Advances in Economics and Econometrics

Theory and Applications, Ninth World Congress, Volume III

Edited by

Richard Blundell
University College London

Whitney K. Newey
Massachusetts Institute of Technology

Torsten Persson
Stockholm University

CAMBRIDGE
UNIVERSITY PRESS

CHAPTER 13

Fixed and Random Effects in Nonlinear Panel Data Model: A Discussion of a Paper by Manuel Arellano and Jinyong Hahn

Tiemen M. Woutersen*

1 INTRODUCTION

It was a pleasure to serve as the discussant for this session. The authors have played a major role in developing the area under discussion. The paper is very good so that my role is to comment rather than criticize. I will discuss a connection between reducing the estimation bias in a fixed effects model versus reducing misspecification bias in a random effects model. For this discussion, I use the framework of Woutersen (2002).

Woutersen (2002) proposes to use a moment that approximately separates the parameter of interest from the individual parameters. Let z_i denote that data on individual i, let α_i be the individual parameter and let θ be the common parameter. Then, for any likelihood model, a moment function $g(\alpha,\theta) = \sum_i g_i(\alpha,\theta,z_i)/N$ can be constructed with the properties (i) $Eg(\alpha_0,\theta_0) = 0$ where $\{\alpha_0,\theta_0\}$ denote the true parameter values and (ii) $Eg_{\alpha_i}(\alpha_0,\theta_0) = 0$ for all i, that is, the partial derivative of the moment function with respect to α_i is zero for all i. Condition (ii) means that the moment function depends somewhat less on α. The moment function $g(\alpha,\theta)$ is then integrated with respect to the likelihood $L^i(\alpha,\theta)$ and the prior $\pi(\alpha_i;\theta)$,

$$g^{i,I}(\theta) = \frac{\int g^i(\alpha_i,\theta)e^{L^i(\alpha_i,\theta)}\pi(\alpha_i;\theta)d\alpha_i}{\int e^{L^i(\alpha_i,\theta)}\pi(\alpha_i;\theta)d\alpha_i}. \qquad (1)$$

Inference is then based on the integrated moment

$$g^I(\theta) = \frac{\sum_i}{NT} g^{i,I}(\theta))$$

by minimizing

$$Q(\theta) = g^I(\theta)' g^I(\theta)$$

with respect to θ. Under conditions given in Woutersen (2002) the asymptotic bias of the resulting estimator for θ is $O(T^{-2})$ in an asymptotic in which N increases as fast or faster than T, where N is the number of individuals and T is the number of observations per individual. Thus, correcting the bias in the moment function reduces the bias in the resulting estimator.

* Department of Economics, Johns Hopkins University, 3400 North Charles Street, Baltimor, E-mail: woutersen@jhu.edu.

The review paper of Arellano and Hahn is for an important part based on the seminal paper of Hahn and Newey (2004). Hahn and Newey (2004) propose to correct for asymptotic bias. Hahn and Newey (2004) assume that the data are independent and identically distributed and allow for GMM estimators. Thus, compared to Woutersen (2002), Hahn and Newey (2004) allow a more general objective function but for less dependence across time periods. In particular, Woutersen (2002) allows for lagged dependence, general predetermined variables and time dummies, but assumes that expectation of the second derivative of the objective function can be calculated. This expectation can be calculated for likelihood models and some other models but it restricts the number of objective functions. The main difference (if a flat prior, $\pi(\alpha_i;\theta) = 1$ is used in equation (1)) between Hahn and Newey (2004) and Woutersen (2002) is that Hahn and Newey (2004) use an average operator (an average over realizations of an individual) where Woutersen (2002) uses an expectation operator. Subsequent papers have tried to combine the advantages of both papers by proposing to use the average operator for models with lagged dependence. A challenge for these subsequent papers is that one only averages over the realizations of one individual. That is, only over T-dependent realizations where T is small (otherwise, there would not be an "incidental parameter" problem to begin with). Averaging over T-dependent realization may very well work for some models but does not work so well for the simulations that I have done or seen.

In their review, Arellano and Hahn do not discuss models with mixing distributions (often called random effects models). In a model with a mixing distribution, one models the mixing or heterogeneity distribution, see Chamberlain (1984) and Baltagi (1995) for reviews. An advantage of fixed effects models is that one does not need to specify the mixing distribution so one is nonparametric for that aspect of the model. However, introducing many parameters into the objective function usually yields an estimation problem that is usually referred to as the incidental parameter problem of Neyman and Scott (1948). An interesting aspect of mixing models is that the problem of misspecifying the mixing distribution is related to the estimation problem of fixed effects. In particular, choosing the wrong mixing distribution also yields an $O(T^{-1})$ bias under general conditions. Moreover, one can remove the sensitivity to the wrong choice of mixing distribution by using approximate parameter separation. In particular, one can interpret $\pi(\alpha_i;\theta)$ in equation (1) as a mixing distribution (i.e., some parameter of the vector θ govern the mixing distribution). This is algebraically the same as interpreting $\pi(\alpha_i;\theta)$ as a prior so that the same tools can be used and the asymptotic bias can be reduced to $O(T^{-2})$. See Woutersen (2002) for details.

Additional References

BALTAGI, B. H. (1995): *Econometric Analysis of Panel Data*, New York: John Wiley and Sons.

CHAMBERLAIN, G. (1984): "Panel Data," in *Handbook of Econometrics*, Vol. 2, edited by Z. Griliches and M. D. Intriligator, Amsterdam: North-Holland.

Name Index