International Journal of Social Science and Economic Research

ISSN: 2455-8834

Volume:02, Issue:01

ESTIMATION OF DYNAMIC PANEL DATA MODELS WHEN REGRESSION COEFFICIENTS AND INDIVIDUAL EFFECTS ARE TIME-VARYING

Chung-ki Min

Professor, Hankuk University of Foreign Studies, Seoul, Korea

ABSTRACT

This study suggests an estimation method for dynamic panel data models when their coefficients and individual effects are time-varying. The quasi-differencing transformation is employed to eliminate the time-varying individual effects. Empirical results from simulated data herein support the performance of the quasi-differencing approach for estimation and test of dynamic panel data models, particularly for small-sized samples.

Keywords: dynamic panel data model; time-varying individual effects; quasi-differencing; orthogonality conditions

1. INTRODUCTION

The objective of this study is to suggest an estimation method for dynamic panel data models when their regression coefficients and individual effects are time-varying. As the regression coefficients are likely to vary over time in reality, particularly during a long period of time, this study is expected to provide applied researchers with useful implications.

After applying the quasi-differencing (QD) transformation to eliminate the time-varying individual effects, we estimate the time-varying parameters by employing the generalized method of moment (GMM) (Arellano and Bond, 1991; Arellano and Bover, 1995; and Blundell and Bond, 1998). We also apply a Wald test to examine whether the regression coefficients varied over time. Empirical results from simulated data herein support the performance of the QD approach for estimation and test of dynamic panel data models, particularly for small-sized samples.

2. ESTIMATION METHODS

ISSN: 2455-8834

Volume:02, Issue:01

The model considered in this study is from a two-variable vector autoregressive regression of lag order one, VAR(1). For cross-sectional unit *i* (= 1,..., *M*) and time period *t* (= 1, ..., *T*), this model allows for time-specific and individual effects.

$$y_{it} = \alpha_t y_{i,t-1} + \beta_t x_{i,t-1} + \delta_t + \psi_t f_i + u_{it} \tag{1}$$

where the error term u_{it} is uncorrelated between units and between time periods, and also satisfies the orthogonality conditions $E[y_{is}u_{it}] = E[x_{is}u_{it}] = 0 \quad (s < t)$. The time-specific effects (δ_t) are common to all cross-sectional units. This model allows individual effects to vary over time as the time-invariant individual effects f_i are multiplied by a time-varying coefficient ψ_t (Holtz-Eakin et al., 1988). Ahn et al. (2001) list various applications of panel data models with such time-varying individual effects.

This study employs the QD transformation used in Chamberlain (1983) and Holtz-Eakin et al. (1988). After multiplying Eq.(1) for time period *t*-1 by $r_t = \psi_t / \psi_{t-1}$, the result is subtracted from the equation for time period *t*.

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

or

$$y_{it} = w_{it}^{'} \gamma_t + v_{it}$$

where $\theta_{1t} = \alpha_t + r_t$, $\theta_{2t} = \beta_t$, $\theta_{3t} = -\alpha_t r_t$, $\theta_{4t} = -\beta_t r_t$, $d_t = \delta_t - r_t \delta_{t-1}$ and $v_{it} = u_{it} - r_t u_{i,t-1}$, $w_{it} = [y_{i,t-1} \;\; x_{i,t-1} \;\; y_{i,t-2} \;\; x_{i,t-2} \;\; 1]'$, and $\gamma_t = [\theta_{1t} \;\; \theta_{2t} \;\; \theta_{3t} \;\; \theta_{4t} \;\; d_t]'$.

The orthogonality conditions in Eq.(1) 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 estimate the parameters of Eq.(2), are included in the following vector.

$$z_{it}^{QD} = [\, y_{i,t-2} \;\cdots\; y_{i1} \;\; x_{i,t-2} \;\cdots\; x_{i1} \;\; 1]'$$

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

ISSN: 2455-8834

Volume:02, Issue:01

$$\frac{1}{M}\sum_{i=1}^{M} z_{it}^{QD} v_{it} \xrightarrow{M\to\infty} 0 \tag{3}$$

If the regression coefficients and ψ_t are constant over time, then Eq.(2) becomes the first-differenced (FD) specification with $r_t = 1$.

$$\Delta y_{it} = \alpha\ \Delta y_{i,t-1} + \beta\ \Delta x_{i,t-1} + \Delta\delta_t + \Delta u_{it} \tag{4}$$

or

$$\Delta y_{it} = h_{it}^{'}\varphi + \Delta u_{it}$$

where Δ denotes the difference between time period *t* and *t*-1, $h_{it} = [\Delta y_{i,t-1}\ \ \Delta x_{i,t-1}\ \ D1_t \cdots DT_t]'$, $\varphi = [\alpha\ \ \beta\ \ \eta_1 \cdots \eta_T]'$, and $(D1_t \cdots DT_t)$ are time dummies with their corresponding coefficients $(\eta_1 \cdots \eta_T)$.

Since the regression coefficients are constant over time, the instrumental variables for the FD specification (z^{FD}) satisfy the orthogonality conditions defined for the entire period, not period by period.

$$\frac{1}{MT}\sum_{i=1}^{M}\sum_{t=1}^{T} z_{it}^{FD}\Delta u_{it} \xrightarrow{M\to\infty} 0 \tag{5}$$

3. EMPIRICAL RESULTS USING SIMULATED DATA

To evaluate the estimation performance of the QD and the FD approach, data are generated using the following VAR(1) specification.

$$\begin{aligned} y_{it} &= \alpha_t y_{i,t-1} + \beta_t x_{i,t-1} + \delta_t + \psi_t f_i + u_{it} \\ x_{it} &= \gamma_1 y_{i,t-1} + \gamma_2 x_{i,t-1} + w_{it} \end{aligned} \tag{6}$$

After data are generated for *t* = -19 to 12, the first 20 observations (*t* = -19 to 0) are discarded to minimize any effects of the starting values. The values assigned for the time-varying regression coefficients are reported in Table 2. Values for the time-specific effects (δ_t's) and for the individual effects (f_i) are independently drawn from uniform distributions $\delta_t \sim Uniform(-0.5, 0.5)$ and $f_i \sim Uniform(-2, 2)$, respectively.

ISSN: 2455-8834

Volume:02, Issue:01

(1) Case 1: time-varying regression coefficients (α_t, β_t)

We apply the QD and FD approaches with assuming constant coefficients in this section. This is to compare how well the two approaches can account for the time variation in (α_t, β_t).

Table 1 shows that as the time variation in (α_t, β_t) increases with a larger σ_η, the FD estimates are severely biased. When the averages of the true values (α_t, β_t) are (0.7, 0.3), the FD estimates are $(\hat{\alpha}, \hat{\beta})$ = (0.631, 0.215) for $\sigma_\eta = 0.01$, (0.081, -0.200) for $\sigma_\eta = 0.03$ and (-0.181, -0.369) for $\sigma_\eta = 0.05$. Even when the QD instrumental variables are used for estimation, the FD estimates are still biased; (0.572, 0.167) for $\sigma_\eta = 0.01$, (0.142, -0.145) for $\sigma_\eta = 0.03$ and (-0.050, -0.252) for $\sigma_\eta = 0.05$.

In contrast, the QD estimates are not biased when the variation in (α_t, β_t) is small with $\sigma_\eta = 0.01$, $(\hat{\alpha}, \hat{\beta})$=(0.684, 0.265). For larger values of σ_η, the QD estimates are much less biased; $(\hat{\alpha}, \hat{\beta})$ = (0.601, 0.192) for $\sigma_\eta = 0.03$ and (0.538, 0.143) for $\sigma_\eta = 0.05$. This is because the QD can partially account for the time variation in (α_t, β_t) through the time-varying individual effects.

Table 1: QD and FD specifications estimated with assuming constant regression coefficients when they varied over time

		QD transformation estimated by Z^{QD}		FD transformation estimated by Z^{FD}		FD transformation estimated by Z^{QD}	
		estimate	se	estimate	se	estimate	se
$\sigma_\eta = 0$ (constant)							
	α	0.665	0.028	0.677	0.030	0.664	0.028
	β	0.286	0.022	0.290	0.024	0.285	0.022
$\sigma_\eta = 0.01$	α	0.684	0.022	0.631	0.032	0.572	0.027
	β	0.265	0.015	0.215	0.027	0.167	0.021
$\sigma_\eta = 0.03$	α	0.601	0.020	0.081	0.045	0.142	0.031
	β	0.192	0.013	-0.200	0.042	-0.145	0.030
$\sigma_\eta = 0.05$	α	0.538	0.020	-0.181	0.038	-0.050	0.026
	β	0.143	0.013	-0.369	0.047	-0.252	0.034

ISSN: 2455-8834

Volume:02, Issue:01

Note: The time-varying regression coefficients were generated by $\alpha_t = 0.7 + \eta_t^{\alpha}$ and $\beta_t = 0.3 + \eta_t^{\beta}$, where η_t^{α} and η_t^{β} were randomly drawn from a normal distribution, $N(0, \sigma_\eta^2)$.
The true values of the time-varying regression coefficients are reported in Table 2.

(2) Estimation and test for time-varying regression coefficients

The time-varying regression coefficients (α_t, β_t) are estimated by the QD approach without assuming constant coefficients. Table 2 shows that for all of the cases, the QD approach correctly estimated the time-varying coefficients; the true coefficients are included in the 95% confidence intervals.

Table 2: Estimates of the time-varying regression coefficients by the QD approach

α_t / β_t	$\sigma_\eta = 0$ (constant)			$\sigma_\eta = 0.01$			$\sigma_\eta = 0.03$			$\sigma_\eta = 0.05$		
	true value	estimate	s.e.	true value	estimate	s.e.	true value	estimate	s.e.	true value	estimate	s.e.
α_4	0.7	0.728	0.066	0.696	0.721	0.063	0.687	0.707	0.058	0.679	0.695	0.053
α_6	0.7	0.646	0.079	0.697	0.642	0.077	0.690	0.634	0.075	0.684	0.627	0.072
α_8	0.7	1.082	0.194	0.695	1.090	0.198	0.685	1.106	0.207	0.676	1.122	0.217
α_{10}	0.7	0.747	0.049	0.724	0.777	0.057	0.772	0.843	0.088	0.819	0.898	0.079
α_{12}	0.7	0.678	0.058	0.714	0.693	0.059	0.741	0.723	0.062	0.769	0.754	0.067
β_4	0.3	0.363	0.048	0.299	0.360	0.045	0.298	0.356	0.041	0.297	0.352	0.038
β_6	0.3	0.245	0.061	0.294	0.239	0.060	0.281	0.228	0.056	0.268	0.216	0.053
β_8	0.3	0.606	0.192	0.294	0.611	0.194	0.282	0.620	0.199	0.270	0.627	0.204
β_{10}	0.3	0.324	0.028	0.304	0.335	0.035	0.313	0.370	0.091	0.321	0.411	0.184
β_{12}	0.3	0.293	0.049	0.297	0.290	0.050	0.291	0.286	0.052	0.285	0.283	0.055
Wald test of a null hypothesis that the regression coefficients (α_t, β_t) are constant over time.												
p-value		0.423			0.342			0.087			0.0003	

See the note in Table 1.

We also applied a Wald test to examine whether the regression coefficients significantly varied over time. When the coefficients were constant $(\sigma_\eta = 0)$ or varying a little $(\sigma_\eta = 0.01)$, the Wald test failed to reject the null hypothesis of constant coefficients; the p-value is 0.423 and 0.342, respectively. As the time variation in (α_t, β_t) increased with a larger σ_η, the null hypothesis was more significantly rejected; the p-value is 0.087 for $\sigma_\eta = 0.03$ and 0.0003 for $\sigma_\eta = 0.05$.

ISSN: 2455-8834

Volume:02, Issue:01

These test results are consistent with the estimation results in Table 1. When the assumption of constant coefficients was not rejected for $\sigma_\eta = 0$ and $\sigma_\eta = 0.01$, the regression coefficients were correctly estimated by the QD approach with assuming constant coefficients; $(\hat{\alpha}, \hat{\beta})$ = (0.665, 0.286) and (0.684, 0.265) as shown in the second column of Table 1. However, for $\sigma_\eta = 0.03$ and $\sigma_\eta = 0.05$ with which the assumption of constant coefficients was rejected by the Wald test, the QD estimates were biased; $(\hat{\alpha}, \hat{\beta})$ = (0.601, 0.192) and (0.538, 0.143). It is because the coefficients were estimated assuming constant when they varied significantly.

These results in Tables 1 and 2 provide useful implications for applied researchers as the regression coefficients are likely to vary over time in reality, particularly during a long period of time. If an economy experiences big changes such as global financial crisis and bad weather, the relations between variables of a model might not be constant. Based on the above results, this study suggests that we first test about the constancy of regression coefficients using the QD estimates of time-varying coefficients, as shown in Table 2. If the time variation in (α_t, β_t) is not significant, the QD approach, together with assuming constant regression coefficients, is expected to produce correct estimates and small standard errors, as shown in Table 1.

4. CONCLUSIONS

In many panel data sets, there are a large number of cross-sectional units with a short period of time. Therefore, how to control for individual effects becomes a main issue. In this study, the quasi-differencing (QD) approach was applied to eliminate the time-varying individual effects. The first-differencing (FD) approach, which is valid only when individual effects and regression coefficients are constant over time, was also applied. Overall, the QD approach produced more reliable estimates than the FD approach in all of the cases considered in this study.

As the regression coefficients in reality are likely to vary over time, this study employed the QD approach to estimate time-varying regression coefficients and to test whether they are constant over time. The results indicate that if the Wald test fails to reject a null hypothesis of constant coefficients, we may assume constant regression coefficients and use the QD approach (but not the FD). Therefore, we conclude that in all of the cases considered in this study, the QD approach dominates over the FD for the estimation and test of panel data models.

ACKNOWLEDGEMENT

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

ISSN: 2455-8834

Volume:02, Issue:01

REFERENCES

Ahn, S.C., Lee, Y.H., and Schmidt, P. (2001). GMM estimation of linear panel data models with time-varying individual effects. *Journal of Econometrics*, 101, 219-255.

Arellano, M., and Bond, S. (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 Bover, O. (1995). Another look at the instrumental variables estimation of error components models. *Journal of Econometrics*, 68(1), 29-52.

Blundell, R., and Bond, S. (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.

Holtz-Eakin, D., Newey, W. and Rosen, H.S. (1988). Estimating vector autoregressions with panel data. *Econometrica*, 56 (6), 1371-1395.

