A comprehensive approach to measurement error in panel data models

DRAFT–PLEASE DO NOT QUOTE

Laura Spierdijk Tom Wansbeek

Faculty of Economics and Business, University of Groningen

May 23, 2011

Abstract

The presence of measurement error is a pervasive phenomenon in regression analysis leading, if untreated, to inconsistent estimators. Yet the topic is commonly neglected. This may be due to the well-known lack of identification of the parameters in the simplest case. We begin by briefly reviewing this case and indicate how identification can be achieved in more complicated cases.

One such case is the linear regression model for panel data. For that case, we derive the full set of moment conditions that can be derived from the third moments of the observations and from the linear restrictions on the covariance matrices of the disturbances and the measurement error, thereby extending the known results in various ways.

The underlying strategy is to generate many instruments and to tackle any weak-instruments problem that may arise by using limited-information maximum likelihood together with bootstrap standard errors.

1 Introduction

Measurement error should be a pervasive issue in econometrics as variables are often measured with error. Also, sometimes the measurement is correct but the observed variable is not entirely the relevant one. One step further, a variable may be entirely conceptual ("latent"), like income, inflation, productivity, and many other economically relevant variables. Wansbeek and Meijer (2000) provide a fairly comprehensive overview of the area.

Untreated measurement error in a regressor leads to inconsistent estimators. There is no quick fix. Hence the issue is usually neglected, due to identification problems in the simplest case. That is, there is no identification of parameters in the model with a single regressor under the conditions of linearity, independence and normality. Usually these conditions constitute the most ideal situation for a researcher but the opposite holds in the case of measurement error. As soon as these conditions do not hold, the scope for identification widens. In the case of panel data, the observations over time per individual will not be independent anymore. Hence, the presence of panel data offers the scope for consistent estimation even if the regressors are subject to measurement error.

At the same time, the need to treat measurement error is larger with panel data than with a cross-section, at least in a model with fixed effects. As Griliches and Hausman (1986) have pointed out in their seminal contribution, the regressor tends to be more strongly correlated over time than the measurement error, and transformation of the data to wipe out the individual effect will reduce the variance of the regressor more than the variance of the measurement error, magnifying the measurement error effect.

At the same time, though, the panel data context with its additional dimension can offer a way to consistent estimation. Griliches and Hausman (1986) show how restrictions on the covariance matrix of the measurement errors can be exploited to obtain a consistent estimator for the regression coefficient. A somewhat more structured approach is given by Wansbeek and Koning (1991), for the case of i.i.d. measurement error. For the slightly more general case of uncorrelated measurement error, Biørn and Klette (1998) present a GMM approach. Biørn (2000) takes also the error in the equation into account. He derives moment restrictions available from structuring both the covariance matrix of the error in the equation and the one of the measurement error. These can be diagonal or band-diagonal or can have the same off-diagonal elements. Wansbeek (2001) presented a general GMM approach based on linear restrictions, of any form, on the measurement error covariance matrix. An elaboration is given by Xiao, Shao, Xu, and Palta (2007). They correct an error in Wansbeek (2001), provide small-sample insight through simulation, and identify cases where a single-step approach in GMM is already optimal and no second step is needed. Shao, Xiao and Xu (2011) extend Wansbeek (2001) to the case of unbalanced panel data. Xiao, Shao and Palta (2010a,b) provide several extensions, including the presence of multiple covariates measured with error.

This literature is based on exploiting conditions on parameter matrices. Another approach is through exploiting higher moments of the data. For the cross-section case, this approach has spawned an elaborate literature over the years, summarized in Wansbeek and Meijer (2000), pp. 144-145. Recent work includes Erickson and Whited (2000, 2002), who pay some attention to the panel data case.

In this paper we review the econometrics of measurement error in panel data and provide a comprehensive approach. To set the stage for the discussion, we give a very brief overview of the measurement error issue in section **??**. This pertains to a single cross section only. In section **??** we investigate how this has to be adapted for the simplest possible panel data model. We discuss identification and briefly discuss instrumental variables in general. We elaborate the latter in section **??** in two ways. We derive the instruments that become available when the third moment of the latent variable is non-zero, and we derive the instruments induced by any linear restrictions that may be present on the covariance matrices of the error in the equation and the measurement error. This is all about a simple model without regressors beyond the single, latent one and without fixed effects. Section panelplus shows how the results for this simple model can be readily adapted to incorporate these two important elements. Section **??** concludes.

2 The basic measurement error model

To set the stage, we recapitulate some issues regarding measurement error in the case of a single cross-section. As a non-essential simplification, we take all variables in deviations from their means. We start with the case of a single regressor, and assume all random variables to be normally distributed for the moment.

For $n = 1, \ldots, N$ the regressor ξ_n is unobservable, and instead a proxy x_n is observed:

$$\begin{aligned} y_n &= \xi_n \beta + \varepsilon_n & (1) \\ x_n &= \xi_n + v_n, & (2) \end{aligned}$$

with $\xi_n \sim N(0, \sigma_\xi^2)$, $\varepsilon_n \sim N(0, \sigma_\varepsilon^2)$ and $v_n \sim N(0, \sigma_v^2)$, independent. Eliminate ξ_n to obtain

$$\begin{aligned} y_n &= x_n \beta + u_n \\ u_n &\equiv \varepsilon_n - v_n \beta \ (= y_n - x_n \beta). \end{aligned}$$

So $\mathrm{E}(x_n u_n) = -\sigma_v^2 \beta \neq 0$, which spells trouble. Since there holds

$$\begin{aligned} \mathrm{E}(x_n y_n) &= \sigma_\xi^2 \beta \\ \mathrm{E}(x_n^2) &= \sigma_\xi^2 + \sigma_v^2, \end{aligned}$$

OLS in the reduced form gives

$$\operatorname*{plim}_{N\to\infty} b_{\text{OLS}} = \operatorname*{plim}_{N\to\infty} \frac{\sum_n x_n y_n}{\sum_n x_n^2} = \frac{\sigma_\xi^2}{\sigma_\xi^2 + \sigma_v^2}\beta < \beta \tag{3}$$

if $\beta > 0$ and otherwise the opposite holds. There is bias towards zero, "attenuation". The results disappoint, coefficients are lower than expected.

How serious is the problem? Since we have assumed normality, the complete sufficient statistic is $\left\{s_y^2, s_{yx}, s_x^2\right\}$. So we get to know (in the limit) the left-hand side of their expectation, respectively given by

$$\begin{aligned} \sigma_y^2 &= \sigma_\xi^2\beta^2 + \sigma_\varepsilon^2 && (4)\\ \sigma_{yx} &= \sigma_\xi^2\beta && (5)\\ \sigma_x^2 &= \sigma_\xi^2 + \sigma_v^2. && (6) \end{aligned}$$

but nothing more. We have three equations in four unknown parameters. None of the parameters is identified. The problem is serious. Reiersøl (1950) has shown that the model is not identified if and only if ξ_n is normal. When the ξ_n are taken to be fixed ("incidental") parameters rather than random elements, the situation does not change (Wald, 1948).

The situation is not entirely hopeless. Another way to obtain information on β from the data is by performing the *reverse* regression, which is the regression of x on y, and invert the result to make it comparable to the result of the direct regression. This yields

$$\operatorname*{plim}_{N\to\infty} b_{\text{REV}} = \operatorname*{plim}_{N\to\infty} \frac{\sum_n y_n^2}{\sum_n x_n y_n} = \frac{\sigma_y^2}{\sigma_{yx}} = \beta + \frac{\sigma_\varepsilon^2}{\beta\sigma_\xi^2}, \tag{7}$$

which has the same sign as β, but is larger in magnitude. Since (**??**) showed that b_{OLS} converges to $\beta\sigma_\xi^2/\sigma_x^2$, also the same sign but smaller in magnitude, we have that in the limit β is bounded by the direct and reverse regression. If the bound is tight, the result is useful.

Usually, the combination of linearity, independence and normality is attractive due to its simplicity. Here it creates trouble. We have a way out as soon as the model is more complicated. Directions for a solution include exact additional information, nonnormality, nonlinearity, additional information through external instrumental variables, latent variable modeling, and dependence of observations.

Due to its relevance later on, we elaborate nonnormality. As soon as the model is nonnormal, we can derive instrumental variables from within the model. Consider

$$\mathrm{E}\left((x_n y_n)x_n\right) = \mathrm{E}\left((\xi_n\beta + \varepsilon_n)(\xi_n + v_n)^2\right) = \mathrm{E}(\xi_n^3)\beta, \tag{8}$$

which, for all practical purposes, is nonzero if the distribution of ξ_n is asymmetric, cf. Meijer (2002). Furthermore,

$$\mathrm{E}\left((x_n y_n)u_n\right) = \mathrm{E}\left((\xi_n + v_n)(\xi_n\beta + \varepsilon_n)(\varepsilon_n - v_n\beta)\right) = 0. \tag{9}$$

Consequently, when the third-order moment of ξ_n does not vanish, $x_n y_n$ is a valid instrument. Since variables in econometrics are seldom normal, this observation opens the way to a simple, virtually assumption-free test for the presence of measurement error in linear regression. To the best of our knowledge, this idea has not been elaborated let alone put to practice.

3 A simple panel data model

We now consider the panel data case and see how the above needs adaptation. Since panel data models tend to become unwieldy because of the extra dimension, we start with the simplest case and do not consider additional regressors and individual effects. That extension has to wait till section **??**.

The model is as in (**??**) and (**??**) but $y_n, \xi_n, \varepsilon_n$ and v_n are now T-vectors with $\xi_n \sim N(0, \Sigma_\xi), \varepsilon_n \sim N(0, \Sigma_\varepsilon)$ and $v_n \sim N(0, \Sigma_v)$, all $T \times T$ matrices. Since we do not impose a structure on these matrices as yet, this specification encompasses the random effects model as it structures Σ_ε. We have

$$\begin{aligned} \mathrm{E}(y_n' y_n) &= \operatorname{tr}\Sigma_\xi \beta^2 + \operatorname{tr}\Sigma_\varepsilon && (10) \\ \mathrm{E}(x_n' y_n) &= \operatorname{tr}\Sigma_\xi \beta && (11) \\ \mathrm{E}(x_n' x_n) &= \operatorname{tr}\Sigma_\xi + \operatorname{tr}\Sigma_v. && (12) \end{aligned}$$

OLS in the reduced form now gives, cf. (**??**),

$$\operatorname*{plim}_{N\to\infty} b_{\mathrm{OLS}} = \operatorname*{plim}_{N\to\infty} \frac{\sum_n x_n' y_n}{\sum_n x_n' x_n} = \frac{\operatorname{tr}\Sigma_\xi}{\operatorname{tr}\Sigma_\xi + \operatorname{tr}\Sigma_v}\beta. \tag{13}$$

So also here there is a bias towards zero. For the reverse regression for this case we have

$$\operatorname*{plim}_{N\to\infty} b_{\mathrm{REV}} = \beta + \frac{\operatorname{tr}\Sigma_\varepsilon}{\beta \operatorname{tr}\Sigma_\xi}, \tag{14}$$

cf. (**??**), which is away from zero. So also in the panel data case, in the limit, β is between the results of the direct and reverse regressions.

Identification

The panel data context does not offer a way out directly. Analogous to (**??**)-(**??**) we have as the complete second-order implications of the model

$$\begin{aligned} \mathrm{E}(y_n y_n') &= \Sigma_\xi \cdot \beta^2 + \Sigma_\varepsilon & (15) \\ \mathrm{E}(y_n x_n') &= \Sigma_\xi \cdot \beta & (16) \\ \mathrm{E}(x_n x_n') &= \Sigma_\xi + \Sigma_v & (17) \end{aligned}$$

As an informal check on identification, notice that from (**??**) we need the information contained in the second moment of y_n to identify Σ_ε since the latter only occurs in (**??**) and we need the information contained in the second moment of x_n to identify Σ_v. This leaves us with the covariance of y_n and x_n to identify both β and Σ_ξ, which occur as a product and cannot be disentangled without further information in some form. Hence dependence of observation per se is insufficient for identification.

Instrumental variables

As usual, the way to proceed is through instrumental variables (IVs). They may be available from outside the model. Under circumstances to be described, they are implied by the model itself. Whatever their origin, IVs are observations R_n such that $\mathrm{E}(R_n' u_n) = 0$ while $\mathrm{E}(R_n' \xi_n) \neq 0$. Cameron and Trivedi (2005) provide a good summary of IV estimation in a panel data context. With $R' \equiv (R_1', \ldots, R_N')'$ and W a weighting matrix, the basic IV estimator is

$$\hat{\beta}_{\mathrm{IV}} = \frac{x'RWR'y}{x'RWR'x}, \tag{18}$$

with the y_n and x_n collected in NT-vectors y and x.

This formulation is more general than apparent at first sight. Assume that there is an observed variable r_n ($T \times 1$) whose elements do not correlate with any of the elements of u_n. Then

$$\mathrm{E}(r_n \otimes u_n) = \mathrm{E}(r_n \otimes I_T) u_n = 0, \tag{19}$$

so for this case of strong exogeneity $R_n = r_n' \otimes I_T$. In case of weak exogeneity some of the rows of (**??**) have to be discarded.

As always with IV, we have to consider the possibility that the instruments are weak, and that holds a fortiori in a panel data setting where the number of IVs increases fast with T like in (**??**). One approach is to condense the instruments. Another approach is to leave the instruments as they are and employ an estimation method like limited-information maximum likelihood (LIML), which outperforms IV when the instruments are weak or the number of observations is not large.

4 Instruments implied by the model

We consider two kinds of instruments that are possibly implied by the model. We first consider the use of third moments of the variables, and next turn to restrictions that may be considered on the covariance matrices of the disturbances and the measurement errors.

Using third-order moments

We can obtain instruments from within the model when we relax the assumed normality of ξ_n. Then, higher-order moments become available. We restrict ourselves here to third moments and exploit any skewness there may be in the distribution of ξ_n. Let

$$\lambda \equiv \mathrm{E}(\xi_n \otimes \xi_n \otimes \xi_n) \neq 0, \tag{20}$$

where the latter means at least one element is non-zero. Now, (**??**) and (**??**) generalize respectively to

$$\begin{aligned} \mathrm{E}(x_n \otimes y_n \otimes x_n) &= \mathrm{E}\left((x_n \otimes y_n \otimes I_T)x_n\right) = \lambda\beta \neq 0 & (21) \\ \mathrm{E}(x_n \otimes y_n \otimes u_n) &= \mathrm{E}\left((x_n \otimes y_n \otimes I_T)u_n\right) = 0, & (22) \end{aligned}$$

with $u_n = y_n - x_n\beta$ as before but now a vector. So, with the subscript "3" denoting that we consider third moments, we have that the matrix

$$R'_{3,n} \equiv x_n \otimes y_n \otimes I_T, \tag{23}$$

of order $T^3 \times T$, correlates with x_n as $\mathrm{E}(R'_{3,n}x_n) \neq 0$ from (**??**) but does not correlate with u_n as $\mathrm{E}(R'_{3,n}u_n) = 0$ from (**??**), thus providing instruments. Let $R_3 \equiv (R'_{3,1}, \ldots, R'_{3,N})'$, of order $NT \times T^3$. Then the IV estimator using third moments is given by (**??**) for $R = R_3$. There are T^3 instruments and also here we run the risk of weak instruments.

In (**??**) we assume λ to have at least one non-zero element. When some of the elements of λ are zero, (**??**) shows that some of the moment conditions do not involve parameters. As Qian and Schmidt (1999) have shown, it is in general worthwhile to maintain these moment conditions in order to improve the asymptotic efficiency of the estimators.

Imposing restrictions on Σ_ε and Σ_v

We now consider linear restrictions that we may be willing to impose on Σ_ε and Σ_v. We need some notation to deal with the redundancy when stacking symmetric matrices. Let L_m of order $m^2 \times m(m+1)/2$ be the elimination matrix that eliminates the elements above the diagonal when stacking a matrix M, say, of order $m \times m$, so $\text{vech}M = L_m' \text{vec}M$. Let D_m of order $m^2 \times m(m+1)/2$ be the duplication matrix with $D_m \text{vech}M = \text{vec}M$ if M is symmetric. There holds $L_m' D_m = I_{T(T+1)/2}$. We use the symbol $\bar{\otimes}$ as shorthand for $w_1 \bar{\otimes} w_2 = L_m'(w_1 \otimes w_2)$ for any m-vectors w_1 and w_2.

The restrictions we consider are linear and can hence be expressed as $\text{vech}\Sigma_\varepsilon = C_\varepsilon \pi_\varepsilon$ and $\text{vech}\Sigma_v = C_v \pi_v$, with C_ε and C_v known and π_ε ($r_\varepsilon \times 1$) and π_v ($r_\pi \times 1$) unknown. For example, when we have the basic random-effects model, $C_\varepsilon = (\text{vech}I_T, \iota_{T(T+1)/2})$. Or, when the measurement errors are uncorrelated over time and have the same variance, $C_v = \text{vech}I_T$. A more complicated form, but still within this general formulation of linearity, arises when we assume that the covariance matrices of ε_n or v_n have a Toeplitz structure.

In stacked form, the moment conditions (**??**)–(**??**) with the restrictions inserted can now, with $\sigma_\xi \equiv \text{vech}\Sigma_\xi$, be written as

$$\mathrm{E}\begin{bmatrix} y_n \otimes y_n - D_T \sigma_\xi \beta^2 - D_T C_\varepsilon \pi_\varepsilon \\ x_n \otimes y_n - D_T \sigma_\xi \beta \\ x_n \otimes x_n - D_T \sigma_\xi - D_T C_v \pi_v \end{bmatrix} = 0. \tag{24}$$

For convenience, and without loss of information, we transform this system by premultiplying it by

$$A \equiv \begin{bmatrix} L_T' & -\beta L_T' & 0 \\ 0 & I_{T^2} & -\beta I_{T^2} \\ 0 & 0 & L_T' \end{bmatrix}, \tag{25}$$

to obtain the alternative set of moment conditions, using for brevity $u_n = y_n - x_n\beta$,

$$\mathrm{E}\begin{bmatrix} u_n \bar{\otimes} y_n - C_\varepsilon \pi_\varepsilon \\ x_n \otimes u_n + D_T C_v \pi_v \beta \\ x_n \bar{\otimes} x_n - \sigma_\xi - C_v \pi_v \end{bmatrix} = 0. \tag{26}$$

The effect of this transformation by A is twofold. First, we eliminate redundancy due to the symmetry of $\mathrm{E}(y_n y_n')$ and $\mathrm{E}(x_n x_n')$. Second, we make the system linear in β.

Now, let $C_\varepsilon^+ \equiv (C_\varepsilon' C_\varepsilon)^{-1} C_\varepsilon'$ and let $C_\varepsilon^{\perp'}$ be an orthogonal complement of C_ε. Let C_v^+ and $C_v^\perp$ be analogous with respect to $D_T C_v$. We multiply the first and second line of (**??**) by

$$\begin{bmatrix} C_\varepsilon^\perp \\ C_\varepsilon^+ \end{bmatrix} \text{ and } \begin{bmatrix} C_v^\perp \\ C_v^+ \end{bmatrix},$$

respectively, to obtain

$$\mathrm{E}\begin{bmatrix} C_\varepsilon^\perp(u_n \bar{\otimes} y_n) \\ C_\varepsilon^+(u_n \bar{\otimes} y_n) - \pi_\varepsilon \\ C_v^\perp(x_n \otimes u_n) \\ C_v^+(x_n \otimes u_n) + \pi_v \beta \\ x_n \bar{\otimes} x_n - \sigma_\xi - C_v \pi_v \end{bmatrix} = 0. \tag{27}$$

The second, fourth and fifth rows of (**??**) just-identify π_ε, π_v and σ_ξ. For estimating β, we have the first and third line of (**??**). This suggests a two-step estimation procedure. After estimating β by GMM based on the first and third line of the sample counterpart of (**??**), the result can be substituted in the other equations and the estimates of the other parameters follow directly by MM.

As to estimating β, we have

$$\begin{aligned} C_\varepsilon^\perp(u_n \bar{\otimes} y_n) &= C_\varepsilon^\perp L_T'(u_n \otimes y_n) = \left[C_\varepsilon^\perp L_T'(I_T \otimes y_n)\right] u_n \equiv R_{\varepsilon,n}' u_n \\ C_v^\perp(x_n \otimes u_n) &= \left[C_v^\perp(x_n \otimes I_T)\right] u_n \equiv R_{v,n}' u_n, \end{aligned}$$

with $R_{\varepsilon,n}$ and $R_{v,n}$ implicitly defined. So the restrictions provide us with the instrument matrices $R_\varepsilon \equiv (R_{\varepsilon,1}', \ldots, R_{\varepsilon,N}')'$, of order $NT \times r_\epsilon$, and $R_v \equiv (R_{v,1}, \ldots, R_{v,N})'$, of order $NT \times r_v$.

5 Extending the simple panel data model

We now extend the model to include additional regressors and individual effects in order to get a full-fledged panel data model,

$$y_n = \xi_n \beta + Z_n \gamma + \iota_T \alpha_n + \varepsilon_n \tag{28}$$

$$x_n = \xi_n + v_n, \tag{29}$$

with Z_n, of order $T \times k$, a matrix of exogenous regressors (which could include the empirically important time dummies) with coefficients γ ($k \times 1$), ι_T is a T-vector of ones, and α_n is the random, mean-zero individual effect. It may correlate with ξ_n (and is hence conventionally called a fixed effect), and we employ the matrix B, of order $T \times (T-1)$ and rank $T-1$, with property $B'\iota_T = 0$, to eliminate it. We put a tilde on a matrix or vector with T rows when it has been premultiplied by B'.

On bounds

We start by considering the bounds on β from OLS and reverse regression, and first assess the effect of the additional regressors only. We collect the y_n, x_n and Z_n in y, x and Z, of order $NT \times 1$, $NT \times 1$ and $NT \times k$, respectively. We project Z out by $M_Z \equiv I_{NT} - Z(Z'Z)^{-1}Z'$, and do OLS and reverse regression on the transformed variables $M_Z y$ and $M_Z x$. Let $\sigma_{Z\xi} \equiv \mathrm{E}(Z_n'\xi_n)$ and $\Sigma_Z \equiv \mathrm{E}(Z_n'Z_n)$. Clearly, the presence of additional regressors, which we assume to be uncorrelated with the error in the equation and the measurement error, leads to an adaptation of (**??**) and (**??**) where $\sigma_{Z\xi}'\Sigma_Z\sigma_{Z\xi}$ has to be subtracted from Σ_ξ while Σ_ε and Σ_v remain unaffected. As a result, the bound widens.

We now assess the effect of individual effects only. We eliminate them by premultiplying the model by B'. Hence, in (**??**) and (**??**), we have to replace $\Sigma_\xi, \Sigma_\varepsilon$ and Σ_v by $\Sigma_{\tilde{\xi}} \equiv B'\Sigma_\xi B$, $\Sigma_{\tilde{\varepsilon}} \equiv B'\Sigma_\varepsilon B$ and $\Sigma_{\tilde{v}} \equiv B'\Sigma_v B$, respectively. In the empirically likely case that the regressor is more correlated over time than the error in the equation and the measurement error, the variance of ξ is reduced much more than the variance of ε and v and hence the bound widens. This occurs a fortiori in the full model (**??**)-(**??**), with both additional regressors and individual effects. Yet, the bound might be tight enough to be useful.

It is of some interest to notice that the bounds are parameters that, by definition, can be consistently estimated. (For readability, avoiding lots of tildes, we look at the case without individual effects.) Let us call these parameters β_{OLS} and β_{REV}. They are estimated from the systems

$$\begin{aligned} \mathrm{E}(x_n'y_n - x_n'x_n\beta_{\mathrm{OLS}} - x_n'Z_n\pi_*) &= 0 \\ \mathrm{E}(Z_n'y_n - Z_n'x_n\beta_{\mathrm{OLS}} - Z_n'Z_n\pi_*) &= 0 \end{aligned}$$

and

$$\begin{aligned} \mathrm{E}(y_n'x_n - y_n'y_n/\beta_{\mathrm{REV}} - y_n'Z_n\pi^*) &= 0 \\ \mathrm{E}(Z_n'x_n - Z_n'y_n/\beta_{\mathrm{REV}} - Z_n'Z_n\pi^*) &= 0, \end{aligned}$$

with π_* and π^* irrelevant parameters. By GMM we can find the asymptotic variance of the upper and lower bound and their difference in a straightforward way.

Adapting the moment conditions

After these preliminaries, we consider the adaptations, to be made to the two kind of moment conditions implied by the model, when there are additional regressors and individual effects. Adding regressors requires a straightforward adaptation. It suffices to read $y_n - x_n\beta - Z_n\gamma$ for u_n instead of $y_n - x_n\beta$ and add the additional moment condition $\mathrm{E}(Z_n'u_n) = 0$ to allow for estimating γ. (If Z_n is correlated with the individual effect, we should rather add the moment condition $\mathrm{E}(Z_n'BB'u_n) = 0$.)

The instruments implied by the third moments are based (**??**). When individual effects α_n are introduced this becomes

$$\mathrm{E}(x_n \otimes y_n \otimes u_n) = \mathrm{E}\left((\xi_n \otimes (\xi_n\beta + \iota_T\alpha_n) \otimes \iota_T)\alpha_n\right).$$

Hence, in order to have proper instruments, we have to redefine $R_{3,n}'$ in (**??**) as $R_{3,n}' \equiv x_n \otimes y_n \otimes B'$.

We now turn to the moment conditions emanating from the restrictions on the covariance matrices. Since y_n depends on α_n and $\tilde{y}_n$ does not, we adapt the treatment by considering moments in $\tilde{y}_n$ rather than in y_n. So we start again from (**??**) but now premultiply the system by

$$\dot{A} \equiv \begin{bmatrix} L_{T-1}'(B' \otimes B') & -\beta L_{T-1}'(B' \otimes B') & 0 \\ 0 & I_T \otimes B' & -\beta(I_T \otimes B') \\ 0 & 0 & L_T' \end{bmatrix}.$$

We add the condition for γ and let $\dot{C}_\varepsilon \equiv L_{T-1}'(B' \otimes B')D_T C_\varepsilon$ and $\dot{C}_v \equiv (I_T \otimes B)D_T C_v$ and obtain

$$\mathrm{E}\begin{bmatrix} \tilde{u}_n \bar{\otimes} \tilde{y}_n - \dot{C}_\varepsilon \pi_\varepsilon \\ x_n \otimes \tilde{u}_n + \dot{C}_v \pi_v \beta \\ x_n \bar{\otimes} x_n - \sigma_\xi - C_v \pi_v \\ Z_n' u_n \end{bmatrix} = 0, \tag{30}$$

cf. (**??**). Analogous to (**??**) this implies, in obvious notation,

$$\mathrm{E}\begin{bmatrix} \dot{C}_\varepsilon^\perp(\tilde{u}_n \bar{\otimes} \tilde{y}_n) \\ \dot{C}_\varepsilon^+(\tilde{u}_n \bar{\otimes} \tilde{y}_n) - \pi_\varepsilon \\ \dot{C}_v^\perp(x_n \otimes \tilde{u}_n) \\ \dot{C}_v^+(x_n \otimes \tilde{u}_n) + \pi_v \beta \\ x_n \bar{\otimes} x_n - \sigma_\xi - C_v \pi_v \\ Z_n' u_n \end{bmatrix} = 0. \tag{31}$$

Again, the second, fourth and fifth rows of (**??**) just-identify $\pi_{\varepsilon}, \pi_{v}$ and σ_{ξ}. The last line just-identifies γ. The moment conditions for β become

$$\mathrm{E}\begin{bmatrix} \dot{C}_{\varepsilon}^{\perp}(B'\bar{\otimes}\tilde{y}_n) \\ \dot{C}_{v}^{\perp}(x_n \otimes B') \end{bmatrix} u_n = 0,$$

from which the instruments readily follow. Notice the use of x_n rather than $\tilde{x}_n$ although we have fixed effects. We still can use the x_n in levels.

6 Conclusions

In this paper we have presented a fairly exhaustive treatment of handling measurement error in a panel data context. We have extended the available approaches in a number of ways and have integrated them in a comprehensive but still fairly simple framework. We allow for restrictions on the covariance matrix of the disturbances in a general. Up till now, the focus was on zero restrictions. Another extension is the consideration of third-order moments and the instruments they imply. Using third moments to solve the measurement error problem has been the subject of a fair number of papers but most of these pertain to a single cross-section. We have also extended the bounds provided by OLS and reverse regression to the panel data case.

There are several aspects that require further work. One concerns the identification of the model. The result by Reiersøl mentioned in section **??** pertains to a single equation and has not been extended to a panel data context. Second, the identifying restrictions on the covariance matrices considered in this paper are of a linear form, and have to be extended to a non-linear formulation where the elements of these matrices are expressed in a number of underlying parameters. This extension e.g. covers the case where the researcher is willing to assume that the disturbances in the model or the measurement errors follow an AR(1) process. A third extension is the dynamic case, where the regressors include the lagged dependent variable.

Another field for further study is the handling of the many instruments that can be generated by the third moments and the restrictions on the covariance matrices. They give a high degree of overidentification for β and then IV estimation can have poor properties in the sense of a small-sample bias and the underestimation of the standard errors. One way out is to combine the moment conditions into a smaller set, but a more promising approach is to use limited-information maximum likelihood, together with bootstrap standard errors.

References

Biørn, E. (2000), "Panel data with measurement errors: instrumental variables and GMM procedures combining levels and differences", *Econometric Reviews*, **19**, 391–424.

Biørn, E. and J.T. Klette (1998), "Panel data with errors-in-variables: essential and redundant orthogonality conditions in GMM estimation", *Economics Letters*, **59**, 275–282.

Cameron, C. and P.K. Trivedi (2005), *Microeconometrics*, Cambrigde University Press.

Erickson, T. and T.M. Whited (2000), "Measurement error and the relationship between investment and *q*,", *Journal of Political Economy*, **108**, 10271057.

Erickson, T. and T.M. Whited (2002), "Two-step GMM estimation of the errors-in-variables model using high-order moments", *Econometric Theory*, **18**, 776–799.

Griliches, Z. and J.A. Hausman (1986), "Errors in variables in panel data", *Journal of Econometrics*, **31**, 93–118.

Meijer, E. (2002), *An asymmetric distribution with zero skewness*, Working paper, University of Groningen.

Qian, H. and P. Schmidt (1999), "Improved instrumental variables and generalized method of moments estimators", *Journal of Econometrics*, **91**, 145–169.

Reiersøl, O. (1950), "Identifiability of a linear relation between variables which are subject to error", *Econometrica*, **18**, 375–389.

Shao, J., Z. Xiao, and R. Xu (2011), "Estimation with unbalanced panel data having covariate measurement error", *Journal of Statistical Planning and Inference*, **141**, 800–808.

Wald, A. (1948), "Estimation of a parameter when the number of unknown parameters increases indefinitely with number of observations", *Annals of Mathematical Statistics*, **19**, 220–227.

Wansbeek, T.J. (2001), "GMM estimation in panel data models with measurement error", *Journal of Econometrics*, **104**, 259–268.

Wansbeek, T.J. and R.H. Koning (1991), "Measurement error and panel data", *Statistica Neerlandica*, **45**, 85–92.

Wansbeek, T.J. and E. Meijer (2000), *Measurement error and latent variables in econometrics*, North-Holland, Amsterdam.

Xiao, Z., J. Shao and M. Palta (2010a), "Instrumental variable and GMM estimation for panel data with measurement error", *Statistica Sinica*, **20**, 1725–1747.

Xiao, Z., J. Shao and M. Palta (2010b), "GMM in linear regression for longitudinal data with multiple covariates measured with error", *Journal of Applied Statistics*, **37**, 791–805.

Xiao, Z., J. Shao, R. Xu, and M. Palta (2007), "Efficiency of GMM estimation in panel data models with measurement error", *Sankhyā*, **69**, 101–111.