OXFORD BULLETIN OF ECONOMICS AND STATISTICS, 49, 4 (1987)
0305-9049 $3.00

PRACTITIONERS' CORNER

Computing Robust Standard Errors for Within-groups Estimators*

M. Arellano

I. THE MODEL AND THE ESTIMATOR

The purpose of this note is to explain how to use standard packages to calculate heteroskedasticity and serial correlation consistent standard errors for within-groups estimators of a linear regression model from panel data. We are concerned to discuss the model

$$y_{it} = x_{it}'\beta + e_i + u_{it} \qquad (t = 1, \ldots, T; i = 1, \ldots, N) \tag{1}$$

where x_{it} is a $k \times 1$ vector of exogenous variables such that

$$E(u_{it} \mid x_{i1}, \ldots, x_{iT}, e_i) = 0$$

and e_i is an unobservable permanent effect potentially correlated with x_{it}. The u_{it} are assumed to be independently distributed across individuals but no restrictions are placed on the form of the autocovariances for a given individual:

$$E(u_{it} \cdot u_{is} \mid x_{i1}, \ldots, x_{iT}, e_i) = \omega_{tsi}$$

thus allowing for heteroskedasticity and serial correlation of arbitrary form. Alternatively model (1) can be written as

$$\underset{T\times 1}{y_i} = \underset{T\times k}{X_i}\underset{k\times 1}{\beta} + e_i l + \underset{T\times 1}{u_i} \qquad (i = 1, \ldots, N) \tag{2}$$

$$u_i \sim \text{i.d.}\ (0, \Omega_i)$$

where l is a $T \times 1$ vector of ones. We assume that T is small and N is large, as is often the case with household or company data, thus considering asymptotic results as $N \to \infty$ for fixed T.

Transforming the variables in (2) into deviations from time means eliminates the e_i's. Let $y_i^+ = Qy_i$, $X_i^+ = QX_i$ and $u_i^+ = Qu_i$, where Q is the deviations from time means operator:

$$Q = I_T - ll'/T$$

the within-groups estimator of β is just the OLS regression of y_i^+ on X_i^+:

* I am grateful to David Hendry for helpful comments.

$$\hat{\beta}_{WG} = (X^{+\prime}X^{+})^{-1}X^{+\prime}y^{+} \tag{3}$$

with $X^{+\prime} = (X_1^{+\prime} \ldots X_N^{+\prime})$ and $y^{+\prime} = (y_1^{+\prime} \ldots y_N^{+\prime})$.

It is well known that under appropriate regularity conditions

$$\sqrt{N}(\hat{\beta}_{WG} - \beta) \underset{a}{\sim} N(0, M^{-1}VM^{-1})$$

where

$$M = p \lim_{N \to \infty} N^{-1}(X^{+\prime}X^{+})$$

and

$$V = p \lim_{N \to \infty} N^{-1} \sum_{i=1}^{N} (X_i^{+\prime}\Omega_i X_i^{+})$$

II. ESTIMATING THE avm OF $\hat{\beta}_{WG}$

A White estimator of V is of the form (see White, 1984, p. 136):

$$\hat{V} = \frac{1}{N} \sum_{i=1}^{N} X_i^{+\prime}\hat{u}_i^{+}\hat{u}_i^{+\prime}X_i^{+}$$

where $\hat{u}_i^{+} = y_i^{+} - X_i^{+}\hat{\beta}_{WG}$. Note that $X_i^{+\prime}\Omega_i X_i^{+} = X_i^{+\prime}\Omega_i^{+}X_i^{+}$ with $\Omega_i^{+} = Q\Omega_i Q$, which suggests the consistency of $\hat{V}$ for V. In effect, this can be proved as a special case of the Theorem 6.3 given by White (1984). Therefore a robust estimator of $\text{avm}(\hat{\beta}_{WG})$ can be obtained as

$$\text{av}\hat{\text{m}}(\hat{\beta}_{WG}) = (X^{+\prime}X^{+})^{-1}\left(\sum_{i=1}^{N} X_i^{+\prime}\hat{u}_i^{+}\hat{u}_i^{+\prime}X_i^{+}\right)(X^{+\prime}X^{+})^{-1} \tag{4}$$

This covers the most general case. Two alternative estimators that have also been used in practice are the following:

$$\text{av}\tilde{\text{m}}(\hat{\beta}_{WG}) = (X^{+\prime}X^{+})^{-1}\left(\sum_{i=1}^{N} X_i^{+\prime}\hat{\Omega}^{+}X_i^{+}\right)(X^{+\prime}X^{+})^{-1} \tag{5}$$

where

$$\hat{\Omega}^{+} = N^{-1} \sum_{i=1}^{N} \hat{u}_i^{+}\hat{u}_i^{+\prime}$$

This estimator has been discussed by Kiefer (1980) and will produce consistent standard errors if x_{it} and u_{it} are fourth order independent and in particular if $\Omega_i = \Omega$ for all i, that is, in the case of homoskedastic arbitrary intertemporal covariance. Finally, if the classical assumption $\Omega_i = \sigma^2 I_T$ holds, $\text{avm}(\hat{\beta}_{WG})$ can be estimated by means of

$$\text{av}\bar{\text{m}}(\hat{\beta}_{WG}) = \hat{\sigma}^2(X^{+\prime}X^{+})^{-1} \tag{6}$$

where

$$\hat{\sigma}^2 = \frac{1}{N(T-1)} \sum_{i=1}^{N} \hat{u}_i^{+\prime} \hat{u}_i^{+}$$

Note, however, that if $N(T-1)$ is replaced by NT or $(NT-k)$ in the definition of σ^2, we obtain an estimator of σ^2 that is not consistent for fixed T (e.g. see Hausman and Taylor, 1981).

III. COMPUTATIONAL COMMENTS

Expression (4) can be easily calculated from a standard package provided the option of White standard errors is available for a multivariate regression with linear restrictions (for example in TSP: option 'ROBUST' with the command 'LSQ'). The idea is that instead of calculating $\hat{\beta}_{WG}$ as the 'stacked regression' of y^+ on X^+ (which would produce standard errors that require degree of freedom corrections even under classical assumptions), we estimate β by least squares in the system of T equations with cross-equation linear restrictions

$$\begin{gathered} y_{i1}^{+} = x_{i1}^{+\prime}\beta + u_{i1}^{+} \\ \vdots \\ y_{iT}^{+} = x_{iT}^{+\prime}\beta + u_{iT}^{+} \end{gathered}$$

The $y_{i1}^{+}, \ldots, y_{iT}^{+}$ are thus treated as 'different' variables on which N observations are available and similarly the x_{it}^{+}. The resulting estimator of β will also be $\hat{\beta}_{WG}$, but the standard errors computed robustly will be calculated according to the formula in (4).

Clearly, once $\widehat{\text{avm}}(\hat{\beta}_{WG})$ is available, one can perform valid hypothesis testing in the presence of heteroskedasticity and serial correlation of unknown form. Also notice that the previous arguments extend to instrumental variable estimators and models that are non-linear in the parameters.

Note, however, that despite our ability to estimate consistently the standard errors for $\hat{\beta}_{WG}$, unless $\Omega_i = \sigma^2 I_T$ holds, more efficient estimators of β are possible. If $\Omega_i = \Omega$ for all i, a computationally simple efficient estimator is SUR-GLS based on the first differences of (1). More generally, under heteroskedasticity of unknown form, efficiency gains through the use of auxiliary variables (e.g. cross-products of the x's) of the type considered by Chamberlain (1982) and Cragg (1983) are possible. Indeed, asymptotically efficient estimators of β can in principle be derived relying on nonparametric estimators of first difference transformations of the Ω_i on the lines of Robinson (1985) and Newey (1986).

Institute of Economics and Statistics,
University of Oxford

Date of Receipt of Final Manuscript: May 1987

REFERENCES

Chamberlain, G. (1982). 'Multivariate Regression Models for Panel Data', *Journal of Econometrics*, Vol. 18, pp. 5-46.

Cragg, J. (1983). 'More Efficient Estimation in the Presence of Heteroskedasticity of Unknown Form', *Econometrica*, Vol. 51, pp. 751-64.

Hausman, J. A. and Taylor, W. E. (1981). 'Panel Data and Unobservable Individual Effects', *Econometrica*, Vol. 49, pp. 1377-398.

Kiefer, N. M. (1980). 'Estimation of Fixed Effect Models for Time Series of Cross-Sections with Arbitrary Intertemporal Covariance', *Journal of Econometrics*, Vol. 14, pp. 195-202.

Newey, W. K. (1986). 'Efficient Estimation of Models with Conditional Moment Restrictions', mimeo.

Robinson, P. (1987). 'Asymptotically Efficient Estimation in the Presence of Heteroskedasticity of Unknown Form', *Econometrica*, forthcoming.

White, H. (1984). *Asymptotic Theory for Econometricians*, Academic Press.