Estimation and Inference on Dynamic Panel Data Models with Stochastic Volatility

Wen Xu

Department of Economics & Oxford-Man Institute
University of Oxford

(Preliminary, Comments Welcome)

Theme

$$y_{it} = B(L)y_{i,t-1} + \gamma x_{it} + \mu_i + \varepsilon_{it}$$

- ε_{it}: martingale difference sequences parameterized through stochastic volatility
- In a macroeconomic context: T is relatively large (>20); N is relatively small (<100)

Contribution

- Efficient estimation by particle filter techniques
  - Frequentist approach: maximization of approximate simulated-likelihood
  - Bayesian approach: particle Metropolis-Hastings sampler
- Two-step LSDV-QML: fast but less efficient
- Ignoring the existence of stochastic volatility might induce systematically false rejection in panel unit root tests
- Monte Carlo studies show that particle-filter based estimators are more precise than other estimators on average in the presence of stochastic volatility and even in the case of homoscedasticity
- Our methodology is straightforwardly applied to panel VAR models

Motivation

Time-varying volatility in macroeconomic time series:

- Evidence of conditional heteroskedasticity in the residuals of many estimated dynamic regression models in macroeconomics: Weiss (1984)
- U.S. real GDP: Kim and Nelson (1999), McConnell and Perez-Quiros (2000), Blanchard and Simon (2001).
- Other macro variables: Stock and Watson (2003), Fernández-Villaverde and Rubio-Ramírez (2013).

Motivation

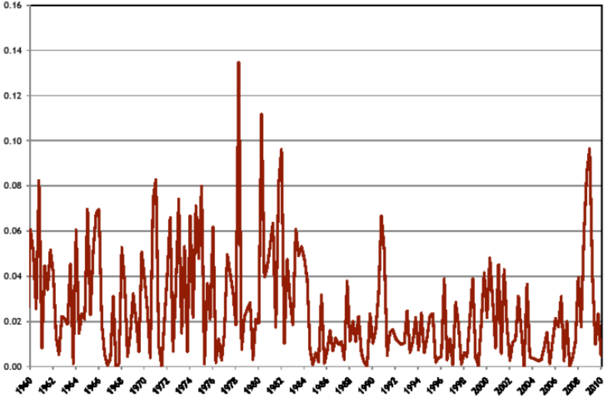

Figure : U.S. Real GDP Growth, Absolute Deviations from Mean (from Fernández-Villaverde and Rubio-Ramírez, 2013)

Motivation

Some attempts in modeling stochastic volatility in macroeconomics

- DSGE: Fernández-Villaverde and Rubio-Ramírez (2007)
- VAR: e.g., Koop and Korobilis (2010)
- Hamilton (2010)
  - Hypothesis tests on the mean might be invalid if the variance is misspecified
  - Statistical efficiency gains can be obtained by incorporating observed features of time-varying volatility into the estimation of the conditional mean.

Motivation

Dynamic panel data models in macroeconomic applications: study common relationships across countries or regions

- purchasing power parity
- mean reversion of interest rates
- growth convergence

Panel unit root tests

- Levin and Lin (1992)
- Harris and Tzavalis (1999):

$$\sqrt{N}(\hat{\beta}_{LSDV} - 1 + \frac{3}{T+1}) \xrightarrow{d} N(0, \frac{3(17T^2 - 20T + 17)}{5(T-1)(T+1)^3})$$

Baseline Model

$$y_{it} = \beta y_{i,t-1} + \mu_i + \varepsilon_{it}$$

$$\varepsilon_{it} = \sigma_{it}\epsilon_{it}$$

$$\log(\sigma_{it}^2) = \mu + \phi(\log(\sigma_{i,t-1}^2) - \mu) + \eta_{it}$$

$$\begin{pmatrix} \epsilon_{it} \\ \eta_{it} \end{pmatrix} \sim N(0, \begin{bmatrix} 1 & 0 \\ 0 & \theta^2 \end{bmatrix})$$

Particle Filters

Nonlinear and non-Gaussian state space models

$$\begin{aligned} y_t &= h_t(x_t, \epsilon_t) \\ x_t &= m_t(x_{t-1}, \eta_t) \end{aligned}$$

Particle filters: simulation-based filtering techniques to approximate the posterior density $p(x_{1:t}|y_{1:t}, \theta)$ by a discrete distribution made of weighted draws $(x_{1:t}^{(j)}, \hat{w}_t^{(j)})$ $(j = 1, \ldots, M)$ termed particles. Algorithms differ mainly in the choices of incremental importance distributions and resampling algorithms which are aimed to improve the level of statistical efficiency in terms of Monte Carlo variation.

- bootstrap filter (Gordon et al., 1993)
- auxiliary particle filters (Pitt and Shephard, 1999)
- mixture Kalman filters (Chen and Liu, 2000)
- .....

Mixture Kalman Filters (Chen and Liu, 2000)

$$y_t = Z(\alpha_{t,1})\alpha_{n,2} + \varepsilon_t$$

$$\alpha_{t,2} = T(\alpha_{t,1})\alpha_{t-1,2} + \eta_t$$

$$\varepsilon_t \sim N(0, H(\alpha_{t,1}))\ \eta_t \sim N(0, Q(\alpha_{t,1}))$$

$$\alpha_{t,1} \text{ follows a first order Markov process}$$

- Particle state variables $\alpha_{t,1}$: simulation
  - M draws or particles per period: $(\alpha_{t,1}^{(j)}, \hat{w}_{i,t|t-1}^{(j)})$ $(j = 1, \ldots, M)$
- Kalman state variables $\alpha_{t,2}$: conditionally normal given $\alpha_{t,1}$, Kalman filters

State Space Forms

$$\begin{aligned}
\Delta y_{it} &= \begin{bmatrix} 1 & 1 & 0 \end{bmatrix} \alpha_{it} \\
\alpha_{it} &= \begin{bmatrix} \beta & \beta & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix} \alpha_{i,t-1} + \begin{bmatrix} 0 \\ 1 \\ -1 \end{bmatrix} \sigma_{it}\epsilon_{it} \\
\log(\sigma_{it}^2) &= \mu + \phi(\log(\sigma_{i,t-1}^2) - \mu) + \eta_{it}
\end{aligned}$$

where the particle state variable is σ_{it} and the Kalman state vector

$$\alpha_{it} = \begin{bmatrix} \beta \Delta y_{i,t-1} \\ \sigma_{it}\epsilon_{it} - \sigma_{i,t-1}\epsilon_{i,t-1} \\ -\sigma_{it}\epsilon_{it} \end{bmatrix}.$$

Variables of interests

- particle state variable $\sigma_{it}^{(j)}$
- conditonal means of Kalman state vector given $\mathcal{Y}_{i,t-1}$ and $\sigma_{it}^{2(j)}$: $m_{it|t-1}^{(j)} = \mathrm{E}(\alpha_{it}|\mathcal{Y}_{i,t-1}, \sigma_{it}^{2(j)})$
- conditonal variances of Kalman state vector $\mathcal{Y}_{i,t-1}$ and $\sigma_{it}^{2(j)}$: $\Sigma_{it|t-1}^{(j)} = \mathrm{Var}(\alpha_{it}|\mathcal{Y}_{i,t-1}, \sigma_{it}^{2(j)})$
- conditonal likelihood of observations $\log[\hat{L}(\Delta y_{it}|\mathcal{Y}_{i,t-1}, \beta, \mu, \phi, \theta)]$

Algorithm

For $i = 1, ..., N$,

1. Set the starting values $\sigma_{i2}^{2(j)}$, $m_{i2|1}^{(j)}$ and $\Sigma_{i2|1}^{(j)}$.

For $t = 2, ..., T$,

2. Draw the particle state variable

$$\log(\sigma_{i,t+1}^{2(j)}) = (1-\phi)\log(\mu) + \phi\log(\sigma_{it}^{2(j)}) + \theta\eta_{i,t+1}^{(j)}$$

3. Compute the conditional likelihood for each particle j. That is,

$$l_{it}^{(j)} = -0.5\log|V_{it}^{(j)}| - 0.5 v_{it}^{(j)}(V_{it}^{(j)})^{-1} v_{it}^{(j)}$$

Algorithm

4. Update the particle weight

$$w_{i,t+1|t}^{(j)} = w_{it|t-1}^{(j)} + l_{it}^{(j)}$$

$$\hat{w}_{i,t+1|t}^{(j)} = \frac{\exp(w_{i,t+1|t}^{(j)})}{\sum_{j=1}^{M} \exp(w_{i,t+1|t}^{(j)})}$$

5. Resample with replacement M particles $\sigma_{i,t+1}^{2(j)}$, $m_{it|t-1}^{(j)}$ and $\Sigma_{it|t-1}^{(j)}$ with the weight $\hat{w}_{i,t+1|t}^{(j)}$ every three increments. After doing this, reset $w_{it|t-1}^{(j)} = 0$ and $\hat{w}_{it|t-1}^{(j)} = \frac{1}{M}$.
6. Update Kalman filter estimates of $m_{it|t-1}^{(j)}$ and $\Sigma_{it|t-1}^{(j)}$.
7. Go back to step 2.

Simulation-based estimate of the joint log-likelihood

$$\log[\hat{L}(\Delta y_{it}|\mathcal{Y}_{i,t-1}, \beta, \mu, \phi, \theta)] = \log[\sum_{j=1}^{M} \hat{w}_{it|t-1}^{(j)} \exp(l_{it}^{(j)})]$$

Resampling

- Weight degeneracy problem

As the series grows over time, one particle's normalized importance weight converges to one while the others converge to zero. In other words, the discrete distribution made of weighted draws would become degenerate. The estimator will be imprecise because it is finally a function of a single draw.

- Resampling: aimed to mitigate the weight degeneracy problem by eliminating the particles which have low importance weights and multiplying the heavily weighted particles.
  - multinomial resampling: Gordon et al. (1993)
    - It draws new particles $\{\tilde{\sigma}_{it}^{2(j)}, \tilde{m}_{it|t-1}^{(j)}, \tilde{\Sigma}_{it|t-1}^{(j)}\}_{j=1}^{M}$ from the point mass distribution $\{\sigma_{it}^{2(j)}, m_{it|t-1}^{(j)}, \Sigma_{it|t-1}^{(j)}, \hat{w}_{it|t-1}^{(j)}\}_{j=1}^{M}$
  - stratified resampling (Kitagawa, 1996)
  - residual resampling (Liu and Chen, 1998)
  - systematic resampling (Carpenter et al., 1999)

Frequentist approach

Direct maximization of the simulated likelihood suffers from discontinuity induced from the generalized inverse operation at the resampling stage, which makes invalid the common gradient-based optimization methods.

- Pitt (2002) overcame the problem of non-smoothness by developing a new resampling method, but his method is valid only when the dimension of state space is one.
- Expectation Maximization algorithm: numerically stable and computationally cheap but only guaranteed to be locally optimal.
- Olsson and Rydén (2008): step functions or B-spline interpolation, and showed consistency and asymptotic normality of the estimators which maximize the approximate likelihood under some assumptions.

Frequentist approach

Olsson and Rydén (2008)

- Discretize the parameter space Ω by a grid $\bar{\Omega} \triangleq \{\omega_g\}_{g=1}^{G} \subseteq \Omega$. Let $[\omega]$ denote the closest point in the grid to $\omega \in \Omega$

$$\log[\hat{L}(\Delta \mathbf{y}_1, \ldots, \Delta \mathbf{y}_N | \omega)] \approx \log[\hat{L}(\Delta \mathbf{y}_1, \ldots, \Delta \mathbf{y}_N | [\omega])]$$

- Maximize approximate simulated-likelihood
- Advantage: theoretical proof of consistency and asymptotic normality, good finite sample performance, fast.
- Limit: asymptotic proof requires compact state space
  - potentially released given new results of uniform convergence properties in time dimension in the general state space models.

Bayesian approach

Andrieu et al. (2010) combined particle filters with standard MCMC algorithms

- Particle independent Metropolis-Hastings sampler
  - A random walk proposal is used for $\log(\beta)$, $\log(\mu)$, $\log(\phi)$ and $\log(\theta)$.
  - Noninformative prior $f(\beta, \mu, \phi, \theta) = 1$.
  - Probability of accepting β^* , μ^* , ϕ^* and θ^* can be written as

$$\min\left\{1, \frac{\hat{L}(\Delta \mathbf{y}_1, \ldots, \Delta \mathbf{y}_N | \beta^*, \mu^*, \phi^*, \theta^*)\beta^*\mu^*\phi^*\theta^*}{\hat{L}(\Delta \mathbf{y}_1, \ldots, \Delta \mathbf{y}_N | \beta, \mu, \phi, \theta)\beta\mu\phi\theta}\right\}.$$

Metropolis-Hastings sampler can work provided the estimated likelihood is unbiased.

LSDV-QML

LSDV

$$\hat{\beta}_{LSDV} = \frac{\sum_{i=1}^{N} \mathbf{y}_{i,(-1)} Q \mathbf{y}_i}{\sum_{i=1}^{N} \mathbf{y}_{i,(-1)} Q \mathbf{y}_{i,(-1)}}$$

where $Q = I - \iota\iota'/T$ and ι is a T-dimension vector of ones.

LSDV-QML

Assumptions

1. $|\beta| < 1$.
2. The initial values of y_{it} follow the steady state distribution $y_{i0} = \frac{\mu_i}{1-\beta} + \sum_{t=0}^{\infty} \beta^t \varepsilon_{i,-t}$.
3. The stochastic volatility is stationary: $|\phi| < 1$.
4. The initial values of $\log(\sigma_{it}^2)$ follow the steady state distribution $\log(\sigma_{i0}^2) = \mu + \sum_{t=0}^{\infty} \beta^t \eta_{i,-t}$.
5. $\{\epsilon_{it}\}$ $(i = 1, \ldots, N;\ t = 1, \ldots T)$ are i.i.d. variables across both time and individuals with $\mathrm{E}(\epsilon_{it}) = 0$ and $\mathrm{Var}(\epsilon_{it}) = 1$ and $\mathrm{E}(\epsilon_{it}^4) = \kappa < \infty$ and independent of μ_i and η_{it} for all i and t.
6. $\{\eta_{it}\}$ $(i = 1, \ldots, N;\ t = 1, \ldots T)$ are i.i.d. variables normally distributed across both time and individuals with $\mathrm{E}(\eta_{it}) = 0$ and $\mathrm{Var}(\eta_{it}) = \theta^2 < \infty$.

LSDV-QML

(LSDV Estimators: Stationary)
Under Assumptions 1-6, as $N \to \infty$ and $T \to \infty$,

$$\sqrt{NT}[\hat{\beta}_{LSDV} - (\beta - \frac{1}{T}(1+\beta))] \xrightarrow{d} N(0, B(\beta, \phi, \theta))$$

where $B(\beta, \phi, \theta) = (1-\beta^2)^2 \sum_{t=1}^{\infty} [\exp(\frac{\theta^2 \phi^t}{1-\phi^2})\beta^{2t-2}]$.

- Stochastic volatility changes asymptotic variances of the estimator while keeps unchanged asymptotic means.
- If $\phi > 0$ and $\theta > 0$, $B(\beta, \phi, \theta) > B(\beta, 0, 0)$ which is equal to $1 - \beta^2$ same as in Alvarez and Arellano (2003).
- When $\beta = 0.7$, for example, asymptotic variance is approximately 0.51 for the model without stochastic volatility, but can be 1.52 when $\phi = 0.9$ and $\theta = 0.5$.

LSDV-QML

Assumptions

7. The data generation process has a unit root: $\beta = 1$.
8. The initial values y_{i0} are fixed.

(LSDV Estimators: Unit Root)

Under Assumptions 3-8, as $N \to \infty$ and T is fixed,

$$\sqrt{N}(\hat{\beta}_{LSDV} - (1 - \frac{3}{T+1})) \xrightarrow{d} N(0, B_u(\phi, \theta))$$

where

$B_u(\phi, \theta) = \frac{36\left(2-5T+2T^2\right)}{5(-1+T)T(1+T)^3} \exp(\frac{\theta^2}{1-\phi^2})\kappa + \sum_{t=1}^{T-1} [\exp(\frac{\theta^2 \phi^t}{1-\phi^2}) C(t)]$

in which

$C(t) =$
$\frac{36\left(-9t^5+30t^4T-5t^3T(2+11T)+5t^2T\left(1+2T+13T^2\right)-2t\left(-2+5T+5T^2+20T^4\right)+T\left(-4+10T+5T^2+9T^4\right)\right)}{5(-1+T)^2T^2(1+T)^4}.$

LSDV-QML

- If $\phi > 0$ and $\theta > 0$, $B_u(\phi, \theta) > B_u(0, 0)$ which is equal to $\frac{3(17T^2-20T+17)}{5(T-1)(T+1)^3}$ as in Harris and Tzavalis (1999).
- When $T = 20$, for example, asymptotic variance is approximately 0.02 for the model without stochastic volatility, but can be 0.07 when $\phi = 0.9$ and $\theta = 0.5$.
- $B_u(\phi, \theta) > B_u(0, 0)$ implies that naively ignoring stochastic volatility would systematically reject Harris and Tzavalis' panel unit root test more frequently than it should be, although the estimates lose only a little statistical efficiency, so stochastic volatility corrected variance is needed.

LSDV-QML

- Estimate the error terms using the first-step estimator, i.e.,

$$\hat{\varepsilon}_{it} = y_{it} - \hat{\beta}_{LSDV} y_{i,t-1} - \hat{\mu}_i$$

where $\hat{\mu}_i = \frac{1}{T} \sum_{t=1}^{T} (y_{it} - \hat{\beta}_{LSDV} y_{i,t-1})$ is a consistent estimator of the individual effect.

- Calculate the quasi-likelihood of $\hat{\varepsilon}_{it}$ via Kalman filters in a state space form (Harvey and Shephard, 1996)

$$\log(\hat{\varepsilon}_{it}^2) = \log(\sigma_{it}^2) + \mathrm{E}(\log(\epsilon_{it}^2)) + \xi_{it}$$

$$\log(\sigma_{it}^2) = \mu + \phi(\log(\sigma_{i,t-1}^2) - \mu) + \eta_{it}$$

where $\mathrm{E}(\log(\epsilon_{it}^2)) = 1.27$ and ξ_{it} is a normal variable with mean zero and variance 4.93.

Panel VAR

$$y_{it} = B(L)y_{i,t-1} + \mu_i + \varepsilon_{it}$$

$$\varepsilon_{it} = V_{it}^{1/2}\epsilon_{it}$$

$$h_{it} = \mu_i + \Phi_i(h_{i,t-1} - \mu_i) + \eta_{it}$$

$$\begin{pmatrix}\epsilon_{it}\\ \eta_{it}\end{pmatrix} \sim N(0, \begin{bmatrix}\Sigma_\epsilon & 0\\ 0 & \Sigma_\eta\end{bmatrix}).$$

Panel VAR

2-variable vector

$$\mathbf{y}_{it} = B\mathbf{y}_{i,t-1} + \mu_i + \varepsilon_{it}$$

$$\begin{aligned}
\Delta \mathbf{y}_{it} &= \begin{bmatrix} 1 & 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 1 & 0 \end{bmatrix} \alpha_{it} \\
\alpha_{it} &= \begin{bmatrix} \beta_{11} & \beta_{12} & \beta_{11} & 0 & \beta_{12} & 0 \\ \beta_{21} & \beta_{22} & \beta_{21} & 0 & \beta_{22} & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} \alpha_{i,t-1} + \begin{bmatrix} 0 & 0 \\ 0 & 0 \\ 1 & 0 \\ -1 & 0 \\ 0 & 1 \\ 0 & -1 \end{bmatrix} \sigma_{it} \begin{bmatrix} \epsilon_{it,1} \\ \epsilon_{it,2} \end{bmatrix}
\end{aligned}$$

where the state vector $\alpha_{it} = \begin{bmatrix} \beta_{11}\Delta y_{i,t-1,1} + \beta_{12}\Delta y_{i,t-1,2} \\ \beta_{21}\Delta y_{i,t-1,1} + \beta_{22}\Delta y_{i,t-1,2} \\ \sigma_{it}\epsilon_{it,1} - \sigma_{it-1}\epsilon_{i,t-1,1} \\ -\sigma_{it}\epsilon_{it,1} \\ \sigma_{it}\epsilon_{it,2} - \sigma_{it-1}\epsilon_{i,t-1,2} \\ -\sigma_{it}\epsilon_{it,2} \end{bmatrix}$.

Monte Carlo Studies

$$y_{it} = \beta y_{i,t-1} + (1-\beta)\mu_i + \varepsilon_{it}$$

$$\mu_i = \sqrt{\tau}\left(\frac{q_i - 1}{\sqrt{2}}\right)\varsigma_i$$

$$\varepsilon_{it} = \sigma_{it}\epsilon_{it}$$

$$\log(\sigma_{it}^2) = \mu + \phi(\log(\sigma_{i,t-1}^2) - \mu) + \eta_{it}$$

where $q_i \sim \chi_1^2$; $\epsilon_{it}, \varsigma_i \sim N(0,1)$ and $\eta_{it} \sim N(0,\theta^2)$; q_i, ϵ_{it}, ς_i and η_{it} are all i.i.d. within series and also independent of each others.

- $\tau = 1$: the degree of cross-section to time-series variation.
- $\mu = \log(0.04)$: long-run mean of the log volatility.
- less persistent $\beta = 0.5$; persistent $\beta = 0.9$; unit root $\beta = 1$.
- persistent stochastic volatility: $\phi = 0.9$, $\theta = 0.5$; homoscedasticity : $\phi = \theta = 0$.

Simulation results of $\hat{\beta}$ when $\beta = 0.5$

T	N	ϕ	LSDV	IV	GMM	SGMM	PF
20	20	0	-0.082	0.005	-0.093	0.059	-0.079
			(0.095)	(0.193)	(0.110)	(0.190)	**(0.093)**
		0.9	-0.077	0.022	-0.093	0.050	-0.075
			(0.102)	(0.176)	(0.123)	(0.130)	**(0.095)**
	50	0	-0.087	0.011	-0.048	0.100	-0.085
			(0.090)	(0.110)	**(0.062)**	(0.131)	(0.088)
		0.9	-0.093	0.012	-0.070	0.047	-0.094
			(0.322)	(0.121)	**(0.095)**	(0.100)	(0.101)
50	20	0	-0.030	0.005	-0.035	-0.057	-0.025
			(0.044)	(0.075)	(0.048)	(0.249)	**(0.044)**
		0.9	-0.033	0.007	-0.039	-0.032	-0.028
			(0.053)	(0.094)	(0.058)	(0.177)	**(0.048)**
	50	0	-0.030	-0.001	-0.032	0.048	-0.025
			(0.035)	(0.047)	(0.037)	(0.106)	**(0.035)**
		0.9	-0.030	0.000	-0.036	0.031	-0.026
			(0.045)	(0.071)	(0.050)	(0.076)	**(0.041)**

Simulation results of $\hat{\beta}$ when $\beta = 0.9$

T	N	ϕ	LSDV	IV	GMM	SGMM	PF
20	20	0	-0.127	-0.227	-0.166	-0.110	-0.127
			(0.134)	(2.111)	(0.178)	(0.168)	**(0.133)**
		0.9	-0.129	-0.035	-0.168	-0.092	-0.117
			(0.141)	(0.803)	(0.185)	**(0.125)**	(0.128)
	50	0	-0.130	0.013	-0.127	-0.034	-0.126
			(0.131)	(0.160)	(0.138)	**(0.053)**	(0.127)
		0.9	-0.135	0.019	-0.150	-0.064	-0.124
			(0.138)	(0.164)	(0.163)	**(0.084)**	(0.127)
50	20	0	-0.046	0.007	-0.053	-0.295	-0.038
			(0.050)	(0.102)	(0.058)	(0.383)	**(0.045)**
		0.9	-0.049	0.010	-0.057	-0.182	-0.031
			(0.057)	(0.119)	(0.065)	(0.244)	**(0.040)**
	50	0	-0.045	-0.001	-0.054	-0.087	-0.039
			(0.046)	(0.058)	(0.056)	(0.108)	**(0.043)**
		0.9	-0.046	0.001	-0.060	-0.065	-0.027
			(0.049)	(0.088)	(0.063)	(0.082)	**(0.034)**

Simulation results of $\hat{\beta}$ when $\beta = 1$

T	N	ϕ	LSDV	GMM	SGMM	PF
20	20	0	-0.152	-0.221	-0.088	-0.148
			(0.155)	(0.230)	**(0.119)**	(0.152)
		0.9	-0.156	-0.233	-0.045	-0.152
			(0.163)	(0.247)	**(0.067)**	(0.158)
	50	0	-0.153	-0.223	-0.062	-0.147
			(0.154)	(0.233)	**(0.080)**	(0.149)
		0.9	-0.160	-0.261	-0.046	-0.157
			(0.164)	(0.273)	**(0.055)**	(0.159)
50	20	0	-0.063	-0.076	-0.235	-0.066
			(0.065)	(0.078)	(0.281)	(0.068)
		0.9	-0.066	-0.082	-0.138	-0.053
			(0.069)	(0.087)	(0.180)	**(0.055)**
	50	0	-0.061	-0.088	-0.082	-0.064
			(0.061)	(0.090)	(0.094)	(0.065)
		0.9	-0.061	-0.098	-0.053	-0.049
			(0.063)	(0.100)	(0.062)	**(0.051)**

Other simulation results

				LSDV-QML			PF		
β	T	N	ϕ	$\hat{\phi}$	$\hat{\theta}$	$\hat{\mu}$	$\hat{\phi}$	$\hat{\theta}$	$\hat{\mu}$
0.5	50	20	0	0.221	0.231	-0.001	0.271	0.116	-0.001
				(0.402)	(0.360)	(0.003)	(0.410)	(0.165)	(0.003)
			0.9	-0.004	-0.025	0.002	-0.027	0.072	0.003
				(0.042)	(0.133)	(0.008)	(0.050)	(0.121)	(0.009)
		50	0	0.251	0.200	-0.001	0.283	0.064	-0.001
				(0.413)	(0.302)	(0.002)	(0.412)	(0.089)	(0.002)
			0.9	0.004	-0.040	0.001	-0.013	0.037	0.001
				(0.030)	(0.087)	(0.004)	(0.028)	(0.066)	(0.006)

Other simulation results

Robustness of $\hat{\beta}$: the shocks have t-distribution with degree of freedom 5.

$\beta = 0.5$

T	N	ϕ	LSDV		PF	
			t	Normal	t	Normal
50	20	0	-0.032	-0.030	-0.023	-0.025
			(0.042)	(0.044)	(0.039)	(0.044)
		0.9	-0.038	-0.033	-0.030	-0.028
			(0.078)	(0.053)	(0.045)	(0.048)
	50	0	-0.032	-0.030	-0.023	-0.025
			(0.036)	(0.035)	(0.034)	(0.035)
		0.9	-0.043	-0.030	-0.027	-0.026
			(0.073)	(0.045)	(0.041)	(0.041)

Other simulation results

$$y_{it} = \beta y_{i,t-1} + \alpha x_{it} + (1-\beta)\mu_i + \varepsilon_{it}$$

$$x_{it} = q x_{i,t-1} + \zeta_{it}$$

$N = 50,\ T = 50,\ \beta = 0.5,\ \alpha = 0.7,\ \phi = 0.9,\ \theta = 0.5,$
$\mu = \log(0.04),\ q = 0.9,\ \zeta_{it} \sim N(0, 0.01)$

	$\hat{\beta}$	$\hat{\alpha}$	$\hat{\phi}$	$\hat{\theta}$	$\hat{\mu}$
LSDV-QML	-0.027	0.025	-0.004	-0.020	0.001
	(0.038)	(0.045)	(0.030)	(0.087)	(0.005)
PF	-0.030	-0.062	-0.025	-0.004	0.008
	(0.042)	(0.071)	(0.046)	(0.070)	(0.009)