UNIVERSITY OF OXFORD

APPLIED ECONOMICS DISCUSSION PAPER

No. 57

AN ALTERNATIVE TRANSFORMATION FOR FIXED EFFECTS MODELS WITH PREDETERMINED VARIABLES

BY

MANUEL ARELLANO

NOVEMBER 1988

INSTITUTE OF ECONOMICS AND STATISTICS
MANOR ROAD, OXFORD.

AN ALTERNATIVE TRANSFORMATION FOR FIXED EFFECTS MODELS WITH PREDETERMINED VARIABLES

Manuel Arellano
Institute of Economics and Statistics, Oxford.

ABSTRACT

The purpose of this note is to propose an alternative transformation of a fixed effects model which (i) eliminates the individual effects, (ii) preserves the orthogonality among the transformed errors, and (iii) it is very useful in order to obtain in a natural way optimal IV or GMM estimators for models with predetermined variables.

1. Introduction

The within–groups or covariance estimator remains the most popular technique for the estimation of linear regression models from panel data: fixed effects are eliminated after individual means are substracted from the observations, and OLS is applied to the transformed equation. An attractive feature of the resulting estimates is that they are consistent for short panels (small T) of large cross–sections (large N) regardless of the correlation between the regressors and the individual effects. However, as it has been repeatedly pointed out in the literature (e.g. Nickell (1981), Chamberlain (1984)), this is only true if the regressors are strictly exogenous. When the explanatory variables are only predetermined (that is, uncorrelated with present and future shocks but correlated to past ones) within–groups estimates are inconsistent as $N \to \infty$ for fixed T. These inconsistencies are of order 1/T if the random errors are white noise but they still are a serious concern for the type of samples usually available. Moreover, since the error term for any one period in the model in deviations from time means depends on the errors for all other periods, it is difficult to find valid moment restrictions for this representation of the model. This has prompted the development of instrumental variables (IV) or generalized method of moments (GMM) estimators based on first or longer difference transformations (e.g. Anderson and Hsiao (1981 and 1982), Griliches and Hausman (1986), Arellano and Bond (1988)). Regrettably, differencing the model introduces serial correlation and one may suspect that often this will lead to inefficient estimators.

The purpose of this note is to propose an alternative transformation of the model which (i) eliminates the individual effects, (ii) preserves the orthogonality among the transformed errors – if the original random errors are independently and identically distributed (iid), the transformed errors are also iid, and (iii) it is very useful in order to obtain in a natural way optimal IV or GMM estimators for models with predetermined variables.

2. The Transformation

The model is given by

$$y_{it} = x_{it}\beta + u_{it}$$

$$u_{it} = \eta_i + v_{it} \qquad (t=1,...,T;\ i=1,...,N)$$

where x_{it} is a 1xk vector of explanatory variables, β is the kx1 vector of coefficients to be estimated, η_i is an unobservable individual effect and v_{it} is a random error assumed to be iid across individuals and time with zero mean and variance σ^2. The explanatory variables may well be correlated with the effects so that $E\left[x_{it}\eta_i\right] \neq 0$. Alternatively we can write the system of T equations for individual i as

$$y_i = X_i\beta + u_i$$

where $y_i = \left[y_{i1}...y_{iT}\right]'$, $X_i = \left[x'_{i1}...x'_{iT}\right]'$ and $u_i = \left[u_{i1}...u_{iT}\right]'$. The within-groups operator is the TxT idempotent matrix Q of rank (T-1) given by

$$Q = I_T - \frac{1}{T}\iota\iota'$$

where I_T is an identity matrix of order T and ι is a Tx1 vector of ones. Thus the transformed errors $\bar{u}_i = Qu_i$ are of the form

$$\bar{u}_{it} = u_{it} - \frac{1}{T}\left[u_{i1}+...+u_{iT}\right] \qquad (t=1,...,T) .$$

Here we propose to remove individual effects by mean of the following transformation

$$u^+_{it} = u_{it} - \frac{1}{(T-t)}\left[u_{i(t+1)}+...+u_{iT}\right] \quad (t=1,...,T-1) ,$$

that is, to each of the first (T–1) observations we substract the mean of the remaining future observations available in the sample. The operator that produces this transformation is a (T–1)xT matrix A^+ of the form

$$A^+ = \begin{bmatrix} 1 & -(T-1)^{-1} & -(T-1)^{-1} & \dots & -(T-1)^{-1} & -(T-1)^{-1} & -(T-1)^{-1} \\ 0 & 1 & -(T-2)^{-1} & \dots & -(T-2)^{-1} & -(T-2)^{-1} & -(T-2)^{-1} \\ \vdots & & & & & & \\ 0 & 0 & 0 & \dots & 1 & -1/2 & -1/2 \\ 0 & 0 & 0 & \dots & 0 & 1 & -1 \end{bmatrix}$$

It is straightforward to verify that $E\left[u^+_{it} u^+_{is}\right] = 0$ for any $t \neq s$. For example, if $t < s$ we have

$$\begin{aligned} E\left[u^+_{it}u^+_{is}\right] &= -\frac{1}{(T-t)}E\left[v_{is}\left[v_{i(t+1)}+\dots+v_{iT}\right]\right] \\ &+ \frac{1}{(T-t)(T-s)}E\left[\left[v_{i(t+1)}+\dots+v_{iT}\right]\left[v_{i(s+1)}+\dots+v_{iT}\right]\right] \\ &= -\frac{\sigma^2}{(T-t)} + \frac{\sigma^2}{(T-t)} = 0 \ . \end{aligned}$$

On the other hand,

$$\mathrm{Var}\left[u^+_{it}\right] = \left[\frac{T-t+1}{T-t}\right]\sigma^2$$

which suggests weighting the u^+_{it} to equalize the variances, thus obtaining

$$u^*_{it} = \left[\frac{T-t}{T-t+1}\right]^{1/2}\left[u_{it} - \frac{1}{(T-t)}\left[u_{i(t+1)}+\dots+u_{iT}\right]\right] \quad (t=1,\dots,T-1)\ .$$

Letting $u^*_i = \left[u^*_{i1}\dots u^*_{i(T-1)}\right]'$, we can write $u^*_i = Au_i$ where A is the (T–1)xT matrix

$$A = \mathrm{diag}\left[\frac{T-1}{T}, \frac{T-2}{T-1}, \ldots, \frac{1}{2}\right]^{1/2} . A^{+}$$

Some properties of this transformation which can be easily verified by direct multiplication are $A\iota = 0$, $AA' = I_{(T-1)}$ and $A'A = Q$. Differencing with respect to the mean of past observations would produce an alternative transformation with similar properties. However since predetermined variables depend on past but not future shocks, the A–transformation turns out to be more useful in order to handle this situation.

3. Some Estimators

Let $y_i^{*} = Ay_i$ and $X_i^{*} = AX_i$. The fact that $A'A = Q$ implies that the OLS regression of y_i^{*} on X_i^{*} will yield the within–groups estimator:

$$\hat{\beta}_{WG} = \left[\sum_{i=1}^{N} X_i^{*\prime} X_i^{*}\right]^{-1} \sum_{i=1}^{N} X_i^{*\prime} y_i^{*} = \left[\sum_i \tilde{X}_i' \tilde{X}_i\right]^{-1} \sum_i \tilde{X}_i' \tilde{y}_i$$

where $\tilde{y}_i = Qy_i$ and $\tilde{X}_i = QX_i$. However $\hat{\beta}_{WG}$ is only consistent when x_{it} is strictly exogenous with respect to v_{it}. That is, when $E\left[x_{it}v_{is}\right] = 0$ for all t and s.

When x_{it} is only predetermined, in the sense that

$$E\left[x_{it}v_{is}\right] \begin{cases} = 0 & \text{for } s \geq t \\ \neq 0 & \text{for } s < t \end{cases}$$

it is still possible to use a different subset of the x_{it}, $t=1,\ldots,T$ as instruments for the transformed equations corresponding to each period. Specifically, since u_{i1}^{*} depends on $v_{i1},\ldots,v_{iT}$ only x_{i1} is a valid instrument in the equation for the first period. For the second period u_{i2}^{*} depends on $v_{i2},\ldots,v_{iT}$ and thus x_{i1} and x_{i2} are both valid instruments, etc. The complete matrix of instruments for the A–transformed system of (T–1) equations is

$$Z_i = \begin{bmatrix} x_{i1} & 0 & 0 & \cdots & 0 & 0 & \cdots & 0 \\ 0 & x_{i1} & x_{i2} & \cdots & 0 & 0 & \cdots & 0 \\ \vdots & & & & & & & \vdots \\ 0 & 0 & 0 & \cdots & x_{i1} & x_{i2} & \cdots & x_{i(T-1)} \end{bmatrix}$$

and the corresponding IV estimator is

$$\hat{\beta}_{IV} = \left[\sum_i X_i^{*\prime} Z_i \left[\sum_i Z_i' Z_i\right]^{-1} \sum_i Z_i' X_i^{*}\right]^{-1} \sum_i X_i^{*\prime} Z_i \left[\sum_i Z_i' Z_i\right]^{-1} \sum_i Z_i' y_i^{*} \ .$$

Since the u^{*}_{it} are iid errors, the standard IV theory ensures the asymptotic efficiency of $\hat{\beta}_{IV}$ within its class. If, however, u_i is heteroskedastic of unknown form, efficiency gains of the type considered by White (1982) are possible: a second step estimator would replace the norm $\left[\Sigma_i \ Z_i' Z_i\right]$ in the formula for $\hat{\beta}_{IV}$ by $\left[\Sigma_i \ Z_i' \hat{u}^{*}_i \hat{u}^{*\prime}_i Z_i\right]$ where the $\hat{u}^{*}_i$ are first step residuals.

References

Anderson, T.W. and C. Hsiao (1981): "Estimation of Dynamic Models with Error Components", Journal of the American Statistical Association, 76, 598–606.

Anderson, T.W. and C. Hsiao (1982): "Formulation and Estimation of Dynamic Models Using Panel Data", Journal of Econometrics, 18, 47–82.

Arellano, M. and S. Bond (1988): "Some Tests of Specification for Panel Data: Monte Carlo Evidence and an Application to Employment Equations", Institute for Fiscal Studies Working Paper 88/4, London.

Chamberlain, G. (1984): "Panel Data", in Z. Griliches and M.D. Intriligator (eds.), Handbook of Econometrics, Volume II, Elsevier Science Publishers.

Griliches, Z. and J.A. Hausman (1986): "Errors in Variables in Panel Data", Journal of Econometrics, 31, 93–118.

Nickell, S.J. (1981): "Biases in Dynamic Models with Fixed Effects, Econometrica, 49, 1417–1426.

White, H. (1982): "Instrumental Variables Regression with Independent Observations", Econometrica, 50, 483–499.

OXFORD APPLIED ECONOMICS DISCUSSION PAPER SERIES

No. 1 Empirical Modelling in Dynamic Econometrics
David Hendry

No. 2 How Rational are the Expectations made by Firms
Colin Mayer and *Matthias Mors*

No. 3 Developments in the Study of Cointegrated Economic Variables
Clive Granger

No. 4 The Short Run Behaviour of Labour Supply
Stephen Nickell

No. 5 Disaggregative versus Aggregative Wage Equations
Wilfred Beckerman and Tim Jenkinson

No. 6 Consequences of a Commodity Boom in a Controlled Economy: Accumulation and Redistribution in Kenya, 1975–83
D.L. Bevan, P. Collier and J.W. Gunning

No. 7 The Role and Occupation in the Determination of Wages
J. de Beyer and J.B. Knight

No. 8 The Development of British Econometrics, 1945–85
Christopher L. Gilbert

No. 9 Post–war Fertility and Female Labour Force Participation Rates
Alison Sprague

No.10 Econometric Evaluation of Linear Macro–Economic Models
Yock Y. Chong and David F. Hendry

No.11 Estimating Contaminated Limited Dependent Variable Models
Manuel Arellano and Olympia Bover

No.12 Liquidity Constraints and Aggregation in the Consumption Function under Uncertainty
John Muellbauer and Olympia Bover

No.13 Some Dynamic Life–Cycle Models of Labour Supply Estimated From Panel Data
Olympia Bover

No.14 Rent Seeking and Bank Lending to Developing Countries: The Implications of Derivative Business
Harold Cataquet

No.15 Why is Wage Inflation in Britain So High ?
Stephen Nickell

No.16 Employment Adjustment in UK Manufacturing
Stephen Nickell

No.17 The Estimation and Hypothesis Testing of Dynamic Models from Panel Data
Manuel Arellano

No.18 Intertemporal Employment and Pricing Decisions Rules in UK Manufacturing
Juan J. Dolado

No.19 Indexation, Contract Duration and Non–Contingent Wage Adjustment: A Study Based on Contract Data,
Louis N. Christofides

No.20 Monte Carlo Experimentation Using PC–NAIVE,
David F. Hendry and Adrian J. Neale

No.21 Portfolio Theory and the Financial Sector of the London Business School (LBS) Model
A.S. Courakis

No.22 On Compatible Trade, Monetary and Exchange Rate Policies
Paul Collier

No.23 Extreme Bounds, Vector Autoregressions and Dynamic Structural Models: Alternative Approaches to Econometric Methodology
Christopher L. Gilbert

No.24 Colliery Results and Closures after the 1984–85 Coal Dispute
Andrew Glyn

No.25 Assertion Without Empirical Basis: An Econometric Appraisal of Monetary Trends...in the United Kingdom by Milton Freidman and Anna J. Schwartz
David F. Hendry and Neil R. Ericsson

No.26 Labour Supply and Hours Constraints
Manuel Arellano and Costas Meghir

No.27 Food Subsidies and Poverty Alleviation
Timothy Besley and Ravi Kanbur

No.28 Mergers, Takeovers and the Enforcement of Profit Maximization
Dieter Helm

No.29 Wages, Prices, Employment and Output in UK Industry
Stephen Nickell and Paul Kong

No.30 The Impact of Exchange Rates and Developing Country Debt on Commodity Prices
Christopher L. Gilbert

No.31 Turnover in UK Manufacturing
Simon M. Burgess and Stephen Nickell

No.32 Intertemporal Rules with Variable Speed of Adjustment: an UK Application to UK Manufacturing Employment
Simon Burgess and Juan Dolado

No.33 Labour Skills and Human Capital in the Explanation of Trade Patterns
Anthony S. Courakis

No.34 The Management of Developing Country Commodity Risks: A New Role for Public Policy
Christopher L. Gilbert and Andrew Powell

No.35 The Erosion of Apartheid in the South African Labour Market: Measures and Mechanisms
J.B. Knight and M.D. McGrath

No.36 A Case for Exchange Controls
Andrew Glyn

No.37 Efficient and Inefficient Employment Outcomes: A Study Based on Canadian Contract Data
Louis N. Christofides and Andrew J. Oswald

No.38 An Empirical Study of the Interrelationship between Employment, Price and Inventory Decisions in UK Manufacturing
Juan J. Dolado

No.39 Cointegration: A Survey of Recent Developments
Juan J. Dolado and Tim Jenkinson

No.40 The Diversity of Unemployment Experience since 1973
Bob Rowthorn and Andrew Glyn

No.41 Real Wages and Unemployment in Britain during the 1930s
N.H. Dimsdale, S.J. Nickell and N. Horsewood

No.42 Employment and Structural Change in Britain: Trends and Policy Options
Christine Greenhalgh

No.43 Work Experience, Earnings and Participation: Evidence from the Women and Employment Survey
Alison Sprague

No.44 Employment and Structural Change in Britain – Quantitative Evidence and Policy Simulations
Christine Greenhalgh, Mary Gregory and Amit Ray

No.45 Tests of the Life Cycle–Permanent Income Hypothesis in the Presence of Random Walks: Asymptotic Theory and Small Sample Interpretations
Anindya Banerjee and Juan Dolado

No.46 Means Testing versus University Provision in Poverty Alleviation Programmes
Tim Besley

No.47 Firms as Portfolios: A Study of Unquoted UK Companies
Donald A. Hay and Helen Louri

No.48 External Debt Behaviour Reconsidered
Nawal Kamel

No.49 An Investigation into the Power of Insiders in Wage Determination
Stephen Nickell and Paul Kong

No.50 Criteria for National Solvency
Nawal Kamel

No.51 Pricing and Employment in the UK an Open Economy Model
V. Bhaskar

No.52 Fiscal Policy and Inflationary Expectations: The Hungarian Tax Pengo Experiment of 1946
P.L. Siklos

No.53 Dynamic Specification with the General Error–Correction Form
Anindya Banerjee, John W. Galbraith, Juan Dolado

No.54 Does GNP have a Unit Root ? A Detailed Examination using Recursive Methods
Juan Dolado, Anindya Banerjee, John W. Galbraith

No.55 Some Tests of Specification for Panel Data: Monte Carlo Evidence and an Application to Employment Equations
Manuel Arellano and Stephen Bond

No.56 Testing for Autocorrelation in Dynamic Random Effects Models
Manuel Arellano

No.57 An Alternative Transformation for Fixed Effects Models with Predetermined Variables
Manuel Arellano

No.58 Exchange Rate Pass-Through When Market Share Matters
Kenneth A. Froot and Paul Klemperer

No.59 An Empirical Study of Unemployment and the Number of Children in Care
Alan Carruth and Andrew Oswald

No.60 Did Policy Activism Work ?
Andrea Boltho

No.61 Price Inertia and Policy Ineffectiveness in the United States, 1890–1984: A Reappraisal
P.L. Siklos

No.62 The Pattern of Female Mortality in Iran and Some of its Causes
Feridoon Koohi–Kamali

Single Copies which are not out of print are available for a nominal charge of £1.