International Journal of Social Science and Economic Research

ISSN: 2455-8834

Volume:02, Issue:02

PANEL DATA MODELS WITH UNOBSERVABLE INDIVIDUAL EFFECTS AND TIME-INVARIANT REGRESSORS

Chung-ki Min*

Faculty of Economics, Hankuk University of Foreign Studies, Korea.

ABSTRACT

This study proposes an estimation method for dynamic panel data models which include unobservable individual effects and time-invariant regressors. If the individual effects are also time-invariant as in the fixed-effects models, they are not separable from the time-invariant regressors. For an identification, this study relaxes the individual effects to vary over time. The model is estimated after the quasi-differencing transformation is applied. Empirical results from simulated data herein show that the coefficients for time-invariant regressors can be correctly estimated with the introduction of time-varying individual effects.

Keywords: panel data, dynaminc model, time-varying individual effects, time-invariant regressors

JEL Classification: C139,C33, C36

1. INTRODUCTION

Panel data models are widely employed because they can account for the differences between cross-sectional units. The differences are called individual effects and are often assumed to be constant over time within each unit as in the fixed-effects models. If the models include time-invariant regressors, their impact cannot be separable from the time-constant individual effects. Previous studies including Hausman and Taylor (1981) and Bhargava and Sargan (1983) estimate the models under a strong assumption that enough number of the regressors are uncorrelated with the individual effects.

Alternatively, this study relaxes the individual effects to vary over time (Holtz-Eakin et al., 1988; and Ahn et al., 2001 and 2013). Thus, without the strong requirement of regressors uncorrelated

* **Acknowledgment:** This work was supported by the Hankuk University of Foreign Studies Research Fund.

ISSN: 2455-8834

Volume:02, Issue:02

with the individual effects, the coefficients for time-invariant regressors can be estimated. Following Chamberlain (1983), we apply the quasi-differencing transformation to eliminate the time-varying individual effects. Because of the endogeneity of the lagged dependent variables in the transformed dynamic model, we estimate the parameters using appropriately-defined instrumental variables and othorgonality conditions (Holtz-Eakin et al., 1988). Empirical results from simulated data herein show that with the help of time-varying individual effects, the time-invariant regressors in dynamic panel data models can be separated from individual effects and their coeffcients be correctly estimated.

In the next section, we discuss the main issues of this study and present related studies. The dynamic panel data model along with instrumental variables and orthogonality conditions are presented in section 3. In section 4, the simulated data and the estimation results are reported and discussed. Concluding remarks are provided in section 5.

2. THE ISSUES AND RELATED STUDIES

Dynamic panel data models are widely employed in empirical research because they can conduct consistent estimation and hypothesis test with controlling for the time-specific and the individual effects. In many applications, particularly for annual data, the time period is short and the time-specific effects can be easily accounted for by a small number of dummy variables. In contrast, the number of cross-sectional units is quite large. Further, use of dummy variables for the cross-sectional units in dynamic models leads to inconsistent estimation when the time period is short. This study focuses on how to control for the individual effects in dynamic panel data models when time-invariant regressors are also included.

The following model is a simple case where time-varying $(x_{i,t-1})$ and time-invariant (z_i) regressors are included. This is just a part of a two-variable VAR (vector autoregressive regression) model of lag order one.[1]

$$y_{it} = \alpha\ y_{i,t-1} + \beta\ x_{i,t-1} + \gamma z_i + \delta_t + f_i + u_{it} \tag{1}$$

where δ_t and f_i represent the time-specific and the individual effects, respectively. The error term u_{it} is uncorrelated between cross-sectional units and between time periods, and also satisfies the orthogonality conditions, $E[z_i u_{it}] = 0$ and $E[y_{is} u_{it}] = E[x_{is} u_{it}] = 0 \quad (s < t)$.

[1] If the current value of x_{it} is not correlated with the contemporaneous disturbance, the estimation method proposed in this study can be applied to models which include x_{it} as a regressor (Arellano and Bond, 1991; Arellano and Bover, 1995; and Blundell and Bond, 1998).

ISSN: 2455-8834

Volume:02, Issue:02

If the time-invariant regressor is not present, the individual effects can be removed by the first-differencing transformation. The endogeneity problem caused by the lagged dependent variable in the first-differenced equation is solved by the instrumental variables of appropriately lagged variables (Arellano and Bond, 1991). However, this method cannot be employed to estimate Eq.(1) because the time-invariant variable (z_i) is also removed by the first-differencing transformation.

Arellano and Bover (1995) and Blundell and Bond (1998) propose to use instrumental variables which are uncorrelated with the individual effects. The particular type of variables of this kind are the first-differences of predetermined variables, i.e., $\Delta y_{i,t-1}$ and $\Delta x_{i,t-1}$ for Eq.(1). It is because $E(\Delta y_{i,t-1} f_i) = E(y_{i,t-1} f_i) - E(y_{i,t-2} f_i) = 0$ if we can assume that the predetermined variable $\{y_{i,t-s}\}$ has a constant correlation with the individual effects. Use of such instrumental variables in differences can estimate the equation in levels. However, the time-invariant variable z_i is still a problem because it is often correlated with the individual effects. The first-differencing approach eliminates not only the individual effects f_i but also the time invariant variable z_i.

Assuming the presence of exogenous variables which are uncorrelated with the individual effects, Hausman and Taylor (1981) and Bhargava and Sargan (1983) estimate the equation in levels, Eq.(1), using instrumental variables both in levels and in first-differences. However, Arellano and Bover (1995) point out that the impact of these models in applied work has been limited because of the difficulty in finding exogenous variables that can be convincingly regarded a priori as being uncorrelated with the individual effects.

Different from the above mentioned studies, Holtz-Eakin et al. (1988) relax the assumption of constant individual effects. The individual effects are expressed as a product of f_i and a time-varying parameter ψ_t.

$$y_{it} = \alpha\ y_{i,t-1} + \beta\ x_{i,t-1} + \delta_t + \psi_t f_i + u_{it} \qquad (2)$$

As the authors focus on the causality test, their model does not include a time-invariant variable, which is the main issue of this study. To eliminate the time-varying individual effects, Holtz-Eakin et al. (1988) employ the quasi-differencing transformation used in Chamberlain (1983).

Ahn et al. (2001) list various applications of such time-varying individual effects. Recently, Ahn et al. (2013) extend the single time-invariant component (f_i) to multiple ones. These studies eliminate the time-varying individual effects using an extended within-transformation, which is however not applicable to dynamic panel data models in this study.

3. THE MODEL AND ESTIMATION

The model considered in this study is a mixture of the above two models. We allow the individual effects to vary over time ($\psi_t f_i$) and include a time-invariant variable (z_i) which is correlated with f_i.

$$y_{it} = \alpha\ y_{i,t-1} + \beta\ x_{i,t-1} + \gamma z_i + \delta_t + \psi_t f_i + u_{it} \tag{3}$$

where the error term u_{it} is uncorrelated between cross-sectional units and between time periods, and also satisfies the orthogonality conditions, $E[z_i u_{it}] = 0$ and $E[y_{is} u_{it}] = E[x_{is} u_{it}] = 0$ $(s < t)$.

The quasi-differencing transformation is applied to eliminate the time-varying individual effects. After multiplying Eq.(3) in time period *t*-1 by $r_t = \psi_t / \psi_{t-1}$, the result is subtracted from the equation in time period *t* (Chamberlain, 1983; and Holtz-Eakin et al., 1988).

$$y_{it} = \theta_{1t} y_{i,t-1} + \theta_{2t} x_{i,t-1} + \theta_{3t} y_{i,t-2} + \theta_{4t} x_{i,t-2} + \theta_{5t} z_i + d_t + v_{it} \tag{4}$$

or

$$y_{it} = w_{it}^{'} B_t + v_{it}$$

where $\theta_{1t} = \alpha + r_t$, $\theta_{2t} = \beta$, $\theta_{3t} = -\alpha\ r_t$, $\theta_{4t} = -\beta\ r_t$, $\theta_{5t} = \gamma(1 - r_t)$, $d_t = \delta_t - r_t \delta_{t-1}$, $v_{it} = u_{it} - r_t u_{i,t-1}$, $w_{it} = [y_{i,t-1}\ \ x_{i,t-1}\ \ y_{i,t-2}\ \ x_{i,t-2}\ \ z_i\ \ 1]'$, and $B_t = [\theta_{1t}\ \ \theta_{2t}\ \ \theta_{3t}\ \ \theta_{4t}\ \ \theta_{5t}\ \ d_t]'$.

The orthogonality conditions in Eq.(3) imply that the error term v_{it} satisfies $E[y_{is} v_{it}] = E[x_{is} v_{it}] = 0$ for $s < t-1$ because of the presence of $u_{i,t-1}$ in v_{it}. Thus, the instrumental variables, which can be used to identify the parameters of Eq.(4), are

$$W_{it} = [\ y_{i,t-2}\ \cdots\ y_{i1}\ \ x_{i,t-2}\ \cdots\ x_{i1}\ \ z_i\ \ 1]'$$

Because of the time-varying coefficients B_t, the orthogonality conditions are defined separately for each *t* (Holtz-Eakin, 1988).

$$\frac{1}{M}\sum_{i=1}^{M} W_{it} v_{it} \xrightarrow{M \to \infty} 0 \tag{5}$$

ISSN: 2455-8834

Volume:02, Issue:02

To account for the serial correlation of the disturbance in Eq.(4), which is of a moving-average form $v_{it} = u_{it} - r_t u_{i,t-1}$, we transform the variables such that the disturbance becomes serially uncorrelated; the transformation is explained in Min (2017).

Unless the individual effects are constant over time, the time-invariant variable will remain after the quasi-differencing transformation because the multiplying factor $r_t = \psi_t / \psi_{t-1}$ is different from one. Ahn et al. (2001, 2013) and Holtz-Eakin et al. (1988) support a possibility of time-varying individual effects. Further, due to a sampling error of small-sized samples in practice, the individual effects could be estimated as varying between time periods although they are in fact constant. Thanks to the time-varying individual effects and the quasi-differencing approach, the coefficients for time-invariant variables can be estimated. In contrast, assuming a presence of time-constant individual effects, Hausman and Taylor (1981) and Bhargava and Sargan (1983) are required to have explanatory variables which are uncorrelated with the individual effects; this is a strong requirement.

4. EMPIRICAL RESULTS FROM SIMULATED DATA

To examine the estimation performance of the quasi-differencing approach suggested in this study, we generate data using the following VAR(1) specification: for $i = 1, \cdots, M$ and $t = 1, \cdots, T$,

$$\begin{aligned} y_{it} &= \alpha y_{i,t-1} + \beta x_{i,t-1} + \theta z_i + \delta_t + \psi_t f_i + u_{it} \qquad (6) \\ x_{it} &= \gamma_1 y_{i,t-1} + \gamma_2 x_{i,t-1} + w_{it} \end{aligned}$$

where δ_t is independently drawn from a normal distribution $N(0,1)$; f_i is from $Unif(0,1)$; ψ_t is the sum of $0.1\delta_t$ and a random number from $Unif(0,1)$; z_i is the sum of f_i and a random number from $Unif(0,2)$; and the disturbances u_{it} and w_{it} are independently generated from a normal distribution $N(0, 0.5^2)$. The assigned values for the parameters are $\alpha = 0.5$, $\beta = 0.3$, $\theta = 0.3$, $\gamma_1 = 0.3$ and $\gamma_2 = 0.5$. After generating data for *t* = -29 to 10, we keep the last 10 observations (*t* = 1 to 10) to minimize any effects of starting values; thus *T*=10. Following the same procedures, we generate observations for *M* cross-sectional units. The number of cross-sectional units is set to *M*=200 and *M*=500 to examine the convergence of the estimators as the number of cross-sectional units increases. This completes one iteration. We repeat the data generating iteration 1,000 times.

ISSN: 2455-8834

Volume:02, Issue:02

Table 1: Summary statistics and correlation coefficients

Variable	mean	s.d.	min	max
y_{it}	2.364	1.565	-0.052	4.499
x_{it}	1.488	1.012	0.137	2.622
z_i	1.500	0.645	1.395	1.587
$\psi_t f_i$	0.250	0.220	0.102	0.402
f_i	0.500	0.288	0.458	0.539
	Correlation coefficients			
	$x_{i,t-1}$	z_i	$\psi_t f_i$	f_i
$y_{i,t-1}$	0.593	0.553	0.307	0.544
$x_{i,t-1}$	-	0.528	0.332	0.547
z_i	-	-	0.292	0.446

After calculating the summary statistics for each iteration, we averaged them over 1,000 iterations.

Table 1 reports the summary statistics and the correlation coefficients of the variables and the parameters associated with the individual effects. It is shown that the regressors and the individual effects are significantly correlated. The time-varying individual effects have correlation coefficients of 0.307 with $y_{i,t-1}$, 0.332 with $x_{i,t-1}$ and 0.292 with the time-invariant regressor z_i. Since all regressors are correlated with the individual effects, it is important to control for the individual effects.

Table 2: Estimation results

	M= 200		*M*= 500	
	estimate	s.e.	estimate	s.e.
α (0.5)	0.303	0.194	0.350	0.084
β (0.3)	0.136	0.176	0.177	0.077
γ (0.3)	0.776	0.396	0.678	0.243

After obtaining the estimates and their standard errors for each iteration, we averaged them over 1,000 iterations. The number in each parenthesis is the true value assigned to the parameter.

The estimation results are reported in Table 2 for two cases of *M*=200 and *M*=500. For both cases the true values of parameters are included in the 95% confidence interval, indicating that the estimation by the quasi-differencing approach is accurate. As the number of cross-sectional units increases from *M*=200 to *M*=500, the estimates become more precise with smaller standard

ISSN: 2455-8834

Volume:02, Issue:02

errors, indicating a convergence of the estimators.

5. CONCLUSIONS

We examined herein whether the coefficients for time-invariant regressors in dynamic panel data models can be estimated when the individual effects need to be accounted for. Different from the fixed effects models, this study allowed the individual effects to vary over time and eliminated them using the quasi-differencing transformation. In doing so, the time-invariant regressors are separated from the time-varying individual effects. Therefore, their coefficients can be estimated.

Applications to real data are planned for future work, particularly to data collected from tourism industry. As the tourism demand is highly subjective to changes in each time period, the individual effects are likely to vary over time.

REFERENCES

Ahn, S.C., Y.H. Lee, and P. Schmidt, 2001, GMM Estimation of Linear Panel Data Models with Time-varying Individual Effects. *Journal of Econometrics* 101, 219-255.

Ahn, S.C., Y.H. Lee, and P. Schmidt, 2013, Panel Data Models with Multiple Time-varying Individual Effects. *Journal of Econometrics* 174, 1-14.

Arellano, M., and S. Bond, 1991, Some Tests of Specification for Panel Data: Monte Carlo Evidence and an Application to Employment Equations. *Review of Economics Studies* 58, 277-297.

Arellano, M., and O. Bover, 1995, Another Look at the Instrumental Variables Estimation of Error Components Models. *Journal of Econometrics* 68, 29-52.

Bhagarva, A., and J.D. Sargan, 1983, Estimating Dynamic Random Effccts Models from Panel Data Covering Short Time Periods. *Econometrica* 51, 1635-1659.

Blundell, R., and S. Bond, 1998, Initial Conditions and Moment Restrictions in Dynamic Panel Data Models. *Journal of Econometrics* 87, 115-143.

Chamberlain, G., 1983, Panel Data. Ch. 22 in *The Handbook of Econometrics Volume II*, Ed. by Z. Griliches and M. Intrilligator, Amsterdam: North-Holland Publishing Company.

ISSN: 2455-8834

Volume:02, Issue:02

Hausman, J.A., and W.E. Taylor, 1981, Panel Data and Unobservable Individual Effects. *Econometrica* 49, 1377-1398.

Holtz-Eakin, D., W. Newey, and H.S. Rosen, 1988, Estimating Vector Autoregressions with Panel Data. *Econometrica* 56, 1371-1395.

Min, C., 2017, How to Control for Unobservable Individual Effects in Dynamic Panel Data Models: A Practical Guide," Working paper.

