Smoothed GMM for SAQR

HU Yanan, QU Xinhao

Department of Statistics, Business School, Zhengzhou University

December 27, 2020

LEAD-IN

Chinese New Year is coming...

LEAD-IN

- Exogenous spatial effect*(SLX) ;
- Endogenous spatial effect (SAR) exemplified through taxing and spending of regional finance (Brueckner, 2003) ;
- Spatial effect between random variables*(SEM) ;
- Particularly interested in SAR and SEM as well as *SAC in Econometrics

The basic setting of spatial autoregression model (SAR) :

$$y_n = \rho W_n y_n + X_n \beta + \varepsilon_n,\ \varepsilon_n \sim N(0, \sigma^2 I_n) \tag{1}$$

- Built upon the spatial lag to show endogenous spatial effect
- Definiting adjacency by rules of Rook or Queen, to name but a few
- Certifying each weight through indicator function, inversed distance function (with or without boundary) and so on
- Conditions for stationarity in Kelejian and Prucha (1998, 1999), Lee (2004)
- * The choice of weight matrix from Zhang and Yu (2017)

The reduced form of SAR :

$$y_n = (I_n - \rho W_n)^{-1} X_n \beta + (I_n - \rho W_n)^{-1} \varepsilon_n \tag{2}$$

- Assumed that $(I_n - \rho W_n)^{-1}$ is nonsingular, same with Lee (2004)
- The rising obstacle of endogeneity as

$$E\left(y_n \varepsilon_n'\right) = (I_n - \rho W_n)^{-1} E(\varepsilon_n \varepsilon_n') \neq 0$$

- Thus leading to bias and inconsistency of the estimation

INTRODUCTION

Intuitive method of instrument variable (IV) by Kelejian (1998) and Lee (2003) :

$$X_n, W_nX_n, W_n^2X_n, W_n(I_n - \rho W_n)^{-1}X_n, \ldots \tag{3}$$

- Fit the hypothesis of being strictly exogenous
- Highly-correlated with the spatial lag
- * Related to the selection of the spatial weight matrix, which could be done through a Mallows-type criterion (Zhang and Yu, 2017)

BLOCK-OUT

The setting of mean regression may be doubtful:

- (Information Loss) Only reflecting the expected information instead of the whole distribution
- (Loss-Robust) Heterogeneity and highly-possible outliers in spatial observations
- (Loss of Generality) Heavy-tailed distribution of ε frequently emerges especially for spatial data
  (Anselin, 1988; Glaeser *et al.*, 1996; LeSage, 1999)
- ......

MODEL-SETTING

Thus leading to SAQR naturally :

$$Q_\tau(y_i \mid x_i, z_i) = \rho(\tau) \sum_{j=1}^{n} w_{ij} y_j + x_i' \beta(\tau), \ i \neq j \in \{1, 2, ..., n\} \tag{4}$$

$$P\{\varepsilon_i \leq 0 \mid x_i, z_i\} = \tau$$

- Accurate depict of the conditional distribution
- Coefficient to be a nonrandom function of τ, more robust shown by Koenker (2005)
- Less constraints to ε leads to being more general
- * ρ remains the same between quantile points (Li and Fang, 2018)
- ......

PARAMETER-ESTIMATION

- Ordinary Method of IVQR (Horowitz, 1998) :

① Applying instrument variable to deal with endogeneity, similar with SAR;

② The loss function, namely check function in QR will be

$$L_\tau(\mu) \equiv \mu(\tau - I_{\mu<0})$$

asymmetric least absolute deviation loss (Chernozhukov and Hansen, 2008), shown to be more robust

③ Applying nonparametric smoothing method due to the existence of indicator function

* Bayesian Method (Yu and Moyeed, 2001)

④ Acting well in small sample while able to handle weak IV (when not highly-correlated with the replaced one)

* Generally leading to increasing asymptotic variance of the estimator

- Another Approach of SGMM

- Moment Conditon

① Based on De Castro *et al.* (2019) , intend to discover a moment condition through the assumption of ε by Law of Iterated Expectations (LIE)

$$E\left[g_i^u(\theta(\tau),\tau)\right] \equiv E\left(z_i\left[I\left\{\varepsilon_i \leq 0\right\} - \tau\right]\right) = 0 \tag{5}$$

The corresponding sample moment function will be

$$\hat{M}_n^u(\theta(\tau),\tau) = \frac{1}{n}\sum_{i=1}^{n} g_i^u(\theta(\tau),\tau) \tag{6}$$

② Applying nonparametric method, similar with smoothing check function

$$\hat{M}_n(\theta(\tau),\tau) = \frac{1}{n}\sum_{i=1}^{n} g_{ni}(\theta(\tau),\tau) = \frac{1}{n}\sum_{i=1}^{n}\left\{z_i\left[\widetilde{I}\left(-\frac{\varepsilon_i}{h_n}\right) - \tau\right]\right\} \tag{7}$$

- *Bandwidth Selection

a. Intuitively from the definition of derivative in Calculus when estimating PDF nonparametricly (Chapter.27 Chen, 2010)

$$\hat{f}\left(x_0\right) = \frac{1}{nh}\sum_{i=1}^{n} K\left[\left(x_i - x_0\right)/h\right] \tag{8}$$

b. Bias and variance are all related to $o(h^r)$

$$Bias = O(h^2) \tag{9}$$

$$Var = o(1/nh)$$

* The selection of optimal bandwidth h^* shows more importance than kernel function $K(\cdot)$

c. Needs $nh \to \infty$ when $n \to \infty, h \to 0$ to satisfy consistency

- Another Approach of SGMM

- Optimal Bandwidth

③ Through Kaplan and Sun (2017) , the choice of bandwidth h_n is based on the minimization of MSE to SEE

$$MSE_{\text{SGMM}} \equiv E\left(m_n' V^{-1} m_n\right) \tag{10}$$

$m_n \equiv n^{-1/2} \sum_{i=1}^{n} g_{ni}(\theta(\tau), \tau)$ while V stands for the asymptotic variance

④ The optimal h_n (fits condition for consistency) is easily shown as

$$h^*_{\text{MSE}} = \arg\min_{h_n} nh_n^{2r}(EA)'(EA) - h_n tr\{E(BB')\} \tag{11}$$

$$A \equiv \left(\frac{1}{r!}\int_{-1}^{1} \widetilde{I}'(v)v^r dv\right) f^{(r-1)}_{\Lambda|z}(0 \mid z_i) V^{-1/2} z_i$$

$$B \equiv \left(1 - \int_{-1}^{1} \widetilde{I}^2(u)du\right)^{1/2} \left[f_{\Lambda|z}(0 \mid z_i)\right]^{1/2} V^{-1/2} z_i$$

* Intuitively, $h^*_{\text{MSE}} = h^*_{\text{CPE}}$ (minimization of higher-order type I error)

- Another Approach of SGMM

- Comparison

  ⑤ Compared with smoothing check function
  a. easier to establish high-order results due to less terms in the first order condition
  b. smaller bias ($r+1$ times, $r \geq 2$)

  ⑥ Compared with unsmoothed moment function
  a. reduced asymptotic variance under integral conditions of $\widetilde{I}(\cdot)$
  b. smaller asymptotic MSE when $h_n = h^*_{\text{MSE}}$

- Another Approach of SGMM

- Large Sample Properties

⑦ Expression could be shown, following the strategy in Chapter.8 Hong (2011)

$$\hat{\theta}_{\text{SGMM}}(\tau) = \arg\min_{\theta\in\Theta} \hat{M}_n(\theta(\tau),\tau)^\top \overline{\Omega}^{-1} \hat{M}_n(\theta(\tau),\tau) \tag{12}$$

$$\overline{\Omega} \equiv \frac{1}{n}\sum_{i=1}^{n} g_{ni}(\overline{\theta}(\tau),\tau)\, g_{ni}(\overline{\theta}(\tau),\tau)^\top \tag{13}$$

⑧ Fitting the properties of consistency, asymptotic normality and validity through Theorem 5.7&5.9 Vaart (1994)

$$\hat{\theta}_{\text{SGMM}}(\tau) \xrightarrow{P} \theta_0(\tau) \tag{14}$$

$$\sqrt{n}(\hat{\theta}_{\text{SGMM}}(\tau) - \theta_0(\tau)) \xrightarrow{d} N(0,\Lambda) \tag{15}$$

Λ stands for $(G^\top\Omega\, G)^{-1}G^\top\Omega\,\Sigma_\tau\Omega\, G\,(G^\top\Omega\, G)^{-1}$

- Data Generating Process (DGP)

① Using measurements of Bias and RMSE

② Establishing spatial weight matrix by rule of Rook

③ Considering the distribution of random variables to be i.i.d. Gaussian or heavy-tailed $t(3)$ with or without heteroscedasticity

表 3 扰动项独立同分布于标准正态分布

		Bias				RMSE			
τ	N	ρ	α	β_1	β_2	ρ	α	β_1	β_2
0.1	20	0.0905	-0.6331	0.0507	0.0959	0.7238	3.3960	0.6293	0.2644
	50	0.0503	-0.3674	0.0339	0.0361	0.3647	1.7152	0.3413	0.1637
	100	-0.0040	-0.0561	0.0142	0.0138	0.2495	1.0882	0.2392	0.1050
	500	0.0035	-0.0389	-0.0017	0.0057	0.0941	0.4103	0.1011	0.0463
0.3	20	0.0586	-0.2460	-0.0129	0.0132	0.5262	2.2964	0.4718	0.1987
	50	0.0405	-0.2003	0.0017	0.0088	0.2690	1.1883	0.2564	0.1205
	100	0.0111	-0.0499	0.0052	0.0014	0.1762	0.7832	0.1811	0.0835
	500	0.0005	-0.0076	0.0001	0.0019	0.0697	0.3107	0.0766	0.0359
0.5	20	0.0584	-0.2291	0.0051	0.0002	0.5269	2.3475	0.4436	0.1937
	50	0.0456	-0.1636	0.0011	-0.0041	0.2518	1.0899	0.2396	0.1162
	100	0.0231	-0.0802	-0.0056	-0.0046	0.1610	0.7077	0.1733	0.0790
	500	0.0004	-0.0057	-0.0001	0.0008	0.0655	0.2951	0.0722	0.0339

- Foreign Trade Agglomeration

Based on Wei (2011), adding spatial lag as well as seven exogenous variables since spatial correlation remains

$$TRA_i = \rho(\tau)\sum_{j=1}^{n} w_{ij} TRA_j + \alpha(\tau) + \beta(\tau) X_i + \varepsilon_i,\ i \neq j \in \{1, 2, ..., n\}$$

Result shows:

① “Siphon Effect” as well as “Mathew Effect” actually remain;
② INDUS and TRANS factors show negative impact at lower quantile points while the PGDP remains the opposite, corresponding with Wei (2011);
③ Sign of FDI may be doubtful in Wei (2011) when comes to middle and lower quantile level;
④ Results related to parameters of GOV and TECH may be wrong from the perspective of Coombs (1987)

CONCLUSION

- Applying SAQR as cross-sectional dependence, heterogeneity and possible outliers frequently appear in spatial data;
- Discovering moment function as well as smoothing it instead of check function since the formal embraces several advantages;
- Choosing the optimal bandwidth under the criterion of MSE;
- Demonstrating good properties of large sample under assumptions;
- Showing faster speed in calculation and verified large sample properties through DGP;
- Application on foreign trade agglomeration comes up with several new angles different from the previous study

Thank you!

Smoothed GMM for SAQR

HU Yanan, QU Xinhao

Department of Statistics, Business School, Zhengzhou University

December 27, 2020