiowetz, N. A. (1985), *A Validation Study of Economic Survey Data,* Ann Arbor, MI: Institute for Social Research.

Hall, R. E. and Mishkin, F. S. (1982), "The Sensitivity of Consumption to Transitory Income: Estimates from Panel Data on Households," *Econometrica,* 50, 461-481.

Mellow, W. and Sider, H. (1983), "Accuracy of Response in Labor Market Surveys: Evidence and Implications," *Journal of Labor Economics,* 1, 331-344.

Newey, W. K. (1985), "Generalized Method of Moments Specification Testing," *Journal of Econometrics,* 29, 229-256.

Sudman, S. and Bradburn, N. M. (1974), *Response Effects in Surveys. A Review and Synthesis,* Chicago, IL: Aldine Publishing Compnay.

Table 1. Sample Statistics: Panel

Variable	Mean	Std. Dev.	Minimum	Maximum
Validation Earnings 82	10.27	0.29	8.71	10.96
Survey Earnings 82	10.28	0.30	8.58	11.00
Measurement Error 82	0.01	0.12	-0.48	0.62
Validation Earnings 86	10.47	0.21	9.76	11.02
Survey Earnings 86	10.47	0.24	9.62	10.97
Measurement Error 86	-0.00	0.11	-0.63	0.27
Highest Grade	13.0	1.94	8	17
Tenure	18.9	7.4	8	40
Age	40.3	8.1	29	63
Female	0.10	0.30	0	1
Salaried	0.50	0.50	0	1

Note: Number of observations is 234.

Table 2. Sample Statistics: 1986 Cross Section

Variable	Mean	Std. Dev.	Minimum	Maximum
Validation Earnings 86	10.46	0.21	9.74	11.02
Survey Earnings 86	10.46	0.23	9.62	11.08
Measurement Error 86	-0.00	0.12	-0.63	0.55
Highest Grade	12.58	2.07	4	17
Tenure	20.30	8.03	8	43
Age	41.67	8.67	27	63
Female	0.08	0.27	0	1
Salaried	0.33	0.47	0	1

Note: Number of observations is 437.

Table 3. Basic Facts About Measurement Error in Earnings

Data set	nobs	var(m)	var(y)	var(x)	b_{my}	b_{mx}
Panel 1986	234	0.0110	0.0441	0.0557	0.007 (0.037)	0.203 (0.032)
Panel 1986 (y > 9)	231	0.0110	0.0446	0.0562	0.006 (0.037)	0.201 (0.032)
Cross Sect. 1986	437	0.0137	0.0440	0.0517	-0.068 (0.036)	0.207 (0.026)
Panel 1982	234	0.0135	0.0860	0.0875	-0.070 (0.038)	0.086 (0.026)
Panel 1982 (y > 9)	231	0.0135	0.0697	0.0697	-0.097 (0.042)	0.097 (0.031)
Cross Sect. 1982	352	0.0142	0.0739	0.0735	-0.099 (0.035)	0.094 (0.023)
Cross Sect. 1982 (y > 9)	351	0.0142	0.0670	0.0654	-0.118 (0.034)	0.096 (0.025)

Note: Heteroskedasticity consistent standard errors in parentheses.

Table 4. Minimum Distance Estimates for Earnings Changes and Measurement Error

Parameter	Estimate	Std. Err.	Parameter	Estimate	Std. Err.
$\sigma_{\eta 81}$	0.106	0.027	σ_{ε}	0.042	0.055
$\sigma_{\eta 82}$	0.139	0.022	$\sigma_{\xi 82}$	0.105	0.011
$\sigma_{\eta 83}$	0.132	0.021	$\sigma_{\xi 86}$	0.096	0.010
$\sigma_{\eta 84}$	0.120	0.033	σ_{μ}	0.034	0.013
$\sigma_{\eta 85}$	0.073	0.021	α	-0.258	0.123
$\sigma_{\eta 86}$	0.089	0.025			
χ^2 -Statistic [dof]			23.870 [17]		
p-value			0.123		

Note: Panel data, number of observations is 231.

Table 5. Estimated Covariance Matrix of Earnings and Measurement Error

	Δy_{i82}	Δy_{i83}	Δy_{i84}	Δy_{i85}	Δy_{i86}	m_{i82}	m_{i86}
Δy_{i82}	0.0322	-0.54	-0.05	-0.18	0.09	-0.21	-0.21
Δy_{i83}	-0.0193	0.0399	-0.45	0.17	0.07	0.22	0.08
Δy_{i84}	-0.0018	-0.0163	0.0335	-0.43	-0.06	0.05	0.01
Δy_{i85}	-0.0049	0.0051	-0.0121	0.0239	-0.30	0.09	0.00
Δy_{i86}	0.0019	0.0016	-0.0012	-0.0055	0.0141	-0.05	-0.18
m_{i82}	-0.0044	0.0051	0.0011	0.0016	-0.0007	0.0135	0.10
m_{i86}	-0.0039	0.0016	0.0001	0.0000	-0.0023	0.0012	0.0110

Note: Covariances are below the diagonal, correlations above the diagonal

Table 6. Fitted Covariance Matrix

	Δy_{i82}	Δy_{i83}	Δy_{i84}	Δy_{i85}	Δy_{i86}	m_{i82}	m_{i86}
Δy_{i82}	0.0322	-0.55	0.00	0.00	0.00	-0.24	0.00
Δy_{i83}	-0.0192	0.0383	-0.48	0.00	0.00	0.22	0.00
Δy_{i84}	0.0000	-0.0173	0.0334	-0.54	0.00	0.00	0.00
Δy_{i85}	0.0000	0.0000	-0.0144	0.0214	-0.30	0.00	0.00
Δy_{i86}	0.0000	0.0000	0.0000	-0.0053	0.0149	0.00	-0.16
m_{i82}	-0.0050	0.0050	0.0000	0.0000	0.0000	0.0135	0.10
m_{i86}	0.0000	0.0000	0.0000	0.0000	-0.0020	0.0012	0.0110

Note: Covariances are below the diagonal, correlations above the diagonal

Table 7. Regression of Measurement Error on Earnings Using Hours Changes as Instruments

	1986 Panel	1982 Panel	1986 Cross Section	1982 Cross Section
Coefficient	-0.31 (0.20)	-0.15 (0.12)	-0.23 (0.93)	-0.54 (0.17)
R^2	0.07	0.07	0.14	0.23
no. of obs.	114	114	168	292

Note: Heteroskedasticity consistent standard errors in parentheses. See text for the definition of the reported R^2.

Table 8. The Biases in Estimated Variances of Earnings

Year	Underrep. of trans. earnings (1)	White noise Error (2)	Total Bias (1) + (2) (3)	$var(\Delta y_{it})$ (4)	Reliability Ratio (5)
			Actual Innovation Variances		
1982	-0.0137	0.0221	0.0084	0.0322	0.79
1986	-0.0060	0.0184	0.0124	0.0141	0.53
			Period t Innovation Variances		
1982	-0.0174	0.0221	0.0047	0.0322	0.87
1986	-0.0071	0.0184	0.0113	0.0141	0.56

Table 9. The Biases in the Estimated First Order Autocovariance of Earnings

Year	Underrep. of trans. earnings (1)	White noise Error (2)	Total Bias (1) + (2) (3)	$cov(\Delta y_t, \Delta y_{t-1})$ (4)	Reliability Ratio (5)
1982	0.0087	-0.0110	-0.0023	-0.0193	0.89
1986	0.0036	-0.0092	-0.0056	-0.0055	0.50

Table 10. Implied Estimates of the Innovation Variances

Year	$\sigma_{\Delta xt}$	$\sigma_{\eta t}$	$\sigma_{\varepsilon t}$	Contrib. of Perm. Comp.
	(1)	(2)	(3)	(4)
1982	0.19	0.15	---	---
	(0.18)	(0.14)	(0.04)	(5.5 %)
1986	0.16	0.11	0.06	12.6 %
	(0.12)	(0.09)	(0.04)	(12.5 %)

Note: Estimates from Table 4 in parentheses.

Table A1. Earnings for Observations Omitted form the Final Sample

ID	Record	1981	1982	1983	1984	1985	1986
20	Validation	***62***	24204	17525	20379	33020	32224
	Reported 83	2400	24000				
	Reported 87	4000	24000	23000	30000	32000	32000
127	Validation	***29379***	51951	55720	57258	57045	54195
	Reported 83	52000	52000				
	Reported 87	50000	50000	55000	55000	55000	60000
191	Validation	***5237***	30653	37565	41557	40803	45327
	Reported 83	26500	36000				
	Reported 87	45000	51000	52000	55000	62000	63500
307	Validation	40730	39691	44020	48307	52063	***11076***
	Reported 87	40000	42000	43000	49000	51000	55000

3248 026

MIT LIBRARIES
3 9080 00832816 0

Date Due

Lib-26-67