Economic Modelling 41 (2014) 9–14

Contents lists available at ScienceDirect

Economic Modelling

journal homepage: www.elsevier.com/locate/ecmod

Moran's I test of spatial panel data model — Based on bootstrap method

Tongxian Ren [a], Zhihe Long [a], Rengui Zhang [a,b,*], Qingqing Chen [a]

[a] School of Economics and Commerce, South China University of Technology, Guangzhou, China
[b] School of Economics, Shenzhen Polytechnic, Shenzhen, China

ARTICLE INFO

Article history:
Accepted 15 April 2014
Available online xxxx

Keywords:
Bootstrap methods
Moran's I test
Spatial panel data
Monte Carlo simulation

ABSTRACT

Under the condition of the finite sample or the unknown distributed error term, testing for spatial dependence in panel data models is an unresolved problem in spatial econometrics. In this paper, a fast double bootstrap (FDB) method is used to construct bootstrap Moran's I tests for Moran's I test in spatial panel data models, and Monte Carlo simulation experiments are used to prove the effectiveness from two aspects including size distortion and power. The experiment results show that, in asymptotic Moran's I test, there is serious size distortion, which could be rectified in bootstrap Moran's I test.

© 2014 Elsevier B.V. All rights reserved.

1. Introduction

Spatial econometric models have been extensively studied in the past thirty years. They have fewer assumptions than classic econometric models. And, to be precise, spatial econometrics can test spatial effects including spatial dependence and spatial heterogeneity. Therefore, spatial econometrics can get more reasonable and realistic conclusions than classic econometrics. To test the existence of spatial dependence in a spatial econometric model has been a core issue. Spatial econometric models include a spatial lag model (SLM) and spatial error model (SEM). The former studies how dependent variables in the vicinity of the behavior affect other parts of the overall system behavior. And the latter, whose spatial dependence exists in the error term, studies the influence of the error shock on neighboring region behavior. Currently, there are many methods to test the existence of spatial dependence in spatial cross-sectional data models, but rarely are there methods to test the existence of spatial dependence in spatial panel data models. The common methods to test the existence of spatial dependence in spatial cross-sectional data models, include Moran's I, LM, LR, Rao's score, etc. Moran's I test of spatial dependence is assumed to be a non-alternative hypothesis model, and, it can test spatial lag dependence and test spatial error dependence. Therefore, Moran's I test has been most commonly used.

As we all know, the panel data contains both the characteristics of cross-sectional data and time series data. Hence, it can provide richer information for regression analysis. The panel data models offer the researchers more opportunities to extend model possibilities as compared to cross-sectional data models. The use of panel data also results in a greater availability of degrees of freedom, and increases efficiency in the estimation. Spatial panel data models lead into spatial effects, considering the individual heterogeneity and the cross-sectional dimensions to test whether the spatial dependence exists. Thus, there are much broader prospects of practical application in actual researches. However, panel data information is more plentiful than cross-sectional data. Hence, testing for spatial dependence of spatial panel data models is a difficult problem in spatial econometrics. Furthermore, the traditional methods of testing spatial dependence depended on the assumptions that the error term is normally distributed or large sample size. However, the real economy is a complex system which is affected by diversity factors. In a large number of empirical studies, the above assumptions can't satisfy strictly. For example, in real economic management analysis, data is non-normally distributed, or its sample size is limited for empirical discussion. Evidences (Lin et al., 2009; Yang, 2011) showed that, in finite samples, the Moran's I test referring to asymptotic critical values may suffer from the problems of size distortion and low power. Bootstrap methods, which have a good property of finite samples, are an effective way to solve the above problems. They are introduced into spatial cross-sectional data models by Lin et al. (2011). But spatial panel data models are more complicated and would affect the bootstrap sampling. Therefore, the bootstrap methods can't be applied to the spatial panel models directly. To our knowledge, under the conditions of finite samples, unknown distributed or heteroscedastic error terms, testing for spatial dependence in panel data models is an unresolved problem in empirical studies of spatial econometrics at present.

* Corresponding author at: School of Economics, Shenzhen Polytechnic, Shenzhen 518055, China. Tel.: +86 13632930040.
E-mail address: mathsren@foxmail.com (R. Zhang).

http://dx.doi.org/10.1016/j.econmod.2014.04.022
0264-9993/© 2014 Elsevier B.V. All rights reserved.

In this paper, we try to apply fast double bootstrap (FDB) method to Moran's I test in spatial panel data models, and then, Monte Carlo simulation experiments are used to prove the effectiveness of the test statistics from size distortion and power.

The rest of the paper proceeds as follows. Spatial panel models and Moran's I test statistics are discussed in Section 2. The applications of the bootstrap methods are presented in Section 3. In Section 4, we report results of Monte Carlo simulation to show size distortion and power performance of bootstrap Moran's I test. And, we will compare results of bootstrap Moran's I test with asymptotic test. Section 5 contains some concluding remarks.

2. Spatial panel data models and Moran's I test statistic

Elhorst (2003, 2010) extended spatial cross-sectional data models to spatial panel models. And, they proposed spatial panel model estimation methods. Like the spatial cross-sectional data models, spatial panel data models can be divided into spatial panel data lag models and spatial panel data error models. The expressions of two models are as follows:

$$\begin{cases} \text{SLM}: y_t = \lambda W y_t + X_t\beta + \varepsilon_t \\ \text{SEM}: y_t = X_t\beta + \varepsilon_t, \varepsilon_t = \rho W \varepsilon_t + v_t \end{cases} t = 1, 2\ldots, T \tag{1}$$

where y_t is a $N \times 1$ vector of a dependent variable for period t, λ is spatial lag dependence coefficient, X_t is a $N \times K$ matrix of non-stochastics for period t, W is spatial weights matrix, β is $K \times 1$ parameter vector, ρ is spatial error dependence coefficient, ε_t is a $N \times 1$ vector of the regression error term for period t, and v_t is a $N \times 1$ vector of the remain error for period t. It is assumed that $|\lambda| < 1, |\rho| < 1, \varepsilon_t \sim N(0, \sigma_\varepsilon^2 I_N), \ \ v_t \sim N(0, \sigma_v^2 I_N)$.

Spatial effects will be tested in spatial panel models before the spatial panel data models are established. At present, the most commonly used method to test spatial dependence is Moran's I test which was proposed by Moran (1948). The test statistics developed for the cross-section were extended to the panel data model by Arbia (2005). Its expression is as follows:

$$I = \frac{e' W_{NT} e}{e'e} \tag{2}$$

where I represents Moran's I test statistic of a spatial panel data model, $W_{NT} = I_T \otimes W$ is the spatial weights matrix, $\otimes$ is the Kronecker product, and e is the residual.

Anselin (1988) demonstrated when the error term does not obey the classic distribution or heteroscedasticity, Moran's I test would be lapse. The bootstrap method is an effective way to solve above problems (Lin et al, 2009; Yang, 2011).

3. Moran's I test methods

Bootstrapping is a method that performs inference using pseudo-datasets created by sampling from observed data (Efron, 1979). It is essentially a Monte Carlo simulation procedure. It does not need to assume that the error terms are independent and normally distributed and a parametric estimate of the variance of the estimate. And, it need not provide an observational data distribution form. Therefore, the bootstrap methods are applied to non-classical error term conditions.

Different bootstrap methods have been developed for different types of regression. Such as residuals bootstrap (Efron, 1979), wild bootstrap (Beran, 1988), block bootstrap (Efron, 1979), pairs bootstrap (Freedman, 1981). The residuals bootstrap is used in cross-section and panel data models (Efron, 1979). The block bootstrap is used in time series models. The wild bootstrap is used to deal with panel data models and heteroscedasticity (Davidson and MacKinnon, 2007). The pairs bootstrap is used to dynamic model or the heteroscedastic model which error term is unknown distributed.

Among the bootstrap methods applied in the panel data model, Chang (2003) applied bootstrap method to unit root tests for dependent panel data model. And they found that bootstrap tests perform better in finite samples than in an asymptotic test. Cerrato and Sarantis (2007) used bootstrap methods to deal with cross-sectional dependence in panel unit root tests of real exchange rates, their results showed that the statistic has good power and no size distortion for moderate and large samples. Godfrey (2009) suggested that the wild bootstrap procedure is well-behaved in finite samples under heteroscedasticity and match the performance less of robust tests under classical assumptions. However, it does not mean that bootstrap tests always perform well in finite samples. Thus, Beran (1988) proposed the double bootstrap (DB). But DB tends to be very computationally demanding. Davidson and MacKinnon (2007) developed a fast double bootstrap (FDB) based on the double bootstrap. The FDB requires no more than about twice as much computation as the general bootstrap, making it feasible while the double bootstrap is not. This paper is the first one which applies the FDB to test the spatial dependence of the spatial panel data models. The most useful way to perform a bootstrap test is to calculate P value of bootstrap. And, Monte Carlo simulation experiments are used to prove more effective than asymptotic test of the test statistics from two aspects including size distortion and power.

The simplest and most informative method to perform a bootstrap test is to calculate a bootstrap P value. We obtain a bootstrap P value of the bootstrap test statistics which are more accurate than the actual test statistic. When this P value is below the significant value, we reject the null hypothesis. The P value of bootstrap is calculated by the following expression:

$$P^* = 2 * \min\left(\frac{1}{B}\sum_{j=1}^{B} I\left(\eta_j^* \le \eta\right), \frac{1}{B}\sum_{j=1}^{B} I\left(\eta_j^* > \eta\right)\right) \tag{3}$$

where P^* is the bootstrap P value, $I(\cdot)$ is a denote function, B represents then bootstrap test frequency, η denotes the test statistic of bootstrap.

The double bootstrap which is proposed by Beran (1988) can be used to calculate P values. It should be more accurate than the general bootstrap in theory. The first step in the double bootstrap is to generate B_1 first-level bootstrap samples which are used to compute bootstrap test statistics $\eta_j^* (j = 1, 2, \ldots, B_1)$. The second step is to get a second-level bootstrap DGP (data generating process) for each first-level bootstrap sample indexed by j. Each second-level bootstrap DGP is used to generate B_2 bootstrap samples that are used to calculate test statistics $\eta_{jl}^* (l = 1, 2, \ldots, B_2)$. In practice, the double bootstrap is very costly in terms of computation. For each of B_1 bootstrap samples, $B_2 + 1$ test statistics should be computed. Thus, the total amount of test statistics will reach up to $1 + B_1 + B_1 B_2$. The double bootstrap is costly because we need to generate B_2 s-level bootstrap samples for every first-level bootstrap sample. It is necessary because the distribution of the η_{jl}^* may be dependent on statistic η_j^*.

Davidson and MacKinnon (2007) Proposed the assumption that the η_j^* distribution is independent of η_{jl}^*. It is called the fast double bootstrap test. And, the calculation of bootstrap test will be greatly reduced. For the FDB test, only one second-level bootstrap statistic η_j^{**}, is computed along with each η_j^*. The expression of FDB P value was given by Davidson and MacKinnon (2006):

$$p_F^{**} = \frac{1}{B}\sum_{j=1}^{B} I\left(\eta_j^{**} > Q_B^{**}(1 - p^*)\right) \tag{4}$$

where p^* is the bootstrap P value, p_F^{**} is FDB P value, B is bootstrap test times, $Q_B^{**}(1 - p^*)$ represents the $(1 - p^*)$ quantile of the η_j^{**}.

Ahlgren and Antell (2008) suggested that the FDB produces a further improvement in cases where the performance of the asymptotic test is unsatisfactory.

Download English Version:

https://daneshyari.com/en/article/5054109

Download Persian Version:

https://daneshyari.com/article/5054109

Daneshyari.com