Biørn, Erik; Klette, Tor Jakob

Working Paper — Digitized Version
Panel Data with Errors-in-Variables: A Note on Essential and Redundant Orthogonality Conditions in GMM-estimation

Discussion Papers, No. 190

Provided in Cooperation with:
Research Department, Statistics Norway, Oslo

Suggested Citation: Biørn, Erik; Klette, Tor Jakob (1997) : Panel Data with Errors-in-Variables: A Note on Essential and Redundant Orthogonality Conditions in GMM-estimation, Discussion Papers, No. 190, Statistics Norway, Research Department, Oslo

This Version is available at:
http://hdl.handle.net/10419/192174

Terms of use:
Documents in EconStor may be saved and copied for your personal and scholarly purposes.
You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.
If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.
Panel Data with Errors-in-Variables: A Note on Essential and Redundant Orthogonality Conditions in GMM-estimation
Erik Biørn and Tor Jakob Klette

Panel Data with Errors-in-Variables: A Note on Essential and Redundant Orthogonality Conditions in GMM-estimation

Abstract: General Method of Moments (GMM) estimation of a linear one-equation model using panel data with errors-in-variables is considered. To eliminate fixed individual heterogeneity, the equation is differenced across one or more than one periods and estimated by means of instrumental variables. With non-autocorrelated measurement error, we show that only the one-period and a few two-period differences are essential, i.e. relevant for GMM-estimation. GMM estimation based on all orthogonality conditions on the basis of a generalized inverse formulation is shown to be equivalent to estimation using only the essential orthogonality conditions

Keywords: Panel Data, Errors-in-Variables, Instrumental Variables, GMM Estimation, Generalized inverse

JEL classification: C23, C33, C12, C13.

Address: Erik Biørn, University of Oslo, Department of Economics, P.O. Box 1095 Blindern, N-0317 Oslo, Norway E-mail: erik.biorn@econ.uio.no

Tor Jakob Klette, Statistics Norway, Research Department, P.O.Box 8131 Dep., N-0033 Oslo, Norway. E-mail: tjk@snb.no
1 Introduction

Estimation and testing of econometric models for panel data by means of the Generalized Method of Moments (GMM) has received considerable attention in recent years; see Baltagi (1995, especially ch. 8) for a survey. GMM estimation can be used for consistent and asymptotically efficient estimation of linear equations with endogenous right-hand side variables, with lagged values of left-hand side variables as right-hand side variables, with random measurement errors in the right-hand side variables, and for some kinds of nonlinear models.

The focus in this paper is on the errors-in-variables problem for panel data in a single linear static equation context, although several of the procedures and the results below may also be made applicable to other, more complex, situations. A primary motivation of the paper is to elaborate the matrix algebra for GMM estimation of differenced equations and the associated orthogonality conditions in a panel data context in more detail than is commonly given in the literature, cf. e.g. Griliches and Hausman (1986). We specifically discuss a rank problem which arise when using GMM estimation of equations expressed as differences in an errors-in-variables context, when the various differenced equations and the associated orthogonality conditions are not linearly independent. This rank problem can be handled by either (i) replacing the standard inverses in the expressions for the GMM estimators by generalized (Moore-Penrose) inverses, or (ii) eliminating the redundant orthogonality conditions from the GMM procedure, as we prove to be equivalent and computationally more attractive.

2 Instrumental variable estimation for panel data with errors-in-variables.

2.1 The basic model

Consider a balanced set of panel data for $N$ units or groups in $T$ successive periods and the relationship between a left-hand side variable $y$ and a right-hand side variable $x$ (both scalars). The analysis is limited to the case with only one regressor for convenience, as did Griliches and Hausman (1986), but the insights can be generalized rather straightforwardly to cases with several regressors. Let $y_{it}$ and $x_{it}$ denote their values for observation unit $i$ (in the following denoted as individual $i$) in period $t$, satisfying

$$y_{it} = x_{it} \beta + a_i + u_{it}, \quad i = 1, \ldots, N; \ t = 1, \ldots, T,$$

where $a_i$ is a fixed effect (including a common constant term), specific to individual $i$, $\beta$ is an unknown scalar constant, and $u_{it}$ is a zero mean, random disturbance/error term. We assume,
owing to endogeneity of \( x_{it} \) or random measurement error in \( x_{it} \), (i) that \( u_{it} \) is correlated with \( x_{it} \), (ii) that \( u_{it} \) is uncorrelated with \( u_{jt} \) for all \( j \neq i \), and (iii) that \( u_{it} \) is uncorrelated with \( x_{is} \) for all \( i \) and some \( s \neq t \). We will mostly focus on the case where \( u_{it} \) is uncorrelated with \( u_{it} \) for all \( t \neq s \), but we will discuss more general cases in section 3.3. It is convenient to rewrite (1) as \( N \) vector equations, one for each individual \( i \):

\[
y_i = \bar{x}_i \beta + a_i \otimes e_T' + u_i, \quad i = 1, \ldots, N.
\]  

where \( y_i, \bar{x}_i \) and \( u_i \) are \((1 \times T)\) vectors (with \( y_{it}, x_{it} \) and \( u_{it} \) from individual \( i \) as elements), \( \otimes \) is the Kronecker product operator, \( a = (a_1 \cdots a_N) \) and \( e_T = (1 \cdots 1)' \) is a \((T \times 1)\) vector of ones.

In order to eliminate \( a_i \) from (2), we will work with observations differenced across time periods. If \( T > 2 \), these differences can be taken across 1, 2, \ldots, \( T - 1 \) periods. To formalize this, we introduce the \textit{differencing vectors}

\[
D_{ts} = \begin{cases} 
(1 \times T) \text{ vector with} \\
\text{element } t = +1, \text{ element } s = -1, \\
\text{and zero otherwise}
\end{cases}, \quad t, s = 1, \ldots, T; \; t > s,
\]  

where \( t > s \) can be assumed without loss of generality. Premultiplying a \((T \times 1)\) vector by \( D_{ts} \) takes a difference between its \( t \)'th and \( s \)'th elements. Since there are \( S = \frac{1}{2} T(T-1) \) different ways of drawing two elements from \( T \), there are \( S \) such \( D_{ts} \) vectors, among which \( T-1 \) take differences across one period, \( T-2 \) across two periods, \ldots, two across \( T-2 \) periods, and one across \( T-1 \) periods. The differencing vectors are not independent, since all \( S - (T-1) = \frac{1}{2} (T-1)(T-2) \) differences over two or more periods can be constructed from the \( T-1 \) one-period differences, formally \( D_{ts} = \sum_{j=s+1}^{T} D_{j,j-1} \), with \( t > s \) and \( t, s = 1, \ldots, T \). Postmultiplying through (2) by \( D_{ts}' \), recalling that \( D_{ts} e_T = 0 \), we get \( y_{it} D_{ts}' = \bar{x}_i D_{ts}' \beta + u_i D_{ts}' \), or

\[
y_{it} - y_{is} = (x_{it} - x_{is}) \beta + (u_{it} - u_{is}), \quad t, s = 1, \ldots, T; \; t > s,
\]  

Defining the stacked \((S \times T)\) differencing matrix

\[
D = \begin{bmatrix} 
D_{21}' & D_{22}' & \cdots & D_{2,T-1}' & D_{31}' & D_{32}' & \cdots & D_{3,T-2}' & \cdots & D_{T1}' 
\end{bmatrix}',
\]  

we can rewrite (4) as:

\[
Y_i = X_i \beta + U_i, \quad i = 1, \ldots, N,
\]  

where

\[
\begin{align*}
Y_i &= D y_i', \\
X_i &= D x_i', \\
U_i &= D u_i',
\end{align*}
\]
may be considered a system of $S$ equations with a common slope coefficient $\beta$ and with $N$ observations of each equation. When different pairs of periods are involved, we always assume that the pairs of periods $(t, s)$ are ordered in the same way as in (5).

2.2 Instrumental variables and the orthogonality conditions

The structural parameter $\beta$ in the model can be estimated using lagged and leaded $x$'s as instrumental variables if the measurement errors in the $x$'s are non-autocorrelated, as was shown by Griliches and Hausman (1986). Specifically, if we consider (4) for two given periods $t$ and $s$, valid instruments are $x_{i\tau}$ for $\tau \neq t, s$ ($\tau = 1, ..., T$). Consequently, we have different instruments for each of the $S$ equations in (6), and consequently a GMM-procedure is called for to estimate $\beta$ from the whole system of equations jointly.

The idea we follow is, for one pair of periods $(t, s)$, to use as IV's all the $T - 2$ elements of $x_i$ for the $T - 2$ periods which are not used in the construction of the differenced variables in (4). A similar general idea has been followed in the literature on dynamic panel data models, and by Griliches and Hausman (1986), for panel data models with errors in variables. We define the $S = \frac{1}{2}T(T - 1)$ selection matrices

$$P_{ts} = \begin{bmatrix} ((T - 2) \times T) & \text{matrix} \\ \text{obtained by deleting} & \end{bmatrix}, \quad t, s = 1, \ldots, T; \quad t > s,$$

and

$$z_{its} = x_i P_{ts}' = \begin{bmatrix} (1 \times (T - 2)) \text{vector} \\ \text{obtained by deleting} & \end{bmatrix}, \quad i = 1, \ldots, N; \quad t, s = 1, \ldots, T; \quad t > s.$$

To carry out GMM estimation of $\beta$ based on the complete system of $S$ equations, we must stack the instruments as follows. Define the $(S \times S(T - 2))$ IV matrix for individual $i$

$$Z_i = \begin{bmatrix} z_{i12} & 0 & \cdots & 0 \\ 0 & z_{i32} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & z_{iT1} \end{bmatrix} = \begin{bmatrix} x_i P_{21}' & 0 & \cdots & 0 \\ 0 & x_i P_{32}' & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & x_i P_{T1}' \end{bmatrix},$$

which can be written as

$$Z_i = (I_S \otimes x_i) P',$$

1This requires, of course, that $\lim_{N \to \infty} \{1/N \sum_{i=1}^N x_{it} (x_{it} - u_{it})\} \neq 0$ in addition to $\lim_{N \to \infty} \{(1/N) \sum_{i=1}^N x_{it} (u_{it} - u_{it})\} \neq 0$.
2See Baltagi (1995, chapter 8) for a survey.
3See also Biørn (1996, section 10.2.3).
where \( P \) is the \( (S(T - 2) \times ST) \) matrix, containing only zeros and ones,

\[
P = \begin{bmatrix}
P_{21} & 0 & \cdots & 0 \\
0 & P_{32} & \cdots & 0 \\
\vdots & \vdots & \ddots & \vdots \\
0 & 0 & \cdots & P_{T1}
\end{bmatrix}.
\] (10)

Premultiplying (6) by \( Z_i' \), we get

\[
Z_i' Y_i = Z_i' X_i \beta + Z_i' U_i, \quad i = 1, \ldots, N.
\] (11)

The orthogonality conditions which are \textit{a priori} relevant to GMM estimation of \( \beta \) based on (11) can be stated:

\[
E(Z_i' U_i) = 0, \quad i = 1, \ldots, N.
\] (12)

where \( 0 \) is a \( (S(T - 2) \times 1) \) vector of zeros, i.e. (12) represents \( S(T - 2) \) scalar orthogonality conditions.

2.3 The GMM-estimator and the generalized inverse

Some of the elements in the (vector) equation (12) are redundant, as they do not contain additional information. This can be seen by noticing that for say \( l < s \) (or \( l > t \)) we have that

\[
E[x_{il}(u_{it} - u_{is})] = E[x_{il}(u_{i,s-1} - u_{is})] + \cdots + E[x_{il}(u_{it,s-1})].
\] (13)

That is, the orthogonality condition \( E[x_{il}(u_{it} - u_{is})] = 0 \) can be constructed as a simple sum of the orthogonality conditions on the right hand side of (13). As a consequence, the variance-covariance matrix needed for GMM estimation on the basis of (11) and (12) does not have full rank. We will elaborate on this point below.

Denote the (reduced rank) variance-covariance matrix associated with (12) by \( \Omega_i \), i.e.

\[
\Omega_i = E(Z_i' U_i U_i' Z_i).
\] (14)

White (1986) has considered efficient estimation based on orthogonality conditions such as (12), in the general case where \( \Omega_i \) might not have full rank. Using Theorem 3.2 in White, we find that the asymptotically efficient GMM estimator \( \hat{\beta} \), based on the orthogonality conditions (12) for \( i = 1, \ldots, N \), can be written:

\[
\hat{\beta} = \left( \sum_i X_i' Z_i \right) \left( \sum_i \Omega_i \right)^+ \left( \sum_i Z_i' X_i \right)^{-1} \left( \sum_i X_i' Z_i \right) \left( \sum_i \Omega_i \right)^+ \left( \sum_i Z_i' Y_i \right),
\] (15)

where the sums cover all individuals and \( (\sum_i \Omega_i)^+ \) is the generalized inverse of \( \sum_i \Omega_i \).
3 Identification of essential orthogonality conditions

3.1 The essential orthogonality conditions

Even though the estimator (15) is efficient in a statistical sense, it is not very efficient in a computational sense. Computationally it is more efficient to consider only the essential orthogonality conditions, which we will now identify. In the case with non-autocorrelated errors uit, we show formally below that all orthogonality conditions can be constructed as simple sums of (i) all admissible orthogonality conditions based on one-period differences and (ii) a specific subset of the admissible orthogonality conditions based on two-period differences. All the other orthogonality conditions can be ignored.

The orthogonality conditions (12) are based on expressions of the form $E[x_{it}(u_{it} - u_{is})] = 0$ where $l \neq t, s$. However, all these orthogonality conditions can be constructed as simple sums of

(i) $E[x_{it}(u_{it} - u_{i,t-1})] = 0, l \neq t, t - 1,$

(ii) $E[x_{it}(u_{i,t+1} - u_{i,t-1})] = 0.$

This is easily seen as follows: First, when $l \notin [s, s + 1, \ldots, t - 1, t]$, any orthogonality condition $E[x_{it}(u_{it} - u_{is})] = 0$ can be constructed on the basis of expressions as in (i), using the identity

$$E[x_{it}(u_{it} - u_{is})] = \sum_{\tau=s+1}^{t} E[x_{it}(u_{it} - u_{i,\tau-1})].$$

Second, when $l \in [s + 1, s + 2, \ldots, t - 2, t - 1]$, expressions both of the forms (i) and (ii) must be combined, using

$$E[x_{it}(u_{it} - u_{is})] = \sum_{\tau=t+2}^{t} E[x_{it}(u_{it} - u_{i,\tau-1})] + E[x_{it}(u_{i,\tau+1} - u_{i,t-1})] + \sum_{\tau=s+1}^{t-1} E[x_{it}(u_{i,\tau} - u_{i,t-1})],$$

since $E[x_{it}(u_{it} - u_{i,t-1})] = 0$ and $E[x_{it}(u_{i,t+1} - u_{it})] = 0$ are inadmissible whereas $E[x_{it}(u_{i,\tau+1} - u_{i,t-1})] = 0$ is admissible. Hence, any orthogonality condition of the form $E[x_{it}(u_{it} - u_{is})] = 0$ where $l \neq (t, s)$, can be constructed from the two kinds of orthogonality conditions (i) and (ii).

It follows that the number of essential orthogonality conditions is $^4 T(T - 2)$, while the total number of orthogonality conditions is $T(T - 1)(T - 2)/2$. Hence, only a fraction $2/(T - 1)$ of the complete set of orthogonality conditions are essential. E.g., for $T = 9$, this fraction is one fourth.

We have shown that only orthogonality conditions based on the one-period and a few two-period differences are essential. We refer to the other orthogonality conditions as redundant. With autocorrelated noise, higher order differences replace the two-period differences; see Bier and Klette (1997).

---

4Among these, $(T - 1)(T - 2)$ are based on one-period differences, and $(T - 2)$ on two-period differences.
3.2 The equivalence of the two GMM-estimators

It follows from section 3.1 that there exists a matrix $H$ of zeros and ones such that

$$[Z'u_i]_R = H [Z'u_i]_E,$$  \hspace{1cm} (16)

where subscripts $R$ and $E$ denote the elements associated with the sets of redundant and essential orthogonality conditions respectively. More generally, we have

$$\begin{bmatrix} (Z'_iY)_R & (Z'_iX_i)_R & (Z'_iU_i)_R \end{bmatrix} = H \begin{bmatrix} (Z'_iY)_E & (Z'_iX_i)_E & (Z'_iU_i)_E \end{bmatrix}.$$  \hspace{1cm} (17)

Define

$$K = \begin{bmatrix} I_{T(T-2)} \\ H \end{bmatrix},$$

where $I_{T(T-2)}$ is the identity matrix of order $T(T-2)$. Stacking $Z'_iU_i$ as follows

$$Z'_iU_i = \begin{bmatrix} (Z'_iU_i)_E \\ (Z'_iU_i)_R \end{bmatrix},$$

and using (16), we have

$$\Omega_i = E [Z'u_iU'_iZ_i] = KE [(Z'_iU_i)_E(U'_iZ_i)_E] K' = K \Omega_{E,i} K'.$$  \hspace{1cm} (18)

where $\Omega_{E,i}$ is defined by the last equality. Using (17) and (18), the GMM estimator in (15) can be rewritten

$$\hat{\beta} = \left\{ \left( \sum_i X'_iZ_i \right)_E K' \left( \sum_i \Omega_{E,i} \right) K' + \left( \sum_i Z'_iX_i \right)_E K' \left( \sum_i \Omega_{E,i} \right) K' \right\}^{-1} \times \left( \sum_i X'_iZ_i \right)_E K' \left( \sum_i \Omega_{E,i} \right) K' \left( \sum_i Z'_iY_i \right)_E.$$  \hspace{1cm} (19)

The definition of the generalized inverse implies

$$(K\Omega_E K')(K\Omega_E K')^+(K\Omega_E K') = (K\Omega_E K'),$$

where we have used $\Omega_E$ as a short-hand notation for $\sum_i \Omega_{E,i}$. Pre- and postmultiplying this equation by $\Omega_E^{-1}(K'K)^{-1}K'$ and $K(K'K)^{-1}\Omega_E^{-1}$ respectively, we find that

$$K'(K\Omega_E K')^+K = \Omega_E^{-1}.$$  \hspace{1cm} (20)

Inserting (20) into (19), we find that

$$\hat{\beta} = \left[ \left( \sum_i X'_iZ_i \right)_E \left( \sum_i \Omega_{E,i} \right)^{-1} \left( \sum_i Z'_iX_i \right)_E \right]^{-1} \left( \sum_i X'_iZ_i \right)_E \left( \sum_i \Omega_{E,i} \right)^{-1} \left( \sum_i Z'_iY_i \right)_E.$$
Hence, the GMM-estimator based on the complete set of orthogonality conditions is equivalent to the GMM-estimator based on only the essential orthogonality conditions.

As remarked above, the essential set of orthogonality conditions constitutes a fraction $2/(T-1)$ of the complete set of orthogonality conditions. Exploiting only the former can reduce the computational burden considerably, in particular related to inverting the variance-covariance matrix $(\sum_i \Omega_i)$. With a moderately long panel such as $T = 9$, using the complete set of orthogonality conditions, this matrix has dimension $(252 \times 252)$, which is reduced to $(63 \times 63)$ when using only the essential orthogonality conditions.

Here we should point out that with more than one regressor, say $G$ regressors, the dimensionality of the IV matrix $Z_i$ and hence the variance-covariance matrix $\Omega_i$ will grow in proportion to $G$, while the fraction $2/(T-1)$ of the complete set of orthogonality conditions that is essential remains the same. Hence, with $G = 3$ and $T = 9$, $\Omega_i$ will have dimension $(756 \times 756)$, while $\Omega_{E,i}$ has dimension $(189 \times 189)$. Whether it is a good idea in practice to use all essential orthogonality conditions (as defined above) with $T = 9$ and $G = 3$ depends on the sample size and the stochastic processes for the regressors. However, these are issues discussed elsewhere under the labels “overfitting” and “weak instruments”; see e.g. Davidson and MacKinnon (1993, chs. 7 and 17) and Staiger and Stock (1994, 1996).

In Biørn and Klette (1997) we elaborate on the analysis above and show, in particular, how to identify the essential orthogonality conditions in cases with autocorrelated measurement errors.

4 Final remarks

This note has examined the orthogonality conditions relevant for GMM estimation of differenced equations from panel data with errors-in-variables, using variables in levels as IVs for differenced variables. We have shown that with non-autocorrelated measurement errors, only a small fraction of the potential orthogonality conditions are essential, namely those based on one-period and a few two-period differences. When only predetermined variables are valid instruments as in autoregressive panel data models, even the two-period differences are inadmissible, and one is left only with the orthogonality conditions based on one-period differences.
References


Recent publications in the series Discussion Papers


52 R. Aaberge and T. Wenne (1990): Non-Stationary Inflow and Duration of Unemployment.


56 N.M. Stulen (1990): Is there a NAIRU in Norway?


66 E. Bowitz and E. Storm (1991): Will Restrictive Demand Policy Improve Public Sector Balance?


S. Kverndokk (1993): Coalitions and Side Payments in International CO2 Treaties


A. Brendemoen and H. Vennemo (1993): The Marginal Cost of Funds in the Presence of External Effects


A. Aaheim and K. Nyborg (1993): "Green National Product": Good Intentions, Poor Device?

K.H. Alfsen, H. Birkelund and M. Aaserud (1993): Testing for Purchasing Power Parity and Interest Rate Parities on Norwegian Data

A. Brendemoen and H. Vennemo (1993): The Marginal Cost of Funds in the Presence of External Effects


A. Brendemoen and H. Vennemo (1993): The Marginal Cost of Funds in the Presence of External Effects

K.H. Alfsen, H. Birkelund and M. Aaserud (1993): Testing for Purchasing Power Parity and Interest Rate Parities on Norwegian Data


K.A. Magnussen (1994): Precautionary Saving and Old-Age Pensions


K.A. Brekke and P. Bering (1994): The Volatility of Oil Wealth under Uncertainty about Parameter Values

M.J. Simpson (1994): Foreign Control and Norwegian Manufacturing Performance

Y. Willansen and T.J. Klette (1994): Correlated Measurement Errors, Bound on Parameters, and a Model of Producer Behavior

D. Wetterwald (1994): Car ownership and private car use. A microeconomic analysis based on Norwegian data


K. Mohn (1994): On Equity and Public Pricing in Developing Countries


E. Bjørn and T.J. Klette (1994): Errors in Variables and Panel Data: The Labour Demand Response to Permanent Changes in Output

I. Svendsen (1994): Do Norwegian Firms Form Extrapositional Expectations?

T.J. Klette and Z. Griliches (1994): The Inconsistency of Common Scale Estimators when Output Prices are Unobserved and Endogenous


L. A. Grünfeld (1994): Monetary Aspects of Business Cycles in Norway: An Exploratory Study Based on Historical Data


K.A. Brekke and H.A. Grønningmyhr (1994): Adjusting NNP for instrumental or defensive expenditures. An analytical approach

T.O. Thoresen (1995): Distributional and Behavioural Effects of Child Care Subsidies


