

Article

Research on the Validity of Bootstrap LM-Error Test in Spatial Random Effect Models

Tongxian Ren ¹, Lin Xu ² and Zhengliang Ren ^{1,*}

¹ School of Economics and Statistics, Guangzhou University, Guangzhou 510006, China; rentongxiangd@gzhu.edu.cn

² School of Economics and Finance, South China University of Technology, Guangzhou 510006, China; linxu@scut.edu.cn

* Correspondence: sesrzl2017@gzhu.edu.cn

Abstract: Under the condition of non-classical distributed errors, the test for spatial dependence in spatial panel data models is still a problem waiting to be solved. In this paper, we apply the FDB (Fast Double Bootstrap) method to spatial panel data models to test spatial dependence. In order to research the validity of the Bootstrap LM-Error test in spatial random effect models under the condition that the error term obeys a normal distribution, heteroscedasticity, or time-series correlation, we construct Bootstrap LM-Error statistics and make use of Monte Carlo simulation from size distortion and power aspects to carry out our research. The Monte Carlo simulation results show that the asymptotic LM-Error test in the spatial random effects model has a large size of distortion when the error term disobeys classical distribution. However, the FDB LM-Error test can effectively correct the size distortion of the asymptotic test with the precondition that there is nearly no loss of power in the FDB test. Obviously, compared to the asymptotic LM-Error test, the FDB LM-Error test is a more valid method to test spatial dependence in a spatial random effects model.

Keywords: spatial random effects model; Bootstrap method; LM-Error test; Monte Carlo simulation

MSC: 91B72



Citation: Ren, T.; Xu, L.; Ren, Z.

Research on the Validity of Bootstrap LM-Error Test in Spatial Random Effect Models. *Axioms* **2024**, *13*, 378. <https://doi.org/10.3390/axioms13060378>

Academic Editor: Hari Mohan Srivastava

Received: 18 March 2024

Revised: 22 May 2024

Accepted: 30 May 2024

Published: 4 June 2024



Copyright: © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Recently, spatial econometrics have gradually become one of the ‘standard’ econometric analysis tools; see Anselin et al. (2010) for the consideration of the spatial effects among research objects, where the spatial econometric model provides more practical results compared to the classical econometric model [1]. There are two commonly used spatial econometric models: the Spatial Lag Model (SLM), also known as the Spatial Autoregressive Model (SAR), and Spatial Error Model (SEM). The SLM studies the influence of dependent variable behaviors in the vicinity on the rest of the whole system. The spatial dependence in SEM is present in the error term, and SEM researches the influences of the error impacts in the vicinity on other areas.

Introducing spatial effects into spatial panel data model has a broad practical application prospect, since it not only controls the individual heterogeneity well, but it also considers the spatial dependence in the cross-sectional dimension. However, because the spatial panel data model includes both features of the spatial econometric model and panel data model, which makes model structures more complex, it is more difficult to test spatial dependence and estimate parameters.

In consideration of the individual effects, the spatial panel econometric model is specified as follows:

Spatial Lag Model:

$$y_t = \lambda W y_t + X_t \beta + \mu + \varepsilon_t \quad (1)$$

Spatial Error Model:

$$y_t = X_t\beta + \mu + \varepsilon_t, \varepsilon_t = \rho W\varepsilon_t + v_t \quad (2)$$

In (1) and (2), y_t and X_t respectively present a dependent variable and an independent variable; W stands for spatial weight matrix. The spatial weight matrix W needs to be standardized based on spatial distance. λ in (1) and ρ in (2) are the spatial correlation coefficient of the model, with $|\lambda| < 1$ and $|\rho| < 1$. We define the error term in (1) by $\varepsilon_t \sim N(0, \sigma_\varepsilon^2 I_N)$ and in (2) by $v_t \sim N(0, \sigma_v^2 I_N)$. According to individual effects μ , we can set the model as a fixed effects model or random effects model. In this paper, we let μ denote random effects, and assume that ε_t and μ are mutually independent, noting that $\varepsilon_t \sim N(0, \sigma_\varepsilon^2 I_N)$.

As we know, the spatial dependence test is the premise of spatial econometric analysis. There are many mature methods to test spatial dependence in the spatial panel data model, including such commonly used methods as Moran's I or the LM test method. Moran's I test statistics can test spatial dependence in a model, but is unable to determine the forms of spatial dependence. Generally, the LM test can be divided into the LM-Lag test and LM-Error test, respectively, and can be applied to the spatial lag model and spatial error model for spatial dependence tests.

The efficiency of the spatial dependence LM test method depends on the i.i.d-normal assumption of error under the condition of large samples. In real economic management research, however, empirical research can only obtain finite samples in most cases since the data are difficult to obtain. Meanwhile, due to the complexity of the practical economy and the diversity of influential factors, the error term of the model usually violates the i.i.d-normal classical assumption. At this point, the spatial dependence LM test is no longer valid.

To solve the above problem, Efron (1979) provided an effective method, namely the Bootstrap method [2]. The Bootstrap method does neither need to assume the overall distribution nor derive the estimator's function beforehand. It just reconstructs samples and calculates estimates continuously. Therefore, the Bootstrap method widely applies to the situation that the error term disobeys in classical distribution. The Residual Bootstrap or Wild Bootstrap method have been very commonly used in spatial dependence tests. More specifically, the Wild Bootstrap is more applicable to solve heteroscedasticity or the autocorrelation of an error term (Davidson [3], 2007). Applying the Bootstrap method to the panel data unit root test or cross-sectional correlation test, etc., in the panel data model, Chang (2003) [4], Cerrato (2006) [5] and Godfrey (2009) [6] et al. showed that the Bootstrap test is obviously better than the asymptotic test.

The Bootstrap method has broad applications in classical econometric research; however, as far as we know, it has few applications in the spatial dependence test. Warren (2008) used the Bootstrap method to research the heteroscedasticity, serial correlation or spatial error autocorrelation in a spatial panel data model [7]. Yang et al. [8] (2011) and Lin et al. [9] (2011) also used the Bootstrap method to test spatial dependence in a spatial cross-sectional data model. Ren et al. [10] (2014) studied the performance of Bootstrap tests for spatial pooled data models, Yang [11] (2015) tested spatial dependence in a spatial cross-sectional data model through LM tests, and Lee [12] (2015) focused on the test statistic for Moran's I test for the spatial dependence of a spatial cross-sectional data model. Ou et al. (2015) proposed a robust LM test for a time-varying spatial weight matrix panel data model. In the case of a longer time and more individuals, the robust LM test of the time-varying spatial weight matrix is more efficient.

Zhai et al. (2022) studied the spatial effect testing problem of a fixed effect variable of a coefficient spatial autoregressive panel data model. Their numerical simulation shows that the proposed method has better performance under limited samples performance.

Nevertheless, the above Bootstrap methods may only apply in spatial cross-sectional data models or pooled data models, since we have not searched the relevant literature about research on spatial dependence LM tests in spatial panel data models. This paper will apply the Fast Double Bootstrap (hereafter referred to as FDB) sampling method, which was proposed by Davidson and MacKinnon (2007) to an LM test in a spatial panel data econometric model,

in order to solve such problems as small samples or a spatial dependence test in a spatial panel data model under non-classical distributed errors [3]. From size distortion and power, we employ Monte Carlo simulation to study the efficiency of the Bootstrap LM test in a spatial panel data model. Due to space limitations, this paper just reports the efficiency of the Bootstrap LM-Error test method under the conditions of a random effect.

The rest of this paper is organized as follows. Section 2 introduces the Bootstrap LM test in the spatial panel data econometric model. In Section 3, we conduct the Monte Carlo simulation and explain its results. Section 4 concludes.

2. Bootstrap LM-Error Test for Spatial Panel Data Econometric Model

The first step to test spatial correlation in a spatial panel data model is to establish test statistics. Then, a spatial panel data error model with random effect is given by (2), and the variable μ denotes the random effect in the model [13]. We are interested in testing the null hypothesis of the LM-Error test, $H_0 : \rho = 0$ (with assumption that $\delta_\mu^2 \geq 0$), against the alternative, $H_1 : \rho \neq 0$, which means the presence of a spatial error autocorrelation. Baltagi [14] et al. (2003) proposed LM-Error test statistics for an error term autocorrelation in a spatial panel data error model:

$$LM_{Error-RE} = \frac{\hat{\sigma}(\rho)^2}{(T-1) + \hat{\sigma}_u^4 / \hat{\sigma}_1^4} \quad (3)$$

where $\hat{D}(\rho) = \frac{1}{2} \hat{u}' \left[\frac{\partial_u^4}{\partial_1^4} (\bar{J}_T \otimes (W' + W)) + \frac{1}{\partial_u^2} (E_T \otimes (W' + W)) \right] \hat{u}$, $\sigma_1^2 = T\sigma_\mu^2 + \sigma_v^2$.

Under the null hypothesis, $\hat{\sigma}_u^2 = \frac{\hat{u}' (E_T \otimes I_N) \hat{u}}{N(T-1)}$ and $\hat{\sigma}_1^2 = \frac{\hat{u}' (\bar{J}_T \otimes I_N) \hat{u}}{N}$, respectively, present the ML estimates of σ_u^2 and σ_1^2 . Similarly, $\hat{u}$ is the ML estimated residual under the null hypothesis [15].

The new samples obtained through general Bootstrap sampling cannot be directly applied to a panel data model, as it does not precisely close the population and consider the multidimensionality of panel data. Thus, Beran [16] (1998) proposed the Double Bootstrap (DB) method to test spatial dependence in a spatial panel data model. By using the DB method, we obtain the distribution of test statistics through sampling and then make it approximate to the real distribution. Therefore, the DB method not only obtains a more accurate p value of statistics than the general Bootstrap sampling method, but it also improves the accuracy of population property in a spatial panel data econometric model.

DB sampling is divided into two steps. B_1 initial Bootstrap samples will be generated by the first sampling, which are used to calculate the Bootstrap statistics τ_j^* , for $j = 1, 2, \dots, B_1$. B_2 Bootstrap samples, which are generated by the second sampling on the basis of each of the first samples, are used to calculate the Bootstrap statistics τ_{jl}^{**} , for $l = 1, 2, \dots, B_2$. However, through this method, the operation time is too long and the computation is too much. For every B_1 Bootstrap sample, the $1 + B_2$ test statistics will be computed, thus $1 + B_1 + B_1 * B_2$ test statistics are needed by using DB sampling.

Davidson and MacKinnon et al. [17] (2002) further put forward a Fast Double Bootstrap (FDB) method (at an earlier time, we employed Monte Carlo simulation to compare the DB method with the FDB method for partial test. And we found that the FDB method can greatly save operation time, and, moreover, its test performances are very similar to the DB method) to optimize the DB sampling method. Using the FDB method to generate B_1 initial Bootstrap samples, we obtain test statistics τ_j^* . And to obtain one test statistic τ_j^* on the basis of a second Bootstrap sampling, we need to calculate $1 + B_1 + B_2$ test statistics. In this way, the FDB method greatly reduces Bootstrap sampling time and largely simplifies the operation. The formula for the p value of FDB is given by MacKinnon [18] (2006), as follows:

$$p_F^{**} = \frac{1}{B} \sum_{j=1}^B I(\tau_j^{**} > Q_B^{**}(1 - p^*)) \quad (4)$$

where p_F^{**} denotes the p value of FDB, p^* denotes p value of ordinary Bootstrap method, and $Q_B^{**}(1 - p^*)$ denotes $(1 - p^*)$ quantile of the statistic τ_j^{**} under the second Bootstrap sampling [19].

In this paper, we employ a residual FDB method and a Wild Bootstrap method [20], including symmetric and asymmetric, for LM-Error test statistics in a spatial panel data model. The steps are as follows:

(1) Based on the observation data set (y, X) , and without spatial dependence, we use a ML method to estimate the model in (2), and then we obtain estimated parameter $\hat{\beta}$ and residual vector $\hat{e} = y - X\hat{\beta}$. Given (3), we calculate asymptotic LM-Error statistics $\hat{L}$ and the p value of the asymptotic test.

(2) For the vector $\hat{e}$, we use Bootstrap sampling with N times replacement to generate a Bootstrap sample $\tilde{e}$, where $\tilde{e}$ is a repeatable random order permutation of each element of $\hat{e}$. If $\hat{e} = \tilde{e}$, we suggest a standard residual Bootstrap method under heteroscedastic errors. If $e_k = e_{1k} * v_k$ (for $k = 1, 2, \dots, N$), the Wild Bootstrap method is better, which can solve heteroscedasticity in the model. There, v_k has two forms:

① Symmetric Wild FDB: $e_k = e_{1k} * v_{k1}$

$$v_{k2} = \begin{cases} -(\sqrt{5}-1)/2, & p = (\sqrt{5}+1)/2\sqrt{5} \\ (\sqrt{5}+1)/2, & p = (\sqrt{5}-1)/2\sqrt{5} \end{cases} \quad (k = 1, 2, \dots, N) \quad (5)$$

② Asymmetric Wild FDB: $e_k = e_{1k} * v_{k2}$

$$v_{k1} = \begin{cases} -1, & p = 1/2 \\ 1, & p = 1/2 \end{cases} \quad (k = 1, 2, \dots, N) \quad (6)$$

In (5) and (6), v_{k1} , v_{k2} , respectively, correspond to the symmetric Wild Bootstrap and asymmetric Wild Bootstrap. In (5), the probability of the random variable v_{k1} , which is equal to 1 or -1 , is $\frac{1}{2}$. Similarly, the probability of the random variable v_{k2} , which is equal to $-\frac{\sqrt{5}-1}{2}$ or $\frac{\sqrt{5}+1}{2}$, is, respectively, $\frac{\sqrt{5}+1}{2\sqrt{5}}$ and $\frac{\sqrt{5}-1}{2\sqrt{5}}$.

(3) As a result, we obtain X , $\hat{\beta}$ and residual e , which are used to define the new dependent variable as $y_1 = X\hat{\beta} + e$. To obtain a new estimated parameter $\hat{\beta}^*$ and residual vector e^* , we apply the OLS method to estimate the model, and based on vector e^* , calculate the first sampling Bootstrap LM-Error test statistics as given in (4), denoting it by τ^* .

(4) Bootstrap sample e^{**} , which is generated via Bootstrap sampling for e^* again, is used to define $y_2 = X\hat{\beta} + e^{**}$. As the preceding step, the new estimated parameter $\hat{\beta}^{**}$ and $e^{**} = y_2 - X\hat{\beta}^{**}$ are given by OLS estimation. According to (4), we obtain the second Bootstrap sampling LM-Error test statistic, denoted as τ_j^{**} .

(5) Repeat step (2)~(4) times. Thus, the B first Bootstrap sampling LM-Error test statistics τ_j^* (for $j = 1, 2, \dots, B$) and the B second Bootstrap sampling LM-Error test statistics τ_j^{**} will be generated.

(6) $P^* = 2 * \min(\frac{1}{B} \sum_{j=1}^B I(\tau_j^* \leq \tau), \frac{1}{B} \sum_{j=1}^B I(\tau_j^* > \tau))$ is the p value of first Bootstrap sampling, which we substitute into (5) to calculate the FDB p value p_F^{**} .

Next, on the basis of Monte Carlo simulation, we prove the efficiency of the spatial dependence LM test in spatial panel data model by using the FDB method.

In the spatial panel data model, the size distortion of Bootstrap LM-Error test is given by the following: (the size distortion of the Bootstrap LM-Error test in the spatial panel data model is referred to as the difference between the test level and the given nominal significance level. The smaller the size distortion is, indicating that the Bootstrap test level is quite close to the nominal significance level, then the more reliable the hypothesis test result will be. The power of the Bootstrap LM-Error test in the spatial panel data model implies the probability of the Bootstrap LM-Error to reject the null hypothesis when the null hypothesis does not hold, i.e., there is spatial dependence in the spatial panel model)

$$\text{Size distortion} = \frac{1}{M} \sum_{i=1}^M I(p_F^{**}(\hat{I}_i) < \alpha) - \alpha \quad (7)$$

where α denotes significance level, as usual $\frac{W_{ij}}{\sum_{j=1}^N W_{ij}} = 0.05$, and W denotes Monte Carlo simulation times. ε_2 is indicator function, equal to 1 when $p_F^{**}(\hat{I}_i) < \alpha$, otherwise, 0.

Furthermore, the power of Bootstrap LM-Error test in spatial panel data model is given by:

$$\text{Power} = \frac{1}{M} \sum_{i=1}^M I(p_F^{**}(\hat{I}_i) < \alpha) \quad (8)$$

In the following section, by taking above steps and Monte Carlo simulation, when the null hypothesis or alternative hypothesis holds, we research the efficiency of the spatial dependence test in the spatial panel data model from two aspects: size distortion and efficiency.

3. Monte Carlo Simulation Experiment

In our context, we researched the size distortion and power of the Bootstrap LM-Error test and asymptotic test in the spatial panel data model, in order to prove the efficiency of the Bootstrap LM-Error test via Monte Carlo simulation experiment. Generally, the smaller the size distortion and the bigger the power is, the more valid the test statistics will be.

3.1. Parameter Setting and Experimental Procedures of Monte Carlo Simulation Experiment

The parameters of simulation experiment are given as follows:

(1) The data generation process is spatial error model given by:

$$y_t = \alpha \iota_N + X_t \beta + \mu + \varepsilon_t, \varepsilon_t = \rho W \varepsilon_t + v_t \quad (9)$$

where,

$$X = 10\text{rndu}(N, T) + 5\text{rndn}(N, T)$$

$\text{rndu}(N, T)$, $\text{rndn}(N, T)$, respectively, denotes the $N \times T$ matrix on even distribution and standard normal distribution. The spatial weight matrix is specified to the Rook matrix and Queen matrix. There are two forms of adjacency matrix: the Queen adjacency matrix and the Rook adjacency matrix. The Queen adjacency matrix determines that two regions are adjacent if they share a common boundary or a common vertex. The Rook adjacency matrix, however, only considers two regions as adjacent if they share a common boundary.

(2) In (9), α is the intercept and ι_N is all the matrixes of the N -dimension. We let $\alpha = 5$, $\beta = 0.5$. And μ indicates random effects with assumption, $\mu \sim N(0, I_N)$. Notice that v_t and μ are mutually independent, and $v_t \sim N(0, \sigma_v^2 I_N)$.

(3) Error term is set to v (v_1, v_2, v_3). Specifically speaking, v_1 denotes the standard normal error imposed $v_1 \sim N(0, 1)$, and v_2 denotes heteroscedastic distribution, which is specified as the product of independent variable and random generating normally distributed variable. v_3 is time-series error, which is generated by $v_{3it} = \theta v_{3i,t-1} + \eta_{it}$, with $\eta_{it} \sim N(0, \delta^2)$, $v_{3i,0} \sim N(0, \delta^2 / (1 - \theta^2))$, and θ presents serial correlation coefficient for time-series.

(4) Study size distortion of LM-Error test in spatial panel data model by setting $\rho = 0$, and the study power when $\rho \neq 0$ under condition of $\rho \in [-0.9, 0.9]$, which interval specified as 0.1.

(5) Take 5000 samples to carry out Monte Carlo simulation. Bootstrap times is given as $B = 199, 299 \dots 999$.

According to the above-mentioned setting, the Monte Carlo simulation experimental procedures are as follows:

(1) Based on computer-generated variables X and $\varepsilon = (I - \rho W)^{-1}v$, we obtain the dependent variable y from (9) and thereby sample (y, X) .

(2) Conduct Bootstrap LM-Error test to obtain $1+B$ spatial dependence statistics LM_{E1} and $LM^*_{E1}, LM^*_{E2}, \dots, LM^*_{EB}$; we then calculate the p value of the FDB LM-Error test and the asymptotic LM-Error test denoted p_F^* and p_i , respectively.

(3) Repeating step (1), (2) M times, we obtain the M p values of the FDB LM-Error test and asymptotic LM-Error test, i.e., p_{Fi}^* and p_i^* (for $i = 1, 2, \dots, M$).

(4) Calculate the size distortion and power of asymptotic LM-Error test and FDB LM-Error test based on (7) and (8).

According to above mentioned steps, we start the simulation experiment via Gauss10.0 programming.

3.2. Monte Carlo Simulation Results

Through Monte Carlo simulation, we study the efficiency of the FDB method when it is applied to the Bootstrap LM-Error test in a spatial panel data model under the random effects condition, with the presence of normal distributed errors, heteroscedastic errors or time-series autocorrelation. We combine the FDB method with asymmetric the Wild Bootstrap method to perform our simulation, and due to limited space, we only use the Rook matrix (one of spatial weight matrix) to show our simulation results.

3.2.1. Size Distortion

Let $\rho = 0$ and individual effect be the random effect. Respectively, under conditions that the error term obeys normal distribution (v_1), or the presence of heteroscedastic errors (v_2) or autocorrelation in time-series (v_3), we are interested in size distortion of FDB LM-Error test under the influence of Bootstrap simulation times ($B = 199, 299 \dots 999$), Rook matrix as spatial weight matrix, and sample size given by $(N, T) = (49, 10), (49, 30), (100, 10)$.

(1) Error term with normal distribution

As shown in Figure 1, the horizontal axis represents the Bootstrap simulation times ($B = 199, 299 \dots 999$), and the longitudinal axis reflects size distortion. The letter 'Theory' stands for theoretical value of the size distortion, i.e., 0. 'Bp' and 'Asy', respectively, show the size distortion of FDB LM-Error test and asymptotic LM-Error test. The smaller the size distortion and the closer the size distortion curve and the theoretical curve are, the more valid the test will be.

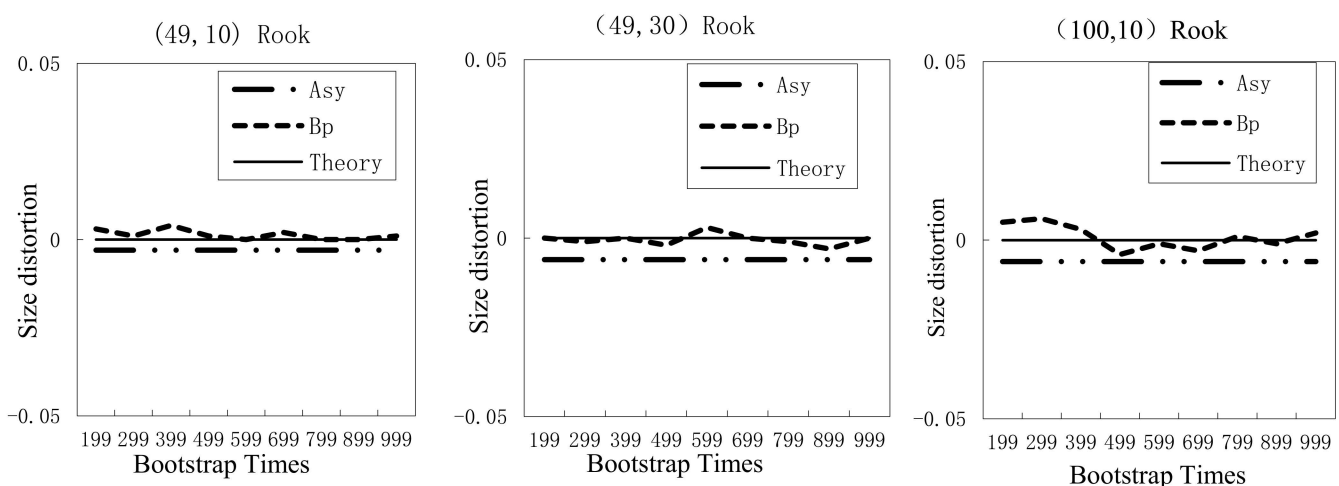


Figure 1. Size distortion of FDB LM-Error test in spatial panel data model with standard normal distributed errors.

Under standard normal distributed errors, the Bp curve gradually closes to the Theory curve with the sample size increasing, i.e., when the size distortion of FDB LM-Error decreases, Asy curve will be move gradually away from the Theory curve, which means that the size distortion of the FDB LM-Error is smaller than the asymptotic test (Asy curve).

Moreover, the size distortion curve of either the FDB LM-Error test or the asymptotic LM-Error test will level off as the Bootstrap simulation times increases.

(2) The presence of heteroscedastic errors

When the error term has heteroscedasticity, we study how the different sample sizes and Bootstrap simulation times will influence the size distortion of the LM-Error test. (See Figure 2).

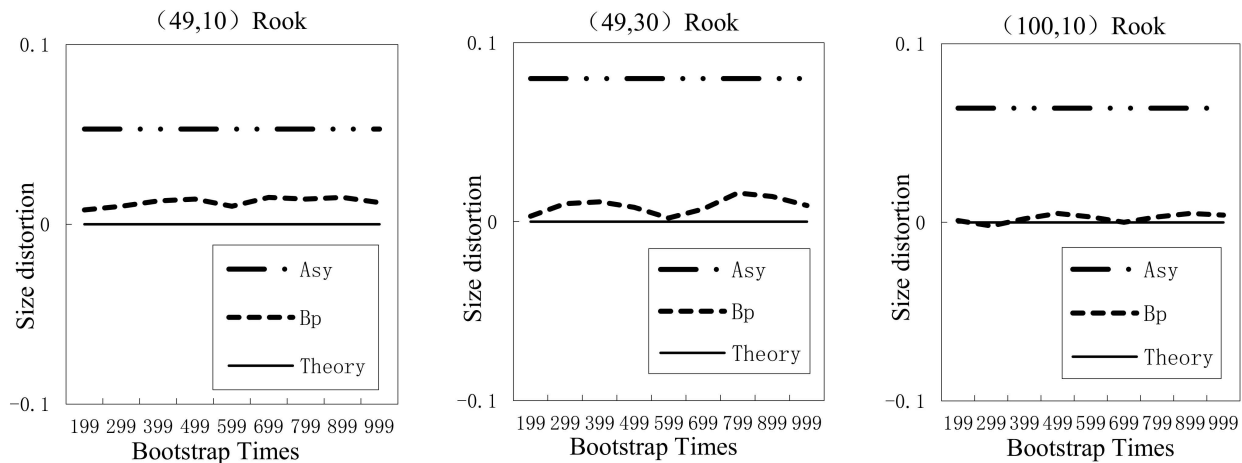


Figure 2. Size distortion of FDB LM-Error test in spatial panel data model with heteroscedastic errors.

Under heteroscedastic errors, the set spatial weight matrix is Rook matrix, and the size distortion of the FDB LM-Error decreases with the increase in sample size. Meanwhile, the Bp curve becomes gradually closer to the Theory curve, while the Asy curve is far from the Theory curve and thus has a bigger size distortion. In conclusion, the asymptotic LM-Error test has a large size distortion, which the FDB LM-Error can remedy, and, in the same situation, performs more validly.

(3) Time serial correlated errors

When we let the correlation coefficient be $\theta = 0.5$, then the different sample size and the Bootstrap simulation times influence the size distortion of the FDB LM-Error test; this can be seen in Figure 3.

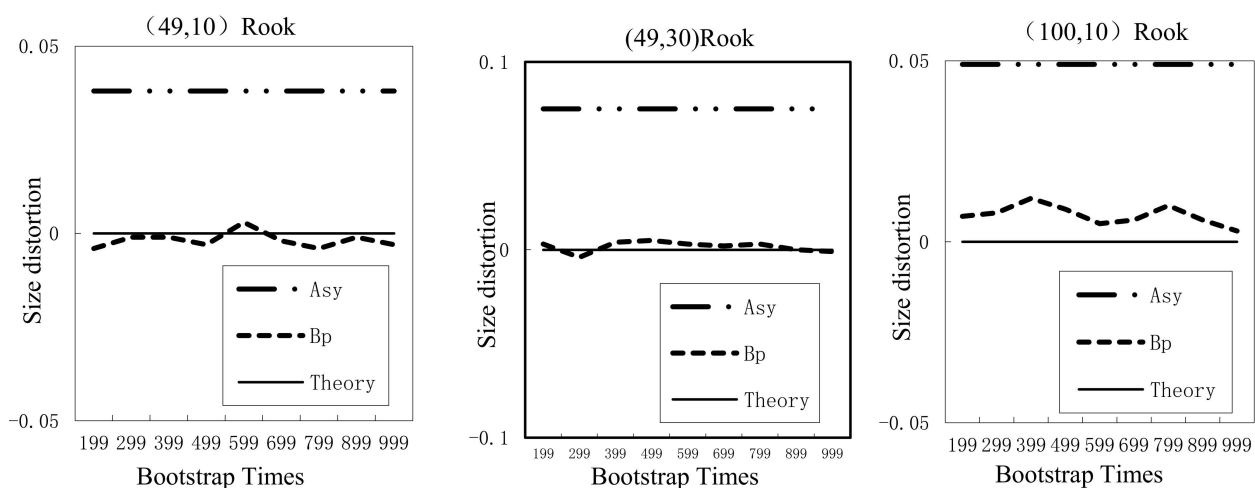


Figure 3. Size distortion of FDB LM-Error test in spatial panel data model with time-series correlated errors.

As shown in Figure 3, under time-series correlated errors, the asymptotic LM-Error test has a large size distortion while the FDB LM-Error test is very small, near to the Theory value, i.e., 0. Moreover, when the sample size increases, the size distortion of the asymptotic LM-Error test will become larger. On the contrary, the FDB LM-Error test will be stable and

converge to the Theory value 0. As the Bootstrap simulation times increases, we also find that the size distortion of the FDB LM-Error test is close to the Theory value 0 and levels off. Therefore, the FDB LM-Error test is a more valid method for a spatial dependence test.

Furthermore, on the basis of the aforementioned Bootstrap simulation, the simulation results tend to be stable when the Bootstrap times reach 399. In the following example, we let the Bootstrap times B equal to 399, taking the sample size as $(49,10)$, $(49,20)$, $(49,30)$ and using Rook matrix. Then, we wanted to know what the influences were on the size distortion of the asymptotic LM-Error test and the FDB LM-Error test when correlation coefficient θ and time T changes.

As we can see from Figure 4, when θ increases, the size distortion of the asymptotic LM-Error test increases and suffers from a huge upward trend, and is thus very far away from the Theory value 0. On the other hand, the FDB LM-Error test performs well, its size distortion is very small and is always close to the Theory value 0. Thus, the FDB LM-Error test can effectively refine the size distortion of the asymptotic LM-Error test. Furthermore, as T changes from 10 to 30, the FDB LM-Error test still levels off, and the asymptotic LM-Error test also becomes larger. This indicates that with the disturbance of errors growing (i.e., correlation coefficient of time-series increasing) under time-series correlation, the asymptotic LM-Error test will deviate more from the Theory value as it cannot deal with such disturbances. However, the FDB LM-Error test can avoid disturbance of errors by resampling. Thereby, its simulation results are always near the Theory value. Consequently, the FDB LM-Error test is more valid than the asymptotic LM-Error test.

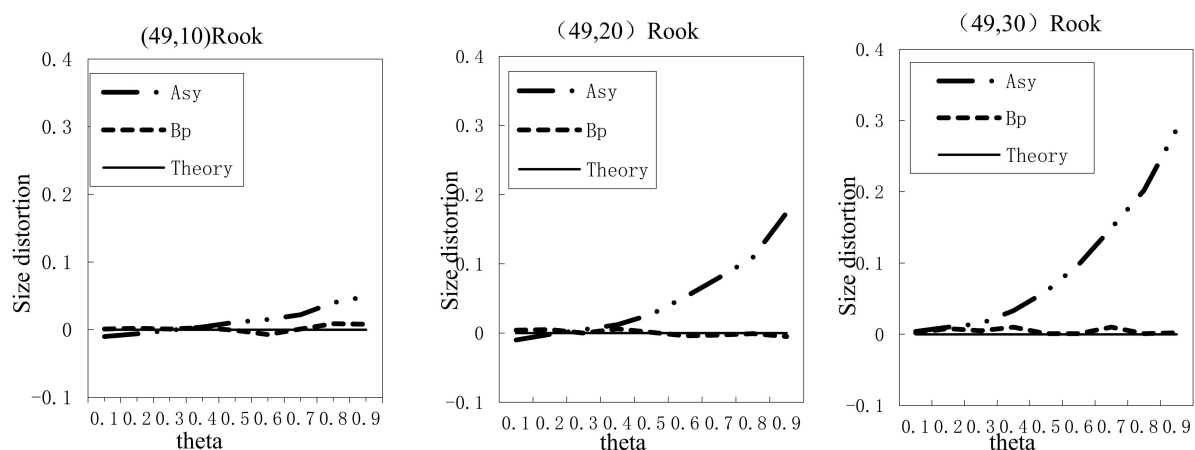


Figure 4. The size distortion of asymptotic LM-Error test and FDB LM-Error test with the influence of θ and T .

3.2.2. Power

When $\rho \in [-0.9, 0.9]$ and $\rho \neq 0$ ($\rho = 0$ meaning test level, covered earlier), letting the Bootstrap simulation times be $B=399$, taking the Rook matrix as spatial weight matrix, and the sample size is respectively $(49,10)$, $(49,30)$, $(100,10)$, we research the power of the FDB LM-Error test for testing the spatial dependence in the spatial panel data model under normal distributed errors, heteroscedastic errors or time-series correlation.

(1) Error term with normal distribution

Setting the conditions as mentioned above, the power of the asymptotic LM-Error test and FDB LM-Error test is shown in Figure 5.

In Figure 5, 'Asy' and 'Bp' still, respectively, stands for asymptotic test and FDB test. The horizontal axis represents the correlation coefficient of spatial error ρ , and the longitudinal axis reflects the power. With the growing sample size, we find that the power of the asymptotic LM-Error test and FDB LM-Error test increases as ρ increases. Furthermore, the power of the FDB LM-Error test is quite close to that of asymptotic LM-Error test. Moreover, when $|\rho| > 0.3$, the power of both are almost equal to 1, which show that when there is the presence of large spatial dependence, both tests are quite valid.

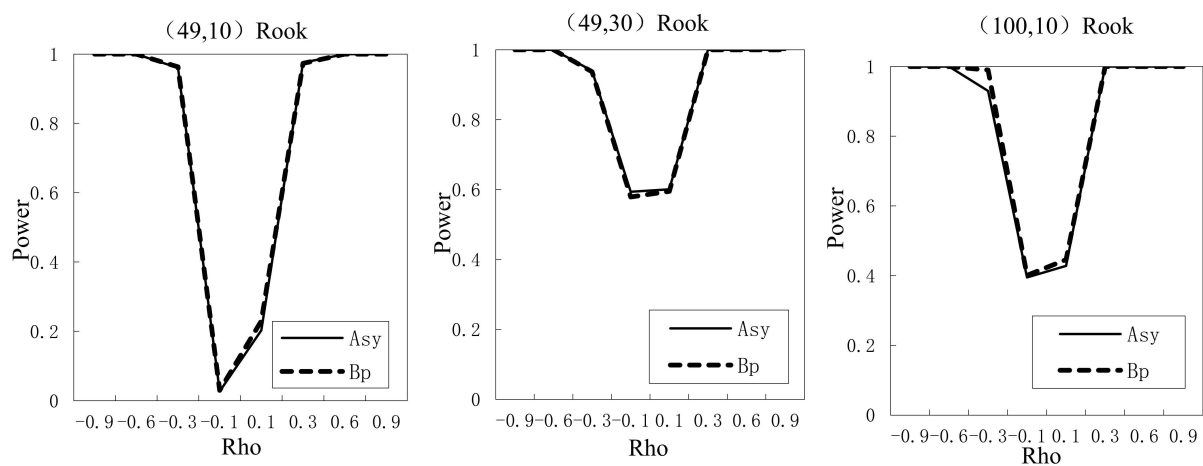


Figure 5. The power of asymptotic LM-Error test and FDB LM-Error test with standard normal distributed errors.

(2) The presence of heteroscedastic errors

The parameters are set as above. We can see the power of asymptotic LM-Error test and FDB LM-Error test in Figure 6.

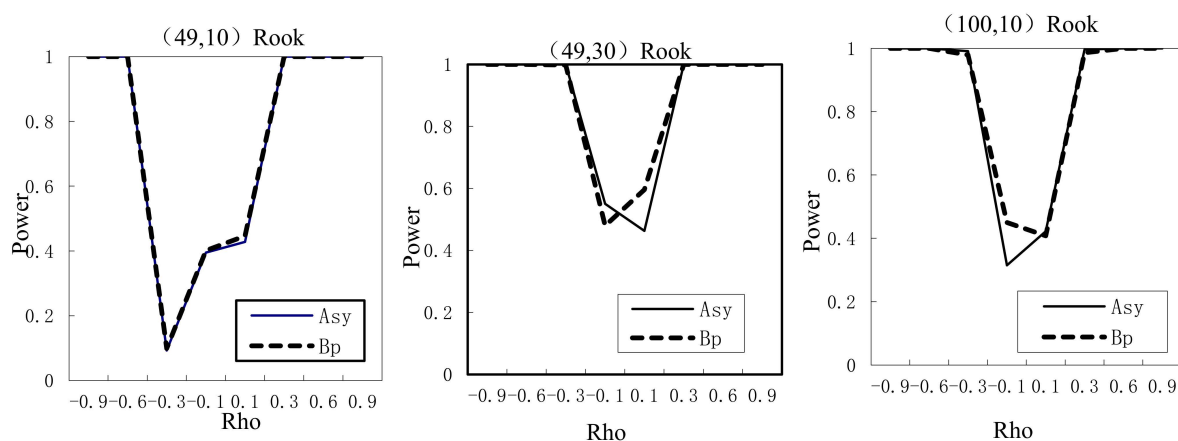


Figure 6. The power of asymptotic LM-Error test and FDB LM-Error test with heteroscedastic errors.

Crosswise, we can easily find that both the power of the asymptotic LM-Error test and the FDB LM-Error test increases as the sample size grows, and the power of the FDB LM-Error test is greater than or equal to that of the asymptotic LM-Error test. Meanwhile, both powers are almost equal to 1 when the condition $|\rho| \geq 0.6$ is satisfied. Thus, we can draw the conclusion that both the FDB LM-Error test and the asymptotic LM-Error test have good performance, although the spatial dependence is large.

(3) Time serial correlated errors

In this section, we set the correlation coefficient as $\theta = 0.5$, and for the others it is the same as previously mentioned. Then, Figure 7 shows the power of the FDB LM-Error test for testing spatial dependence in the spatial panel data model with random effects, under the condition that the error term has a time-series correlation.

From the horizontal visual angle, with the increase in sample size, both the powers of FDB LM-Error test and asymptotic LM-Error test increase, and the former is greater than or equal to the latter. When $|\rho| \geq 0.3$, they are almost equal to 1.

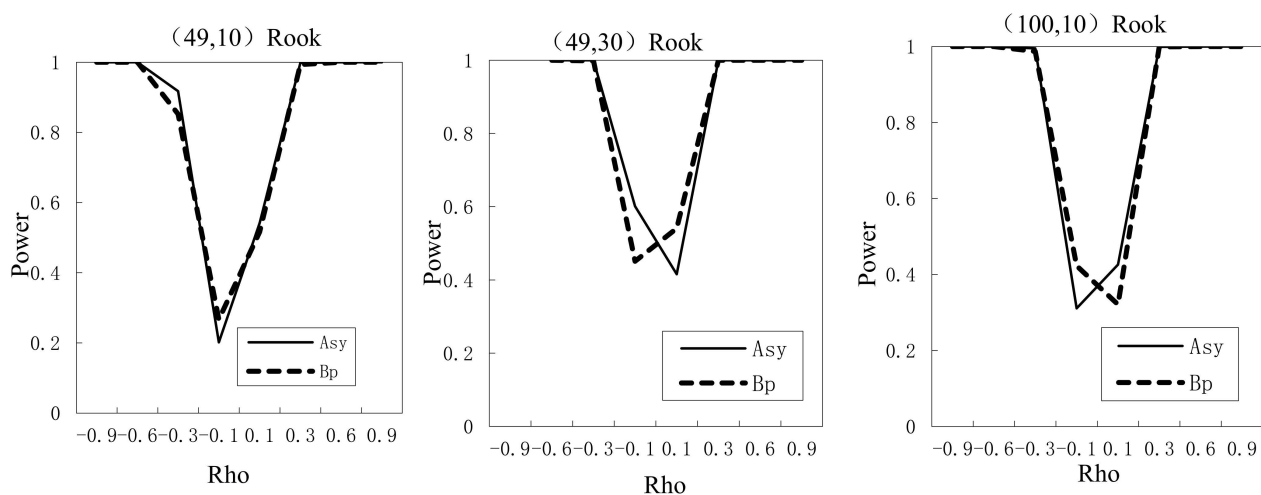


Figure 7. The power of asymptotic LM-Error test and FDB LM-Error test with time-series correlated errors.

We further focus on the influence of the correlation coefficient θ and time dimension T on the power of asymptotic LM-Error test and FDB LM-Error test. By setting the Rook matrix as the spatial weight matrix, Bootstrap times $B = 399$, and taking three sample size as $(49,10)$, $(49,20)$, $(49,30)$, which different from aforementioned sample size, and letting $\rho = 0.3$, $\theta \in [0.1, 0.9]$, the power of the two tests can be seen in Figure 8.

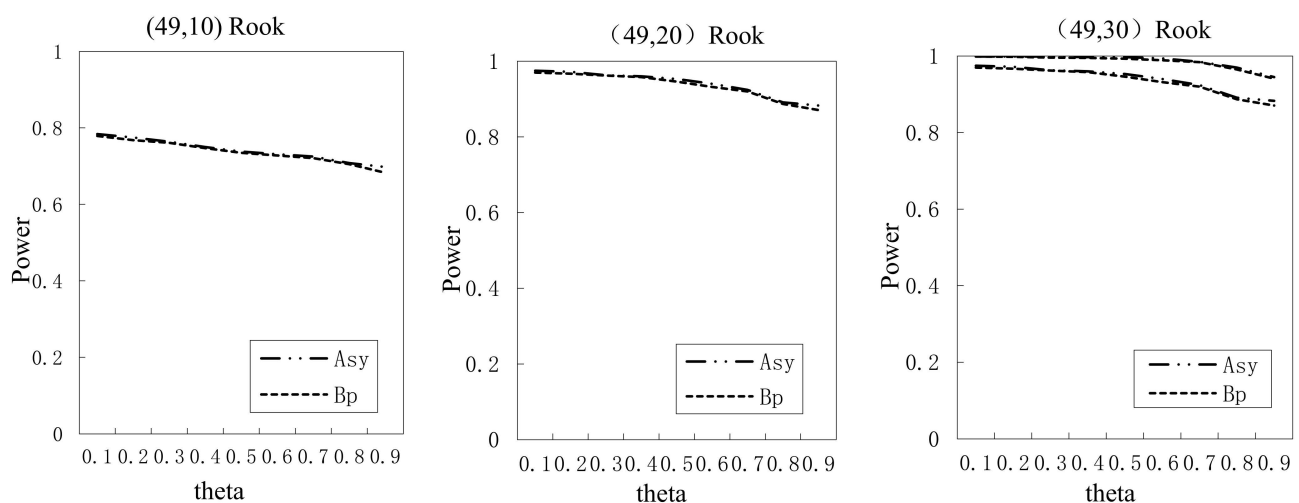


Figure 8. The power of asymptotic LM-Error test and FDB LM-Error test with the influence of θ and T .

It is not very difficult to find that both powers of the FDB LM-Error test and the asymptotic LM-Error present a decreasing trend while θ increases, and they are very close. As T changes from 10 to 30, the power of the two tests both increase and their curve nearly overlaps.

In summary, the Monte Carlo simulation results indicate that the asymptotic LM-Error test has a large size distortion when the error term violates the classical normal distribution in the spatial panel data model, while the FDB LM-Error test allows a small size distortion under the premise that the power of this test has essentially no loss. Therefore, the FDB LM-Error test is a more valid method for the spatial dependence test in the spatial panel data model.

4. Conclusions

This paper applied the FDB method into the LM-Error test for spatial dependence in a spatial panel data model, and based on a large amount of the Monte Carlo simulation, we researched the efficiency of the Bootstrap LM-Error test method for spatial dependence in the spatial panel data random effect model from the size distortion and power two

aspects, under the conditions that the error term is normal distributed, heteroscedastic or time-series correlated.

In brief, since the error term violates classical normal distribution, through Monte Carlo simulation, we find that the asymptotic LM-Error test has a large size distortion, while the FDB LM-Error test can effectively refine the size distortion of an asymptotic LM-Error test on the premise that the power has almost no loss. In conclusion, we prove that the FDB LM-Error test is a more valid method than the asymptotic LM-Error test for testing spatial dependence in a spatial panel data model.

In the future, we will conduct research on the applications of various Bootstrap methods in spatial econometric analysis and carry out empirical studies based on Chinese economic data.

Author Contributions: Conceptualization, T.R.; Methodology, L.X.; Writing—original draft, T.R.; Supervision, Z.R. All authors have read and agreed to the published version of the manuscript.

Funding: This paper is supported by the Natural Science Foundation of China's General Program "Research on Rubin's Causal Paradigm Based on Supervised Machine Learning: Model Construction and Specification Testing", with the project number (72073079). MOE (Ministry of Education in China) Project of Humanities and Social Science, with the project number (19YJC790163).

Data Availability Statement: This paper uses Monte Carlo simulation method, and all the data used in the simulation are randomly generated according to the set conditions.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Anselin, L. Thirty years of spatial econometrics. *Pap. Reg. Sci.* **2010**, *89*, 4–25. [\[CrossRef\]](#)
2. Efron, B. Bootstrap methods: Another look at the jackknife. *Ann. Stat.* **1979**, *7*, 101–118. [\[CrossRef\]](#)
3. Davidson, R.; MacKinnon, J.G. Improving the reliability of bootstrap tests with the fast double bootstrap. *Comput. Stat. Data Anal.* **2007**, *51*, 3259–3281. [\[CrossRef\]](#)
4. Chang, Y. Bootstrap unit root tests in Panels with cross-sectional dependence. *J. Econom.* **2003**, *110*, 263–293.
5. Cerrato, M.; Sarantis, N.A. Bootstrap Panel unit root test under cross-sectional dependence with an application to PPP. *Comput. Stat. Data Anal.* **2007**, *51*, 4028–4037. [\[CrossRef\]](#)
6. Godfrey, L. *Bootstrap Tests for Regression Models*; Palgrave Macmillan: London, UK, 2009.
7. Warren, D.E. Bootstrap Methods for Inference with Spatial Panel Data Models. In Proceedings of the Moriarty Graduate Paper Competition at the 47th Annual Meeting of the Southern Regional Science Association, Arlington, VA, USA, 27–29 March 2008.
8. Yang, Z. A robust LM test for spatial error components. *Reg. Sci. Urban Econ.* **2010**, *40*, 299–310. [\[CrossRef\]](#)
9. Lin, K.P.; Long, Z.; Ou, B. The size and power of bootstrap tests for spatial dependence in a linear regression model. *Comput. Econ.* **2011**, *38*, 153–171. [\[CrossRef\]](#)
10. Ren, T.; Long, Z.; Zhang, R.; Chen, Q. Moran's I Test of Spatial Panel Data Model Based on Bootstrap Method. *Econ. Model.* **2014**, *41*, 9–14. [\[CrossRef\]](#)
11. Yang, Z. LM tests of spatial dependence based on bootstrap critical values. *J. Econom.* **2015**, *185*, 33–59. [\[CrossRef\]](#)
12. Jin, F.; Lee, L.F. On the bootstrap for Moran's I test for spatial dependence. *J. Econom.* **2015**, *184*, 295–314. [\[CrossRef\]](#)
13. Jin-feng, Z. Testing Random Effect for Panel Data Model with Spatial Error Correlation. *J. Quant. Tech. Econ.* **2012**, *10*, 127–138.
14. Baltagi, B.H.; Song, S.H.; Koh, W. Testing panel data regression models with spatial error correlation. *J. Econom.* **2003**, *117*, 123–150. [\[CrossRef\]](#)
15. Ling, O.B.; Bing, L.; Feitong, Z. Study on the validity of robust LM test for time-varying spatial weight matrix panel data model. *Stat. Res.* **2015**, *10*, 98–105.
16. Beran, R. Prepivoting test statistics: A bootstrap view of asymptotic refinements. *J. Am. Stat. Assoc.* **1988**, *83*, 687–697. [\[CrossRef\]](#)
17. Davidson, R.; MacKinnon, J.G. Fast double bootstrap tests of nonnested linear regression models. *Econom. Rev.* **2022**, *21*, 417–427. [\[CrossRef\]](#)
18. Davidson, R.; MacKinnon, J.G. The power of bootstrap and asymptotic tests. *J. Econom.* **2006**, *133*, 421–441. [\[CrossRef\]](#)
19. Shufen, Z.; Yunan, S.; Chuanhua, W. Spatial effect test of fixed effect variable coefficient spatial autoregressive panel data model. *Appl. Math.* **2022**, *4*, 947–955.
20. Burridge, P.B.; Fingleton, B. Bootstrap inference in spatial econometrics: The J-test. *Spat. Econ. Anal.* **2010**, *5*, 93–119. [\[CrossRef\]](#)

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.