Lecture 5: Bootstrap LM Tests of Spatial Dependence in Spatial Linear Regression Models

Zhenlin Yang

School of Economics, Singapore Management University
zlyang@smu.edu.sg
http://www.mysmu.edu.sg/faculty/zlyang/

ECON747 Spatial Econometrics, Term I 2024-25

5.1. Introduction

To test the existence of spatial dependence in an econometric model, a convenient test is the Lagrange Multiplier (LM) test (Anselin, 1988a,b, 2001; Anselin and Bera, 1998). This is because

- LM test requires only the estimation of null model, and thus
- it is computationally much easier than other type of tests, in particular when the test concerns the 'nonlinear parameters'.

However, evidence shows that, in finite sample,

- the true sizes of the LM test referring to the asymptotic critical values can be quite different from their nominal sizes, and
- more so with denser spatial weight matrices and one-sided tests.
- As a result, the LM tests in such circumstances may have low power in detecting a 'negative' or 'positive' spatial dependence.
- LM tests may not be robust to misspecification in error distribution.

In this lecture,

- Residual-based bootstrap methods are introduced for a 2nd-order approximation to the finite sample critical values of LM statistics.
- Conditions for their validity are clearly laid out and formal justifications are given in general, and in details under,
- several popular spatial LM tests: LM tests for SE, LM tests for SL, and LM tests for spatial error components (SEC).
- Further demonstrations are given using: joint LM test for SL and SE, LM test of SE allowing SL and vise versa.

Major conclusions:

- With the unrestricted estimates and residuals, bootstrap is able to provide critical values that agree with the true ones to the 2nd-order; but it is not in general with restricted estimates and/or residuals.
- However, use of restricted estimates/residuals sometimes provides partial asymptotic refinements, leading to improved results over the large sample approximations.

The proposed bootstrap methods are applicable to a wide class of LM testing situations, not only the LM tests for spatial dependence.

We demonstrate that for these methods to work well, it is important to obey the following general guidelines:

(*i*) *The bootstrap DGP resembles the null model;*

(*ii*) *The LM statistic is asymptotically pivotal under the null or its robustified or standardized version must be used;*

(*iii*) *The estimates of the nuisance parameters, to be used as parameters in the bootstrap world, are consistent whether or not the null hypothesis is true;*

(*iv*) *The empirical distribution function of the residuals to be resampled estimates consistently the error distribution whether or not the null hypothesis is true;*

(*v*) *The calculations of the bootstrapped values of the LM statistic are done under the null hypothesis.*

Following discussions are helpful before introducing the methods formally.

Standardization (Koenker, 1981; Robinson, 2008; Yang, 2010; Baltagi and Yang, 2013, etc.) is a popular method that

- makes the LM tests robust against the distributional misspecification,
- and alleviates the problems of size distortion for two-sided tests.

But,

- standardization does not solve the problem of size distortion for one-sided tests, and the problem of lower power.
- The reason is that the finite sample distribution of the LM test is often skewed, due to the spatial effect.

We demonstrate that **standardization coupled with bootstrap** provide a satisfactory solution to these problems.

This lecture draws heavily on work of Yang (2015, JOE).

Bootstrap method is able to provide asymptotic refinements on the critical values of a test statistic if this statistic is **asymptotically pivotal** under the null hypothesis:

- Beran, 1988;
- Hall, 1992;
- Horowitz, 1994, 1997;
- Hall and Horowitz, 1996;
- Davidson and MacKinnon, 1999, 2006;
- van Giersbergen and Kiviet, 2002;
- Godfrey, 2009, etc.

However, as pointed out by Davidson, and reiterated by Godfrey:

- it is not always the case that the asymptotic analysis seems to provide a good explanation of what is observed in finite samples;
- there are unsettled issues on the choices of residuals and the parameter estimates to be used to set up the bootstrap DGP.

Suppose that the model can be written as,

$$q(Y_n, X_n, W_n;\ \theta, \delta) = e_n, \tag{5.1}$$

- e_n: $n \times 1$ vector of errors, iid (0, 1), with true CDF $\mathcal{F}$,
- δ: the parameters of interest,
- θ: the nuisance parameters,
- Y_n: $n \times 1$ response vector,
- X_n: $n \times k$ matrix of regressors,
- W_n: $n \times n$ spatial weight matrix.

Suppose that the model can be inverted to give

$$Y_n = h(X_n, W_n;\ \theta, \delta;\ e_n). \tag{5.2}$$

Consider a general hypothesis

$$H_0 : \delta = \delta_0 \text{ versus } H_a : \delta \neq \delta_0 \text{ (or in scalar case }, < \delta_0,\ > \delta_0).$$

The most interesting test corresponds to $\delta_0 = 0$, e.g., test of no spatial effect.

Let $\mathrm{LM}_n(\delta_0)$ be the LM test for testing H_0, derived under a 'specified' error distribution, e.g., $N(0, 1)$, which may or may not be the true $\mathcal{F}$. Let,

- $\mathcal{G}_n(\cdot, \theta, \mathcal{F})$: the finite sample null distribution of $\mathrm{LM}_n(\delta_0)$;
- $c_n(\alpha; \theta, \mathcal{F})$, $\alpha \in (0, 1)$: the finite sample critical value (FCV) of $\mathrm{LM}_n(\delta_0)|_{H_0}$;
- $\mathcal{G}(\cdot, \theta, \mathcal{F})$: the limiting null distribution of $\mathrm{LM}_n(\delta_0)$.

Note that, typically,

- $\mathrm{LM}_n(\delta_0)$ is not a pivotal quantity as $\mathcal{G}_n(\cdot, \theta, \mathcal{F})$ depends on $(\theta, \mathcal{F})$;
- but it is an asymptotic pivotal quantity **if** $\mathcal{F}$ is correctly specified, and in this case $\mathcal{G}(\cdot, \theta, \mathcal{F})$ is free of parameters $(\theta, \mathcal{F})$, typically standard normal if δ is a scalar, or chi-square if δ is a vector.
- If $\mathcal{F}$ is misspecified, e.g., specified distribution for $e_{n,i}$ is $N(0, 1)$ but true $\mathcal{F}$ is not, $\mathrm{LM}_n(\delta_0)$ may not even be an asymptotic pivotal quantity.

Denote

- $\tilde{\theta}_n$: the restricted estimate of θ under H_0,
- $(\hat{\theta}_n, \hat{\delta}_n)$: the unrestricted estimates of (θ, δ),
- The observable counterpart of e_n is referred to as residuals,
- $\tilde{e}_n = q(Y_n, X_n, W_n; \tilde{\theta}_n, \delta_0)$: the restricted residuals,
- $\hat{e}_n = q(Y_n, X_n, W_n; \hat{\theta}_n, \hat{\delta}_n)$: the unrestricted residuals,
- $\tilde{\mathcal{F}}_n$: the empirical distribution function (EDF) of $\tilde{e}_n$,
- $\hat{\mathcal{F}}_n$: the EDF of $\hat{e}_n$.

Note that under the LM framework

- only the estimation of the null model is required, and
- the null model is determined by the pair $\{\theta, \mathcal{F}\}$.
- In order to approximate the finite sample null distribution (in particular the critical values) of $\text{LM}_n(\delta_0)$, the bootstrap world must be set up so that it is able to **mimic the real world at the null**.

Thus, the bootstrap DGP should take the following form

$$Y_n^* = h(X_n, W_n; \ddot{\theta}_n, \delta_0; e_n^*), \quad e_n^* \overset{iid}{\sim} \ddot{\mathcal{F}}_n, \tag{5.3}$$

– $\ddot{\theta}_n$: the bootstrap parameter vector,
– $\ddot{\mathcal{F}}_n$: the bootstrap error distribution.

The steps for finding the bootstrap critical values (BCV) for $\mathrm{LM}_n(\delta_0)|_{H_0}$ is summarized as follows:

General Bootstrap Algorithm 1 (GBA-1):

(a) Draw a bootstrap sample e_n^* from $\ddot{\mathcal{F}}_n$;

(b) Compute $Y_n^* = h(X_n, W_n; \ddot{\theta}_n, \delta_0;\ e_n^*)$ to obtain the bootstrap data $\{Y_n^*, X_n, W_n\}$;

(c) Estimate the **null model** based on $\{Y_n^*, X_n, W_n\}$, and then compute a bootstrapped value $\mathrm{LM}_n^b(\delta_0)$ of $\mathrm{LM}_n(\delta_0)|_{H_0}$;

(d) Repeat (a)-(c) B times to obtain the EDF of $\{\mathrm{LM}_n^b(\delta_0)\}_{b=1}^B$, and its α-quantile gives a bootstrap critical value $c_n(\alpha; \ddot{\theta}_n, \ddot{\mathcal{F}}_n)$ for $\mathrm{LM}_n(\delta_0)|_{H_0}$.

When does $c_n(\alpha; \ddot{\theta}_n, \ddot{\mathcal{F}}_n)$ give a better approximation to $c_n(\alpha; \theta, \mathcal{F})$, the true α-quantile of $\mathrm{LM}_n(\delta_0)|_{H_0}$, than does the asymptotic critical value, say $c(\alpha)$?

In reality, one does not know whether or not H_0 is true.

Thus it incurs an important issue: the choice of the pair $\{\ddot{\theta}_n, \ddot{\mathcal{F}}_n\}$ for setting up the bootstrap DGP.

It leads to four resampling schemes, to adopt the similar terms as in Godfrey (2009):

- RS_{uu} : *unrestricted resampling scheme*, $\{\ddot{\theta}_n, \ddot{\mathcal{F}}_n\} = \{\hat{\theta}_n, \hat{\mathcal{F}}_n\}$,
- RS_{rr} : *restricted resampling scheme*, $\{\ddot{\theta}_n, \ddot{\mathcal{F}}_n\} = \{\tilde{\theta}_n, \tilde{\mathcal{F}}_n\}$
- RS_{ur} : *hybrid resampling scheme* 1, $\{\ddot{\theta}_n, \ddot{\mathcal{F}}_n\} = \{\hat{\theta}_n, \tilde{\mathcal{F}}_n\}$
- RS_{ru} : *hybrid resampling scheme* 2, $\{\ddot{\theta}_n, \ddot{\mathcal{F}}_n\} = \{\tilde{\theta}_n, \hat{\mathcal{F}}_n\}$.

Alternative to RS_{uu}, consider the bootstrap analog of H_0,

$$H_0^* : \delta = \hat{\delta}_n.$$

The corresponding bootstrap procedure for finding the critical values is:

General Bootstrap Algorithm 2 (GBA-2):

(a) Draw a bootstrap sample $\hat{e}_n^*$ from the EDF $\hat{\mathcal{F}}_n$ of $\hat{e}_n$,

(b) Compute $Y_n^* = h(X_n, W_n; \hat{\theta}_n, \hat{\delta}_n; \hat{e}_n^*)$ to obtain the bootstrap data $\{Y_n^*, X_n, W_n\}$,

(c) Conditional on $\hat{\delta}_n$, estimate the model based on $\{Y_n^*, X_n, W_n\}$, and then compute $\mathrm{LM}_n(\hat{\delta}_n)$ and denote its value as $\mathrm{LM}_n^b(\hat{\delta}_n)$,

(d) Repeat (a)-(c) B times to obtain the EDF of $\{\mathrm{LM}_n^b(\hat{\delta}_n)\}_{b=1}^B$, and the quantiles of it give the bootstrap critical values of $\mathrm{LM}_n(\delta_0)|_{H_0}$.

This resampling scheme is denoted as RS_{uf}.

A major difference from the other schemes is that RS_{uf} requires the construction of an LM-type test at a general δ value (more complicated).

Note that in general, $c(\alpha) - c_n(\alpha;\theta,\mathcal{F}) = O(n^{-1/2})$.

We argue that with a proper choice of the pair $(\ddot{\theta}_n, \ddot{\mathcal{F}}_n)$, it can be that

$$c_n(\alpha;\ddot{\theta}_n,\ddot{\mathcal{F}}_n) - c_n(\alpha;\theta,\mathcal{F}) = O(n^{-1}).$$

$\Rightarrow$ BCV provides a higher-order approximation to the finite sample critical value (FVC) of $\mathrm{LM}_n(\delta_0)|_{H_0}$ than does the asymptotic critical value (ACV).

To this end,

- we need some general conditions on the LM test statistic $\mathrm{LM}_n(\delta_0)$ and its finite sample null distribution $\mathcal{G}_n(\cdot,\theta,\mathcal{F})$ at the true $(\theta,\mathcal{F})$.
- Let $\mathcal{N}_{\theta,\mathcal{F}}$ denote a neighborhood of $(\theta,\mathcal{F})$.
- When the 'specified' CDF for $e_{n,i}$ (i.e., the CDF under which $\mathrm{LM}_n(\delta_0)$ is developed) is the same as $\mathcal{F}$, we say $\mathcal{F}$ is correctly specified, otherwise misspecified.

Assumption G1. *$\mathcal{F}$ is correctly specified such that (i) $\mathrm{LM}_n(\delta_0)$ developed under $\mathcal{F}$ is asymptotically pivotal when H_0 is true; (ii) $(\tilde{\theta}_n, \tilde{\mathcal{F}}_n)$ is $\sqrt{n}$-consistent for $(\theta, \mathcal{F})$ under H_0; and (iii) $(\hat{\theta}_n, \hat{\mathcal{F}}_n)$ is $\sqrt{n}$-consistent for $(\theta, \mathcal{F})$ whether or not H_0 is true.*

Assumption G2. *$\mathcal{F}$ is misspecified but Assumptions G1(ii)-(iii) remain. Furthermore, either $\mathrm{LM}_n(\delta_0)$ is robust (i.e., it remains to be asymptotically pivotal at H_0) or its robust version, denoted as $\mathrm{SLM}_n(\delta_0)$, is used.*

Assumption G3. *For $(\vartheta, F) \in \mathcal{N}_{\theta,\mathcal{F}}$, the null CDF $\mathcal{G}_n(\cdot, \vartheta, F)$ converges weakly to $\mathcal{G}(\cdot, \vartheta, F)$ as n increases, and admits the following asymptotic expansion uniformly in t and locally uniformly for $(\vartheta, F) \in \mathcal{N}_{\theta,\mathcal{F}}$:*

$$\mathcal{G}_n(t, \vartheta, F) = \mathcal{G}(t, \vartheta, F) + n^{-\frac{1}{2}} g(t, \vartheta, F) + O(n^{-1}), \tag{5.4}$$

where $\mathcal{G}(\cdot, \vartheta, F)$ is differentiable and strictly monotone over its support, and $g(t, \vartheta, F)$ is a functional of (t, ϑ, F) differentiable in (ϑ, F).

Assumptions 1 & 2 are standard for ML or QML estimation. Assumption 3 is adapted from Beran (1988), with θ containing only the nuisance parameters.

In an important special case where δ is a scalar and $\mathrm{LM}_n(\delta_0)|_{H_0} \overset{a}{\sim} N(0,1)$, the asymptotic expansion (5.4) at $(\theta, \mathcal{F})$ reduces to

$$\mathcal{G}_n(t, \theta, \mathcal{F}) = \Phi(t) + n^{-\frac{1}{2}}\phi(t)p(t, \theta, \mathcal{F}) + O(n^{-1}), \qquad (5.5)$$

where Φ and ϕ are the CDF and pdf of $N(0,1)$, **if** that the jth cumulant $\boldsymbol{\kappa}_{j,n} \equiv \boldsymbol{\kappa}_{j,n}(\theta, \mathcal{F})$ of $\mathrm{LM}_n(\delta_0)|_{H_0}$ can be expanded as a power series in n^{-1}:

$$\boldsymbol{\kappa}_{j,n} = n^{-\frac{j-2}{2}}(k_{j,1} + n^{-1}k_{j,2} + n^{-2}k_{j,3} + \cdots).$$

from which one has $p(t, \theta, \mathcal{F}) = -k_{1,2} + \frac{1}{6}k_{3,1}(1 - t^2)$.

- See Hall (1992, Sec. 2.3) for details on Asymptotic Expansions.
- This constitutes one of the most important scenarios in constructing bootstrap LM tests.
- The proof of validity of the BLM test for spatial dependence depends critically on the validity of this expansion.

Proposition 5.1 *Under Assumptions G1 and G3, the bootstrap methods under* RS_{uu} *and* RS_{uf} *are generally valid in that they are both able to provide full asymptotic refinements on the critical values of the LM tests, with an error of approximation of order* $O(n^{-1})$.

Proposition 5.2 *Under Assumptions G2 and G3, if further* $\frac{\partial}{\partial \mathcal{F}} g(t, \theta, \mathcal{F}) = O(n^{-\frac{1}{2}})$*, then* $\tilde{\mathcal{F}}_n$ *can be used in place of* $\hat{\mathcal{F}}_n$*, and thus the bootstrap method with* RS_{ur} *is also valid.*

Proposition 5.3 *Under Assumption G1 or G2, and Assumption G3, if either* $\tilde{\theta}_n$ *is also consistent when* H_0 *is false or* LM *or* SLM *test is invariant of* θ*, then* $\tilde{\theta}_n$ *can be used in place of* $\hat{\theta}_n$ *and thus the bootstrap method with* RS_{ru} *is also valid.*

Proposition 5.4 *Under Assumptions G2 and G3, if the conditions for both Propositions 5.2 and 5.3 hold, then all the five bootstrap methods are valid.*

Proof. Under the real world null DGP: $Y_n = h(X_n, W_n, \theta, \delta_0; e_n)$,

$$\begin{aligned} \mathrm{LM}_n(\delta_0)|_{H_0} &\equiv \mathrm{LM}_n(Y_n, X_n, W_n; \delta_0) \\ &= \mathrm{LM}_n[h(X_n, W_n, \theta, \delta_0; e_n), X_n, W_n; \delta_0] \\ &\equiv \mathrm{LM}_n(X_n, W_n, \theta, \delta_0; e_n). \end{aligned}$$

The bootstrap DGP that mimics the real world null DGP is $Y_n^* = h(X_n, W_n; \ddot{\theta}_n, \delta_0; e_n^*)$, where $e_n^* \stackrel{iid}{\sim} \ddot{\mathcal{F}}_n$. Based on the bootstrap data (Y_n^*, X_n, W_n), estimating the null model and computing the bootstrap analogue of $\mathrm{LM}_n(\delta_0)$, we have

$$\begin{aligned} \mathrm{LM}_n^*(\delta_0) &\equiv \mathrm{LM}_n(Y_n^*, X_n, W_n; \delta_0) \\ &= \mathrm{LM}_n[h(X_n, W_n, \ddot{\theta}_n, \delta_0; e_n^*), X_n, W_n; \delta_0] \\ &\equiv \mathrm{LM}_n(X_n, W_n, \ddot{\theta}_n, \delta_0; e_n^*). \end{aligned}$$

Thus, $\mathrm{LM}_n^*(\delta_0)$ is identical in structure to $\mathrm{LM}_n(\delta_0)|_{H_0}$, suggesting that the *bootstrap CDF* of $\mathrm{LM}_n^*(\delta_0)$ has the form $\mathcal{G}_n(\cdot, \ddot{\theta}_n, \ddot{\mathcal{F}}_n)$, identical in form to the finite sample CDF $\mathcal{G}_n(\cdot, \theta, \mathcal{F})$ of $\mathrm{LM}_n(\delta_0)|_{H_0}$.

Under Assumption G3, expansion (5.4) holds for $(\theta, \mathcal{F})$, giving

$$\mathcal{G}_n(t,\theta,\mathcal{F}) = \mathcal{G}(t,\theta,\mathcal{F}) + n^{-\frac{1}{2}} g(t,\theta,\mathcal{F}) + O(n^{-1}). \tag{5.6}$$

Assume $\text{plim}_{n\to\infty}(\ddot{\theta}_n, \ddot{\mathcal{F}}_n) \in \mathcal{N}_{\theta,\mathcal{F}}$. As (5.4) holds locally uniformly for any $(\vartheta, F) \in \mathcal{N}_{\theta,\mathcal{F}}$, the bootstrap CDF admits the asymptotic expansion:

$$\mathcal{G}_n(t,\ddot{\theta}_n,\ddot{\mathcal{F}}_n) = \mathcal{G}(t,\ddot{\theta}_n,\ddot{\mathcal{F}}_n) + n^{-\frac{1}{2}} g(t,\ddot{\theta}_n,\ddot{\mathcal{F}}_n) + O_p(n^{-1}). \tag{5.7}$$

Comparing (5.7) with (5.6), the scenarios under which the bootstrap is able to provide asymptotic refinements on the critical values are clear.

For Proposition 5.1, as $\mathcal{F}$ is correctly specified, $\mathcal{G}(t,\theta,\mathcal{F}) = \mathcal{G}(t)$. Further, as (5.6) holds locally uniformly in $(\theta, \mathcal{F})$, $\mathcal{G}(t,\hat{\theta}_n,\hat{\mathcal{F}}_n) = \mathcal{G}(t)$. Then, we have

$$\mathcal{G}_n(t,\theta,\mathcal{F}) - \mathcal{G}_n(t,\hat{\theta}_n,\hat{\mathcal{F}}_n) = n^{-\frac{1}{2}}[g(t,\theta,\mathcal{F}) - g(t,\hat{\theta}_n,\hat{\mathcal{F}}_n)] + O_p(n^{-1}) = O_p(n^{-1}),$$

- It follows that $c_n(\alpha,\hat{\theta}_n,\hat{\mathcal{F}}_n) - c_n(\alpha,\theta,\mathcal{F}) = O_p(n^{-1})$.
- However, $c_n(\alpha,\theta,\mathcal{F}) - c(\alpha) = O_p(n^{-\frac{1}{2}})$, where $c(\alpha)$ is the ACV of $\text{LM}_n(\delta_0)|_{H_0}$,
- $\Rightarrow$ BCV gives a higher-order approximation to the FCV of $\text{LM}_n(\delta_0)|_{H_0}$.

To prove Proposition 5.2, we have in view of (5.7),

$$\mathcal{G}_n(t, \hat{\theta}_n, \tilde{\mathcal{F}}_n) = \mathcal{G}(t, \hat{\theta}_n, \tilde{\mathcal{F}}_n) + n^{-\frac{1}{2}} g(t, \hat{\theta}_n, \tilde{\mathcal{F}}_n) + O_p(n^{-1}).$$

The fact that $\mathrm{LM}_n(\lambda_0)|_{H_0}$ (or its robust version) is asymptotically pivotal even if $\mathcal{F}$ is misspecified implies

$$\mathcal{G}(t, \theta, \mathcal{F}) = \mathcal{G}(t) \text{ and } \mathcal{G}(t, \hat{\theta}_n, \tilde{\mathcal{F}}_n) = \mathcal{G}(t).$$

Since $\frac{\partial}{\partial \mathcal{F}} g(t, \theta, \mathcal{F}) = O(n^{-\frac{1}{2}})$ and $\hat{\theta}_n$ is consistent, it follows that

$$g(t, \hat{\theta}_n, \tilde{\mathcal{F}}_n) - g(t, \theta, \mathcal{F}) = O_p(n^{-\frac{1}{2}}).$$

The result of Proposition 5.2 thus follows. Proofs of the rest are evident.

Remark 5.1. When $\mathcal{F}$ is misspecified, $\mathrm{LM}_n(\lambda_0)$ is not robust and its limit null CDF $\mathcal{G}(t, \theta, \mathcal{F})$ depends on $(\theta, \mathcal{F})$. Thus, $\mathcal{G}(t, \ddot{\theta}_n, \ddot{\mathcal{F}}_n) - \mathcal{G}(t, \theta, \mathcal{F})$ is $O_p(n^{-\frac{1}{2}})$ if $(\ddot{\theta}_n, \ddot{\mathcal{F}}_n)$ is consistent, otherwise the difference can be $O_p(1)$.

Remark 5.2. Standardization is like prepivoting, giving an LM-type of test robust against distributional misspecification (Beran, 1988).

In this section, we present bootstrap LM (BLM) tests for spatial dependence in several popular spatial linear regression (SLR) models:

- BLM test for SE effect in SE model
- BLM test for SL effect in SL model
- BLM test for SL and SE effects in SLE model
    - Joint BLM test for SL and SE
    - Marginal BLM test for SL allowing SE
    - Marginal BLM test for SE allowing SL
- BLM test for the existence of spatial error component

Validity of each BLM test (i.e., being able to provide a second order approximation) is critically discussed.

Assumption S1. *The innovations* $\{e_{n,i}\}$ *are iid random draws from* $\mathcal{F}$ *with mean zero, variance 1, and finite cumulants* $\kappa_j \equiv \kappa_j(\mathcal{F}), j = 3, 4, 5, 6$.

Assumption S2. *The elements of* X_n *are uniformly bounded for all n, and* $\lim_{n\to\infty} \frac{1}{n} X_n' X_n$ *exists and is nonsingular. (These conditions are to be replaced by their stochastic versions if* X_n *is stochastic. The results are then interpreted conditionally, given the exogenous* X_n*.)*

Assumption S3. *The elements* $\{w_{n,ij}\}$ *of* W_n *are at most of order* h_n^{-1} *uniformly for all i, j, with the rate sequence* $\{h_n\}$ *satisfying* $h_n/n \to 0$ *as* $n \to \infty$. $\{W_n\}$ *are uniformly bounded in both row and column sums with* $w_{n,ii} = 0$ *and* $\sum_j w_{n,ij} = 1, \forall i$.

Furthermore:

- E^*, Var^*, $\xrightarrow{D^*}$, $\xrightarrow{p^*}$, $o_{p^*}(\cdot)$, etc., correspond to the bootstrap error CDF $\ddot{\mathcal{F}}_n$,
- to distinguish from the usual notation corresponding to $\mathcal{F}$.
- Assume throughout $\ddot{\mathcal{F}}_n$ has a zero mean and a unit variance (achievable through centering and scaling), and jth cumulant $\ddot{\kappa}_{jn} \equiv \kappa_j(\ddot{\mathcal{F}}_n), j = 3, 4, 5, 6$.

Recall the linear regression model with SE dependence:

$$Y_n = X_n\beta + u_n, \qquad u_n = \rho W_n u_n + \epsilon_n, \qquad \epsilon_n = \sigma e_n. \tag{5.8}$$

We are interested in testing the hypothesis H_0: $\rho = 0$.

Consider the LM test statistics given in (3.32) and the standardized LM test statistic given in (3.34), written in a slightly different forms:

$$\mathrm{LM}_{\mathrm{SE}}^{\mathrm{FI}} = \frac{n}{\sqrt{K_n}} \frac{\tilde{\epsilon}_n' W_n \tilde{\epsilon}_n}{\tilde{\epsilon}_n' \tilde{\epsilon}_n}, \tag{5.9}$$

$$\mathrm{SLM}_{\mathrm{SE}}^{\circ} = \frac{n}{\sqrt{K_n^{\dagger} + \tilde{\kappa}_{4n} a_n' a_n}} \frac{\tilde{\epsilon}_n' W_n^{\circ} \tilde{\epsilon}_n}{\tilde{\epsilon}_n' \tilde{\epsilon}_n}, \tag{5.10}$$

where $K_n = \mathrm{tr}(W_n' W_n + W_n W_n)$, $W_n^{\circ} = W_n - \frac{1}{n-k}\mathrm{tr}(W_n M_n) I_n$, $a_n = \mathrm{diagv}(\mathcal{A}_n)$, $\mathcal{A}_n = M_n W_n^{\circ} M_n$, $K_n^{\dagger} = \mathrm{tr}[\mathcal{A}_n(\mathcal{A}_n + \mathcal{A}_n')]$, $M_n = I_n - X_n(X_n' X_n)^{-1} X_n'$, and $\tilde{\kappa}_n$ is the excess sample kurtosis of OLS residuals $\tilde{\epsilon}_n$.

Baltagi and Yang (2013) show that $\mathrm{LM}_{\mathrm{SE}}^{\mathrm{FI}}|_{H_0} \overset{a}{=} \mathrm{SLM}_{\mathrm{SE}}^{\circ}|_{H_0} \overset{a}{\sim} N(0, 1)$.

To see the validity of the various bootstrap methods presented in Section 5.2, we concentrate on $\mathrm{LM}_{\mathtt{SE}}^{\mathtt{FI}}$.

Under the real world null DGP: $Y_n = X_n\beta + \sigma e_n$, $\tilde{\epsilon}_n = \sigma M_n e_n$, and

$$\mathrm{LM}_{\mathtt{SE}}^{\mathtt{FI}}|_{H_0} = \frac{n}{\sqrt{K_n}} \frac{e_n' M_n W_n M_n e_n}{e_n' M_n e_n}. \tag{5.11}$$

Clearly, $\mathrm{LM}_{\mathtt{SE}}^{\mathtt{FI}}|_{H_0}$ is free of β and σ. Further, it is easy to show that it is

- an exact pivot if $\mathcal{F}$ is known;
- an asymptotic pivot if $\mathcal{F}$ is unknown, i.e., $\mathrm{LM}_{\mathtt{SE}}^{\mathtt{FI}}|_{H_0} \xrightarrow{D} N(0, 1)$.

In case of a known $\mathcal{F}$, one can simply use Monte Carlo method to find the finite sample critical values of $\mathrm{LM}_{\mathrm{SEC}}|_{H_0}$ to any level of accuracy.

In case of an unknown $\mathcal{F}$, $\mathrm{LM}_{\mathrm{SED}}|_{H_0}$ is not an exact pivot and thus the Monte Carlo method does not work. Bootstrap methods need to be called for asymptotically refined approximations to the FCVs of $\mathrm{LM}_{\mathrm{SED}}|_{H_0}$.

In the bootstrap world, the bootstrap DGP that mimics the real world null DGP is $Y_n^* = X_n\ddot{\beta}_n + \ddot{\sigma}_n e_n^*$, where e_{ni}^* are iid draws from $\ddot{\mathcal{F}}_n$.

Based on the bootstrap data (Y_n^*, X_n), compute OLS estimates of $(\ddot{\beta}_n, \ddot{\sigma}_n)$, OLS residuals and LM test (5.9), to give bootstrap analogue of $\mathrm{LM}_{\mathrm{SED}}|_{H_0}$:

$$\mathrm{LM}_{\mathrm{SE}}^* = \frac{n}{\sqrt{K_n}} \frac{e_n^{*\prime} M_n W_n M_n e_n^*}{e_n^{*\prime} M_n e_n^*}, \tag{5.12}$$

showing that $\mathrm{LM}_{\mathrm{SE}}^*$ is invariant of $\ddot{\beta}_n$ and $\ddot{\sigma}_n^2$. It can be further shown that $\mathrm{LM}_{\mathrm{SE}}^*|_{H_0} \xrightarrow{D^*} N(0,1)$, where D^* denotes 'w.r.t. $\ddot{\mathcal{F}}_n$'.

Comparing (5.12) with (5.11), it is intuitively quite clear that if e_n^* are drawn from an EDF $\ddot{\mathcal{F}}_n$ that consistently estimates $\mathcal{F}$ whether or not H_0 is true, then the EDF of $\mathrm{LM}_{\mathrm{SED}}^*$ offers a consistent estimate of the finite sample distribution of $\mathrm{LM}_{\mathrm{SED}}|_{H_0}$.

This is just like the Monte Carlo approach under a known $\mathcal{F}$. However, with $\ddot{\mathcal{F}}_n$ the finite sample distribution of $\mathrm{LM}_{\mathrm{SED}}|_{H_0}$ is estimated nonparametrically. With this in mind, the attractiveness of the bootstrap approach becomes clearer.

Proposition 5.5. *Suppose Model (5.8) satisfies Assumptions S1-S3. If* (i) $\hat{\rho}_n$ *is* $\sqrt{n}$*-consistent, and* (ii) $|\text{LM}_{\text{SE}}^{\text{FI}}|_{H_0}| \leq U$ *a.e., and* $E(U^4)$ *exists, then the bootstrap methods under* $\text{RS}_{uu}, \text{RS}_{uf}$ *and* RS_{ru} *are valid for* $\text{LM}_{\text{SE}}^{\text{FI}}$, $\forall \mathcal{F}$. *If, in addition,* $\gamma = 0$*, the bootstrap methods under* RS_{ur} *and* RS_{rr} *are valid as well. The same conclusions apply to* $\text{SLM}_{\text{SE}}^{\circ}$.

Proof: Key arguments are given here and details are in Yang (2015).

- $\text{LM}_{\text{SE}}^{\text{FI}}|_{H_0} \xrightarrow{D} N(0,1), \forall \mathcal{F}$, and its CDF admits Edgeworth expansion:

$$\mathcal{G}_n(t, \mathcal{F}) = \Phi(t) + n^{-\frac{1}{2}} c_0^{-\frac{3}{2}} \phi(t)\, p(t, \mathcal{F}) + O(n^{-1}), \tag{5.13}$$

  where $p(t, \mathcal{F}) = -c_0 c_1 + (\frac{1}{6}\kappa_3^2 T_4 + T_5)(1 - t^2)$.

- $\text{LM}_{\text{SE}}^{*} \xrightarrow{D^*} N(0,1), \forall \ddot{\mathcal{F}}_n$, and its bootstrap CDF admits asymptotic expansion:

$$\mathcal{G}_n(t, \ddot{\mathcal{F}}_n) = \Phi(t) + n^{-\frac{1}{2}} c_0^{-\frac{3}{2}} \phi(t)\, p(t, \ddot{\mathcal{F}}_n) + O_p(n^{-1}), \tag{5.14}$$

  where $p(t, \ddot{\mathcal{F}}_n) = -c_0 c_1 + (\frac{1}{6}\ddot{\kappa}_{3n}^2 T_4 + T_5)(1 - t^2)$.

- Taking difference between (5.14) and (5.13) gives,

$$\mathcal{G}_n(t, \ddot{\mathcal{F}}_n) - \mathcal{G}_n(t, \mathcal{F}) = \tfrac{1}{6} n_r^{-\frac{1}{2}} c_0^{-\frac{3}{2}} T_4\, \phi(t)(1 - t^2)(\ddot{\kappa}_{3n}^2 - \kappa_3^2) + O_p(n^{-1}).$$

Remark 5.3: When the error distribution is skewed, the bootstrap methods under RS_{ur} and RS_{rr}, though not strictly valid, improve the asymptotic method as the main second-order terms involving $c_0 c_1$ and T_5 are captured by the bootstrap methods, resulting partial asymptotic refinements.

Remark 5.4: The detailed proof given in Appendix B, Yang (2015), shows that the first three cumulants of $\mathrm{LM}_{\mathrm{SED}}|_{H_0}$ are:

$$\begin{aligned}
\boldsymbol{\kappa}_{1,n} &= n_r^{-\frac{1}{2}} c_0^{-\frac{1}{2}} c_1 + O(n_r^{-\frac{3}{2}}), \\
\boldsymbol{\kappa}_{2,n} &= 1 + O(n_r^{-1}), \quad \text{and} \\
\boldsymbol{\kappa}_{3,n} &= n_r^{-\frac{1}{2}} c_0^{-\frac{3}{2}} (\kappa_3^2 T_4 + 6T_5) + O(n_r^{-\frac{3}{2}}),
\end{aligned}$$

from which we see precisely the reason why the finite sample distribution of a spatial LM test deviates more from its limiting distribution with a denser spatial weight matrix.

- The other three variants of the LM tests, $\mathrm{I}^{\circ}_{\mathrm{SE}}$, $\mathrm{LM}^{\mathrm{MD}}_{\mathrm{SE}}$, and $\mathrm{SLM}^{\mathrm{MD}}_{\mathrm{SE}}$, introduced in Lecture 3 can also be considered.
- Theorem 3.1 shows that under H_0 the five test statistics are asymptotically equivalent and are robust against nonnormality. Thus, any of them can be used for bootstrapping.
- Clearly, $\mathrm{LM}^{\mathrm{FI}}_{\mathrm{SE}}$ is the simplest, and is thus used.
- **Conclusion:** To bootstrap the FCVs, the three resampling schemes with unrestricted residuals are recommended in particular RS_{uu}, although the two with restricted residuals are partially valid.

The advantages of using a bootstrap LM test are:

(i) finite sample performance of the test statistic to be bootstrapped is unimportant as long as it is asymptotically pivotal under the null, and

(ii) second-order refinements are achieved for both two-sided tests and one-sided tests (if test is univariate).

The Monte Carlo experiments are carried out based on the following DGP:

$$Y_n = \beta_0 1_n + X_{n1}\beta_1 + X_{n2}\beta_2 + u_n, \quad u_n = \rho W_n u_n + \sigma e_n,$$

where the parameter values are set at $\beta = \{5, 1, 1\}'$ and $\sigma = 1$ or 2.

- Four different sample sizes are considered, i.e., $n = 50, 100, 200$, and 500.
- Each set of results is based on $M = 2,000$ Monte Carlo samples, and $B = 699$ bootstrap samples for each Monte Carlo sample.
- Details on spatial layouts, error distributions, and fixed regressors' values are given in Appendix C, Yang (2015).
- For $\rho = \{-0.75, -0.5, -0.25, 0, 0.25, 0.5, 0.75\}$, two types of Monte Carlo results are recorded:

  (a) the means and standard deviations of the bootstrap critical values, and

  (b) the rejection frequencies of the LM and SLM tests.

General observations from Monte Carlo results are:

- The (average) BCVs are all very close to the 'true' finite sample critical values (obtained by Monte Carlo simulation), but can all be far from their asymptotic critical values (ACV) which are ± 1.6449 and ± 1.96;
- The BCVs for both LM and SLM under RS_{uf}, RS_{uu} and RS_{ru} are all very stable, and those under RS_{ur} and RS_{rr} change with ρ slightly, which confirms the Remark 5.3.
- The standard deviations of BCVs are all small;
- Use of BCVs significantly improves size and power of LM tests;
- Regressors and spatial weight matrices affect the finite sample behavior of the test. In all these scenarios, standardization helps and bootstrap methods work well in improving LM tests.
- However, standardization does not play a major role as far as BCV is concerned as the original LM is asymptotically robust against distributional misspecification (DM).

Table 5.1. Bootstrap and MC Critical Values for Burridge's LM Test of SE Dependence

spatial Layout: Group Interaction with $g = n^{0.5}$; $H_0 : \rho = 0; \sigma = 2$

Method	ρ	$n = 100$ 2.5%	5%	95%	97.5%	$n = 200$ 2.5%	5%	95%	97.5%
Normal Error									
RS_{rr}	0.0	-1.863	-1.717	1.097	1.509	-1.883	-1.729	1.166	1.580
	0.5	-1.862	-1.717	1.103	1.516	-1.883	-1.729	1.166	1.582
RS_{uu}	0.0	-1.863	-1.717	1.096	1.510	-1.883	-1.730	1.165	1.580
	0.5	-1.862	-1.716	1.103	1.518	-1.882	-1.729	1.166	1.580
MC		-1.872	-1.722	1.113	1.505	-1.851	-1.692	1.222	1.648
Normal Mixture Error									
RS_{rr}	0.0	-1.878	-1.701	1.039	1.446	-1.877	-1.714	1.130	1.538
	0.5	-1.879	-1.704	1.047	1.456	-1.879	-1.717	1.132	1.539
RS_{uu}	0.0	-1.878	-1.701	1.039	1.448	-1.877	-1.715	1.130	1.539
	0.5	-1.880	-1.701	1.039	1.450	-1.878	-1.715	1.127	1.537
MC		-1.9158	-1.7221	1.0062	1.403	-1.855	-1.691	1.195	1.575
Log-Normal Error									
RS_{rr}	0.0	-1.824	-1.651	1.001	1.440	-1.816	-1.659	1.087	1.522
	0.5	-1.834	-1.663	1.010	1.447	-1.828	-1.669	1.096	1.527
RS_{uu}	0.0	-1.826	-1.652	1.003	1.441	-1.816	-1.660	1.087	1.521
	0.5	-1.823	-1.651	1.000	1.442	-1.818	-1.660	1.088	1.527
MC		-1.863	-1.664	0.984	1.449	-1.786	-1.624	1.113	1.542

RS_{rr} and RS_{uu}: Average bootstrap critical values based on $M = 2,000$ and $B = 699$;

MC: Monte Carlo critical values based on $M = 30,000$; Regressors generated according to XVal-B

Table 5.2a. Rejection Frequencies for One-Sided LM Test of Spatial Error Dependence

spatial Layout: Group Interaction with $g = n^{0.5}$; $H_0 : \rho = 0$

$\lvert\rho\rvert$	$n = 100$				$n = 200$			
	L2.5%	L5%	R5%	R2.5%	L2.5%	L5%	R5%	R2.5%
Normal Error								
ACV: LM test referring to ACV								
0.00	0.0155	0.0690	0.0175	0.0110	0.0200	0.0740	0.0200	0.0115
0.25	0.0700	0.2440	0.2295	0.1760	0.0930	0.2740	0.3245	0.2625
0.50	0.1925	0.4735	0.7865	0.7400	0.3005	0.6000	0.8860	0.8585
RS_{rr}								
0.00	0.0270	0.0530	0.0485	0.0215	0.0290	0.0565	0.0445	0.0235
0.25	0.1165	0.2000	0.3550	0.2520	0.1370	0.2270	0.4450	0.3410
0.50	0.2715	0.4105	0.8570	0.8050	0.3875	0.5245	0.9280	0.8915
RS_{uu}								
0.00	0.0265	0.0530	0.0470	0.0210	0.0300	0.0555	0.0435	0.0235
0.25	0.1170	0.2020	0.3595	0.2560	0.1375	0.2235	0.4430	0.3390
0.50	0.2740	0.4060	0.8565	0.8030	0.3845	0.5275	0.9280	0.8915
ACV*: SLM test referring to ACV								
0.00	0.0015	0.0170	0.0705	0.0420	0.0300	0.0555	0.0435	0.0235
0.25	0.0145	0.0825	0.4035	0.3440	0.1375	0.2235	0.4430	0.3390
0.50	0.0555	0.2180	0.8785	0.8465	0.3845	0.5275	0.9280	0.8915

Note: L = Left tail ($\rho < 0$), R = Right tail ($\rho > 0$); Regressors generated according to XVal-B

Table 5.2b. Rejection Frequencies for One-Sided LM Test of Spatial Error Dependence

spatial Layout: Group Interaction with $g = n^{0.5}$; $H_0 : \rho = 0$

$\|\rho\|$	$n = 100$ L2.5%	L5%	R5%	R2.5%	$n = 200$ L2.5%	L5%	R5%	R2.5%
Normal Mixture Error								
ACV: LM test referring to ACV								
0.00	0.0165	0.0605	0.0150	0.0090	0.0155	0.0540	0.0205	0.0135
0.25	0.0710	0.2110	0.2305	0.1735	0.0945	0.2635	0.3325	0.2625
0.50	0.2045	0.4460	0.7815	0.7390	0.2940	0.5850	0.9020	0.8705
RS_{rr}								
0.00	0.0250	0.0530	0.0480	0.0230	0.0215	0.0415	0.0520	0.0245
0.25	0.0930	0.1730	0.3705	0.2710	0.1355	0.2255	0.4575	0.3600
0.50	0.2580	0.3915	0.8690	0.8110	0.3595	0.5290	0.9410	0.9120
RS_{uu}								
0.00	0.0245	0.0510	0.0475	0.0235	0.0200	0.0415	0.0515	0.0245
0.25	0.0925	0.1780	0.3700	0.2735	0.1365	0.2235	0.4570	0.3575
0.50	0.2550	0.3935	0.8675	0.8105	0.3665	0.5335	0.9410	0.9105
ACV*: SLM test referring to ACV								
0.00	0.0045	0.0170	0.0600	0.0370	0.0065	0.0165	0.0690	0.0425
0.25	0.0325	0.0790	0.4070	0.3410	0.0260	0.0995	0.5050	0.4270
0.50	0.0960	0.2145	0.8855	0.8535	0.0860	0.3045	0.9480	0.9325

Note: L = Left tail ($\rho < 0$), R = Right tail ($\rho > 0$); Regressors generated according to XVal-B

Table 5.2c. Rejection Frequencies for One-Sided LM Test of Spatial Error Dependence

spatial Layout: Group Interaction with $g = n^{0.5}$; $H_0 : \rho = 0$

$\lvert\rho\rvert$	$n = 100$				$n = 200$			
	L2.5%	L5%	R5%	R2.5%	L2.5%	L5%	R5%	R2.5%
Log-Normal Error								
ACV: LM test referring to ACV								
0.00	0.0125	0.0490	0.0180	0.0090	0.0150	0.0530	0.0210	0.0110
0.25	0.0735	0.1975	0.2190	0.1630	0.0820	0.2605	0.3115	0.2395
0.50	0.2120	0.4350	0.7910	0.7340	0.2805	0.5605	0.9180	0.8900
RS_{rr}								
0.00	0.0295	0.0440	0.0485	0.0240	0.0250	0.0495	0.0660	0.0285
0.25	0.1155	0.1950	0.3600	0.2605	0.1525	0.2485	0.4540	0.3460
0.50	0.2860	0.4235	0.8870	0.8165	0.4090	0.5460	0.9525	0.9255
RS_{uu}								
0.00	0.0290	0.0445	0.0490	0.0250	0.0255	0.0495	0.0635	0.0290
0.25	0.1155	0.1965	0.3650	0.2580	0.1560	0.2520	0.4550	0.3470
0.50	0.2905	0.4255	0.8865	0.8170	0.4110	0.5525	0.9530	0.9230
ACV*: SLM test referring to ACV								
0.00	0.0045	0.0140	0.0535	0.0375	0.0015	0.0165	0.0720	0.0470
0.25	0.0310	0.0760	0.3915	0.3140	0.0255	0.0870	0.4825	0.4025
0.50	0.1170	0.2185	0.9015	0.8630	0.0985	0.2925	0.9570	0.9430

Note: L = Left tail ($\rho < 0$), R = Right tail ($\rho > 0$); Regressors generated according to XVal-B

Recall linear regression with SL dependence or SAR model:

$$Y_n = \lambda W_n Y_n + X_n\beta + \epsilon_n, \;\; \epsilon_n = \sigma e_n. \tag{5.15}$$

The LM test statistic for testing H_0: $\lambda = 0$ is:

$$\mathrm{LM}_{\mathrm{SL}}^{\mathrm{FI}} = \frac{\tilde{\epsilon}_n' W_n Y_n}{\tilde{\sigma}_n^2 \sqrt{\tilde{\eta}_n' M_n \tilde{\eta}_n + K_n}}, \tag{5.16}$$

where $K_n = \mathrm{tr}(W_n' W_n + W_n W_n)$, $\tilde{\eta}_n = \frac{1}{\tilde{\sigma}_n} W_n X_n \tilde{\beta}_n$, $\tilde{\beta}_n$ and $\tilde{\sigma}_n^2$ are the OLS estimates from regressing Y_n on X_n, and $\tilde{\epsilon}_n$ the OLS residuals.

A standardized version of $\mathrm{LM}_{\mathrm{SLD}}$, having better finite sample properties and more robust to spatial layouts, is given in Yang and Shen (2011):

$$\mathrm{SLM}_{\mathrm{SL}}^{\circ} = \frac{\tilde{\epsilon}_n' W_n^{\circ} Y_n}{\tilde{\sigma}_n^2 \sqrt{\tilde{\eta}_n' M_n \tilde{\eta}_n + K_n^{\dagger} + \tilde{\kappa}_n d_n' d_n + 2\tilde{\gamma}_n \tilde{\eta}_n' M_n d_n}}, \tag{5.17}$$

$W_n^{\circ} = W_n - \frac{1}{n-k}\mathrm{tr}(W_n M_n) I_n$, $K_n^{\dagger} = \mathrm{tr}[\mathcal{A}_n(\mathcal{A}_n + \mathcal{A}_n')]$, $a_n = \mathrm{diagv}(\mathcal{A}_n)$, $\mathcal{A}_n = M_n W_n^{\circ}$, $d_n = \mathrm{diagv}(W_n^{\circ})$, and $\tilde{\gamma}_n$ and $\tilde{\kappa}_n$ are 3rd and 4th cumulants of $\tilde{e}_n = \tilde{\sigma}_n^{-1}\tilde{\varepsilon}_n$.

- Yang and Shen (2011) show that both $\mathrm{LM}_{\mathtt{SL}}^{\mathtt{FI}}$ and $\mathrm{SLM}_{\mathtt{SL}}^{\circ}$ have limiting null distribution $N(0,1)$, whether or not $\mathcal{F}$ is correctly specified, showing that both are asymptotically robust against distributional misspecification.
- To implement the bootstrap method under the resampling scheme $\mathtt{RS}_{uf}$, more general LM statistics for a nonzero λ, $\mathrm{LM}_{\mathtt{SL}}^{\mathtt{FI}}(\lambda)$, and its standardized version, $\mathrm{SLM}_{\mathtt{SL}}^{\circ}(\lambda)$, can be found in Yang and Shen (2011).
- The OPMD variants of $\mathrm{LM}_{\mathtt{SL}}^{\mathtt{FI}}$ and $\mathrm{SLM}_{\mathtt{SL}}^{\circ}$ can be developed and BCVs can be obtained and used for a refined inference.

To study the validity of various resampling schemes, we concentrate on $\mathrm{LM}_{\mathtt{SL}}^{\mathtt{FI}}$. Under the real world null DGP: $Y_n = X_n\beta + \sigma e_n$, we have

$$\mathrm{LM}_{\mathtt{SL}}^{\mathtt{FI}}|_{H_0} = \frac{\sqrt{n}(e_n' M_n W_n e_n + e_n' M_n \eta_n)}{(e_n' M_n e_n)^{\frac{1}{2}} \left\{\eta_n' M_n \eta_n + Q(e_n) + 2e_n' P_n W_n M_n \eta_n\right\}^{\frac{1}{2}}}, \quad (5.18)$$

$Q(e_n) = n^{-1} K_n e_n' M_n e_n + e_n' P_n' W_n' M_n W_n P_n e_n$, $\eta_n = \sigma^{-1} W_n X_n \beta$, and $P_n = I_n - M_n$.

- This shows that $\mathrm{LM}_{\mathtt{SL}}^{\mathtt{FI}}|_{H_0} = f(e_n, X_n, W_n, \beta, \sigma)$, meaning that $\mathrm{LM}_{\mathtt{SLD}}|_{H_0}$ is not an exact pivot whether or not $\mathcal{F}$ is known.
- This stands in difference from $\mathrm{LM}_{\mathtt{SE}}^{\mathtt{FI}}|_{H_0}$ considered earlier.
- The dependence of $\mathrm{LM}_{\mathtt{SL}}^{\mathtt{FI}}|_{H_0}$ on (β, σ^2) is expected to impose constraints on the choices of their estimates to be used as parameters in the bootstrap DGP.
- On the other hand, the limiting distribution of $\mathrm{LM}_{\mathtt{SL}}^{\mathtt{FI}}|_{H_0}$ does not depend on (β, σ^2) and $\mathcal{F}$, suggesting (as in Section 5.3.1) that bootstrap methods can be applied to provide asymptotically refined critical values for $\mathrm{LM}_{\mathtt{SL}}^{\mathtt{FI}}|_{H_0}$.

Under the bootstrap world, the bootstrap DGP that mimics the real world null DGP takes the form: $Y_n^* = X_n\ddot{\beta}_n + \ddot{\sigma}_n e_n^*$, where $e_{ni}^* \stackrel{iid}{\sim} \ddot{\mathcal{F}}_n$.

Based on bootstrap data (Y_n^*, X_n), estimating bootstrap model and computing test statistic (5.16) lead to bootstrap analogue of $\mathrm{LM}_{\mathtt{SL}}^{\mathtt{FI}}|_{H_0}$:

$$\mathrm{LM}_{\mathtt{SL}}^* = \frac{\sqrt{n}(e_n^{*\prime} M_n W_n e_n^* + e_n^{*\prime} M_n \ddot{\eta}_n)}{(e_n^{*\prime} M_n e_n^*)^{\frac{1}{2}} \left\{\ddot{\eta}_n' M_n \ddot{\eta}_n + Q(e_n^*) + 2e_n^{*\prime} P_n W_n M_n \ddot{\eta}_n\right\}^{\frac{1}{2}}}, \qquad (5.19)$$

where $\ddot{\eta}_n = \ddot{\sigma}_n^{-1} W_n X_n \ddot{\beta}_n$.

Comparing (5.19) with (5.18), it is intuitively clear that for bootstrap to provide a higher-order approximation to the FCVs of $\mathrm{LM}_{\mathtt{SLD}}|_{H_0}$, it is necessary that $\ddot{\beta}_n, \ddot{\sigma}_n^2$, and $\ddot{\mathcal{F}}_n$ are consistent whether or not H_0 is true.

Following proposition summarizes the results.

Proposition 5.6. *Suppose Model* (5.15) *satisfies Assumptions S1-S3. If* (*i*) $\hat{\lambda}_n$ *is* $\sqrt{n}$*-consistent, and* (*ii*) $|\mathrm{LM}_{\mathtt{SL}}^{\mathtt{FI}}|_{H_0}| \leq U$ *a.e., and* $E(U^4)$ *exists, then the bootstrap methods under* $\mathtt{RS}_{uu}$ *and* $\mathtt{RS}_{uf}$ *are valid for* $\mathrm{LM}_{\mathtt{SL}}^{\mathtt{FI}}$, $\forall \mathcal{F}$*. Further, if* $\gamma = 0$ *and conditions in (A.3) (Yang 2015) hold, the bootstrap methods under* $\mathtt{RS}_{ur}$ *is also valid. The same conclusions apply to* $\mathrm{SLM}_{\mathtt{SL}}^{\circ}$*.*

Proof: Key arguments are given here and details are in Yang (2015).

First, $\mathrm{LM}_{\mathtt{SL}}^{\mathtt{FI}}|_{H_0} \xrightarrow{D} N(0,1), \forall \mathcal{F}$; its finite sample CDF admits Edgeworth expansion:

$$\mathcal{G}_n(t,\theta,\mathcal{F}) = \Phi(t) + n^{-\frac{1}{2}} c_0(\theta)^{-\frac{3}{2}} \phi(t)\, p(t,\theta,\mathcal{F}) + O(n^{-1}), \tag{5.20}$$

$$p(t,\theta,\mathcal{F}) = -\, c_0(\theta) c_1 + [\tfrac{1}{6}\kappa_3^2 T_4 + T_5 + \tfrac{1}{6}\kappa_3(S_3(\theta) + 2S_5(\theta)) + \tfrac{1}{3}S_4(\theta)](1-t^2).$$

Similarly, $\mathrm{LM}_{\mathtt{SLD}}^{*} \xrightarrow{D^*} N(0,1)$; its bootstrap CDF admits asymptotic expansion:

$$\mathcal{G}_n(t,\ddot{\theta}_n,\ddot{\mathcal{F}}_n) = \Phi(t) + n^{-\frac{1}{2}} c_0(\ddot{\theta}_n)^{-\frac{3}{2}} \phi(t)\, p(t,\ddot{\theta}_n,\ddot{\mathcal{F}}_n) + O_p(n^{-1}), \tag{5.21}$$

$$p(t,\ddot{\theta}_n,\ddot{\mathcal{F}}_n) = -\, c_0(\ddot{\theta}_n) c_1 + [\tfrac{1}{6}\ddot{\kappa}_{3n}^2 T_4 + T_5 + \tfrac{1}{6}\ddot{\kappa}_{3n}(S_3(\ddot{\theta}_n) + 2S_5(\ddot{\theta}_n)) + \tfrac{1}{3}S_4(\ddot{\theta}_n)](1-t^2).$$

With these two expansions, the conclusions reached in Proposition 5.6 are clear.

- $\mathcal{G}_n(t, \ddot{\theta}_n, \ddot{\mathcal{F}}_n) - \mathcal{G}_n(t, \theta, \mathcal{F}) = O_p(n_r^{-1})$ only when $\ddot{\theta}_n = \hat{\theta}_n$ and $\ddot{\mathcal{F}}_n = \hat{\mathcal{F}}_n$.
- Similar to the SE model, $p(t, \theta, \mathcal{F})$ depends on $\mathcal{F}$ only through γ, thus $\hat{\mathcal{F}}_n$ can be replaced by $\tilde{\mathcal{F}}_n$ if additional conditions hold, leading to the validity of $\mathtt{RS}_{ur}$.
- Finally, the same set of results are obtained for $\mathrm{SLM}^{\circ}_{\mathrm{SL}}$.

Remark 5.5: When the error distribution is skewed, the bootstrap method under $\mathtt{RS}_{ur}$, though not strictly valid, improves upon the asymptotic method as the main second-order terms involving T_5, $c_0(\theta)c_1$ and $S_4(\theta)$ are captured by bootstrap due to the consistency of $\hat{\theta}_n$, leading to the so-called partial asymptotic refinements.

This explains why the Monte Carlo results (not reported for brevity) under $\mathtt{RS}_{ur}$ are very similar to these under $\mathtt{RS}_{uu}$ even when the errors are skewed.

Remark 5.6: Again, the cumulants of $\mathrm{LM}_{\mathrm{SLD}}|_{H_0}$ given in Appendix B, Yang (2015) , show clearly the effect of spatial weight matrix on the finite sample distribution of $\mathrm{LM}_{\mathrm{SLD}}|_{H_0}$.

We note that for the SL model,

- Restricted estimates $\tilde{\beta}_n$ and $\tilde{\sigma}_n^2$ are not consistent if H_0 is false,
- Neither $\mathrm{LM}_{\mathtt{SL}}^{\mathtt{FI}}$ nor $\mathrm{SLM}_{\mathtt{SL}}^{\mathtt{FI}}$ is invariant of β and σ^2,
- Both $\mathrm{LM}_{\mathtt{SL}}^{\mathtt{FI}}$ and $\mathrm{SLM}_{\mathtt{SL}}^{\mathtt{FI}}$ are asymptotic $N(0,1)$ for any $\mathcal{F}$.

Conclusion:

1. The resampling methods with restricted estimates, $\mathtt{RS}_{\mathtt{rr}}$ and $\mathtt{RS}_{\mathtt{ru}}$, are not valid for both tests;
2. All resampling methods with unrestricted estimates, $\mathtt{RS}_{\mathtt{ur}}$, $\mathtt{RS}_{\mathtt{uu}}$ and $\mathtt{RS}_{\mathtt{uf}}$, are recommended, in particular the $\mathtt{RS}_{\mathtt{uu}}$ scheme,
3. Standardization does not play a major role in bootstrapping critical values as the regular LM test is asymptotically robust against DM,
4. although it does in improving regular LM tests (two-sided).

The finite sample performance of $\text{LM}^{\mathtt{FI}}_{\mathtt{SL}}$ and $\text{SLM}^{\mathtt{FI}}_{\mathtt{SL}}$ for testing H_0: $\lambda = 0$ vs H_a: $\lambda < 0$ or H_a: $\lambda > 0$, when referring to ACV and BCV are investigated, based on the following data generating process:

$$Y_n = \lambda W_n Y_n + \beta_0 1_n + X_{n1}\beta_1 + X_{n2}\beta_2 + \epsilon_n,$$

under various resampling schemes, in terms of

- accuracy and stability of the BCVs with respect to the true value of λ,
- and size and power of the tests.

In a similar way as for the SE model, the parameter values are chosen, and W_n, X_n and ϵ_n are generated.

For BCVs, Monte Carlo results clearly reveal the following:

- BCVs are not affected much by the type of residuals used, consistent with the discussions given in Remark 5.5;
- BCVs can be quite different from the corresponding ACVs, showing the necessity of using finite sample critical values for testing the existence of spatial lag dependence in a linear regression model;
- BCVs based on RS_{rr} (and RS_{ru}) vary significantly with λ. This suggests that, if when H_0 is true the BCVs and sizes are accurate (indeed they are), then when H_0 is false, the BCVs cannot be accurate and therefore the powers cannot be reliable;
- BCVs based on RS_{uu} are very stable with respect to λ, and are very accurate as they agree very well with the corresponding Monte Carlo critical values obtained by imposing H_0 and using $M = 30,000$, and with the BCVs under RS_{rr} and H_0 (considered as an ideal situation). The same holds when $|\lambda|$ further increases from 0.5.
- BVCs do not depend much on the error distribution.

For size and power, Monte Carlo results reveal the following:

- The tests referring to ACVs can have severe size-distortion, and more so with heavier spatial dependence. Referring to BCVs effectively remove the size distortions under any resampling method, but this is unachievable with the restricted estimates as in practice whether H_0 is true or false is unknown.
- BCVs of LM statistic based on restricted estimates tend to increase **in magnitude** as λ increases. As a result, power tends to be lower for a right-tailed test, and higher for a left-tailed test, compared with the power of the tests based on the unrestricted estimates.
- BCVs of SLM statistic based on restricted estimates decrease as λ increases. As a result, power of both left- and right-tailed tests tends to be higher than that based on unrestricted estimates. However, the former corresponds to a larger size due to smaller underlining BCVs.
- As the original LM test is already asymptotically pivotal and robust, standardization does not provide further improvements on the methods in that the use of restricted estimates still lead to BCVs that vary with λ.

Table 5.3a. Bootstrap Critical Values for LM and SLM Tests of Spatial Lag Dependence

Spatial Layout: Group Interaction with $g = n^{0.5}$; $n = 100$; $\sigma = 1$; XVal-B

λ	LM Test L2.5%	LM Test L5%	LM Test U5%	LM Test U2.5%	SLM Test L2.5%	SLM Test L5%	SLM Test U5%	SLM Test U2.5%
Normal Error								
RS_{rr}								
-0.5	-2.0718	-1.8294	1.2718	1.6270	-1.8282	-1.5691	1.7465	2.1265
-0.3	-2.0872	-1.8313	1.2960	1.6438	-1.8529	-1.5813	1.7331	2.1033
0.0	-2.1064	-1.8372	1.3469	1.6844	-1.8904	-1.6090	1.7195	2.0722
0.3	-2.1144	-1.8318	1.4030	1.7303	-1.9238	-1.6322	1.7031	2.0407
0.5	-2.1135	-1.8245	1.4375	1.7608	-1.9383	-1.6417	1.6994	2.0307
RS_{uu}								
-0.5	-2.1034	-1.8378	1.3510	1.6849	-1.8908	-1.6133	1.7145	2.0635
-0.3	-2.1030	-1.8312	1.3507	1.6870	-1.8905	-1.6072	1.7121	2.0638
0.0	-2.1064	-1.8363	1.3559	1.6924	-1.8949	-1.6127	1.7163	2.0682
0.3	-2.1099	-1.8376	1.3563	1.6908	-1.8982	-1.6139	1.7183	2.0667
0.5	-2.1049	-1.8366	1.3578	1.6898	-1.8929	-1.6132	1.7184	2.0655
MC								
0	-2.1190	-1.8415	1.3262	1.6512	-1.9018	-1.6117	1.7002	2.0447

RS_{rr} and RS_{uu}: Average bootstrap critical values based on $M = 2,000$ and $B = 699$;

MC: Monte Carlo critical values based on $M = 30,000$; Regressors generated according to XVal-B

Table 5.3b. Bootstrap Critical Values for LM and SLM Tests of Spatial Lag Dependence

Spatial Layout: Group Interaction with $g = n^{0.5}$; $n = 100$; $\sigma = 1$; XVal-B

	LM Test				SLM Test			
λ	L2.5%	L5%	U5%	U2.5%	L2.5%	L5%	U5%	U2.5%
Normal Mixture Error								
	RS_{rr}							
-0.5	-2.0640	-1.8098	1.2502	1.6027	-1.8228	-1.5513	1.7074	2.0825
-0.3	-2.0809	-1.8167	1.2730	1.6198	-1.8494	-1.5695	1.6954	2.0620
0.0	-2.0941	-1.8170	1.3308	1.6675	-1.8818	-1.5923	1.6900	2.0411
0.3	-2.1066	-1.8191	1.3962	1.7254	-1.9197	-1.6235	1.6859	2.0250
0.5	-2.1095	-1.8175	1.4302	1.7542	-1.9361	-1.6367	1.6885	2.0196
	RS_{uu}							
-0.5	-2.0972	-1.8206	1.3424	1.6743	-1.8888	-1.6003	1.6899	2.0362
-0.3	-2.1001	-1.8210	1.3401	1.6761	-1.8918	-1.6008	1.6887	2.0385
0.0	-2.0959	-1.8175	1.3414	1.6763	-1.8872	-1.5971	1.6898	2.0389
0.3	-2.0978	-1.8204	1.3428	1.6777	-1.8900	-1.6009	1.6899	2.0368
0.5	-2.0975	-1.8229	1.3425	1.6761	-1.8886	-1.6023	1.6913	2.0389
	MC							
0.0	-2.1175	-1.8320	1.3125	1.6077	-1.9059	-1.6033	1.6781	1.9927

RS_{rr} and RS_{uu}: Average bootstrap critical values based on $M = 2,000$ and $B = 699$;

MC: Monte Carlo critical values based on $M = 30,000$; Regressors generated according to XVal-B

Table 5.3c. Bootstrap Critical Values for LM and SLM Tests of Spatial Lag Dependence

Spatial Layout: Group Interaction with $g = n^{0.5}$; $n = 100$; $\sigma = 1$; XVal-B

λ	LM Test L2.5%	LM Test L5%	LM Test U5%	LM Test U2.5%	SLM Test L2.5%	SLM Test L5%	SLM Test U5%	SLM Test U2.5%
Log-Normal Error								
	RS_{rr}							
-0.5	-2.0232	-1.7734	1.2626	1.6337	-1.7806	-1.5159	1.6860	2.0766
-0.3	-2.0374	-1.7797	1.2960	1.6574	-1.8064	-1.5353	1.6806	2.0586
0.0	-2.0556	-1.7869	1.3500	1.6995	-1.8455	-1.5663	1.6759	2.0381
0.3	-2.0807	-1.7979	1.4160	1.7513	-1.8982	-1.6079	1.6794	2.0233
0.5	-2.0947	-1.8026	1.4362	1.7671	-1.9235	-1.6251	1.6797	2.0169
	RS_{uu}							
-0.5	-2.0612	-1.7899	1.3612	1.7118	-1.8549	-1.5735	1.6780	2.0391
-0.3	-2.0592	-1.7883	1.3631	1.7083	-1.8530	-1.5722	1.6782	2.0348
0.0	-2.0608	-1.7884	1.3581	1.7057	-1.8545	-1.5721	1.6764	2.0344
0.3	-2.0667	-1.7921	1.3664	1.7162	-1.8626	-1.5780	1.6790	2.0388
0.5	-2.0614	-1.7901	1.3601	1.7104	-1.8553	-1.5743	1.6762	2.0373
	MC							
0.0	-2.0276	-1.7597	1.3454	1.6944	-1.8154	-1.5290	1.6663	2.0354

RS_{rr} and RS_{uu}: Average bootstrap critical values based on $M = 2,000$ and $B = 699$;

MC: Monte Carlo critical values based on $M = 30,000$; Regressors generated according to XVal-B

Table 5.4a. Rejection Frequencies for LM Tests of Spatial Lag Dependence

`Spatial Layout: Group Interaction with` $g = n^{0.5}$; $\sigma = 1$; `XVal-B`

$\lvert\lambda\rvert$	$n = 50$ L2.5%	$n = 50$ L5%	$n = 50$ U5%	$n = 50$ U2.5%	$n = 100$ L2.5%	$n = 100$ L5%	$n = 100$ U5%	$n = 100$ U2.5%
Normal Error								
ACV								
0.0	0.0435	0.0970	0.0190	0.0085	0.0430	0.0875	0.0235	0.0095
0.1	0.1010	0.1905	0.0905	0.0550	0.1405	0.2300	0.1240	0.0805
0.2	0.2150	0.3510	0.2885	0.1985	0.2955	0.4400	0.4510	0.3430
0.3	0.3585	0.5420	0.6110	0.4990	0.4705	0.6410	0.8535	0.7690
RS_{rr}								
0.0	0.0285	0.0565	0.0485	0.0260	0.0305	0.0540	0.0445	0.0235
0.1	0.0655	0.1220	0.1640	0.0975	0.1045	0.1725	0.1960	0.1190
0.2	0.1555	0.2455	0.3975	0.2890	0.2405	0.3505	0.5495	0.4310
0.3	0.2870	0.4175	0.7135	0.6055	0.4075	0.5390	0.8920	0.8340
RS_{uu}								
0.0	0.0270	0.0575	0.0555	0.0280	0.0290	0.0555	0.0475	0.0245
0.1	0.0605	0.1195	0.1715	0.1030	0.0995	0.1755	0.2015	0.1255
0.2	0.1415	0.2440	0.4070	0.3020	0.2325	0.3500	0.5590	0.4420
0.3	0.2610	0.4025	0.7260	0.6220	0.3955	0.5350	0.8935	0.8410

Table 5.4b. Rejection Frequencies for SLM Tests of Spatial Lag Dependence

Spatial Layout: Group Interaction with $g = n^{0.5}$; $\sigma = 1$; XVal-B

$\lvert\lambda\rvert$	$n = 50$				$n = 100$			
	L2.5%	L5%	U5%	U2.5%	L2.5%	L5%	U5%	U2.5%
Normal Error								
ACV								
0.0	0.0230	0.0475	0.0635	0.0355	0.0235	0.0495	0.0510	0.0280
0.1	0.0520	0.1100	0.1865	0.1160	0.0915	0.1610	0.2065	0.1350
0.2	0.1265	0.2290	0.4345	0.3325	0.2050	0.3340	0.5740	0.4575
0.3	0.2275	0.3755	0.7500	0.6505	0.3500	0.5055	0.9050	0.8510
RS_{rr}								
0.0	0.0280	0.0565	0.0520	0.0280	0.0300	0.0535	0.0450	0.0240
0.1	0.0655	0.1215	0.1695	0.0980	0.1050	0.1715	0.1955	0.1190
0.2	0.1540	0.2440	0.4005	0.2950	0.2380	0.3505	0.5525	0.4355
0.3	0.2865	0.4145	0.7190	0.6120	0.4075	0.5350	0.8925	0.8340
RS_{rr}								
0.0	0.0235	0.0520	0.0515	0.0235	0.0250	0.0525	0.0450	0.0220
0.1	0.0575	0.1105	0.1650	0.0925	0.0960	0.1690	0.1930	0.1175
0.2	0.1340	0.2270	0.4005	0.2920	0.2240	0.3395	0.5485	0.4220
0.3	0.2485	0.3910	0.7165	0.6060	0.3820	0.5230	0.8905	0.8320

Kelejian and Robinson (1995) proposed an SEC model:

$$Y_n = X_n\beta + u_n, \quad \text{with } u_n = W_n v_n + \epsilon_n, \text{and } \epsilon_n = \sigma e_n. \tag{5.22}$$

The error components v_n and ε_n are assumed to be independent, with iid elements of mean zero and variances σ_v^2 and σ^2, respectively.

- The SEC model provides a useful alternative to the traditional spatial models with a SAR or SMA error process, in particular
- when spatial autocorrelation is constrained to close neighbors, e.g., spatial spillovers in the productivity of infrastructure investments (Kelejian and Robinson, 1997; Anselin and Moreno, 2003).

Let $\lambda = \sigma_v^2/\sigma^2$ and $\Omega_n(\lambda) = I_n + \lambda W_n W_n'$. Then, Var($u_n$) = $\sigma^2\Omega_n(\lambda)$.

If v_n and ϵ_n are normal, then $u_n \sim N(0,\ \sigma^2\Omega_n(\lambda))$ and Gaussian likelihood for $\theta = (\beta', \sigma^2, \lambda)'$ can be formulated. QML estimation of θ proceeds.

We are interested in testing H_0: $\lambda = 0$ vs H_a: $\lambda > 0$.

Or equivalently, H_0: $\sigma_v = 0$ vs H_a: $\sigma_v > 0$. The alternative can only be one-sided as σ_v is non-negative. Anselin (2001) derived an LM test based on the assumptions that errors are normally distributed:

$$\mathrm{LM}_{\mathtt{SEC}}^{\mathtt{FI}} = \frac{n}{\sqrt{K_n}} \frac{\tilde{\epsilon}_n' H_n \tilde{\epsilon}_n}{\tilde{\epsilon}_n' \tilde{\epsilon}_n}, \tag{5.23}$$

$H_n = W_n W_n' - \frac{1}{n}\mathrm{tr}(W_n W_n') I_n$, $K_n = 2\mathrm{tr}(H_n^2)$, and $\tilde{\varepsilon}_n$ is the vector of OLS residuals. The limiting null distribution of $\mathrm{LM}_{\mathtt{SEC}}^{\mathtt{FI}}$ is $N(0,1)$ when $\mathcal{F} = \Phi$.

Yang (2010) provided a robust/standardized LM test:

$$\mathrm{SLM}_{\mathtt{SEC}}^{\circ} = \frac{n}{\sqrt{K_n^{\dagger} + \tilde{\kappa}_n a_n' a_n}} \frac{\tilde{\epsilon}_n' H_n^{\dagger} \tilde{\epsilon}_n}{\tilde{\epsilon}_n' \tilde{\epsilon}_n}, \tag{5.24}$$

where $H_n^{\dagger} = W_n W_n' - \frac{1}{n-k}\mathrm{tr}(W_n W_n' M_n) I_n$, $K_n^{\dagger} = 2\mathrm{tr}(\mathcal{A}_n^2)$, $a_n = \mathrm{diagv}(\mathcal{A}_n)$, $\mathcal{A}_n = M_n H_n^{\dagger} M_n$, and $\tilde{\kappa}_n$ is the 4th cumulant of $\tilde{e}_n = \tilde{\sigma}_n^{-1} \tilde{\varepsilon}_n$.

Yang (2010) showed that $\mathrm{SLM}_{\mathtt{SEC}}^{\circ}|_{H_0} \xrightarrow{D} N(0,1), \forall \mathcal{F}$.

Under H_0, $\tilde{\epsilon}'_n = M_n \epsilon_n = \sigma M_n e_n$, and the statistics $\mathrm{LM}^{\mathtt{FI}}_{\mathtt{SEC}}$ can be written as

$$\mathrm{LM}^{\mathtt{FI}}_{\mathtt{SEC}}|_{H_0} = \frac{n}{\sqrt{K_n}} \frac{e'_n M_n H_n M_n e_n}{e'_n M_n e_n}. \tag{5.25}$$

$\Rightarrow \mathrm{LM}^{\mathtt{FI}}_{\mathtt{SEC}}|_{H_0}$ is invariant of the nuisance parameters and thus a pivot if $\mathcal{F}$ is known. Monte Carlo method can be used to find its FCVs.

However, when $\mathcal{F}$ is unknown and possibly misspecified, $\mathrm{LM}^{\mathtt{FI}}_{\mathtt{SEC}}|_{H_0}$ is not even an asymptotic pivot. Indeed, Lemma A2 (Yang, 2015) leads to

$$(1+\kappa c_0)^{-\frac{1}{2}} \mathrm{LM}^{\mathtt{FI}}_{\mathtt{SEC}}|_{H_0} \xrightarrow{D} N(0,1), \ \forall \mathcal{F},$$

where $c_0 = \lim_{n\to\infty} K_n^{-1} b'_n b_n$ with $b_n = \mathrm{diagv}(M_n H_n M_n)$, and κ is the 4th cumulant of e_{ni} which is non-zero if e_{ni} is non-normal.

Thus, the limit distribution of $\mathrm{LM}^{\mathtt{FI}}_{\mathtt{SEC}}|_{H_0}$ depends on $\mathcal{F}$ through κ. Then, what is the consequence of ignoring this when conducting bootstrap?

The bootstrap DGP is again: $Y_n^* = X_n\ddot{\beta}_n + \ddot{\sigma}_n e_n^*$. Based on bootstrap data (Y_n^*, X_n), compute OLS estimates, OLS residuals, and statistic (5.23). Some algebra leads to the bootstrap analogue of (5.25):

$$\mathrm{LM}^*_{\mathrm{SEC}} = \frac{n}{\sqrt{K_n}} \frac{e_n^{*\prime} M_n H_n M_n e_n^*}{e_n^{*\prime} M_n e_n^*}. \tag{5.26}$$

Similarly, Lemma A2 (Yang, 2015) gives $(1+\ddot{\kappa}_n c_0)^{-\frac{1}{2}}\mathrm{LM}^*_{\mathrm{SEC}} \xrightarrow{D^*} N(0,1)$.

Therefore, the leading terms in the asymptotic expansions of bootstrap CDF of $\mathrm{LM}^*_{\mathrm{SEC}}$ and CDF of $\mathrm{LM}^{\mathrm{FI}}_{\mathrm{SEC}}|_{H_0}$ are such that,

$$\Phi(t/\sqrt{1+\ddot{\kappa}_n c_0}) - \Phi(t/\sqrt{1+\kappa c_0}) = o_p(1), \text{ if } \ddot{\mathcal{F}}_n = \hat{\mathcal{F}}_n; \; O_p(1) \text{ if } \ddot{\mathcal{F}}_n = \tilde{\mathcal{F}}_n.$$

This clearly shows that when $\mathcal{F}$ is misspecified the bootstrap method is unable to provide an improved approximation to the FCVs of $\mathrm{LM}^{\mathrm{FI}}_{\mathrm{SEC}}|_{H_0}$ over the ACVs even if the unrestricted residuals are used, and that the use of the restricted residuals worsens the approximation.

Therefore, one must use a robust LM statistic for bootstrap to provide a refined test. Similar algebra as for $\mathrm{LM}_{\mathtt{SEC}}^{\mathtt{FI}}|_{H_0}$ and $\mathrm{LM}_{\mathtt{SEC}}^{*}$ gives

$$\mathrm{SLM}_{\mathtt{SEC}}^{\circ}|_{H_0} = \frac{n}{\sqrt{K_n^{\dagger} + \kappa(e_n) a_n' a_n}} \frac{e_n' M_n H_n^{\dagger} M_n e_n}{e_n' M_n e_n}, \tag{5.27}$$

and its bootstrap analogue

$$\mathrm{SLM}_{\mathtt{SEC}}^{*} = \frac{n}{\sqrt{K_n^{\dagger} + \kappa(e_n^*) a_n' a_n}} \frac{e_n^{*\prime} M_n H_n^{\dagger} M_n e_n^*}{e_n^{*\prime} M_n e_n^*}, \tag{5.28}$$

where $\kappa(e_n)$ and $\kappa(e_n^*)$ are the 4th cumulants of $\frac{\sqrt{n} M_n e_n}{\sqrt{e_n' M_n e_n}}$ and e_n^*.

- Similar to $\mathrm{SLM}_{\mathtt{SEC}}^{\circ}|_{H_0} \xrightarrow{D} N(0,1), \forall \mathcal{F}$, $\mathrm{SLM}_{\mathtt{SEC}}^{*} \xrightarrow{D^*} N(0,1)$ for both $\hat{\mathcal{F}}_n$ and $\tilde{\mathcal{F}}_n$.
- The implication of these results is that when bootstrapping the standardized LM test given in (5.24), using either unrestricted residuals or restricted residuals leads to bootstrap critical values that are correct asymptotically.
- However, as stated in the following proposition, only the use of unrestricted residuals leads to full asymptotic refinements.

How can we get the unrestricted estimate of e_n in (5.22)?

- The SEC model is not the standard model considered in Sec. 5.2.
- The unrestricted estimate $\hat{u}_n = Y_n - X_n\hat{\beta}_n$ cannot be decomposed into $\hat{v}_n$ and $\hat{e}_n$ to give a consistent $\hat{\mathcal{F}}_n$ directly based $\hat{e}_n$, unless v_n and ε_n are normal.
- Note that $e_n^\circ = \sigma^{-1}\Omega_n^{-\frac{1}{2}}(\lambda)u_n \sim (0, I_n)$, which is normal if v_n and ε_n are.
- We consider its unrestricted estimate, $\hat{e}_n^\circ = \hat{\sigma}^{-1}\Omega_n^{-\frac{1}{2}}(\hat{\lambda}_n)\hat{u}_n$, as an approximation to $\hat{e}_n$, and draw bootstrap samples from $\hat{e}_n^\circ$.

Proposition 5.7. *Suppose Assumptions S1-S3 hold for Model (5.22) with $u_n = \Omega_n^{\frac{1}{2}}(\lambda)\epsilon_n$. If (i) $\hat{\lambda}_n$ is $\sqrt{n/h_n}$-consistent, and (ii) $|\mathrm{SLM}_{\mathrm{SEC}}|_{H_0}| \leq U$ a.e., and $E(U^4)$ exists, then the bootstrap methods under the resampling schemes RS_{uu}, and RS_{ru} are valid for $\mathrm{SLM}^\circ_{\mathrm{SEC}}$. The results remain if instead $u_n = W_n v_n + \epsilon_n$ such that the jth sample cumulant of $\sigma^{-1}\Omega_n^{-\frac{1}{2}}(\lambda)u_n \xrightarrow{p} \kappa_j$, $j = 1, \ldots, 6$.*

Proof: See Yang (2015).

Some final remarks are as follows.

Remark 5.7. Under $\text{SLM}_{\texttt{SEC}}$, use of $\hat{\mathcal{F}}_n$ leads to bootstrap critical values in error of order $o_p(n_r^{-\frac{1}{2}})$, whereas use of $\tilde{\mathcal{F}}_n$ leads to bootstrap critical values in error of order $O_p(n_r^{-\frac{1}{2}})$. This means that at least in theory the bootstrap critical values based on the restricted residuals offer no improvement over the asymptotic ones. However, a closer examination on the Edgeworth expansion shows that the bootstrap based on $\tilde{\mathcal{F}}_n$ can still do a better job as the main second-order effect, term involving $\frac{1}{3}(2T_3 - T_1 + 3T_5)$, is captured by the bootstrap. Our Monte Carlo results given below confirm this point.

Remark 5.8. The point that a denser weight matrix makes the finite sample null distribution of the test statistic deviate more from the limiting distribution is once again demonstrated by the first three cumulants of $\text{LM}_{\texttt{SEC}}|_{H_0}$, which are derived as those of $\text{SLM}_{\texttt{SEC}}|_{H_0}$ given in Appendix B.

Monte Carlo results for SEC model

The finite sample performance of $\text{LM}_{\texttt{SEC}}^{\texttt{FI}}$ and $\text{SLM}_{\texttt{SEC}}^{\circ}$ for testing $H_0 : \lambda = 0$ vs $H_a : \lambda > 0$, when referring to the asymptotic critical values and the bootstrap critical values under various resampling schemes, are investigated in terms of the accuracy and stability of the bootstrap critical values with respect to the true value of λ, and the size and power of the tests. The Monte Carlo experiments are carried out based on the following data generating process:

$$Y_n = \beta_0 1_n + X_{n1}\beta_1 + X_{n2}\beta_2 + W_n v_n + \epsilon_n$$

where $\{v_{n,i}\}$ are iid draws from $\sqrt{0.6}t_5$, and the methods for generating W_n, X_n and ε_n are described in Appendix C, Yang (2015). The regressors are treated as fixed in the experiments. The parameter values are set at $\beta = \{5, 1, 1\}'$ and $\sigma = 1$, and sample sizes used are $n = (54, 108, 216, 513)$. All results reported below are based on $M = 2,000$ Monte Carlo samples, and $B = 699$ bootstrap samples for each Monte Carlo sample generated. The bootstrap critical values are bench-marked against the Monte Carlo ($\texttt{MC}$) critical values obtained based on $M = 50,000$ Monte Carlo samples under H_0.

Table 5.5a. Bootstrap Critical Values **for LM Test** of Spatial Error Components

Group Sizes $\{2, 3, 4, 5, 6, 7\}$, **m = 8**, $\sigma = 1$, XVAL-B

	Normal Error			Normal Mixture			Lognormal		
λ	10%	5%	1%	10%	5%	1%	10%	5%	1%
RS_{rr}									
0.0	1.0763	1.4706	2.2198	1.6682	2.3600	3.6943	2.1365	3.2755	5.5770
0.5	1.0766	1.4684	2.2162	1.5400	2.1660	3.3711	1.9092	2.9030	4.8534
1.0	1.0784	1.4699	2.2376	1.4653	2.0475	3.1833	1.8069	2.7365	4.5543
1.5	1.0836	1.4811	2.2416	1.4126	1.9668	3.0347	1.6942	2.5609	4.2301
2.0	1.0935	1.4932	2.2571	1.3744	1.9066	2.9449	1.6207	2.4350	4.0063
RS_{uu}									
0.0	1.0754	1.4690	2.2184	1.6453	2.3256	3.6383	2.0866	3.1835	5.3784
0.5	1.0738	1.4649	2.2097	1.5392	2.1640	3.3659	1.9024	2.8723	4.7672
1.0	1.0709	1.4609	2.2217	1.4829	2.0749	3.2285	1.8312	2.7751	4.5934
1.5	1.0710	1.4632	2.2140	1.4439	2.0192	3.1225	1.7438	2.6375	4.3598
2.0	1.0732	1.4657	2.2190	1.4137	1.9705	3.0440	1.6968	2.5611	4.2373
MC									
0.0	1.0772	1.4737	2.2308	1.7310	2.4793	4.0564	2.2162	3.4827	7.4663

MC: Monte Carlo Critical values based on $M = 50,000$.

Table 5.5b. Bootstrap Critical Values **for SLM Test** of Spatial Error Components

Group Sizes $\{2, 3, 4, 5, 6, 7\}$, **m = 8**, $\sigma = 1$, XVAL-B

	Normal Error			Normal Mixture			Lognormal		
λ	10%	5%	1%	10%	5%	1%	10%	5%	1%
RS_{rr}									
0.0	1.3219	1.7255	2.4923	1.3693	1.8443	2.7315	1.4028	1.9818	2.9503
0.5	1.3204	1.7213	2.4860	1.3578	1.8204	2.6939	1.3953	1.9451	2.8880
1.0	1.3181	1.7185	2.4993	1.3520	1.8043	2.6625	1.3877	1.9264	2.8542
1.5	1.3175	1.7208	2.4910	1.3498	1.7939	2.6297	1.3729	1.8944	2.8019
2.0	1.3218	1.7272	2.4974	1.3463	1.7834	2.6192	1.3654	1.8749	2.7717
RS_{uu}									
0.0	1.3215	1.7251	2.4921	1.3675	1.8399	2.7248	1.3998	1.9700	2.9357
0.5	1.3202	1.7212	2.4856	1.3581	1.8205	2.6921	1.3954	1.9418	2.8843
1.0	1.3176	1.7182	2.4977	1.3543	1.8077	2.6717	1.3900	1.9348	2.8701
1.5	1.3169	1.7186	2.4882	1.3529	1.8049	2.6488	1.3783	1.9076	2.8291
2.0	1.3197	1.7224	2.4938	1.3505	1.7988	2.6390	1.3748	1.8983	2.8169
MC									
0.0	1.3189	1.7238	2.5153	1.3714	1.8843	2.8192	1.3823	2.0921	3.1531

MC: Monte Carlo Critical values based on $M = 50,000$.

Table 5.5c. Bootstrap Critical Values **for LM Test** of Spatial Error Components

Group Sizes $\{2, 3, 4, 5, 6, 7\}$, **m=19**, $\sigma = 1$, XVAL-B

	Normal Error			Normal Mixture			Lognormal		
λ	10%	5%	1%	10%	5%	1%	10%	5%	1%
RS_{rr}									
0.0	1.1502	1.5338	2.2687	1.8199	2.4772	3.7743	2.6510	4.0188	6.6391
0.5	1.1526	1.5372	2.2731	1.6764	2.2758	3.4514	2.3728	3.6083	5.9152
1.0	1.1558	1.5418	2.2794	1.5888	2.1559	3.2507	2.1695	3.2701	5.2968
1.5	1.1612	1.5485	2.2905	1.5333	2.0724	3.1247	2.0451	3.0671	4.9338
2.0	1.1710	1.5607	2.3023	1.4866	2.0145	3.0414	1.9375	2.8884	4.6049
RS_{uu}									
0.0	1.1499	1.5333	2.2678	1.8084	2.4606	3.7472	2.6015	3.9204	6.4371
0.5	1.1495	1.5341	2.2673	1.6880	2.2929	3.4773	2.3740	3.5833	5.8370
1.0	1.1489	1.5332	2.2617	1.6219	2.2022	3.3299	2.2332	3.3549	5.4274
1.5	1.1473	1.5295	2.2640	1.5833	2.1467	3.2393	2.1538	3.2382	5.2166
2.0	1.1525	1.5362	2.2629	1.5435	2.0858	3.1581	2.0716	3.1130	4.9923
MC									
0.0	1.1569	1.5445	2.2472	1.8325	2.5278	3.9093	2.6464	4.1103	8.5357

MC: Monte Carlo Critical values based on $M = 50,000$.

Table 5.5d. Bootstrap Critical Values **for SLM Test** of Spatial Error Components

Group Sizes $\{2,3,4,5,6,7\}$, **m = 19**, $\sigma = 1$, XVAL-B

	Normal Error			Normal Mixture			Lognormal		
λ	10%	5%	1%	10%	5%	1%	10%	5%	1%
RS_{rr}									
0.0	1.3026	1.6901	2.4312	1.3369	1.7769	2.6263	1.3836	1.9416	2.9125
0.5	1.3029	1.6909	2.4326	1.3286	1.7583	2.5835	1.3722	1.9220	2.8692
1.0	1.3003	1.6882	2.4274	1.3259	1.7512	2.5635	1.3661	1.8927	2.8153
1.5	1.3011	1.6880	2.4293	1.3239	1.7435	2.5493	1.3579	1.8737	2.7870
2.0	1.3048	1.6925	2.4299	1.3194	1.7361	2.5346	1.3549	1.8583	2.7535
RS_{uu}									
0.0	1.3024	1.6899	2.4311	1.3360	1.7745	2.6227	1.3820	1.9352	2.9025
0.5	1.3026	1.6911	2.4319	1.3287	1.7610	2.5867	1.3742	1.9187	2.8649
1.0	1.3010	1.6895	2.4243	1.3274	1.7571	2.5783	1.3696	1.9025	2.8324
1.5	1.3000	1.6862	2.4279	1.3280	1.7526	2.5666	1.3657	1.8932	2.8177
2.0	1.3045	1.6926	2.4266	1.3238	1.7442	2.5579	1.3643	1.8821	2.8027
MC									
0.0	1.3033	1.6967	2.4031	1.3209	1.7774	2.6576	1.3432	2.0206	3.0694

MC: Monte Carlo Critical values based on $M = 50,000$.

Table 5.6a. Rejection Frequencies **for LM Test** of Spatial Error Components

Group Sizes $\{2,3,4,5,6,7\}$, $m=8$, $\sigma=1$, XVAL-B

	Normal Error			Normal Mixture			Lognormal		
λ	10%	5%	1%	10%	5%	1%	10%	5%	1%
ACV									
0.0	0.0690	0.0330	0.0070	0.1480	0.1075	0.0575	0.1790	0.1420	0.0960
0.5	0.5845	0.4640	0.2490	0.5550	0.4590	0.2980	0.5795	0.5015	0.3665
1.0	0.9005	0.8460	0.6780	0.8540	0.7870	0.6470	0.8110	0.7635	0.6525
1.5	0.9815	0.9695	0.9180	0.9650	0.9415	0.8680	0.9140	0.8890	0.8205
2.0	0.9960	0.9910	0.9665	0.9850	0.9750	0.9340	0.9530	0.9375	0.9010
RS_{rr}									
0.0	0.1010	0.0465	0.0120	0.1045	0.0555	0.0135	0.1180	0.0625	0.0180
0.5	0.6560	0.5215	0.2760	0.4890	0.3505	0.1505	0.4735	0.3275	0.1610
1.0	0.9330	0.8720	0.7045	0.8140	0.6980	0.4520	0.7190	0.5945	0.3850
1.5	0.9890	0.9765	0.9310	0.9535	0.9045	0.7405	0.8600	0.7820	0.6115
2.0	0.9960	0.9935	0.9720	0.9805	0.9545	0.8570	0.9215	0.8560	0.7130
RS_{uu}									
0.0	0.1010	0.0480	0.0115	0.1065	0.0605	0.0205	0.1215	0.0685	0.0395
0.5	0.6570	0.5230	0.2840	0.4835	0.3490	0.1540	0.4690	0.3245	0.1580
1.0	0.9330	0.8740	0.7055	0.8090	0.6850	0.4320	0.7145	0.5820	0.3670
1.5	0.9890	0.9775	0.9300	0.9520	0.8980	0.7020	0.8555	0.7640	0.5780
2.0	0.9960	0.9930	0.9715	0.9795	0.9490	0.8295	0.9160	0.8395	0.6735

Table 5.6b. Rejection Frequencies **for SLM Test** of Spatial Error Components

Group Sizes $\{2,3,4,5,6,7\}$, $m = 8$, $\sigma = 1$, XVAL-B

	Normal Error			Normal Mixture			Lognormal		
λ	10%	5%	1%	10%	5%	1%	10%	5%	1%
ACV									
0.0	0.1025	0.0525	0.0130	0.1090	0.0660	0.0255	0.1160	0.0795	0.0440
0.5	0.6640	0.5465	0.3210	0.4985	0.3875	0.2085	0.4685	0.3735	0.2225
1.0	0.9340	0.8845	0.7485	0.8210	0.7320	0.5420	0.7160	0.6320	0.4745
1.5	0.9890	0.9780	0.9445	0.9540	0.9185	0.8155	0.8480	0.8005	0.6925
2.0	0.9965	0.9950	0.9805	0.9810	0.9590	0.9010	0.9045	0.8725	0.7900
RS_{rr}									
0.0	0.1015	0.0465	0.0120	0.0970	0.0515	0.0105	0.1040	0.0590	0.0170
0.5	0.6535	0.5205	0.2755	0.4730	0.3380	0.1435	0.4400	0.3105	0.1585
1.0	0.9330	0.8715	0.7050	0.8045	0.6825	0.4510	0.6870	0.5705	0.3765
1.5	0.9885	0.9765	0.9310	0.9490	0.8980	0.7290	0.8335	0.7590	0.6120
2.0	0.9960	0.9935	0.9715	0.9785	0.9510	0.8530	0.8955	0.8335	0.7080
RS_{uu}									
0.0	0.1000	0.0485	0.0110	0.0975	0.0525	0.0110	0.1035	0.0595	0.0210
0.5	0.6550	0.5220	0.2780	0.4750	0.3405	0.1460	0.4420	0.3105	0.1560
1.0	0.9320	0.8730	0.7020	0.8050	0.6795	0.4470	0.6875	0.5670	0.3715
1.5	0.9890	0.9770	0.9295	0.9480	0.8970	0.7240	0.8325	0.7560	0.5995
2.0	0.9960	0.9935	0.9715	0.9780	0.9485	0.8475	0.8940	0.8330	0.7010

In this section, we further illustrate these methods using a more general model: the linear regression with both SL and SE, or the SLE model:

$$Y_n = \lambda W_{1n} Y_n + X_n\beta + u_n, \quad u_n = \rho W_{2n} u_n + \epsilon_n, \quad \epsilon_n = \sigma e_n, \tag{5.29}$$

where all quantities are defined as in (5.8) and (5.15). W_{1n} and W_{2n} can be the same. Clearly, (5.29) has the form of the general model given in (5.1): $\sigma^{-1} B_n(\rho)[A_n(\lambda) Y_n - X_n\beta] = e_n$.

The following hypotheses are of interest:

$H_0^{\mathrm{SLE}}: \delta_0 = (\lambda_0, \rho_0)' = 0,$ standard liner regression model suffices,
$H_0^{\mathrm{SL|SE}}: \lambda_0 = 0,$ SE model suffices,
$H_0^{\mathrm{SE|SL}}: \rho_0 = 0,$ SL model suffices.

The corresponding LM tests can be found in Anselin et al. (1996) and can be written as (assuming $W_{1n} = W_{2n} = W_n$): for testing $H_0^{\texttt{SLE}}$,

$$\mathrm{LM}_{\texttt{SLE}}^{\texttt{FI}} = \frac{(\tilde{\epsilon}_n' W_n Y_n - \tilde{\epsilon}_n' W_n \tilde{\epsilon}_n)^2}{\tilde{\sigma}_n^4 \tilde{\eta}_n' M_n \tilde{\eta}_n} + \frac{(\tilde{\epsilon}_n' W_n \tilde{\epsilon}_n)^2}{\tilde{\sigma}_n^4 K_n}, \tag{5.30}$$

where all quantities are defined in (5.9) and (5.16); for testing $H_0^{\texttt{SE|SL}}$,

$$\mathrm{LM}_{\texttt{SE|SL}}^{\texttt{FI}} = \frac{\tilde{\epsilon}_n' W_n \tilde{\epsilon}_n}{\tilde{\sigma}_n^2 [K_n - \tilde{S}_{1n}^2 / (\tilde{\eta}_n' M_n \tilde{\eta}_n + \tilde{S}_{2n})]^{1/2}}, \tag{5.31}$$

where $\tilde{S}_{1n} = \mathrm{tr}[(W_n + W_n')\tilde{F}_n]$, $\tilde{S}_{2n} = \mathrm{tr}[(\tilde{F}_n^{\circ} + \tilde{F}_n^{\circ\prime})\tilde{F}_n^{\circ}]$, $\tilde{F}_n = W_n A_n^{-1}(\tilde{\lambda}_n)$, and $\tilde{F}_n^{\circ} = \tilde{F}_n - \frac{1}{n}\mathrm{tr}(\tilde{F}_n) I_n$; and for testing $H_0^{\texttt{SL|SE}}$,

$$\mathrm{LM}_{\texttt{SL|SE}}^{\texttt{FI}} = \frac{\tilde{\epsilon}_n' \tilde{B}_n W_n Y_n}{\tilde{\sigma}_n^2 [\tilde{S}_{3n} + \tilde{\eta}_n' \tilde{B}_n' \tilde{B}_n \tilde{\eta}_n + \tilde{h}_n' \tilde{J}_n^{-1} \tilde{h}_n]^{1/2}}, \tag{5.32}$$

where $\tilde{S}_{3n} = \mathrm{tr}(W_n^2 + \tilde{G}_n' \tilde{B}_n' \tilde{B}_n \tilde{G}_n)$, $\tilde{h}_n = \{\tilde{\sigma}_n^{-1} X_n' \tilde{B}_n' \tilde{B}_n \tilde{\eta}_n,\ 0,\ \mathrm{tr}((\tilde{G}_n' \tilde{B}_n + W_n)\tilde{G}_n)\}'$, $\tilde{J}_n = J_n(\tilde{\theta}_n)$ given in (B.1) of Yang (2015), $\tilde{B}_n = \tilde{B}_n(\tilde{\rho}_n)$, and $\tilde{G}_n = W_n \tilde{B}_n^{-1}$.

The bootstrap methods can be implemented in the same manner. The bootstrap DGPs that mimic the real world null DGPs are,

$$Y_n^* = X_n\ddot{\beta}_n + \ddot{\sigma}_n e_n^*,$$
$$Y_n^* = A_n^{-1}(\ddot{\lambda}_n)(X_n\ddot{\beta}_n + \ddot{\sigma}_n e_n^*), \text{ and}$$
$$Y_n^* = X_n\ddot{\beta}_n + \ddot{\sigma}_n B_n^{-1}(\ddot{\rho}_n)e_n^*,$$

respectively, for the three hypotheses. For example, to obtain BCVs of $\mathrm{LM}^{\mathtt{FI}}_{\mathtt{SL|SE}}|_{H_0}$, based on the unrestricted estimates/residuals:

(a) Compute the unrestricted QMLEs $(\hat{\beta}_n, \hat{\sigma}_n^2, \hat{\lambda}_n, \hat{\rho}_n)$ based on Model (5.29);

(b) Compute $\hat{e}_n = \hat{\sigma}_n^{-1}B_n(\hat{\rho}_n)[A_n(\hat{\lambda}_n)Y_n - X_n\hat{\beta}_n]$, and standardize, to give $\hat{\mathcal{F}}_n$;

(c) Draw a bootstrap sample e_n^* from $\hat{\mathcal{F}}_n$; compute $Y_n^* = X_n\hat{\beta}_n + \hat{\sigma}_n B_n^{-1}(\hat{\rho}_n)e_n^*$;

(d) Estimate null model $Y_n = X_n\beta + u_n$, $u_n = \rho W_n u_n + \varepsilon_n$, based on bootstrap data (Y_n^*, X_n, W_n); then compute a bootstrap value $\mathrm{LM}^{b}_{\mathtt{SL|SE}}$ of $\mathrm{LM}^{\mathtt{FI}}_{\mathtt{SL|SE}}$;

(e) Repeat $(c) - (d)$ B times to obtain EDF of $\{\mathrm{LM}^{b}_{\mathtt{SL|SE}}\}_{b=1}^{B}$ and its α-quantile — a bootstrap estimate of the true finite sample α-quantile of $\mathrm{LM}_{\mathtt{SL|SE}}|_{H_0}$.

Monte Carlo results for SLE model

Table 5.7. Rejection Frequencies for $\text{LM}^{\texttt{FI}}_{\texttt{SE|SL}}$, H_0: $\rho = 0$

Group Interaction with $g = n^{0.5}$, $\sigma = 1$, $n = 100$, $\lambda = 0.25$, XVal-B

		Normal Error				Lognormal Error			
Method	$\lvert\rho\rvert$	L2.5%	L5%	U5%	U2.5%	L2.5%	L5%	U5%	U2.5%
ACR	0.00	0.0415	0.1030	0.0180	0.0080	0.0325	0.0750	0.0145	0.0090
	0.25	0.1815	0.3080	0.1370	0.0920	0.1095	0.2360	0.1225	0.0770
	0.50	0.4105	0.5720	0.5275	0.4300	0.2965	0.4920	0.5235	0.4285
RS_{rr}	0.00	0.0225	0.0450	0.0520	0.0235	0.0310	0.0490	0.0420	0.0185
	0.25	0.3215	0.4275	0.2550	0.1645	0.1110	0.1780	0.2510	0.1540
	0.50	0.3215	0.4275	0.6765	0.5745	0.3050	0.4175	0.6980	0.5855
RS_{ur}	0.00	0.0225	0.0425	0.0515	0.0230	0.0255	0.0470	0.0415	0.0170
	0.25	0.1155	0.1880	0.2590	0.1680	0.1010	0.1745	0.2505	0.1385
	0.50	0.3140	0.4250	0.6815	0.5815	0.2910	0.4115	0.7010	0.5640
RS_{ru}	0.00	0.0245	0.0470	0.0500	0.0225	0.0325	0.0525	0.0405	0.0180
	0.25	0.3295	0.4260	0.2505	0.1650	0.1160	0.1815	0.2555	0.1520
	0.50	0.3140	0.4250	0.6690	0.5605	0.3120	0.4300	0.6965	0.5855
RS_{uu}	0.00	0.0230	0.0440	0.0530	0.0240	0.0290	0.0465	0.0400	0.0175
	0.25	0.1195	0.1890	0.2585	0.1715	0.1005	0.1725	0.2495	0.1405
	0.50	0.3140	0.4250	0.6820	0.5840	0.2920	0.4170	0.7025	0.5710

L = Left tail ($\rho = -0.25, -0.5$, in the rejection frequencies), R = Right tail ($\rho = 0.25, 0.5$);
MC: Monte Carlo critical values based on $M = 30,000$.

Table 5.8. Bootstrap Critical Values for $\mathrm{LM}^{\mathtt{I}}_{\mathtt{SE|SL}}$, H_0: $\rho = 0$

Group Interaction with $g = n^{0.5}$, $\sigma = 1$, $n = 100$, $\lambda = 0.25$, XVal-B

Method	ρ	Normal Error				Lognormal Error			
		L2.5%	L5%	U5%	U2.5%	L2.5%	L5%	U5%	U2.5%
		Bootstrap Critical Values							
RS_{rr}	-0.50	-2.1615	-1.9302	1.1107	1.4812	-1.9531	-1.7562	1.0984	1.5787
	-0.25	-2.1735	-1.9390	1.1115	1.4823	-1.9577	-1.7561	1.0847	1.5536
	0.00	-2.1857	-1.9459	1.1178	1.4897	-1.9903	-1.7667	1.0758	1.5207
	0.25	-2.2003	-1.9557	1.1326	1.4991	-2.0973	-1.7993	1.0647	1.4808
	0.50	-2.2453	-1.9832	1.1639	1.5186	-2.3318	-1.8836	1.0706	1.4511
RS_{ur}	-0.50	-2.1766	-1.9397	1.1182	1.4870	-2.0037	-1.7737	1.0731	1.5107
	-0.25	-2.1856	-1.9463	1.1174	1.4863	-2.0271	-1.7780	1.0706	1.5140
	0.00	-2.1912	-1.9489	1.1212	1.4916	-2.0460	-1.7848	1.0741	1.5158
	0.25	-2.1881	-1.9474	1.1236	1.4931	-2.0702	-1.7895	1.0752	1.5193
	0.50	-2.1914	-1.9508	1.1256	1.4921	-2.0858	-1.7998	1.0786	1.5218
RS_{ru}	-0.50	-2.1485	-1.9194	1.1051	1.4742	-1.9382	-1.7440	1.0948	1.5791
	-0.25	-2.1633	-1.9297	1.1059	1.4727	-1.9514	-1.7511	1.0795	1.5523
	0.00	-2.1816	-1.9425	1.1140	1.4866	-1.9825	-1.7638	1.0727	1.5146
	0.25	-2.2076	-1.9619	1.1365	1.5033	-2.0891	-1.8010	1.0653	1.4805
	0.50	-2.2990	-2.0294	1.1910	1.5552	-2.3591	-1.9098	1.0844	1.4731
RS_{uu}	-0.50	-2.1756	-1.9394	1.1190	1.4897	-1.9881	-1.7672	1.0728	1.5149
	-0.25	-2.1850	-1.9457	1.1173	1.4841	-2.0160	-1.7759	1.0696	1.5167
	0.00	-2.1913	-1.9499	1.1204	1.4922	-2.0265	-1.7810	1.0733	1.5110
	0.25	-2.1887	-1.9476	1.1244	1.4928	-2.0675	-1.7890	1.0740	1.5138
	0.50	-2.1928	-1.9509	1.1271	1.4915	-2.0794	-1.7905	1.0732	1.5178
MC	0.00	-2.1641	-1.9271	1.1382	1.4957	-2.1120	-1.8134	1.0815	1.4698

L = Left tail ($\rho = -0.25, -0.5$, in the rejection frequencies), R = Right tail ($\rho = 0.25, 0.5$);
MC: Monte Carlo critical values based on $M = 30,000$.

References