A Spacial Autoregressive Stochastic Frontier Model with Inefficiency Heterogeneity for Panel Data

Shirong Zhao

July 16, 2019

Outline

Introduction

The global financial crisis has given rise to reexamine the performance of the banking sector.

- Pre-crisis: bigger is better
- Post-crisis: too-big-to-fail

Previous literature ignored the spacial spillover effect.

Research Question

A bank holding company (BHC) is a company that controls one or more banks, but does not necessarily engage in banking itself.

- Does there exist spacial spillovers in the U.S. BHC during global financial crisis?
- Controlling for the spacial spillovers, what determines the cost/profit inefficiency of BHC?

Stochastic Frontier Model

First introduced independently by Aigner et al. (1977), and Meeusen and van Den Broeck (1977).

$$\begin{aligned} y_i &= \boldsymbol{x_i}\boldsymbol{\beta} + \nu_i - u_i, \\ \nu_i &\sim N(0, \sigma^2), \\ i &= 1, 2 \ldots, N \end{aligned} \tag{SFM}$$

- y_i: the (logged) cost/profit
- x_i: the (logged) output, input prices and or environmental variables
- ν_i: two-sided, idiosyncratic error
- u_i: one-sided error, measuring efficiency ($\exp(-u_i)$): exponential, truncated normal, half-normal, gamma, ...

SFM with Heterogeneous Inefficiency for Panel Data

$\boldsymbol{z_{it}}$ are observed covariates vector that affects inefficiency.

$$
\begin{aligned}
y_{i,t} &= \boldsymbol{x_{i,t}}\boldsymbol{\beta} + \nu_{i,t} - u_{i,t}, \\
\nu_{i,t} &\sim N(0, \sigma^2), \\
u_{i,t} &\sim \exp(\lambda), \\
\lambda &= \exp(\boldsymbol{z_{i,t}}\gamma) \\
i = 1, 2, \ldots, N, &\quad t = 1, 2, \ldots, T,
\end{aligned}
\tag{SFMH}
$$

SFM with Spacial Autoregressive Effect

Glass et al. (2016) proposed the spacial autoregressive stochastic frontier model, similar as following :

$$
\begin{aligned}
y_{i,t} = \rho \sum_{j=1}^{N} W_{i,j} y_{j,t} + \boldsymbol{x_{i,t}} \boldsymbol{\beta} + \nu_{i,t} - u_{i,t}, \\
\nu_{i,t} \sim N(0, \sigma^2), \\
u_{i,t} \sim \exp(\lambda) \\
i = 1, 2, \ldots, N, \quad t = 1, 2, \ldots, T,
\end{aligned} \tag{SAR}
$$

W: a neighborhood structure, measuring geographic or economic distances

I propose the following SAR model with heterogeneous inefficiency for panel data:

$$
\begin{aligned}
y_{i,t} = \boldsymbol{x_{i,t}}\boldsymbol{\beta} + \rho \sum_{j=1}^{N} W_{i,j} y_{j,t} + \nu_{i,t} - u_{i,t}, \\
\nu_{i,t} \sim N(0, \sigma^2), \\
u_{i,t} \sim \exp(\lambda), \\
\lambda = \exp(\boldsymbol{z_{i,t}}\gamma) \\
i = 1, 2, \ldots, N, \quad t = 1, 2, \ldots, T,
\end{aligned}
\tag{SARH}
$$

Estimation Method

Bayesian inference techniques organized around Markove Chain Monte Carlo (MCMC) are applied.
The model is completed by specifying the following priors:

$$\begin{aligned} \boldsymbol{\beta} &\sim N(\mathbf{0}, \sigma^2 \boldsymbol{R}), \\ \rho &\sim Unif(\frac{1}{r_{min}}, 1), \\ \sigma^{-2} &\sim \Gamma(a_s, b_s), \\ \gamma &\sim N(\mathbf{0}, \boldsymbol{R_1}) \end{aligned} \tag{1}$$

Where r_{min} is the most negative real characteristic root of $W = (W_{i,j})$ to ensure the above equation (SARH) stable.

Posterior Distribution

$$
\begin{aligned}
p(\boldsymbol{\beta}, \sigma^2, \rho, u, \gamma | X, y) &\propto p(X, y | \boldsymbol{\beta}, \sigma^2, \rho, u, \gamma) p(u|\gamma) p(\boldsymbol{\beta}|\sigma^2) p(\sigma^2) p(\rho) p(\gamma) \\
&\propto (\sigma^{-2})^{\frac{NT}{2}} \exp[-\frac{(y^* - X\beta + u)'(y^* - X\beta + u)}{2\sigma^2}] \\
&\times (\exp(z_{it}\gamma))^{NT} \exp(-\exp(z_{it}\gamma) u' 1_{NT}) I(u \geq 0) \\
&\times (\sigma^{-2})^{\frac{p}{2}} \exp(-\frac{\beta' \boldsymbol{R}^{-1} \beta}{2\sigma^2}) \\
&\times (\sigma^{-2})^{a_s - 1} \exp(-\sigma^{-2} b_s) \\
&\times I(\rho \in (\frac{1}{r_{min}, 1})) \\
&\times \exp(-\frac{\gamma' \boldsymbol{R_1}^{-1} \gamma}{2})
\end{aligned} \tag{2}
$$

where $y^* = (I_{NT} - \rho I_T \otimes W_N) y$, and $\otimes$ is kronecker product.

Conditional Posterior Distribution

$$\boldsymbol{\beta}|else \sim N((X'X + R^{-\mathbf{1}})^{-\mathbf{1}} X'(y+u), \sigma^{\mathbf{2}}(X'X + R^{-\mathbf{1}})^{-\mathbf{1}}) \tag{3}$$

$$\sigma^{-\mathbf{2}}|else \sim \Gamma(\frac{NT+p}{2} + a_s, \frac{(y^* - X\beta + u)'(y^* - X\beta + u)}{2} + b_s + \frac{\beta' R^{-\mathbf{1}}\beta}{2}) \tag{4}$$

$$u|else \sim N(-(y^* - X\beta + \sigma^{\mathbf{2}} \exp(z_{it}\gamma)\mathbf{1}_{NT}), \sigma^{\mathbf{2}}\mathbf{1}_{NT}) \tag{5}$$

$$\rho|else \sim N(-(y'(IW)'(IW)y)^{-\mathbf{1}} y'(IW)'(y - X\beta + u), (y'(IW)'(IW)y)^{-\mathbf{1}}\sigma^{\mathbf{2}}) I(\rho \in (\frac{1}{r_{min}}, 1)) \tag{6}$$

$$p(\gamma|else) \propto (\exp(z_{it}\gamma))^{NT} \exp(-\exp(z_{it}\gamma)u'\mathbf{1}_{NT}) I(u \geq 0) \exp(-\frac{\gamma' R_{\mathbf{1}}\gamma}{2}) \tag{7}$$

where $IW = I_T \otimes W_N$.

Monte Carlo Experiments

Simulation data are generated according to Equation (SARH).

		— **SARH** —	
Param	True	Estimate	95%HPD
β_1	1	1.1893	[0.7146, 1.5705]
β_2	4	4.0469	[3.8602, 4.2376]
ρ	0.5	0.5048	[0.4949, 0.5168]
σ^2	1	0.9128	[0.5039, 1.4246]
γ_1	-2	-2.0264	[-2.1191, -1.9155]
γ_2	1	0.9719	[0.8835, 1.0754]

Monte Carlo Experiments

		— SAR —		— SFMH —		— SFM —	
Param	True	Estimate	95%HPD	Estimate	95%HPD	Estimate	95%HPD
β_1	1	2.1784	[1.8159, 2.5630]	-7.6708	[-8.6454,6.7861]	-5.0684	[-5.8480,-3.9995]
β_2	4	4.0403	[3.7775, 4.2874]	4.5026	[3.8493,5.1452]	4.6834	[3.9951,5.3196]
ρ	0.5	0.5091	[0.5002, 0.5189]	NA	NA	NA	NA
σ^2	1	0.4372	[0.1750, 0.7045]	15.9938	[10.5209,20.8060]	8.7776	[4.9596,12.7367]
γ_1	-2	NA	NA	-2.4489	[-2.5855,-2.3336]	NA	NA
γ_2	1	NA	NA	0.7995	[0.6833,0.8993]	NA	NA
λ	NA	0.0708	[0.0645, 0.0769]	NA	NA	0.0511	[0.0466,0.0563]

Data and Variables

Source: Federal Reserve FR Y-9C Bank Holding Company Reports

- Quarterly data from 1986 to 2018
- Effective March 31, 2006, the FRY-9C filing threshold was increased from $150 million to $500 million or more
- I use yearly end data over 2006-2014
- Remove BHC with NA and negative equity/profit

Year	No. of BHC
2006	290
2007	277
2008	218
2009	203
2010	230
2011	249
2012	268
2013	285
2014	298

- No. of BHC in banlanced panel: 143

Variable Description

X_1	Sum of deposits, federal funds purchased, trading liabilities and other borrowed money
X_2	Number of full-time equivalent employees at end of current period
X_3	Premises and fixed assets (including capitalized leases)
Y_1	Loans and lease financing receivables
Y_2	Securities: held-to-maturity and available-for-sale
Y_3	Total non-interest income
Z_1	Total equity capital
Z_2	Total assets
Z_3	Allowance for loan and lease losses (NPL)
W_1X_1	Total interest expense
W_2X_2	Salaries and employee benefits
W_3X_3	Expenses of premises and fixed assets
TC	$W_1X_1 + W_2X_2 + W_3X_3$
TP	Net income (loss) attributable to holding company

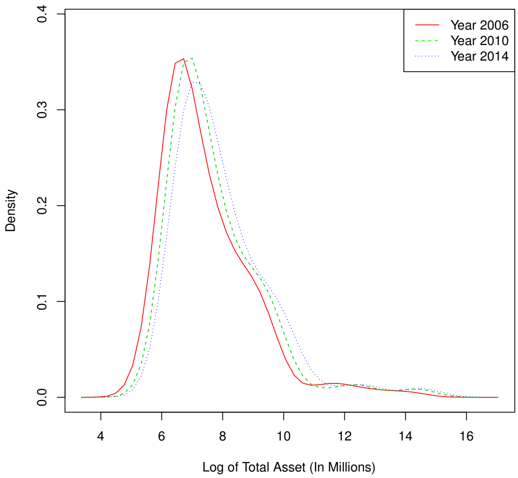
Inefficiency Determinants

The inefficiencies are assumed to depend upon covariates as in Tecles and Tabak (2010), they are:

- Market share of loans (MS)
- Non-performing loans/Asset (NPL)
- Equity/Asset (EQUITY)
- Asset size (SIZE)
- ...

$$\lambda = \exp(\gamma_1 MS + \gamma_2 NPL + \gamma_3 EQUITY + \gamma_4 SIZE + \dots) \tag{8}$$

Density of (log) Total Assets



Summary Statistics for Year 2014 (Dollar Amount in Millions)

Statistic	N	Mean	St. Dev.	Min	Pctl(25)	Pctl(75)	Max
Funds and Deposits (X_1)	143	35,279.47	217,535.10	379.56	824.74	5,072.30	2,134,102.00
Labor (X_2)	143	5,581.61	30,121.13	95	224	1,353.5	264,452
Fixed Asset (X_3)	143	295.91	1,355.71	2.63	17.39	87.91	12,936.00
Loans (Y_1)	143	18,602.25	100,025.80	158.28	560.06	3,790.52	865,677.00
Securities (Y_2)	143	7,444.89	39,989.74	2.53	259.54	1,715.29	346,901.00
Non-interest income (Y_3)	143	850.48	5,489.55	1.97	7.99	75.62	51,394.00
Equity (Z_1)	143	4,278.75	25,136.01	45.04	103.63	775.45	231,727.00
Asset (Z_2)	143	41,664.43	259,375.40	434.24	933.69	5,826.67	2,572,773.00
NPL (Z_3)	143	287.98	1,618.71	1.44	8.67	47.02	14,185.00
Total Cost (TC)	143	844.03	5,089.92	7.91	22.51	123.30	46,005.00
Total Profit (TP)	143	451.51	2,698.06	1.72	9.78	74.25	23,057.00

Neighborhood Matrix W

We have two possible definitions:

- distance between pairs of BHC (Glass and Kenjegalieva (2019)))
- k nearest neighbor (KNN)

Row sums of W are equal to 1.

SARH with Time-varying Spacial Spillovers

The changes of ρ_t can be interpreted as contagion.

$$\begin{aligned} y_{i,t} &= \boldsymbol{x_{i,t}}\beta + \rho_t \sum_{j=1}^{N} W_{i,j} y_{j,t} + \nu_{i,t} - u_{i,t}, \\ \nu_{i,t} &\sim N(0, \sigma^2), \\ u_{i,t} &\sim \exp(\lambda), \\ \lambda &= \exp(\boldsymbol{z_{i,t}}\gamma) \\ i = 1, 2, \ldots, N, \quad t &= 1, 2, \ldots, T, \end{aligned} \tag{9}$$

SAR with Inefficiency Spillovers

Inefficiency could also be spatially correlated. That is to say, inefficiency of one unit could also be affected by its neighbours.

$$
\begin{aligned}
y_{i,t} &= \boldsymbol{x_{i,t}}\boldsymbol{\beta} + \rho \sum_{j=1}^{N} W_{i,j} y_{j,t} + \nu_{i,t} - u_{i,t}, \\
\nu_{i,t} &\sim N(0, \sigma^2), \\
u_{i,t} &= \theta \sum_{j=1}^{N} W_{i,j} u_{j,t} + \epsilon_{i,t} \\
\epsilon_{i,t} &\sim \exp(\lambda), \\
i = 1, 2, \dots, N, &\quad t = 1, 2, \dots, T
\end{aligned}
\tag{10}
$$

Spatial-Temporal Stochastic Frontier Model

Inefficiency could also be time correlated. That is to say, there is adjustment cost for firms.

$$\begin{aligned} y_{i,t} &= \boldsymbol{x_{i,t}}\boldsymbol{\beta} + \nu_{i,t} - u_{i,t}, \\ \nu_{i,t} &\sim N(0, \sigma^2), \\ u_{i,t} &= \exp(\boldsymbol{z_{i,t}}\boldsymbol{\lambda} + \epsilon_{i,t}), \\ \boldsymbol{\epsilon_t} &= (\epsilon_{1,t}, \epsilon_{2,t}, \ldots, \epsilon_{N,t})^T, \\ \boldsymbol{\epsilon_t} = \theta\boldsymbol{\epsilon_{t-1}} + \boldsymbol{\phi_t}, \quad \text{for} \quad t &= 2, 3, \ldots, T, \\ \boldsymbol{\epsilon_1} &= \boldsymbol{\phi_1}, \\ \boldsymbol{\phi_t} \sim N(0, \tau^2(\boldsymbol{D} - \rho\boldsymbol{W})^{-1}), \quad \text{for} \quad t &= 1, 2, \ldots, T \end{aligned} \tag{11}$$

Aigner, D., C. Lovell, and P. Schmidt (1977), Formulation and estimation of stochastic frontier production function models, *Journal of Econometrics* 6, 21–37.
URL: *http://linkinghub.elsevier.com/retrieve/pii/0304407677900525*

Glass, A. J. and K. Kenjegalieva (2019), A spatial productivity index in the presence of efficiency spillovers: Evidence for U.S. banks, 1992–2015, *European Journal of Operational Research* 273, 1165–1179.
URL: *http://www.sciencedirect.com/science/article/pii/S0377221718307574*

Glass, A. J., K. Kenjegalieva, and R. C. Sickles (2016), A spatial autoregressive stochastic frontier model for panel data with asymmetric efficiency spillovers, *Journal of Econometrics* 190, 289–300.
URL: *https://linkinghub.elsevier.com/retrieve/pii/S0304407615001803*

Meeusen, W. and J. van Den Broeck (1977), Efficiency Estimation from Cobb-Douglas Production Functions with Composed Error, *International Economic Review* 18, 435–444.
URL: *https://www.jstor.org/stable/2525757*

Tecles, P. L. and B. M. Tabak (2010), Determinants of bank efficiency: The case of Brazil, *European Journal of Operational Research* 207, 1587–1598.
URL: *http://www.sciencedirect.com/science/article/pii/S0377221710004133*