Skewness-based test diagnosis of technical inefficiency in spatial autoregressive stochastic frontier models

Ming-Yu Deng[1], Levent Kutlu[2*], Mingxi Wang[1]

[1]School of International Trade and Economics, University of International Business and Economics, Beijing, 100029, China.
[2]Department of Economics, University of Texas Rio Grande Valley, TX, USA.

*Corresponding author(s). E-mail(s): levent.kutlu@utrgv.edu; Contributing authors: my_deng@stu.ecnu.edu.cn; mxwang@uibe.edu.cn;

Abstract

In the spatial autoregressive (SAR) stochastic frontier model, the inefficiency term, as the salient feature that differs from the SAR model, can capture the effects of technical inefficiency. For whether the inefficiency exists significantly, this paper proposes a skewness-based test, which does not rely on the disturbance's normality specification and allows the inefficiency to follow an unknown one-sided distribution. Monte Carlo simulations show that our test is robust against non-normal disturbances and performs satisfactorily in different one-sided distributions for inefficiency. An application to the data on 137 dairy farms in Northern Spain is considered to illustrate the existence of technical inefficiency in production by our test.

Keywords: Spatial stochastic frontier, Technical inefficiency, Hypothesis testing, Skewness

JEL Classification: C12 , C21 , R32

1 Introduction

Stochastic frontier (SF) analysis, established almost simultaneously by Aigner et al (1977) and Meeusen and van den Broeck (1977) for measuring productivity, has been widely applied to handle different problems, and one can see Fried et al (2008) for a survey of the literature. Actually, firms usually tend to concentrate in clusters to facilitate imitation and improvement in production (Baptista, 2000; Galli, 2023). Therefore, by introducing the spatial lag term of the dependent variable (Cliff and Ord, 1973, 1981, named SAR term) that captures spatial correlation across production units, the distribution-based SAR SF (SARSF) model is developed by Glass et al (2016) to analyse (in)efficiency (spillovers), and later amplified by Kutlu (2018).[1]

Recently, there is a growing interest in the SARSF model (Kutlu et al, 2020; Lai and Tran, 2022; Kutlu, 2022, 2023), which takes the following form,

$$Y_n = \lambda W_n Y_n + X_n \beta - U_n + V_n, \tag{1}$$

where n is the number of producers, $Y_n \in \mathbb{R}^n$ is the observation on the objective variable, λ and $W_n \in \mathbb{R}^{n\times n}$ are successively a spatial lag parameter and spatial weight matrix, $X_n \in \mathbb{R}^{n\times k}$ is the observation on frontier variables with coefficient $\beta \in \mathbb{R}^k$, $V_n = [v_{n1}, \cdots, v_{nn}]^\top$ is the random disturbance that intends to capture the effects of statistical noise, and $U_n = [u_{n1}, \cdots, u_{nn}]^\top \geq \mathbf{0}_n$ is the inefficiency term ($\mathbf{0}_n$ denotes the n-dimensional zero-vector).

Notably, the inefficiency term U_n, intending to capture the effects of technical inefficiency specific to each unit (Meintanis and Papadimitriou, 2022), is the salient feature of the SARSF model that differs from the conventional SAR one. Therefore, testing the significant level of such a term in model (1) is meaningful, which is equivalent to testing whether all firms are fully efficient. Actually, such hypothesis testing is also valuable in SF analysis, pioneered by Schmidt and Lin (1984), followed by Coelli (1995). The introduction of the SAR term $\lambda W_n Y_n$ in (1), however, makes this topic in spatial SF analysis different from that in SF counterpart.

In view of this, this paper mainly concentrates on testing whether the inefficiency term is present in (1), which aims to complement the growing literature on the SARSF model.[2] To the best of our knowledge, Jin and Lee (2020) is the first to consider this problem, by proposing the score test.[3] However, their score test seems to rely on both the existence of the intercept term and the normal specification for the disturbance term strongly, and our subsequent simulations portray that the test is usually invalid once anyone is not satisfied. Moreover, besides the half-normal in Aigner et al (1977) that is specified

[1] Indeed, spatial SF analysis, as an emerging spatial regression technique in the last two decades, started with Druska and Horrace (2004), where a spatial panel error SF model is developed with fixed effects to calculate time-invariant efficiency.

[2] As stated in Glass et al (2016) and Kutlu (2018), the related literature is rather sparse since the establishment of the distribution-based SARSF model, especially in terms of statistical inference.

[3] Jin and Lee (2020) also derive the likelihood ratio test, whose statistic is complex computationally since it involves simultaneously estimation of the restricted and unrestricted SARSF models.

as what the inefficiency in model (1) follows, u_{ni}'s also can be suppose to follow the Exponential (Meeusen and van den Broeck, 1977), Gamma (Greene, 1990), and truncated-normal (Stevenson, 1980) distributions, even a mixture of them. For example, consider the case where u_{ni}'s follow a gamma distribution, say $\sim Ga(\alpha, \gamma)$, (or a truncated-normal distribution, say $u_{ni} \sim N^+(\mu, \sigma_u^2)$), and a test of the hypothesis that there is no technical inefficiency refers to the joint test that $\alpha = \gamma = 0$ (or $\mu = \sigma_u = 0$, more details can be found in Coelli (1995)). This indicates that different one-sided distributions correspond to different score tests within the maximum likelihood framework.

In view of this, this paper suggests a general test based on the skewness of the composed error, which is adaptive to different one-sided distributions in the SARSF model and therefore provides a useful testing tool in the empirical study. In our methodology, on the one hand, the distributional assumption regarding the disturbance term is relaxed from normality to symmetry around zero, which can avoid the severity of the distributional specification giving rise to the composed error (Kopp and Mullahy, 1990). on the other hand, the inefficiency term is allowed to follow any one-sided distribution. Therefore, our setting is general and nests distribution-based SARSF models (Glass et al, 2016; Kutlu et al, 2020; Lai and Tran, 2022) as a special case. Especially, we conclude that only in the scenario where the disturbance is normal and an intercept is included in model (1), does the derived statistic degenerates to what the one in SF analysis is.

The remainder of this paper is organized as follows. Section 2 introduces the skewness-based test and model specifications. Section 3 establishes the asymptotic theory of our proposed test statistic. Section 4 investigates the finite sample performance of the proposed test via Monte Carlo simulations. Section 5 provides an empirical example to illustrate how our test works and Section 6 concludes. Appendix provides more technical details.

2 The skewness-based test and model specifications

In this section, we mainly consider a skewness-based test for whether all firms are fully efficient significantly in the SARSF model, i.e., the null and alternative hypotheses are

$$H_0 : U_n = \mathbf{0}_n \text{ in (1)} \quad \text{v.s.} \quad H_1 : U_n > \mathbf{0}_n \text{ in (1)}. \tag{2}$$

In what follows, define the composed error as $\epsilon_n = [\epsilon_{n1}, \cdots, \epsilon_{nn}]^\top$ with $\epsilon_{ni} = -u_{ni} + v_{ni}$. u_{ni}'s and v_{ni}'s are generally supposed to be, respectively, half-normal and normal (Glass et al, 2016; Kutlu, 2018; Jin and Lee, 2020; Kutlu, 2023). Here we relax both as follows.

Assumption 1. (i) v_{ni}'s follow a symmetric i.i.d. with mean zero and variance $\sigma_{v,o}^2$. Moreover, $\mathrm{E}|v_{ni}|^{6+\iota} < \infty$ for some $\iota > 0$. (ii) $U_n \geq \mathbf{0}_n$ follows an unknown one-sided distribution. (iii) u_{ni} and v_{ni} are independent of each other and of $\mathbf{x}_{i,n}$, which is the i^{th} observations on frontier variables.

The reason for we relax the distributional assumption regarding V_n from normality to symmetry in Assumption 1(i) are 2-fold. Besides the aforementioned one in the introduction, Assumption 1(i) makes the calculation of the third moment $\mathrm{E}[(v_{ni} - u_{ni})^3]$ irrelevant to the distributional parameter of v_{ni}. Accordingly, H_0 means $\epsilon_{ni} = v_{ni}$ is symmetric. Equivalently, H_0 makes the distributional skewness of ϵ_{ni} equal zero, and vice versa. The relaxation in Assumption 1(ii) means that the half-normal distribution is not necessary for U_n, which can follow anyone of the aforementioned one-sided distributions, even the endogeneity resulted from the correlation between U_n and V_n arises (Kutlu et al, 2020). Assumption 1(iii) summarizes the exogeneity of explanatory variables.

Under Assumption 1, to test H_0 we consider a test based on the skewness of ϵ_{ni}. Unfortunately, ϵ_n is not observable, so the statistic based on the skewness is not feasible in practise. This motivates us to predict ϵ_n by the restricted residual, say $\tilde{\epsilon}_n$, generated from the restricted model (1) under H_0, i.e., the SAR model $Y_n = \lambda W_n Y_n + X_n\beta + V_n$. The estimation theory on the SAR counterpart is approaching maturity, such as (quasi-)maximum likelihood estimation (Lee, 2004, MLE), and the methods of 2-stage least square (Kelejian and Prucha, 1998, 2SLS) and GMM (Lee, 2007), among others. Particularly, these strategies involve the quadratic/linear forms in ϵ_{ni}. Here we no longer discuss this aspect but directly suppose that $\theta_* = [\lambda_*, \beta_*^\top]^\top$, the restricted value of $\theta = [\lambda, \beta^\top]^\top$, is consistently estimated by $\tilde{\theta}_n = [\tilde{\lambda}_n, \tilde{\beta}_n^\top]^\top$, which satisfies the following assumption.

Assumption 2. (i) $n^{1/2}(\tilde{\theta}_n - \theta_*) = g_n(\theta_*) + o_p(1)$, where $g_n(\theta_*)$ is quadratic forms in ϵ_{ni} such that $g_n(\theta_*) \sim AN(0, \Sigma_*)$, where Σ_* is positive definite. (ii) The parameter space Θ of θ is compact.

Given $\tilde{\theta}_n$, the sample skewness of $\tilde{\epsilon}_n = Y_n - \tilde{\lambda}_n W_n Y_n - X_n\tilde{\beta}_n$ is

$$\begin{aligned} S_n(\tilde{\epsilon}_n) &= \frac{1}{n}\sum_{i=1}^{n}\left(\epsilon_{ni} - \frac{1}{n}\sum_{i=1}^{n}\epsilon_{ni}\right)^3 \cdot \left[\frac{1}{n}\sum_{i=1}^{n}(\epsilon_{ni} - \frac{1}{n}\sum_{i=1}^{n}\epsilon_{ni})^2\right]^{-3/2} \\ &\triangleq \phi\left(n^{-1}\sum_{i=1}^{n}\tilde{\epsilon}_{ni}, n^{-1}\sum_{i=1}^{n}\tilde{\epsilon}_{ni}^2, n^{-1}\sum_{i=1}^{n}\tilde{\epsilon}_{ni}^3\right), \end{aligned} \tag{3}$$

where $S_n(\cdot)$ and $\phi(a,b,c) = (c - 3ab - 2a^3)/(b - a^2)^{3/2}$. Then the skewness-based test statistic can be constructed as

$$S_n(\tilde{\epsilon}_n)/\mathrm{Asy.var}[S_n(\tilde{\epsilon}_n)]^{1/2}, \tag{4}$$

where the asymptotic variance $\mathrm{Asy.var}[S_n(\tilde{\epsilon}_n)]^{1/2}$ indicates obtaining the joint limiting distribution $[n^{-1}\sum_{i=1}^n \tilde{\epsilon}_{ni}, n^{-1}\sum_{i=1}^n \tilde{\epsilon}_{ni}^2, n^{-1}\sum_{i=1}^n \tilde{\epsilon}_{ni}^3]$ is a prerequisite. After that, the former follows from the well-known Delta method (Billingsley, 2012). To this end, the Laws of Large Numbers (LLN) in (Jenisha and Prucha, 2012, Theorem 1) for spatial mixing and near-epoch dependence (NED) processes is useful (see Appendix A for details.), thereby the following basic regular assumptions are necessary.

Assumption 3. Individual units in the economy are located or living in a region $D_n \subset D \subset \mathbb{R}^d$, where D is a (possibly) unevenly spaced lattice, the cardinality $|D_n|$ of a finite set D_n satisfies $\lim_{n\to\infty} |D_n| = \infty$. The distance $d(i,j)$ between any two different individuals i and j is larger than or equal to a constant, which is assumed to be 1 for convenience.

Assumption 4. $c_1 \equiv \lambda_m \sup_n \|W_n\|_\infty < 1$, and $[-\lambda_m, \lambda_m]$ is the compact parameter space of λ on the real line. In addition, at least one of the following two conditions (a) and (b) holds:

(a) Only individuals whose distances are less than or equal to some positive constant d_o may affect each other directly, i.e., $w_{n,ij} \neq 0$ only if $d(i,j) \leq d_o$.
(b) (i) For every n, the number of columns $\mathbf{w}_{n,j}$ of W_n with $|\lambda_o| \sum_{i=1}^n |w_{n,ij}| > c_1$ is less than or equal to some fixed nonnegative integer that does not depend on n. (ii) There exists an $\alpha > d$ and a constant c_2 such that $|w_{n,ij}| \leq c_2/d(i,j)^\alpha$.

Assumption 5. (i) $\{\mathbf{x}_{i,n}\}_{i=1}^n$ is α-mixing random field with α-mixing coefficient $\alpha(u,v,s) \leq (u+v)^{c_3}\hat{\alpha}(s)$ for some $c_3 > 0$, where $\hat{\alpha}(s)$ satisfies $\sum_{s=1}^\infty s^{d-1}\hat{\alpha}(s) < \infty$. (ii) $\sup_{1\leq j\leq k,i,n} \mathrm{E}|x_{ij,n}|^{4+\iota} < \infty$ for some $\iota > 0$.

Assumption 3, originated in Jenisha and Prucha (2009, 2012), allows the spaced lattice to be high-dimensional as a subset of $\mathbb{R}^d$, because the distance between any two spatial units can be a geometrical distance, an economic distance, or a mixture of both. Assumption 4, originated in Xu and Lee (2015), considers two different kinds of W_n. Assumption 4(a) only allows direct interaction for any two units when their distance is not bigger than the specific number d_o. For any different units i and j, Assumption 4(b)(ii) requires their interactions to decline geometrically fast as the interactions exist. Moreover, by imposing a constraint on the column sums of W_n in absolute value, Assumption 4(b)(i) guarantees that the number of spatial units having large aggregated effects on other units is fixed. The mixing coefficient for the random field $\{\mathbf{x}_{i,n}\}$ in Assumption 5 does not only depend on the distance between two separate subsets of spatial units but also their sizes. Indeed, Assumptions 3–5, being standard for specifying increasing domain asymptotics in spatial NED processes, is natural in regional research, such as Jenisha and Prucha (2009, 2012), Xu and Lee (2015, 2018), Liu and Lee (2019), Jin and Lee (2020), and Jin and Wang (2023), among others.

3 Asymptotic theory

In this section, we derive the joint asymptotic distribution, $n^{1/2}[n^{-1}\sum_{i=1}^n \tilde{\epsilon}_{ni}, n^{-1}\sum_{i=1}^n \tilde{\epsilon}_{ni}^2, n^{-1}\sum_{i=1}^n \tilde{\epsilon}_{ni}^3]$, whereby the limiting distribution of $n^{1/2}S_n(\tilde{\epsilon}_n)$ also can be derived by the Delta method.

We first show the probability limits of $n^{-1}\sum_{i=1}^n \tilde{\epsilon}_{ni}$, $n^{-1}\sum_{i=1}^n \tilde{\epsilon}_{ni}^2$, and $n^{-1}\sum_{i=1}^n \tilde{\epsilon}_{ni}^3$ under H_0. Let $G_n = W_n(I_n - \lambda_* W_n)^{-1}$ and $\sigma_{\epsilon,o}^2 = \mathrm{E}[\epsilon_{ni}^2]$. By the expansion, we have

$$\frac{1}{n}\sum_{i=1}^n \tilde{\epsilon}_{ni} = \frac{1}{n}\sum_{i=1}^n \epsilon_{ni} + \frac{1}{n}\sum_{i=1}^n \mathbf{z}_{i,n}^\top (\theta_* - \tilde{\theta}_n), \tag{5}$$

where $\mathbf{z}_{i,n} = [\mathbf{w}_{i,n}Y_n, \mathbf{x}_{i,n}^\top]$ with $\mathbf{w}_{i,n}$ being the i^{th} row of W_n. By the Lindeberg-Lévy CLT, $n^{-1}\sum_{i=1}^n \epsilon_{ni} = O_p(n^{-1/2})$ since $\mathrm{E}[\epsilon_{ni}] = \mathrm{E}[v_{ni}] = 0$ under H_0. By Lemma 2 (in Appendix A), $\mathbf{w}_{i,n}Y_n$ is uniformly L_r-norm bounded for some $r = 4 + \iota > 4$ and is uniformly L_2-NED on $\{\mathbf{x}_{i,n}, \epsilon_{ni}\}_{i=1}^n$ under Assumptions 1–5, and so does $\{\mathbf{z}_{i,n}\}$. This means that conditions in the LLN (Lemma 1 in Appendix A) of Jenisha and Prucha (2012) hold,[4] thereby $n^{-1}\sum_{i=1}^n \mathbf{z}_{i,n} = \mathrm{E}[n^{-1}\sum_{i=1}^n \mathbf{z}_{i,n}] + o_p(1) = O_p(1)$. Working with $\theta_* - \tilde{\theta}_n = o_p(1)$, it yields that $n^{-1}\sum_{i=1}^n \mathbf{z}_{i,n}^\top(\theta_* - \tilde{\theta}_n) = o_p(1)$. Therefore, the probability limit of $n^{-1}\sum_{i=1}^n \tilde{\epsilon}_{ni}$ is same as that of $n^{-1}\sum_{i=1}^n \epsilon_{ni}$, and those of $n^{-1}\sum_{i=1}^n \tilde{\epsilon}_{ni}^2$ and $n^{-1}\sum_{i=1}^n \tilde{\epsilon}_{ni}^3$ also can be proved. Based on these, we have the proposition as shown below.

Proposition 1. *Under Assumptions 1–5 and H_0, we have $n^{-1}\sum_{i=1}^n \tilde{\epsilon}_{ni} = n^{-1}\sum_{i=1}^n \tilde{\epsilon}_{ni}^3 = o_p(1)$, and $n^{-1}\sum_{i=1}^n \tilde{\epsilon}_{ni}^2 - \sigma_{\epsilon,o}^2 = o_p(1)$.*

After obtaining the three probability limits in Proposition 1, we can analyse the joint asymptotic distribution, $\left[n^{-1}\sum_{i=1}^n \tilde{\epsilon}_{ni}, n^{-1}\sum_{i=1}^n \tilde{\epsilon}_{ni}^2, n^{-1}\sum_{i=1}^n \tilde{\epsilon}_{ni}^3\right]$. Under H_0, it follows from the mean value that

$$\begin{aligned}
\frac{1}{n}\sum_{i=1}^n \tilde{\epsilon}_{ni} &= \frac{1}{n}\sum_{i=1}^n \epsilon_{ni} - \frac{1}{\sqrt{n}}\left[\frac{\mathbf{1}_n^\top G_n X_n \beta_*}{n}, \frac{\mathbf{1}_n^\top X_n}{n}\right]\sqrt{n}(\tilde{\theta}_n - \theta_*) + o_p(\frac{1}{\sqrt{n}}), \\
\frac{1}{n}\sum_{i=1}^n \tilde{\epsilon}_{ni}^2 &= \frac{1}{n}\sum_{i=1}^n \epsilon_{ni}^2 - \frac{2\sigma_{\epsilon,o}^2}{\sqrt{n}}\left[\frac{1}{n}\mathrm{tr}(G_n), \mathbf{0}_{1\times k}\right]\sqrt{n}(\tilde{\theta}_n - \theta_*) + o_p(\frac{1}{\sqrt{n}}), \\
\frac{1}{n}\sum_{i=1}^n \tilde{\epsilon}_{ni}^3 &= \frac{1}{n}\sum_{i=1}^n \epsilon_{ni}^3 - \frac{3\sigma_{\epsilon,o}^2}{\sqrt{n}}\left[\frac{\mathbf{1}_n^\top G_n X_n \beta_*}{n}, \frac{\mathbf{1}_n^\top X_n}{n}\right]\sqrt{n}(\tilde{\theta}_n - \theta_*) + o_p(\frac{1}{\sqrt{n}}),
\end{aligned} \tag{6}$$

which shall be checked later. Proof for multivariate normality needs the Cramér–Wold device. Let $\mathbf{c} = [c_1, c_2, c_3]^\top \in \mathbb{R}^3$, and it is equivalent to consider the asymptotic distribution of

$$\begin{aligned}
\mathbf{c}^\top n^{-1}\left[\sum_{i=1}^n \tilde{\epsilon}_{ni}, \sum_{i=1}^n \tilde{\epsilon}_{ni}^2, \sum_{i=1}^n \tilde{\epsilon}_{ni}^3\right] =& \frac{c_3}{n}\sum_{i=1}^n \epsilon_{ni}^3 + \frac{c_2}{n}\sum_{i=1}^n \epsilon_{ni}^2 + \frac{c_1}{n}\sum_{i=1}^n \epsilon_{ni} \\
&+ \frac{1}{\sqrt{n}}\left[\mathbf{f}(c_1, c_2, c_3)\sqrt{n}(\tilde{\theta}_n - \theta_*) + o_p(1)\right],
\end{aligned} \tag{7}$$

[4] L_p-NED also implies L_q-NED for any $q < p$.

where $\mathbf{f}(\cdot) \in \mathbb{R}^{1\times(k+1)}$ is defined as

$$\mathbf{f}(c_1, c_2, c_3) = -\frac{1}{n}(c_1 + 3c_3\sigma^2_{\epsilon,o}) \left[\mathbf{1}_n^\top G_n X_n \beta_* + \frac{2c_2\sigma^2_{\epsilon,o}\mathrm{tr}(G_n)}{c_1 + 3c_3\sigma^2_{\epsilon,o}}, \mathbf{1}_n^\top X_n \right].$$

With the hep of Proposition 1 and Assumption 2, the asymptotic distribution of (7) can be investigated, and so does the following joint asymptotic distribution.

Proposition 2. *Under Assumptions 1–5 and H_0,*

$$\sqrt{n}\left(n^{-1}\left[\sum_{i=1}^n \tilde{\epsilon}_{ni}, \sum_{i=1}^n \tilde{\epsilon}^2_{ni}, \sum_{i=1}^n \tilde{\epsilon}^3_{ni} \right]^\top - \left[0, \sigma^2_{\epsilon,o}, 0\right]^\top \right) \sim AN(0, \lim_{n\to\infty} \Sigma_n), \quad (8)$$

where the limit of Σ_n (in Appendix B) is assumed to exist.

Although Σ_n seems to be complicated computationally, in practice, we take empirical numbers that are easy to compute instead of Σ_n. Proposition 2 can help derive the asymptotic distribution of $S_n(\tilde{\epsilon}_n)$ via the Delta approach.

Proposition 3. *Under Assumptions 1–5 and H_0, $n^{1/2}S_n(\tilde{\theta}_n) \sim AN\left(0, \lim_{n\to\infty} \mathbf{r}^\top \Sigma_n \mathbf{r}\right)$, where $\mathbf{r} = [-3/\sigma_{\epsilon,o}, 0, 1/\sigma^3_{\epsilon,o}]^\top$.*

The above proposition specifies the test statistic in (4) as follows, meanwhile a special form corresponding to the normal case is provided.

Corollary 1. *Under Assumptions 1–5 and H_0, the skewness-based test statistic in (4) becomes*

$$\mathrm{T} = \sqrt{n}S_n(\tilde{\epsilon}_n) \cdot (\tilde{\mathbf{r}}^\top \Sigma_n \tilde{\mathbf{r}})^{-1/2}, \quad (9)$$

which is distributed as $N(0, 1)$, where $\tilde{\mathbf{r}} = [-3/\tilde{\sigma}_\epsilon, 0, 1/\tilde{\sigma}^3_\epsilon]^\top$ with $\tilde{\sigma}_\epsilon = (n^{-1}\sum_{i=1}^n \tilde{\epsilon}^2_{ni})^{1/2}$. Especially, Suppose that v_{ni}'s are normal and the regressor X_n includes an intercept. Then, (9) simplifies to

$$\mathrm{T} = \frac{S_n(\tilde{\epsilon}_n)}{\sqrt{6/n}}. \quad (10)$$

Corollary 1 derive a general skewness-based test statistic, regard less of the one-sided distribution that U_n follows and of whether V_n is normal. (10) show us that only in the case where an intercept is included in model (1) and V_n is normal, does such a statistic is identically to that in SF analysis (Coelli, 1995).

Remark 1. Note that the sum of a symmetric distribution minus a one-sided (positive) counterpart must suffer from negative skewness (lack of distribution symmetry) in particular. Therefore, the test should be left-sided because $S_n(\tilde{\epsilon}_n) < 0$ under H_1 theoretically. However, $S_n(\tilde{\epsilon}_n)$ may be positive in practise, especially in the case of small sample size. This indicates the "wrong skewness" problem appears.

4 Monte Carlo simulations

We set the first data generated process (DGP) of the SARSF model as

- DGP-I: $Y_n = (I_n - \lambda_o W_n)^{-1}(\mathbf{1}_n\beta_{0,o} + \mathbf{x}_1\beta_{1,o} + \mathbf{x}_2\beta_{2,o} - U_n + V_n)$,

where n is either 100 or 400, $\mathbf{x}_1$ and $\mathbf{x}_2$ are generated from $N(\mathbf{0}_n, I_n)$, and slopes $(\beta_{0,o}, \beta_{1,o}, \beta_{2,o}) = (0.5, 1, 1.5)$. Given that the distribution of u_{ni} is one-sided and that of v_{in} is symmetric, we consider four cases listed in Table 1, where the spatial weight matrix W_n and the spatial lag parameter λ_o are specified. To test H_0, we construct three statistics from (9), in which θ is successively estimated by the quasi-MLE (QMLE), 2SLS, and GMM methods. Denote the corresponding tests as SK-QMLE, SK-GMM, and SK-2SLS tests. Meanwhile, we compare them with the score test (Jin and Lee, 2020, Proposition 2.4(b)). Moreover, to investigate the finite sample performance as the intercept is removed in model (1), we also consider the following DGP-II in Case I (normality)

- DGP-II: $Y_n = (I_n - \lambda_o W_n)^{-1}(\mathbf{x}_1\beta_{1,o} + \mathbf{x}_2\beta_{2,o} - U_n + V_n)$,

where except the intercept, all the other settings are fully identical to those in Case I of Table 1. Set $U_n = \mathbf{0}_n$ for obtaining these tests' sizes, and let κ_u be chosen gradually away from zero for getting the related powers (the nominal size is set to 5%). Nominal and empirical sizes are plotted in Figures 1–2, and empirical powers are plotted in Figures 3–4. All results are based on 5000 replications for each case, and the regressors are randomly redrawn for each repetition.

Table 1 Settings for DGP-I and -II.

Case	Disturbance: v_{in}	Inefficiency: u_{in}	λ_o	W_n
I	$N(0, \sqrt{0.2})$	$N^+(0, \kappa_u)$	-0.5	Queen criterion
II	$0.5 \cdot T(4)$	$Exp(\kappa_u)$	-0.5	Queen criterion
III	$U(-\sqrt{3}, \sqrt{3})$	$Ga(\kappa_u, 1)$	0.5	Rook criterion
IV	$N(0, \sqrt{0.5}) + 0.5 \cdot T(4)$ $+\sqrt{0.5} \cdot U(-\sqrt{3}, \sqrt{3})$	$N^+(1, \kappa_u)$	0.5	Rook criterion

Note: κ_u is a parameter. T, U, and Exp signify Student's t, Uniform,and Exponential distributions. Queen and Rook criteria can be found in Kelejian and Piras (2017, p. 8).

4.1 Size of the test

Figure 1 report the sizes for DGP-I. Our main findings are as follows. We can see from Figure 1 that in either case, plots of the SK-QMLE, SK-GMM, and SK-2SLS tests' sizes are very close to the diagonal line, which indicates their empirical sizes are approximately equal to the nominal sizes (say α). In addition, (i) only in Case I (normality) does the score test have similar

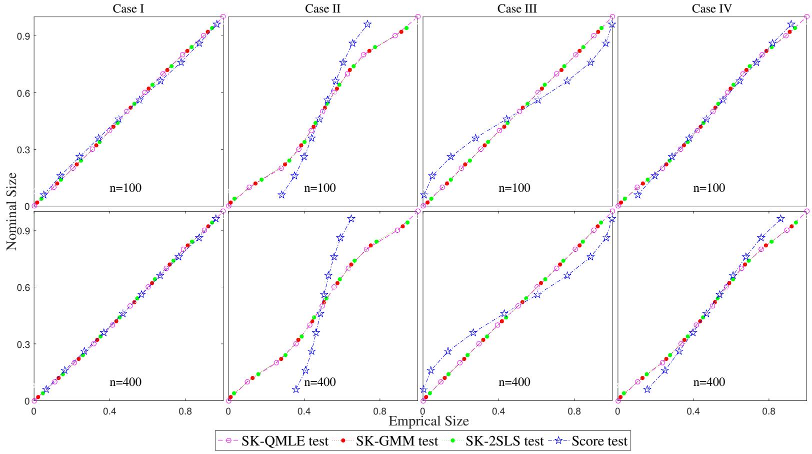

Fig. 1 Plots of Empirical and nominal sizes for DGP-I.

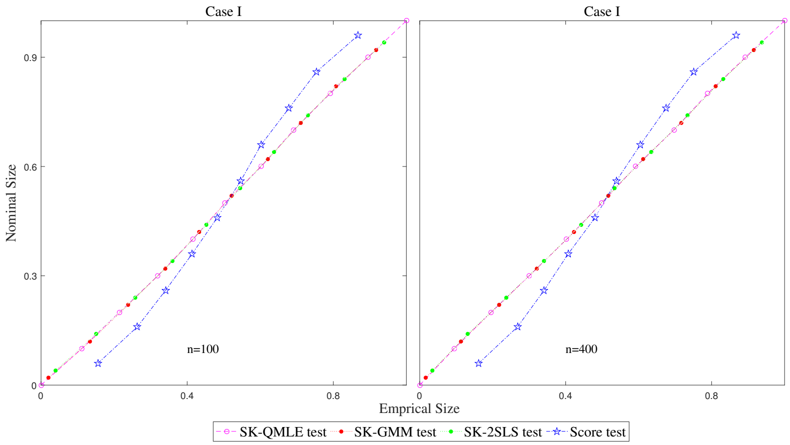

Fig. 2 Plots of Empirical and nominal sizes for DGP-II.

performance to our tests; (ii) in Cases II–III, two plots for the score test strongly deviate from the diagonal line with different deviation directions; (iii) for $n = 100$ of Case IV, our tests also outperform the score one, whose plot lies slightly below the diagonal line meaning slightly inflated type I errors when $\alpha \in (0, 0.5)$, and lies slightly up the diagonal line indicating slightly deflated type I errors when $\alpha \in (0.5, 1)$; the degree of the deviation in Case IV seems to be larger as the sample size n gets larger from 100 to 400.

Turning to Figure 2, which plots the sizes for DGP-II. It is found from the first column figures that plot lies below the diagonal line corresponding to inflated type I errors when $\alpha \in (0, 0.5)$, and lies up the diagonal line corresponding to deflated type I errors when $\alpha \in (0.5, 1)$. This indicates that the test is invalid when the intercept term is removed in the SARSF model, even if v_{ni}'s are normal.

4.2 Power of the test

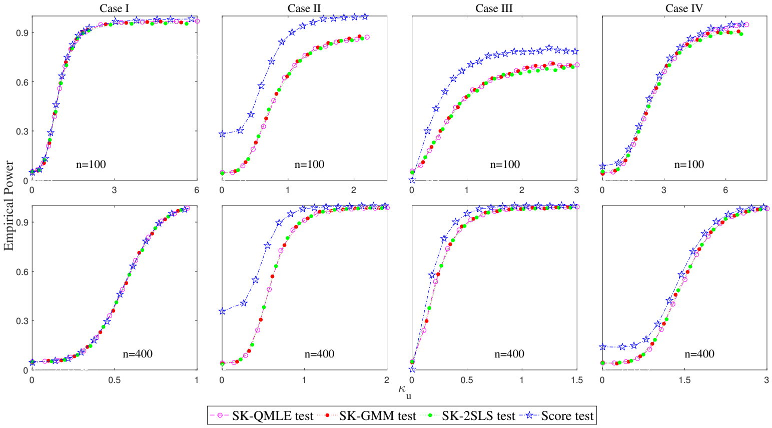

Fig. 3 Plots of Empirical powers for DGP-I.

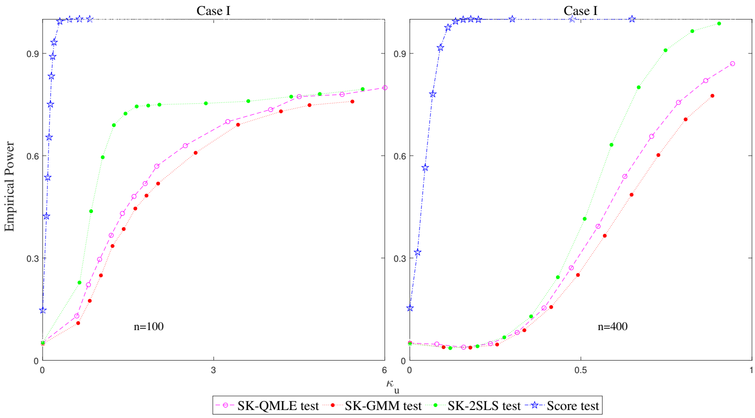

Fig. 4 Plots of Empirical powers for DGP-II.

For the empirical powers, the correct starting points (i.e., $\kappa_u = 0$) of all plots in Figures 3–4 should correspond to 5% of type I error because we default $\alpha = 0.05$. Figure 3 summarizes that in Case I, the score test has the right plot of empirical powers, because the initial point of such a plot corresponds to $\alpha = 5\%$, and so do the triple skewness-based tests. Moreover, the figure also portrays that the score test tends to over-reject in Cases II and IV since the initial point corresponds to $\alpha > 0.05$, and to reject not enough in Cases III since the initial point corresponds to $\alpha < 0.05$. Therefore, the pattern in Figure 3 conforms to that in Figure 1. From Figure 4, it turns out that the score test tends to over-reject since the initial point corresponds to $\alpha > 0.05$, which is in line with the pattern in Figure 2.

In summary, to test the existence of technical inefficiency in the SARSF model, Our skewness-based tests are robust against non-normal disturbances and adaptive for different one-sided distributions that the inefficiency term follows. However, the score test's performance is only satisfactory in the case where the intercept exists and the disturbance follows a normal distribution.

5 An empirical illustration

Using the dataset of Amsler et al (2016), we analyzed a balanced panel data on 137 dairy farms in Northern Spain over a 12-year period (1999-2010). We employed translog production function where the output (MILK) is milk production in kiloliters, while the five inputs are LABOR (number of man-equivalent units), COWS (number of milking cows), FEED (total amount of feedstuffs fed to the dairy cows in tonne), LAND (hectares of land devoted to pasture and crops), and ANIMALEXP (animal expenses in 1000 Euros: veterinary, milking costs, medicines, etc.). The Our production function also includes 11 dummy variables for years, in additional to an intercept. The descriptive statistics for the data is given in Table 2.

Table 2 Descriptive statistics.

Variables	N	Mean	Std. Dev.	25^{th} Perc	Median	75^{th} Perc
MILK (y)	1,644	330.64	207.28	186.53	274.83	425.58
LABOR (x_1)	1,644	1.79	0.80	1.13	1.96	2.00
COWS (x_2)	1,644	42.20	21.18	28.00	38.00	52.00
FEED (x_3)	1,644	153.67	102.66	85.82	126.08	197.66
LAND (x_4)	1,644	19.31	8.85	14.00	18.00	24.00
ANIMALEXP(x_5)	1,644	16.39	14.71	7.50	12.85	20.95

The spatial weight matrix W_n is calculated using a negative exponential distance model with row-normalization. That is, the weights are calculated as follows: $w_{ij} = \exp\{-d_{ij}\}/\sum_{k\neq j}\exp\{-d_{kj}\}$ if $i \neq j$ and $w_{ii} = 0$, where d_{ij} is the distance (in 100km) between the zones of farm i and j.

The test statistics and corresponding p-values (in parenthesis) for MLE, GMM, and 2SLS are -1.956 (0.025), -1.960 (0.025), and -1.989 (0.023), respectively. Hence, at 5% statistical significance level, we conclude that inefficiency is present. The results are robust to imposing 25km, 50km, 100km, 200km, and 400km thresholds distances. We also ran the tests for the W_n obtained by scaling the un-normalized weighting matrix by its maximum eigenvalue. The test statistics and corresponding p-values (in parenthesis) for MLE, GMM, and 2SLS are -2.100 (0.018), -2.100 (0.018), and -2.101 (0.018), respectively. This is in line with our findings based on the row-normalized weight matrix.

Based on our test results, we estimated the following spatial stochastic frontier production model:

$$\begin{aligned}\ln(y_{it}) =& \lambda \sum_j w_{n,ij} \ln(y_{jt}) + \beta_0 + \delta_t + \sum_k \beta_k \ln(x_{k,it}) \\ &+ \frac{1}{2} \sum_{k,l} \beta_k \ln(x_{k,it}) \ln(x_{l,it}) - u_{it} + v_{it},\end{aligned}$$

where y_{it} is the output (MILK_{it}) for farm $i = 1, \cdots, N$ at time $t = 1, \cdots, T$, $x_{k,it}$ is the input k (in the following order: LABOR_{it}, COWS_{it}, FEED_{it}, LAND_{it}, ANIMALEXP_{it}), $w_{n,ij}$ is the $(i, j)^{\text{th}}$ component of the weight matrix W_n, λ is the spatial lag parameter, β's are the frontier parameters, δ_t represents the time fixed-effect, $v_{it} \sim N(0, \sigma_v^2)$ is the statistical noise and $u_{it} \sim N^+(0, \sigma_u^2)$ is the inefficiency term has half-normal distribution . The estimates for the inefficiency term are calculated by:

$$\begin{aligned}\hat{u}_{it} =& \mathrm{E}\left[u_{it} \mid \hat{\varepsilon}_{it}\right] \\ \hat{\varepsilon}_{it} =& \ln(y_{it}) - \Bigg[\hat{\lambda} \sum_j w_{n,ij} \ln(y_{jt}) + \hat{\beta}_0 + \hat{\delta}_t + \sum_k \hat{\beta}_k \ln(x_{k,it}) \\ &+ \frac{1}{2} \sum_{k,l} \hat{\beta}_k \ln(x_{k,it}) \ln(x_{l,it})\Bigg],\end{aligned}$$

where the notation employed for denoting the estimate of a corresponding parameter involves placing a hat symbol on top of the parameter. The spatial lag parameter is statistically significant at any conventional significance level (p-value=0.0000) and equals 0.15416. Hence, spatial interactions are present for the farms in our dataset. The histogram of the inefficiency estimates, $\hat{u}_{it}$, from this model is provided in Figure 5. The mean, median, and standard deviation of $\hat{u}$ are 0.1009, 0.0872, and 0.0516, respectively.

In Figure 6, we also present the spillover corrected efficiency estimates (Kutlu, 2018). The mean, median, and standard deviation of the spillover-corrected efficiency estimates are 90.46%, 91.60%, and 4.49%, respectively. Hence, while the efficiency estimates are high, we still have some inefficiency, which is in line with our test results.

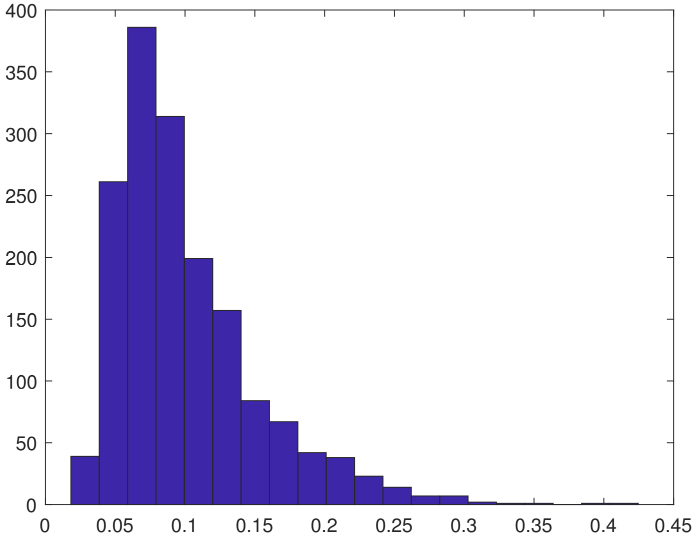

Fig. 5 Histogram of Inefficiency Term (u) Estimates.

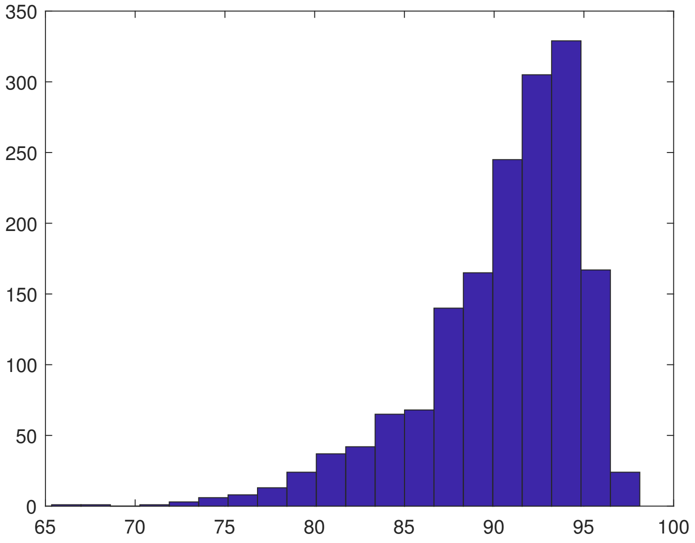

Fig. 6 Histogram of Spillover-corrected Efficiency Estimates.

6 Concluding remark

For distribution-based SARSF models, this paper derives a skewness-based test statistic for whether the inefficiency term exists. Our settings relax the distributional assumption regarding the disturbance, and allow the inefficiency term to follow an unknown one-sided distribution. Monte Carlo simulations suggest that our test performs satisfactorily in the presence of non-normal disturbances and different one-sided distributions for inefficiency. These may push forward the application of SARSF models in spatial SF analysis.

What one-sided distribution for the inefficiency component can fit empirical data best is a classical topic in SF analysis, and there are a series of literature on this topic, such as Wang et al (2011), Chen and Wang (2012), Amsler et al (2019) and Meintanis and Papadimitriou (2022), among others. But this goodness of fit test remains to be researched in the SARSF model, which is also our future direction.

Acknowledgments

This work is supported by the National Natural Science Foundation of China (No. 71971062), the Natural Science Foundation of Beijing (No. 9202015), and the Excellent Young Scholar Foundation of UIBE (No. 21JQ01).

Appendix A. Useful definition and lemma

Definition 1 (L_p-norm and L_p-NED). *For any random variable t with a finite p^{th} absolute moment, where $p \geq 1$, denote its L_p-norm by $\|t\|_p = [\mathrm{E}|t|^p]^{1/p}$. Let $g = \{g_{ni}, i \in D_n, n \geq 1\}$ and $v = \{v_{ni}, i \in D_n, n \geq 1\}$ be two random fields, where D_n satisfies Assumption 3. Assume that $\sup_{i,n} \|g_{ni}\|_p < \infty$. The random field g is said to be L_p-NED on v if*

$$\|g_{ni} - \mathrm{E}\left(g_{ni} \mid \mathcal{F}_{ni}(s)\right)\|_p \leq d_{ni}\psi(s)$$

for some arrays of finite positive constants $\{d_{ni}, i \in D_n, n \geq 1\}$ and for some sequence $\psi(s) \geq 0$ such that $\lim_{s\to\infty} \psi(s) = 0$, where $\mathcal{F}_{ni}(s)$ is the σ-field generated by the random variables v_{nj} 's with units j 's located within the ball $B_i(s)$ that is centered at i with radius s, and $\psi(s)$ is the NED coefficient. If we further have $\sup_n \sup_{i\in D_n} d_{ni} < \infty$, then g is said to be uniformly L_p-NED on v.

Lemma 1 (Theorem 1 in Jenisha and Prucha (2012)). *Let $\{D_n\}$ be a sequence of arbitrary finite subsets of D such that $|D_n| \to \infty$ as $n \to \infty$, where $D \subset \mathbb{R}^d, d \geq 1$ is as in Assumption 3, and let T_n be a sequence of subsets of D such that $D_n \subseteq T_n$. Suppose further that $Z = \{Z_{i,n}, i \in D_n, n \geq 1\}$ is L_1-NED on $\varepsilon = \{\varepsilon_{i,n}, i \in T_n, n \geq 1\}$ with the scaling factors $d_{i,n}$. Suppose that Z and ε satisfy: (a) there exist nonrandom positive constants $\{c_{i,n}, i \in D_n, n \geq 1\}$ such that $Z_{i,n}/c_{i,n}$ is uniformly L_p-bounded for some*

$p > 1$, i.e., $\sup_n \sup_{i \in D_n} E\left|Z_{i,n}/c_{i,n}\right|^p < \infty$, and (b) the α-mixing coefficients of the input field ε satisfy $\bar{\alpha}(u,r,v) \leq \varphi(u,v)\hat{\alpha}(r)$ for some function $\varphi(u,v)$ which is nondecreasing in each argument, and some $\widehat{\alpha}(r)$ such that $\sum_{r=1}^{\infty} r^{d-1}\widehat{\alpha}(r) < \infty$. Then

$$\frac{1}{M_n\left|D_n\right|}\sum_{i \in D_n}\left(Z_{i,n} - EZ_{i,n}\right) \xrightarrow{L_1} 0,$$

where $M_n = \max_{i \in D_n} \max\left(c_{i,n}, d_{i,n}\right)$.

Lemma 2 (Proposition 1 in Xu and Lee (2015))**.** *Suppose the parameter space Θ of θ is compact. Consider the SAR model $Y_n = \lambda W_n Y_n + X_n\beta + V_n$.*
(1) Under Assumption 2, if $\sup_{1 \leqslant j \leqslant k, i, n} \mathrm{E}\left|x_{ij,n}\right|^p < \infty$ and $\sup_{i,n} \mathrm{E}\left|\epsilon_{ni}\right|^p < \infty$ for some $p \geqslant 1$, then $\{y_{ni}\}_{i=1}^n$ and $\{\mathbf{w}_{i,n}Y_n\}_{i=1}^n$ are uniformly L_p bounded.
(2) Under Assumptions 3-4(a) and 5, $\{y_{ni}\}_{i=1}^n$ is geometrically L_2-NED on $\{\mathbf{x}_{i,n}, v_{ni}\}_{i=1}^n$: $\left\|y_{ni} - \mathrm{E}\left[y_{ni} \mid \mathcal{F}_{ni}(s)\right]\right\|_2 \leqslant C\left(\zeta^{1/\bar{d}_0}\right)^s$ for some $C > 0$ that does not depend on i and n. The same conclusion also holds for $\{\mathbf{w}_{i,,n}Y_n\}_{i=1}^n$.
(3) Under Assumptions 3-4(b) and 5, $\{y_{ni}\}_{i=1}^n$ is L_2-NED on $\{\mathbf{x}_{i,n}, v_{ni}\}_{i=1}^n$: $\left\|y_{ni} - \mathrm{E}\left[y_{ni} \mid \mathcal{F}_{ni}(s)\right]\right\|_2 \leqslant C/s^{\alpha-d}$ for some $C > 0$ that does not depend on i and n. The same conclusion also holds for $\{\mathbf{w}_{i,n}Y_n\}_{i=1}^n$.

Lemma 3 (Generalization of Corollary 4.3(b) in Gallant and White (1988))**.** *If $\{Y_{i,n}\}$ and $\{Z_{i,n}\}$ are both uniformly L_{2r} bounded for some $r > 2$, and uniformly and geometrically L_2-NED, then $\{Y_{i,n}Z_{i,n}\}$ is uniformly and geometrically L_2-NED.*

Appendix B. Technical details

- Elements of Σ_n

Σ_n is a symmetric 3×3 matrix, whose $(i,j)^{\text{th}}$ element is denoted as $\delta_{ij,n}$ for $i,j \in \{1,2,3\}$. For any $j \in \{1,2,3\}$, define $C_{\epsilon,jn} = \mathrm{E}\left[n^{-1/2}\sum_{i=1}^n \epsilon_{ni}^j \cdot n^{1/2}(\tilde{\theta}_n - \theta_*)\right]$ as the $(k+1)$-dimensional expectation vector. Recall that $\sigma_{\epsilon,o}^2 = \mathrm{E}[\epsilon_{ni}^2]$ and $\mu_{\epsilon,s} = \mathrm{E}[\epsilon_{ni}^s]$ for $s = 4, 6$. For the diagonal elements of Σ_n,

$$
\begin{aligned}
\delta_{11,n} =& \sigma^2_{\epsilon,o} + \left[\frac{1}{n}\mathbf{1}_n^\top G_n X_n \beta_*, \frac{1}{n}\mathbf{1}_n^\top X_n\right] \Sigma_* \left[\frac{1}{n}\mathbf{1}_n^\top G_n X_n \beta_*, \frac{1}{n}\mathbf{1}_n^\top X_n\right]^\top \\
&- 2\left[\frac{1}{n}\mathbf{1}_n^\top G_n X_n \beta_*, \frac{1}{n}\mathbf{1}_n^\top X_n\right] C_{\epsilon,1n}, \\
\delta_{22,n} =& \mu_{\epsilon,4} - \sigma^2_{\epsilon,2} + 4\sigma^4_{\epsilon,o}\left[\frac{1}{n}\mathrm{tr}(G_n), \mathbf{0}_{1\times k}\right] \Sigma_* \left[\frac{1}{n}\mathrm{tr}(G_n), \mathbf{0}_{1\times k}\right]^\top \\
&- 4\sigma^2_{\epsilon,o}\left[\frac{1}{n}\mathrm{tr}(G_n), \mathbf{0}_{1\times k}\right] C_{\epsilon,2n}, \\
\delta_{33,n} =& \mu_{\epsilon,6} + 9\sigma^4_{\epsilon,o}\left[\frac{1}{n}\mathbf{1}_n^\top G_n X_n \beta_*, \frac{1}{n}\mathbf{1}_n^\top X_n\right] \Sigma_* \left[\frac{1}{n}\mathbf{1}_n^\top G_n X_n \beta_*, \frac{1}{n}\mathbf{1}_n^\top X_n\right]^\top \\
&- 6\sigma^2_{\epsilon,o}\left[\frac{1}{n}\mathbf{1}_n^\top G_n X_n \beta_*, \frac{1}{n}\mathbf{1}_n^\top X_n\right] C_{\epsilon,3n}.
\end{aligned}
$$

For the off-diagonal elements,

$$
\begin{aligned}
\delta_{21,n} = \delta_{12,n} =& 2\sigma^2_{\epsilon,o}\left[\frac{1}{n}\mathbf{1}_n^\top G_n X_n \beta_*, \frac{1}{n}\mathbf{1}_n^\top X_n\right] \Sigma_* \left[\frac{1}{n}\mathrm{tr}(G_n), \mathbf{0}_{1\times k}\right]^\top \\
&- \left[\frac{1}{n}\mathbf{1}_n^\top G_n X_n \beta_*, \frac{1}{n}\mathbf{1}_n^\top X_n\right] C_{\epsilon,2n} - 2\sigma^2_{\epsilon,o}\left[\frac{1}{n}\mathrm{tr}(G_n), \mathbf{0}_{1\times k}\right] C_{\epsilon,1n}, \\
\delta_{31,n} = \delta_{13,n} =& \mu_{\epsilon,4} + 3\sigma^2_{\epsilon,o}\left[\frac{1}{n}\mathbf{1}_n^\top G_n X_n \beta_*, \frac{1}{n}\mathbf{1}_n^\top X_n\right] \Sigma_* \left[\frac{1}{n}\mathbf{1}_n^\top G_n X_n \beta_*, \frac{1}{n}\mathbf{1}_n^\top X_n\right]^\top \\
&- \left[\frac{1}{n}\mathbf{1}_n^\top G_n X_n \beta_*, \frac{1}{n}\mathbf{1}_n^\top X_n\right] C_{\epsilon,3n} - 3\sigma^2_{\epsilon,o}\left[\frac{1}{n}\mathbf{1}_n^\top G_n X_n \beta_*, \frac{1}{n}\mathbf{1}_n^\top X_n\right] C_{\epsilon,1n}, \\
\delta_{32,n} = \delta_{23,n} =& 6\sigma^5_{\epsilon,o}\left[\frac{1}{n}\mathrm{tr}(G_n), \mathbf{0}_{1\times k}\right] \Sigma_* \left[\frac{1}{n}\mathbf{1}_n^\top G_n X_n \beta_*, \frac{1}{n}\mathbf{1}_n^\top X_n\right]^\top \\
&- 2\sigma^2_{\epsilon,o}\left[\frac{1}{n}\mathrm{tr}(G_n), \mathbf{0}_{1\times k}\right] C_{\epsilon,3n} - 3\sigma^2_{\epsilon,o}\left[\frac{1}{n}\mathbf{1}_n^\top G_n X_n \beta_*, \frac{1}{n}\mathbf{1}_n^\top X_n\right] C_{\epsilon,2n}.
\end{aligned}
$$

- Main proofs.

Proof of Proposition 1. The probability limit of $n^{-1}\sum_{i=1}^n \tilde{\epsilon}_{ni}$ is shown in p. 5–6. We now analyse those of $n^{-1}\sum_{i=1}^n \tilde{\epsilon}^3_{ni}$ and $n^{-1}\sum_{i=1}^n \tilde{\epsilon}^2_{ni}$. By the expansion, it follows that

$$
\begin{aligned}
\frac{1}{n}\sum_{i=1}^n \tilde{\epsilon}^3_{ni} =& \frac{1}{n}\sum_{i=1}^n \epsilon^3_{ni} + \frac{3}{n}\sum_{i=1}^n \epsilon^2_{ni}\mathbf{z}_{i,n}^\top(\theta_* - \tilde{\theta}_n) \\
&+ \frac{3}{n}\sum_{i=1}^n \epsilon_{ni}\left[\mathbf{z}_{i,n}^\top(\theta_* - \tilde{\theta}_n)\right]^2 + \frac{1}{n}\sum_{i=1}^n \left[\mathbf{z}_{i,n}^\top(\theta_* - \tilde{\theta}_n)\right]^3,
\end{aligned} \tag{B.1}
$$

On the one hand, by the Lindeberg-Lévy CLT, $n^{-1}\sum_{i=1}^n \epsilon_{ni}^3 = O_p(n^{-1/2})$ since $\mathrm{E}[\epsilon_{ni}^3] = \mathrm{E}[v_{ni}^3] = 0$ under H_0. On the other hand, under Assumptions 1–5, $\{\mathbf{z}_{i,n}\}$ is uniformly L_r-norm bounded for some $r = 4+\iota > 4$ and is uniformly L_2-NED on $\{\mathbf{x}_{i,n}, \epsilon_{ni}\}_{i=1}^n$, as the argument below (5). Assumption 1 tells us that $\mathrm{E}|v_{ni}|^{6+\iota} < \infty$ for some $\iota > 0$, so $\{\epsilon_{ni}\}$ must be uniformly L_r-norm bounded for some $r > 4$ under H_0. Hence, according to Lemma 3 (in Appendix A), $\{\epsilon_{ni}\mathbf{z}_{i,n}\}$ is uniformly L_2-NED on $\{\mathbf{x}_{i,n}, \epsilon_{ni}\}_{i=1}^n$ under H_0. Similarly, under H_0, all of $\epsilon_{ni}^2\mathbf{z}_{i,n}$, $\epsilon_{ni}\mathbf{z}_{i,n}\mathbf{z}_{i,n}^\top$ and $\{z_{ij,n}\mathbf{z}_{i,n}\mathbf{z}_{i,n}^\top\}$ are also uniformly L_2-NED on $\{\mathbf{x}_{i,n}, \epsilon_{ni}\}_{i=1}^n$. Besides, let $h_{n,ij}$ be either 1 or any element of $\mathbf{z}_{i,n}^\top$ for $j = 1, 2, 3$. According to the generalized Hölder's inequality, under Assumptions 1–5,

$$\sup_{i,n} \mathrm{E}|h_{n,i1}h_{n,i2}h_{n,i3}| \le \sup_{i,n}\left\{(\mathrm{E}|h_{n,i1}|^3)^{1/3}(\mathrm{E}|h_{n,i2}|^3)^{1/3}(\mathrm{E}|h_{n,i3}|^3)^{1/3}\right\} < \infty$$

and

$$\sup_{i,n} \mathrm{E}|v_{ni}^k h_{n,i1}h_{n,i2}| \le \sup_{i,n}\left\{(\mathrm{E}|v_{ni}|^{3k})^{1/3}(\mathrm{E}|h_{n,i1}|^3)^{1/3}(\mathrm{E}|h_{n,i2}|^3)^{1/3}\right\} < \infty$$

hold for $k = 0, 1, 2$. By making use of the LLN in Jenisha and Prucha (2012) again, $n^{-1}\sum_{i=1}^n \epsilon_{ni}^2\mathbf{z}_{i,n}^\top = \mathrm{E}[n^{-1}\sum_{i=1}^n \epsilon_{ni}^2\mathbf{z}_{i,n}^\top] + o_p(1) = O_p(1)$. Similarly, $n^{-1}\sum_{i=1}^n \epsilon_{ni}\mathbf{z}_{i,n}\mathbf{z}_{i,n}^\top = O_p(1)$ and $n^{-1}\sum_{i=1}^n \mathbf{z}_{i,n}\mathbf{z}_{i,n}^\top z_{n,ij} = O_p(1)$. Working with $\theta_* - \tilde{\theta}_n = o_p(1)$, it yields that $n^{-1}\sum_{i=1}^n \epsilon_{ni}^2\mathbf{z}_{i,n}^\top(\theta_* - \tilde{\theta}_n) = o_p(1)$, $n^{-1}\sum_{i=1}^n \epsilon_{ni}[\mathbf{z}_{i,n}^\top(\theta_* - \tilde{\theta}_n)]^2 = o_p(1)$, and $n^{-1}\sum_{i=1}^n[\mathbf{z}_{i,n}^\top(\theta_* - \tilde{\theta}_n)]^3 = o_p(1)$. Therefore, $n^{-1}\sum_{i=1}^n \tilde{\epsilon}_{ni}^3 = o_p(1)$. In addition, $n^{-1}\sum_{i=1}^n \tilde{\epsilon}_{ni}^2 - \sigma_{\epsilon,o}^2 = o_p(1)$ can be proved by a similar argument for the expansion

$$\frac{1}{n}\sum_{i=1}^n \tilde{\epsilon}_{ni}^2 = \frac{1}{n}\sum_{i=1}^n \epsilon_{ni}^2 + \frac{2}{n}\sum_{i=1}^n \epsilon_{ni}\mathbf{z}_{i,n}^\top(\theta_* - \tilde{\theta}_n) + \frac{1}{n}\sum_{i=1}^n\left[\mathbf{z}_{i,n}^\top(\theta_* - \tilde{\theta}_n)\right]^2.$$

The proof is accomplished. □

Proof of Proposition 2. To check the correctness of (6), we take the term regarding $\frac{1}{n}\sum_{i=1}^n \tilde{\epsilon}_{ni}^3$ as the example. Applying the mean-value theorem to $\frac{1}{n}\sum_{i=1}^n \tilde{\epsilon}_{ni}^3 = \frac{1}{n}\sum_{i=1}^n (y_{ni} - \mathbf{z}_{i,n}'\tilde{\theta}_n)^3 \triangleq \frac{1}{n}\sum_{i=1}^n \epsilon_{ni}^3(\tilde{\theta}_n)$ at θ_*, it yields that

$$\begin{aligned}\frac{1}{n}\sum_{i=1}^n \tilde{\epsilon}_{ni}^3 &= \frac{1}{n}\sum_{i=1}^n \epsilon_{ni}^3 - \frac{3}{n^{3/2}}\sum_{i=1}^n \epsilon_{ni}^2(\bar{\theta}_n)[\mathbf{w}_{i,n}Y_n, \mathbf{x}_{i,n}^\top]\cdot\sqrt{n}(\theta_* - \tilde{\theta}_n)\\ &= \frac{1}{n}\sum_{i=1}^n \epsilon_{ni}^3 - \frac{3}{n^{1/2}}\left[\frac{1}{n}\sum_{i=1}^n \epsilon_{ni}^2(\bar{\theta}_n)\mathbf{w}_{i,n}Y_n, \frac{1}{n}\sum_{i=1}^n \epsilon_{ni}^2(\bar{\theta}_n)\mathbf{x}_{i,n}^\top\right]\cdot\sqrt{n}(\theta_* - \tilde{\theta}_n)\end{aligned}$$

where $\bar{\theta}_n$ satisfies $\|\bar{\theta}_n - \theta_*\| \leq \|\tilde{\theta}_n - \theta_*\|$ so $\bar{\theta}_n - \theta_* = o_p(1)$. By expansion we can obtain

$$\begin{aligned}\frac{1}{n}\sum_{i=1}^{n}\epsilon_{ni}^2(\bar{\theta}_n)\mathbf{w}_{i,n}Y_n =& \frac{1}{n}\sum_{i=1}^{n}\epsilon_{ni}^2\mathbf{w}_{i,n}Y_n + \frac{1}{n}\sum_{i=1}^{n}\left[\mathbf{z}_{i,n}^\top(\theta_* - \bar{\theta}_n)\right]^2\mathbf{w}_{i,n}Y_n \\ &+ \frac{2}{n}\sum_{i=1}^{n}\epsilon_{ni}\left[\mathbf{z}_{i,n}^\top(\theta_* - \bar{\theta}_n)\right]\mathbf{w}_{i,n}Y_n.\end{aligned}$$

As mentioned above, any element of $\mathbf{z}_{i,n}^\top = [\mathbf{w}_{i,n}Y_n, \mathbf{x}_{i,n}^\top]$ are uniformly L_{2r}-norm bounded for some $r > 2$, and uniformly L_2-NED on $\{\mathbf{x}_{i,n}, \epsilon_{ni}\}$. Based on this, we can further conclude that $\epsilon_{ni}^2\mathbf{w}_{i,n}Y_n$ is uniformly L_4-norm bounded, and uniformly L_2-NED on $\{\mathbf{x}_{i,n}, \epsilon_{ni}\}$, thereby $n^{-1}\sum_{i=1}^{n}\epsilon_{ni}^2\mathbf{w}_{i,n}Y_n - \mathrm{E}[n^{-1}\mathbf{1}_n^\top G_nX_n\beta_*] = o_p(1)$ by the LLN in Jenisha and Prucha (2012) and the fact $\mathrm{E}[n^{-1}\sum_{i=1}^{n}\epsilon_{ni}^2\mathbf{w}_{i,n}Y_n] = \mathrm{E}[n^{-1}\sum_{i=1}^{n}\epsilon_{ni}^2\mathbf{w}_{i,n}S^{-1}X_n\beta_*] + \mathrm{E}[n^{-1}\sum_{i=1}^{n}\epsilon_{ni}^2\mathbf{w}_{i,n}S^{-1}\epsilon_n] = \mathrm{E}[n^{-1}\mathbf{1}_n^\top G_nX_n\beta_*]$. Ditto for $n^{-1}\sum_{i=1}^{n}z_{ij,n}^\top\mathbf{w}_{i,n}Y_n$ and $n^{-1}\sum_{i=1}^{n}\epsilon_{ni}z_{ij,n}\mathbf{w}_{i,n}Y_n$. Taking with $\theta_* - \bar{\theta}_n = o_p(1)$ together, it follows that $n^{-1}\sum_{i=1}^{n}\left[\mathbf{z}_{i,n}^\top(\theta_* - \bar{\theta}_n)\right]^2\mathbf{w}_{i,n}Y_n$ and $n^{-1}\sum_{i=1}^{n}\epsilon_{ni}^\dagger\left[\mathbf{z}_{i,n}^\top(\theta_* - \bar{\theta}_n)\right]\mathbf{w}_{i,n}Y_n$ are $o_p(1)$. Therefore, $n^{-1}\sum_{i=1}^{n}e_{ni}^2(\bar{\theta}_n)\mathbf{w}_{i,n}Y_n - \mathrm{E}[n^{-1}\mathbf{1}_n^\top G_nX_n\beta_*] = o_p(1)$. By a similar argument for

$$\begin{aligned}\frac{1}{n}\sum_{i=1}^{n}\epsilon_{ni}^2(\bar{\theta}_n)\mathbf{x}_{i,n} =& \frac{1}{n}\sum_{i=1}^{n}\epsilon_{ni}^2\mathbf{x}_{i,n} + \frac{1}{n}\sum_{i=1}^{n}\left[\mathbf{z}_{i,n}^\top(\theta_* - \bar{\theta}_n)\right]^2\mathbf{x}_{i,n} \\ &+ \frac{2}{n}\sum_{i=1}^{n}\epsilon_{ni}\mathbf{z}_{i,n}^\top(\theta_* - \bar{\theta}_n)\mathbf{x}_{i,n},\end{aligned}$$

the last two terms on the right-hand side satisfy $n^{-1}\sum_{i=1}^{n}\left[\mathbf{z}_{i,n}^\top(\theta_* - \bar{\theta}_n)\right]^2\mathbf{x}_{i,n} = o_p(1)$ and $\frac{2}{n}\sum_{i=1}^{n}\epsilon_{ni}\left[\mathbf{z}_{i,n}^\top(\theta_* - \bar{\theta}_n)\right]\mathbf{x}_{i,n} = o_p(1)$. Working with $n^{-1}\sum_{i=1}^{n}\epsilon_{ni}^2\mathbf{x}_{i,n} - \mathrm{E}[n^{-1}\sigma_{\epsilon,o}^2\mathbf{1}_n^\top X_n] = o_p(1)$ by the Chebyshev LLN, it yields that $n^{-1}\sum_{i=1}^{n}e_{ni}^2(\bar{\theta}_n)\mathbf{x}_{i,n} = \mathrm{E}[n^{-1}\sigma_{\epsilon,o}^2\mathbf{1}_n^\top X_n] + o_p(1)$. Therefore, we obtain

$$\begin{aligned}\frac{1}{n}\sum_{i=1}^{n}\tilde{\epsilon}_{ni}^3 =& \frac{1}{n}\sum_{i=1}^{n}\epsilon_{ni}^3 - \frac{1}{\sqrt{n}}3\sigma_{\epsilon,o}^2\left[\frac{1}{n}\mathbf{1}_n^\top G_nX_n\beta_*, \frac{1}{n}\mathbf{1}_n^\top X_n\right]\sqrt{n}(\tilde{\theta}_n - \theta_*) \\ &+ o_p(n^{-1/2}).\end{aligned}$$

Similarly, the other two equalities in (6) hold.

To derive the asymptotic distribution of (7), it is equivalent to derive that of

$$F_n \triangleq \frac{c_3}{n}\sum_{i=1}^{n}\epsilon_{ni}^3 + \frac{c_2}{n}\sum_{i=1}^{n}\epsilon_{ni}^2 + \frac{c_1}{n}\sum_{i=1}^{n}\epsilon_{ni} + \frac{1}{\sqrt{n}}\mathbf{f}(c_1, c_2, c_3)\sqrt{n}(\tilde{\theta}_n - \theta_*), \quad \text{(B.2)}$$

which depends on the limiting distribution of $\sqrt{n}(\tilde{\theta}_n - \theta_*)$. By Assumption 2, we know $n^{1/2}(\tilde{\theta}_n - \theta_*) = g_n(\theta_*) + o_p(1)$, where $g_n(\theta_*)$ is quadratic forms in ϵ_{ni}. Define $A_n^d = A_n - \mathrm{tr}(A_n) \cdot I_n$. Therefore, without lose of generality, we can assume $g_n(\theta_*) = n^{-1/2} D[\epsilon_n^\top P_1^d \epsilon_n, \cdots, \epsilon_n^\top P_s^d \epsilon_n, \epsilon_n^\top Q]^\top \in \mathbb{R}^{k+1}$, where $D = [d_1, \cdots, d_s, \mathbf{d}_q] \in \mathbb{R}^{(k+1)\times(s+q)}$, $P \in \mathbb{R}^{n\times n}$, and $Q \in \mathbb{R}^{n\times q}$ are constant matrices with q and s being arbitrary constants. Accordingly, (B.2) becomes

$$\begin{aligned} F_n =& \frac{c_3}{n}\sum_{i=1}^n \epsilon_{ni}^3 + \frac{\epsilon_n^\top}{n}\left[c_2 I_n + \mathbf{f}(c_1,c_2,c_3)\sum_{j=1}^s d_j P_j^d\right]\epsilon_n \\ &+ \frac{1}{n}\left[c_1 \mathbf{1}_n^\top + \mathbf{f}(c_1,c_2,c_3)\mathbf{d}_q Q^\top\right]\epsilon_n \\ \triangleq& \frac{c_3}{n}\sum_{i=1}^n \epsilon_{ni}^3 + \frac{1}{n}\epsilon_n^\top A_n \epsilon_n + \frac{1}{n}\mathbf{b}_n^\top \epsilon_n. \end{aligned}$$

The central limit theorem for cubic forms in ϵ_{ni}[5] tells us that

$$\sqrt{n} F_n \sim AN(0, \lim_{n\to\infty} \mathbf{c}^\top \Sigma_n \mathbf{c}), \tag{B.3}$$

where Σ_n is defined as in Appendix B. Then, taking (7), (B.2) and (B.3) together, (8) follows from the Slutsky theorem. The proof is accomplished. □

Proof of Proposition 3. Employing the Delta method for (8), we can obtain

$$\begin{aligned} n^{1/2} S_n(\tilde{\theta}_n) =& n^{1/2}\left[\phi\left(\frac{1}{n}\sum_{i=1}^n \tilde{\epsilon}_{ni}, \frac{1}{n}\sum_{i=1}^n \tilde{\epsilon}_{ni}^2, \frac{1}{n}\sum_{i=1}^n \tilde{\epsilon}_{ni}^3\right) - \phi\left(0, \sigma_{\epsilon,o}^2, 0\right)\right] \\ \sim& AN\left(0, \frac{\partial \phi(0, \sigma_{\epsilon,o}^2, 0)}{\partial [a,b,c]} \lim_{n\to\infty} \Sigma_n \frac{\partial \phi(0, \sigma_{\epsilon,o}^2, 0)}{\partial [a,b,c]^\top}\right). \end{aligned}$$

Then the asymptotic distribution of $S_n(\tilde{\epsilon}_n)$ follows from the facts that $\phi\left(0, \sigma_{\epsilon,o}^2, 0\right) = 0$, $\phi\left(n^{-1}\sum_{i=1}^n \tilde{\epsilon}_{ni}, n^{-1}\sum_{i=1}^n \tilde{\epsilon}_{ni}^2, n^{-1}\sum_{i=1}^n \tilde{\epsilon}_{ni}^3\right) = S_n(\tilde{\epsilon}_n)$, and

$$\frac{\partial \phi(0, \sigma_{\epsilon,o}^2, 0)}{\partial [a,b,c]} = \left[-\frac{3}{\sigma_{\epsilon,o}}, 0, \frac{1}{\sigma_{\epsilon,o}^3}\right] \triangleq \mathbf{r}_n^\top.$$

The proof is accomplished. □

Proof of Corollary 1. It follows from Proposition 3 that under Assumptions 1–5 and H_0,

$$n^{1/2} S_n(\tilde{\theta}_n) / \left(\mathbf{r}^\top \Sigma_n \mathbf{r}\right)^{1/2} \sim AN(0, 1),$$

where $\mathbf{r} = [-3/\sigma_{\epsilon,o}, 0, 1/\sigma_{\epsilon,o}^3]^\top$. Then, since $\tilde{\sigma}_\epsilon - \sigma_{\epsilon,o} = o_p(1)$, (9) follows from the Slutsky theorem.

[5] Available on https://github.com/Dr-Man-go/Cubic-CLT.

Recall that $C_{\epsilon,jn} = \mathrm{E}\left[n^{-1/2}\sum_{i=1}^{n}\epsilon_{ni}^{j}\cdot n^{1/2}(\tilde{\theta}_n - \theta_*)\right]$ for $j = 1, 2, 3$. By direct calculation (here we omit the tedious process),

$$C_{\epsilon,1n} = \Sigma_* \left[\frac{1}{n}\mathbf{1}_n^\top G_n X_n \beta_*, \frac{1}{n}\mathbf{1}_n^\top X_n\right]^\top$$

and

$$C_{\epsilon,3n} = 3\sigma_{\epsilon,o}^2 \Sigma_* \left[\frac{1}{n}\mathbf{1}_n^\top G_n X_n \beta_*, \frac{1}{n}\mathbf{1}_n^\top X_n\right]^\top .$$

Noticing the fact

$$\left[\frac{1}{n}\mathbf{1}_n^\top G_n X_n \beta_*, \frac{1}{n}\mathbf{1}_n^\top X_n\right] \Sigma_* \left[\frac{1}{n}\mathbf{1}_n^\top G_n X_n \beta_*, \frac{1}{n}\mathbf{1}_n^\top X_n\right]^\top = \sigma_{\epsilon,o}^2, \text{[6]}$$

it follows that $\delta_{11,n} = \delta_{13,n} = 0$, $\delta_{33,n} = 6\sigma_{\epsilon,o}^6$, whereby we can derive

$$\mathbf{r}^\top \Sigma_n \mathbf{r} = \delta_{11,n} r_1^2 + \delta_{33,n} r_3^2 + \delta_{13,n} r_1 r_3 = 6\sigma_{\epsilon,o}^6 (1/\sigma_{\epsilon,o}^3)^2 = 6.$$

Then (10) follows. The proof is accomplished. □

References

Aigner D, Lovell CAK, Schmidt P (1977) Formulation and estimation of stochastic frontier production function models. J Econometrics 6(1):21–37

Amsler C, Prokhorov A, Schmidt P (2016) Endogenous stochastic frontier models. J Econometrics 190:280–288

Amsler C, Schmidt P, Tsay WJ (2019) Evaluating the cdf of the distribution of stochastic frontier composed error. J Prod Anal 52:29–35

Baptista R (2000) Do innovations diffuse faster within geographical clusters? Int J Ind Organ 18:515–535

Billingsley P (2012) Probability and Measure, Anniversary edn. John Wiley & Sons, Inc., Hoboken, New Jersey

Chen YT, Wang HJ (2012) Centered-residuals-based moment estimator and test for stochas tic frontier models. Economet Rev 31(6):625–653

Cliff A, Ord J (1973) Spatial autocorrelation. Pion, London

[6] We can see from (**??**) that $[n^{-1}\mathbf{1}_n^\top G_n X_n \beta_*, n^{-1}\mathbf{1}_n^\top X_n]^\top$ is the second column of the inverse matrix of $\Sigma_*/\sigma_{\epsilon,o}^2$, so $(1/\sigma_{\epsilon,o}^2)\Sigma_*[n^{-1}\mathbf{1}_n^\top G_n X_n \beta_*, n^{-1}\mathbf{1}_n^\top X_n]^\top = [0, 1, 0 \cdots, 0]^\top$. Because regressors have an intercept (the first column of X_n must be $\mathbf{1}_n$), the second element of $[n^{-1}\mathbf{1}_n^\top G_n X_n \beta_*, n^{-1}\mathbf{1}_n^\top X_n]$ is equal to 1, whereby $(1/\sigma_{\epsilon,o}^2)[n^{-1}\mathbf{1}_n^\top G_n X_n \beta_*, n^{-1}\mathbf{1}_n^\top X_n]\Sigma_*[n^{-1}\mathbf{1}_n^\top G_n X_n \beta_*, n^{-1}\mathbf{1}_n^\top X_n]^\top = 1$.

Cliff A, Ord J (1981) Spatial process: models and applications. Pion, London

Coelli TJ (1995) Estimators and hypothesis tests for a stochastic frontier function: a Monte Carlo analysis. J Prod Anal 6:247–268

Druska V, Horrace WC (2004) Generalized moments estimation for spatial panel data: Indonesian rice farming. Amer J Agric Econom 86(1):185–198

Fried HO, Lovell CK, Schmidt SS (2008) The Measurement of Productive Efficiency and Productivity Growth. Oxford University Press, New York

Gallant AR, White H (1988) A Unified Theory of Estimation and Inference for Nonlinear Dynamic Models. Basil Blackwell, New York

Galli F (2023) A spatial stochastic frontier model including both frontier and error-based spatial cross-sectional dependence. Spatial Economic Analysis 18(2):239–258

Glass AJ, Kenjegalieva K, Sickles RC (2016) A spatial autoregressive stochastic frontier model for panel data with asymmetric efficiency spillovers. J Econometrics 190(2):289–300

Greene WH (1990) A Gamma-distributed stochastic frontier model. J Econometrics 46(1–1):141–163

Jenisha N, Prucha IR (2009) Central limit theorems and uniform laws of large numbers for arrays of random fields. J Econometrics 150:86–98

Jenisha N, Prucha IR (2012) On spatial processes and asymptotic inference under near-epoch dependence. J Econometrics 170:178–190

Jin F, Lee Lf (2020) Asymptotic properties of a spatial autoregressive stochastic frontier model. J Spat Econometrics 1(1)

Jin F, Wang Y (2023) Consistent non-Gaussian pseudo maximum likelihood estimations of spatial autoregressive models. Economet Theor 0:1–39

Kelejian H, Piras G (2017) Spatial econometrics. Academic Press

Kelejian H, Prucha I (1998) A generalized spatial two-stage least squares procedure for estimating a spatial autoregressive model with autoregressive disturbances. J Real Estate Finance Econ 17(1):99–121

Kopp RJ, Mullahy J (1990) Moment-based estimation and testing of stochastic frontier models. J Econometrics 46:165–183

Kutlu L (2018) Estimating efficiency in a spatial autoregressive stochastic frontier model. Econom Lett 163:155–157

Kutlu L (2022) Spatial stochastic frontier model with endogenous weighting matrix. Empir Econ 63:1947–1968

Kutlu L (2023) Calculating efficiency for spatial autoregressive stochastic frontier model. Econom Lett 225

Kutlu L, Tran KC, Tsionas MG (2020) A spatial stochastic frontier model with endogenous frontier and environmental variables. European J Oper Res 286:389–399

Lai Hp, Tran KC (2022) Persistent and transient inefficiency in a spatial autoregressive panel stochastic frontier model. J Prod Anal 58(1):1–13

Lee Lf (2004) Asymptotic distributions of quasi-maximum likelihood estimators for spatial autoregressive models. Econometrica 72(6):1899–1925

Lee Lf (2007) GMM and 2SLS estimation of mixed regressive, spatial autoregressive models. J Econometrics 137:489–514

Liu T, Lee Lf (2019) A likelihood ratio test for spatial model selection. J Econometrics 213:434–458

Meeusen W, van den Broeck J (1977) Efficiency estimation from Cobb-Douglas production functions with composed error. Internat Econom Rev 18(2):435–444

Meintanis SG, Papadimitriou CK (2022) Goodness–of–fit tests for stochastic frontier models based on the characteristic function. J Prod Anal 57:285–296

Schmidt P, Lin TF (1984) Simple tests of alternative specifications in stochastic frontier models. J Econometrics 24:349–361

Stevenson R (1980) Likelihood functions for generalized stochastic frontier estimation. J Econometrics 13(1):57–66

Wang WS, Amsler C, Schmidt P (2011) Goodness of fit tests in stochastic frontier models. J Prod Anal 35:95–118

Xu X, Lee Lf (2015) Maximum likelihood estimation of a spatial autoregressive Tobit model. J Econometrics 188:164–280

Xu X, Lee Lf (2018) Sieve maximum likelihood estimation of the spatial autoregressive Tobit model. J Econometrics 203:96–112