

Anselin, Luc; Moreno, Rosina

**Conference Paper**

## Properties of tests for spatial error components

41st Congress of the European Regional Science Association: "European Regional Development Issues in the New Millennium and their Impact on Economic Policy", 29 August - 1 September 2001, Zagreb, Croatia

**Provided in Cooperation with:**

European Regional Science Association (ERSA)

*Suggested Citation:* Anselin, Luc; Moreno, Rosina (2001) : Properties of tests for spatial error components, 41st Congress of the European Regional Science Association: "European Regional Development Issues in the New Millennium and their Impact on Economic Policy", 29 August - 1 September 2001, Zagreb, Croatia, European Regional Science Association (ERSA), Louvain-la-Neuve

This Version is available at:

<https://hdl.handle.net/10419/115233>

**Standard-Nutzungsbedingungen:**

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

**Terms of use:**

*Documents in EconStor may be saved and copied for your personal and scholarly purposes.*

*You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.*

*If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.*

# Properties of Tests for Spatial Error Components\*

LUC ANSELIN

Regional Economics Applications Laboratory (REAL)

and

Department of Agricultural and Consumer Economics

University of Illinois, Urbana-Champaign

Urbana, IL 61801

USA

anselin@uiuc.edu

ROSINA MORENO

Analisi Quantitativa Regional (AQR)

Department of Econometrics, Statistics and Spanish Economy

University of Barcelona

08034 Barcelona

Spain

rmore@eco.ub.es

April 20, 2001

## Abstract

In spatial econometrics, the typical alternative of spatial autocorrelation is expressed in the form of a spatial autoregressive process. While the bulk of the literature is devoted to specification tests and estimation methods for this model, alternatives have been suggested as well. In this paper, we consider an alternative that takes the form of the spatial error components formulation proposed by Kelejian and Robinson (1995). We consider a number of specification tests against this alternative, based on both a maximum likelihood framework as well as on a general method of moments estimation approach. We compare the performance of these tests in a series of Monte Carlo simulation experiments for a range of different spatial layouts and under a number of different error distributions.

*Key Words:* spatial econometrics; specification tests; spatial error components.

---

\*Anselin's research was supported in part by NSF Grant BCS-9978058 to the Center for Spatially Integrated Social Science (CSISS). Rosina Moreno is grateful to the Regional Research Institute at West Virginia University for its hospitality when carrying out part of this work. Also, she gratefully acknowledges financial support from the DGICYT SEC99-0700. Anil Bera made several useful suggestions, which are greatly appreciated. Other comments were provided by participants in the spatial econometrics sessions at the 40th Western Regional Science Association/Applied Econometrics Association Meetings in Palm Springs, CA, Feb. 2001, and in seminar presentations at the University of Barcelona and Texas A&M University.

## 1 Introduction

A large number of diagnostics for spatial error dependence have been suggested in the spatial econometric literature. These can be broadly categorized as tests against an unspecified alternative of spatial correlation, or tests against specific spatial processes. Examples of the former are Moran's  $I$  (Cliff and Ord 1981) and the Kelejian-Robinson test (Kelejian and Robinson 1992). Examples of the latter are tests based on the maximum likelihood principle, such as Wald (W) tests, likelihood ratio (LR) tests and Lagrange Multiplier (LM) tests (for a review, see Anselin and Bera 1998, Anselin 2001).

Most of the tests against spatial error processes have been formulated for spatial autoregressive (SAR) or spatial moving average (SMA) processes as the alternative. In a recent paper, Kelejian and Robinson (1995) suggested a different type of spatial process that combines a location-specific or local error component with a *regional* or spillover component in what they refer to as a *spatial error component process*.<sup>1</sup>

The spatial error components model seems particularly appropriate when the range of spatial autocorrelation is constrained to close neighbors, and it has been applied in studies of local public finance, such as the role of spatial spillovers in the productivity of infrastructure investments (Kelejian and Robinson 1997). In this paper, we focus on the properties of *specification tests* against alternatives of the spatial error components form. Kelejian and Robinson (1993, 1995) suggest two testing strategies for this case. One consists of an application of the original Kelejian-Robinson test (Kelejian and Robinson 1992), the other of a test for the significance of a coefficient in an auxiliary regression (Kelejian and Robinson 1993). In Anselin (2001), a Lagrange Multiplier statistic is outlined. To date, very little is known about the properties of these tests, and in particular about their performance in situations typically encountered in empirical practice. The objective of our paper is to shed some light on this issue by examining the size and power of the tests in a number of Monte Carlo simulation experiments.

In the next section, we first review the model and highlight some of its properties relative to other specifications for spatial error correlation. We then outline four test statistics in some detail. This is followed by a description of the design of the simulation experiments. The results are discussed next, and we close with some concluding remarks and practical recommendations.

## 2 The Spatial Error Components Model

The spatial error components model (SEC) incorporates a *local* and a *spillover* element in the variance-covariance matrix of the error term in a linear regression model. Formally, the model is

$$y = X\beta + \varepsilon, \quad (1)$$

with

$$\varepsilon = W\psi + \xi, \quad (2)$$

---

<sup>1</sup>Such a process should not be confused with the traditional notion of error components in panel data analysis, where random effects yield an error variance that is decomposed into a location-specific variance, a time-specific variance and a remainder variance.

where  $y$  is an  $n$  by 1 vector of observations on the dependent variable,  $X$  is a  $n$  by  $k$  matrix of observations on the explanatory variables, and  $\beta$  is a  $k$  by 1 vector of regression coefficients. In the error specification (2),  $W$  is a spatial weights matrix of dimension  $n$  by  $n$ ,  $\psi$  is an  $n$  by 1 vector of errors that incorporate the spillover across neighbors, and  $\xi$  is an  $n$  by 1 vector of location-specific disturbance terms. Each vector of errors is assumed to consist of i.i.d. terms and the two components are taken to be uncorrelated:<sup>2</sup>

$$E[\psi] = 0, E[\xi] = 0 \quad (3)$$

$$E[\psi\psi'] = \sigma_\psi^2 I, E[\xi\xi'] = \sigma_\xi^2 I, \quad (4)$$

$$E[\psi_i \xi_j] = 0, \forall i, j. \quad (5)$$

Given the uncorrelatedness between the spillover component and the location-specific component in (5), the complete error variance-covariance matrix is obtained as

$$E[\varepsilon\varepsilon'] = \Omega_{SEC} = \sigma_\psi^2 WW' + \sigma_\xi^2 I, \quad (6)$$

where  $\sigma_\psi^2$  is the variance component associated with the spatial spillovers, and  $\sigma_\xi^2$  is the usual variance term. The variance-covariance matrix (6) is positive definite for  $\sigma_\psi^2 \geq 0$  and  $\sigma_\xi^2 > 0$ , since  $WW'$  is positive semidefinite. This avoids a singularity problem associated with the SAR process (see Kelejian and Robinson 1995, for details).

It is interesting to compare the variance-covariance structure of a spatial error components specification to the variance-covariance matrix that results from a spatial moving average process for the errors. With

$$\varepsilon = \lambda W\xi + \xi, \quad (7)$$

the error variance-covariance matrix follows as

$$\Omega_{SMA} = \sigma_\xi^2 [I + \lambda(W + W') + \lambda^2 WW'] . \quad (8)$$

We can rewrite equation (6) as

$$\Omega_{SEC} = \sigma_\xi^2 [I + \theta WW'] , \quad (9)$$

where  $\theta = \sigma_\psi^2 / \sigma_\xi^2$  is the ratio of the variances of the two error components. This variance-covariance matrix is nearly identical to the SMA error variance (8), except that it lacks the term in  $W + W'$ . Note that since equation (9) cannot be obtained by imposing parameter constraints on equation (8) (setting  $\lambda$  to zero would remove both spatial terms in equation 8), the two specifications are in fact non-nested.<sup>3</sup> The nature of the induced range of the non-zero covariances is similar however, in that both models

<sup>2</sup>Note that the model considered here is the specification presented in Kelejian and Robinson (1995). Additional forms of misspecification may be introduced as well, in the form of heteroskedasticity in either  $\psi$  or  $\xi$ , or both (provided the proper identification conditions are satisfied). Also, higher order processes could be considered. We leave these complications for future work.

<sup>3</sup>They may also differ in terms of higher order moments.

yield zero covariances beyond the second order neighbors.<sup>4</sup> In other words, both SMA and SEC can be considered as models for *local* spatial autocorrelation, as opposed to the SAR specification, which induces *global* autocorrelation, including more remote neighbors as well through a distance decay effect (for technical details, see Anselin and Bera 1998). Note also that this interpretation differs from the original suggestion by Kelejian and Robinson (1995), who saw the SEC model primarily as an alternative to a spatial autoregressive specification (Kelejian and Robinson 1995, p. 89). The latter is only valid approximately and for very small values of the autoregressive parameter, such that higher order lag terms in the covariance matrix (higher than second order) can be ignored. Typically, this is not the case, and the spatial covariance induced by a SAR process goes beyond the second order neighbors.

### 3 Test Statistics

For each of the test statistics that follow, the null hypothesis is that of the classic regression model with uncorrelated and homoskedastic errors,  $E[\varepsilon\varepsilon'] = \sigma^2 I$ . More specifically, for the SEC model in (9), this implies a constraint on the parameters of the form

$$H_0 : \theta = 0, \quad (10)$$

or

$$H_0 : \sigma_\psi^2 = 0. \quad (11)$$

We now turn to three different approaches for testing this null hypothesis against a spatial error components alternative.

#### 3.1 Kelejian-Robinson Test

The original Kelejian and Robinson (1992) test, hereafter referred to as KR, does not require a complete specification of the generating process for the error term.<sup>5</sup> Instead, the alternative is that the covariance between two *neighboring* error terms is non-zero, or that

$$\text{Cov}[\varepsilon_i, \varepsilon_j] = \sigma_{ij} = z_{ij}\gamma, \quad (12)$$

where  $z_{ij}$  is a 1 by  $q$  vector of covariates, typically taken to be a function of the original explanatory variables at  $i$  and  $j$ , with  $i$  and  $j$  as “contiguous” locations in a general spatial ordering of the observations. For example, the  $z_{ij}$  could be constructed from the cross products of  $x_i$  and  $x_j$ . The  $q$  by 1 coefficient vector  $\gamma$  indicates the degree to which the covariates in  $z$  can explain the *non-zero* covariance in (12). Intuitively, the absence

<sup>4</sup>Strictly speaking, this remark pertains to the situation where  $W$  is simple first order contiguity, but it has a similar generalization for other weights specifications.

<sup>5</sup>The KR test is developed for a very general setting and also does not require normality or linearity. However, in this paper, we will limit our comments to the context of a linear regression model. A later version, presented in Kelejian and Robinson (1998) deals with a situation where both spatial autocorrelation and additional heteroskedasticity are present. We do not consider this alternative here and focus solely on the spatial autocorrelation and heteroskedasticity induced by the SEC model as such. The heteroskedasticity follows when the diagonal terms in  $WW'$  are not constant, which is typically the case for irregular lattices.

of spatial autocorrelation should *not* yield a significant relationship between  $Cov[\varepsilon_i, \varepsilon_j]$  and  $z_{ij}$ , or, the estimates for the coefficients  $\gamma$  should not be significant. More formally, the KR test boils down to a test on the null hypothesis  $H_0 : \gamma = 0$  in (12).

The test is implemented by regressing  $h_n$  cross-products of ordinary least squares (OLS) residuals for neighboring locations (where there are  $h_n$  such unique pairs),  $\hat{C}_{ij} = e_i \cdot e_j$ , on matching cross-products of explanatory variables,  $z_{ij} = x_i \cdot x_j$ . With  $\hat{\gamma} = (Z'Z)^{-1}Z'\hat{C}$  as the OLS estimates in this regression (where  $Z$  and  $\hat{C}$  are, respectively, a  $h_n$  by  $q$  matrix and  $h_n$  by 1 vector), the test is constructed as

$$KR = \frac{\hat{\gamma}'Z'Z\hat{\gamma}}{\hat{\sigma}^4}. \quad (13)$$

In (13),  $\hat{\sigma}^4$  is a consistent estimator for  $\sigma^4$ . Under the null hypothesis, the KR statistic is asymptotically distributed as  $\chi^2(q)$ , where  $q$  corresponds to the number of columns in the matrix  $Z$ .<sup>6</sup> Kelejian and Robinson (1992) proposed two asymptotically equivalent estimators for  $\hat{\sigma}^4$ . In what follows, we will use  $\hat{\sigma}^4 = \hat{\alpha} = (\hat{C} - Z\hat{\gamma})'(\hat{C} - Z\hat{\gamma})/h_n$ .

In Kelejian and Robinson (1995, p. 89, fn. 16) it is outlined under what conditions the KR test can be applied to alternatives of the SEC form. In general, these conditions are satisfied in practice for spatial weights based on contiguity, where the number of neighbors for each location is bounded. For more complex weights, they need to be verified in each individual case.

Even though the KR test does not explicitly refer to a spatial weights matrix  $W$ , there is an underlying notion of a “spatial ordering.” In Kelejian and Robinson (1992), this ordering is taken to be equivalent to the notion of first order neighbors, which corresponds to the non-zero elements in the upper (or lower) triangular part of a contiguity-based spatial weights matrix.<sup>7</sup> The rationale behind the selection of the  $i, j$  pairs is to identify those pairs that correspond to nonzero covariances (12). While it is often assumed that this is the same as the non-zero elements of  $W$  (i.e., the neighbors), this is not correct. Most spatial processes *induce* non-zero covariances for location pairs beyond the immediate neighbors (see, for example, the induced covariance structure for a spatial moving average process in equation 8). For the SEC model, non-zero covariances are not present for the first order neighbors, but pertain to all non-zero elements in the upper (or lower) triangular part of  $WW'$ . We therefore believe that considering both first and second order neighbors in the KR test may yield higher power. In the simulation experiments, we consider two forms of the KR test: one bases the pairs  $i, j$  on the non-zero elements in  $W$ ,  $KR_W$ ; the other on the non-zero elements in both  $W$  and  $WW'$ ,  $KR_{WW}$ .<sup>8</sup>

<sup>6</sup>For a technical derivation and the necessary regularity conditions, see Kelejian and Robinson (1992). A slightly different exposition is given in Anselin and Bera (1998, pp. 268–269).

<sup>7</sup>Kelejian and Robinson (1992) argue that their test does not require the specification of a spatial weights matrix. This is indeed the case when the weights are not simple contiguity weights. However, in the latter situation, which is predominant in empirical practice, the information contained in  $W$  and the information used to construct the  $i, j$  pairs in the KR test is exactly the same.

<sup>8</sup>One could argue that a “proper” specification should only consider the non-zero elements in  $WW'$  and exclude the first order neighbors. We kept the latter in order to enhance comparability with the original KR statistic.

### 3.2 GMM-Based Test

A second testing strategy, outlined in Kelejian and Robinson (1993) (see also Kelejian and Robinson 1997, for an application), is based on general method of moments (GMM) estimation. A GMM estimator for the variance component parameters in (6) can be obtained from the moment condition for the diagonal elements of the variance-covariance matrix  $\Omega_{SEC}$ ,

$$E[\varepsilon_i^2] = \sigma_\xi^2 + \sigma_\psi^2 [WW']_{ii}, \quad (14)$$

or, with  $d_i = [WW']_{ii}$  as the diagonal elements in the matrix  $WW'$ ,

$$E[\varepsilon_i^2] = \sigma_\xi^2 + \sigma_\psi^2 d_i. \quad (15)$$

With  $\tilde{\varepsilon}_i$  as the residual based on a consistent estimator for the regression parameters (such as the OLS residuals),  $\phi = (\sigma_\xi^2, \sigma_\psi^2)'$ , and with a  $n$  by 2 matrix of auxiliary variables  $Q = (1, d_i)$ , a consistent estimator for the variance components follows as

$$\hat{\phi} = (Q'Q)^{-1}Q'\tilde{\varepsilon}^2. \quad (16)$$

A test against spatial error components, further referred to as  $KR_{GMM}$ , is based on a t-statistic for the null hypothesis  $H_0 : \sigma_\psi^2 = 0$ . In practice, this is obtained from the estimate  $\hat{\sigma}_\psi^2$  and its estimated standard error in the OLS regression (16).<sup>9</sup> Due to the fact that the parameter space is constrained such that  $\sigma_\psi^2 \geq 0$ , the t-test must be one-sided. In certain empirical situations, it is possible that negative estimates are obtained for  $\hat{\sigma}_\psi^2$ .<sup>10</sup> By using a one-sided test, such values are ignored, or, more precisely, interpreted as evidence in favor of the null hypothesis.<sup>11</sup>

### 3.3 Lagrange Multiplier Test

Both the original KR statistic and the GMM based test are asymptotic and do not require distributional assumptions. When such assumptions are reasonable (for example, an assumption of normality), a maximum likelihood framework may be used as the basis for inference. A test for spatial error components can then be formulated as a special case of testing constraints on the parameters in regression models with non-spherical error terms. With the variance-covariance matrix  $\Omega_{SEC} = \sigma_\xi^2 [I + \theta WW']$ , the null hypothesis is of the form  $H_0 : \theta = 0$ .

Unlike the standard maximum likelihood case, this null hypothesis reflects a true parameter value on the boundary of the parameter space, since  $\theta = \sigma_\psi^2 / \sigma_\xi^2 \geq 0$ . This does not satisfy the basic regularity conditions for ML inference and thus precludes the

<sup>9</sup>See Kelejian and Robinson (1993, 1997) for technical details and derivations.

<sup>10</sup>For example, Table 2, p. 122 in Kelejian and Robinson (1997) reports a negative and significant  $\hat{\sigma}_\psi^2$  for their “basic” model.

<sup>11</sup>An intuitive way to interpret negative variance component estimates is to consider these to be a consequence of “uncooperative” data, and hence as not providing support for a model with spatial error components.

application of Wald or Likelihood Ratio tests without further corrections.<sup>12</sup> A Lagrange Multiplier or Rao Score test however does not suffer from this problem. Such a test statistic can be based on the general principles for testing in spatial models outlined in Anselin (1988), and is derived in Anselin (2001). The test only requires the residuals of an OLS regression. With  $\hat{\sigma}^2 = e'e/n$  as the estimate of error variance based on the OLS residuals  $e$ ,  $T_1 = \text{tr}WW'$  (and  $\text{tr}$  as the matrix trace operator), and  $T_2 = \text{tr}WW'WW'$ , the statistic is obtained as

$$\text{LM}_{SEC} = \left[ \frac{e'WW'e}{\hat{\sigma}^2} - T_1 \right]^2 / 2 \left[ T_2 - \frac{(T_1)^2}{n} \right]. \quad (17)$$

Under the null hypothesis, this statistic is asymptotically distributed as  $\chi^2(1)$ .

## 4 Design of Monte Carlo Experiments

The performance of the four statistics outlined in sections 3.1 to 3.3 ( $\text{LM}_{SEC}$ ,  $\text{KR}_W$ ,  $\text{KR}_{WW}$ , and  $\text{KR}_{GMM}$ ) is evaluated in a series of Monte Carlo experiments for a number of different spatial layouts, as well as for different underlying error distributions. They are also compared to the performance of the more traditional Moran's  $I$  ( $I$ ) and LM-Error ( $\text{LM}_{ERR}$ ) statistics. The simulation design follows the same general format as in earlier work such as Anselin and Rey (1991) and Anselin and Florax (1995).

In total, twelve different spatial layouts are considered, evenly divided among regular and irregular lattice structures.<sup>13</sup> The six regular lattice structures are 7 by 7 ( $n = 49$ ), 9 by 9 ( $n = 81$ ), 11 by 11 ( $n = 121$ ), 16 by 16 ( $n = 256$ ), 20 by 20 ( $n = 400$ ) and 32 by 32 ( $n = 1024$ ). For all these, the rook criterion of contiguity is used to construct the spatial weights. The irregular layouts match as closely as possible the number of observations for the regular lattices, although their internal connectedness structure will differ. They are obtained by taking spatial groupings of Western U.S. counties for dimensions  $n = 46, 80, 124, 264, 413$  and  $1013$ .<sup>14</sup>

A summary of the connectedness characteristics of the spatial weights is given in Table 1 in terms of the average number of neighbors for each location and the percent non-zero elements, or sparseness. The regular lattices are always sparser (and with fewer neighbors) than their irregular counterparts for similar sample size. Since the rook contiguity criterion is used to define neighbors, one would expect the average for

<sup>12</sup>See Anselin (2001) for a more elaborate treatment. Also, note that the regularity conditions for the "spatial" model are not the same as the classic ML regularity conditions. Additional constraints need to be imposed to limit the degree of dependence and heterogeneity of the underlying process, and use must be made of triangular arrays to establish consistency and asymptotic normality. A technical treatment of these issues as well as the fundamental theorems and lemmas can be found in Kelejian and Prucha (1998, 1999).

<sup>13</sup>A limited study of the performance of the  $\text{LM}_\theta$  statistic was carried out in Anselin (2001) for five regular lattice structures, with  $n = 25, 39, 81, 121$  and  $400$ . The earlier study was limited to normally distributed errors only, and did not consider the other test statistics. To retain compatibility with the results for the other tests, all simulations were recomputed for the current paper.

<sup>14</sup>The spatial groupings are made up of the counties in Utah and Nevada for  $n = 46$ ; Montana and Wyoming for  $n = 80$ ; Montana, Wyoming and Idaho for  $n = 124$  (S124); S124 plus Colorado, Arizona, New Mexico and Utah for  $n = 264$  (S264); S264 plus Nevada, California, Oregon and Washington for  $n = 413$  (S413); and S413 plus North Dakota, South Dakota, Nebraska, Kansas, Oklahoma and Texas for  $n = 1013$ .



the regular lattices to be close to 4. As the sample size grows, this is indeed the case, the lower values in the smaller data sets reflecting the influence of the boundary cells. The degree of sparseness ranges from 11.11% for the irregular  $n = 46$  to 0.38% for the regular 32 by 32 lattice. It is important to keep in mind that in the assessment of the properties of the test statistics, sample size is not the sole determining factor, but other characteristics of the “layout” of the observations play a role as well. Of the four tests,  $KR_{GMM}$  may be most susceptible to this, since it relies on the variation in the diagonal elements of  $WW'$ , which is much higher in the irregular than in the regular lattices. Were it not for boundary locations, the regular lattice structure would violate one of the basic assumptions underlying  $KR_{GMM}$ , namely that the  $WW'_{ii}$  should not be constant. For constant diagonal elements (e.g., as would be the case for a weights matrix based on  $k$ -nearest neighbors), the auxiliary regression (16) breaks down.

As in Anselin and Florax (1995), the regression part of the model is  $\alpha + \beta x_i$ , where  $\alpha$  and  $\beta$  are set to 1, and the  $x_i$  are  $U[0, 10]$ . In order to minimize the random variation associated with the simulations, the  $x_i$  are kept fixed in all replications. We also experimented with values for  $X$  that incorporated spatial autocorrelation, but no qualitative difference in the results was found.

Following the design in Kelejian and Prucha (1999), three different error structures are considered: a standard normal, a log-normal and a mixture of two normals. The two variates that make up the mixture are a standard normal and an i.i.d. normal with variance 100. As in Kelejian and Prucha (1999), the log-normal and mixture errors are transformed to obtain an overall variance under the null hypothesis (of  $\theta = 0$ ) that is the same as for the standard normal. For the lognormal, the error variate is obtained as

$$\varepsilon_i = [\exp(v_i) - \exp(0.5)] / [\exp(2) - \exp(1)]^{0.5}, \quad (18)$$

where  $v_i$  is standard normal. The mixed errors are constructed as

$$\varepsilon_i = [\kappa_i v_i + (1 - \kappa_i) \zeta_i] / (5.95)^{0.5}, \quad (19)$$

with  $v_i$  as in (18) and the  $\zeta_i$  are independent normal with mean zero and variance 100. This structure induces “outliers” beyond what can be expected in a single normal distribution. The random variate  $\kappa_i$  is an i.i.d. Bernouilly random variable with  $\text{Prob}[\kappa_i = 1] = 0.95$ .

Under the alternative hypothesis, the spatial error components variance-covariance structure is obtained by applying a Choleski decomposition to the matrix  $\Omega_{SEC} = \sigma_\xi^2 [I + \theta WW']$  and transforming the original uncorrelated variates. Three different values for the parameter  $\theta$  are considered:  $\theta = 1, 4$  and  $8$ . As shown in Kelejian and Robinson (1995), the parameter  $\theta$  indicates the relative importance of the spillover variance component to the location-specific component. Clearly, the larger this variance, the more important the spatial error components. In order to facilitate the interpretation of the results, we set  $\sigma_\xi^2 = 1$ , such that in effect  $\theta = \sigma_\psi^2$ .<sup>15</sup> In Anselin (2001), results are

<sup>15</sup>Under the null hypothesis, the “fit” of the model is therefore roughly the same between the three error distributions. The actual fit will vary slightly due to the extra heterogeneity in the mixture distribution. Under the alternatives, an additional source of error is added to  $\sigma_\xi^2$ , thus decreasing the fit relative to the model under the null. However, since we are primarily interested in the *relative* performance of the tests, this will not affect our overall conclusions.

also reported for  $LM_{SEC}$  using values for  $\theta < 1$ . These results are qualitatively similar to what is found in the current paper, hence to keep the scope of the study within reason, such values were not considered here. Moreover, small values of  $\theta$  are not very meaningful in practice, since they induce minimal spatial covariance. For example, a typical case can be illustrated using the contiguity structure for the third and fourth observation from the often used Columbus data set in Anselin (1988).<sup>16</sup> The corresponding value of  $WW'_{34}$  is 0.125 (the two locations have neighbors 2 and 5 in common, each with weight 0.25). With  $\sigma_\xi^2 = 1$  and  $\theta = 0.01$  (the smallest value considered in Anselin 2001), this yields a spatial covariance between locations 3 and 4 of 0.00125, which is unlikely to affect OLS inference to any great extent.

## 5 Empirical Results

### 5.1 Size of the Tests

We first consider the empirical rejection frequencies under the null hypothesis, i.e., in the absence of spatial autocorrelation. Tables 2 to 4 provide an initial assessment for a nominal size of  $\alpha = 0.05$ .<sup>17</sup> The rejection frequencies correspond to the number of times the null hypothesis is rejected, using the 5% critical value of the asymptotic  $\chi^2_1$  distribution. These results are based on 10,000 replications, which yields two standard deviation intervals around the nominal rejection levels of 0.0466–0.0534.

For normal error terms, the idealized condition reflected in Table 2, all tests except  $KR_{GMM}$  perform “properly” in the largest sample sizes, achieving the correct size (allowing for the randomness of the experiment) for the regular lattices with  $n = 400$  and 1024, and for the largest irregular lattice, with  $n = 1023$ .  $KR_{GMM}$  is never within the acceptable range, and over-rejects the null for all spatial layouts. However, for the irregular lattices (with greater variance for the  $d_i$ ), the degree of over-rejection seems to decrease as the sample size grows (or the degree of sparseness increases). This may indicate that the examples considered are not sufficiently large for the asymptotic properties of the  $KR_{GMM}$  test to be reflected.

For the smaller data sets, the results are less encouraging and highlight the differences between regular and irregular lattices. For example,  $KR_W$  over-rejects in the regular lattices for  $n < 400$ , but is within the acceptable range for all irregular samples. In contrast,  $KR_{WW}$  under-rejects in all but the three most sparse settings. A similar pattern is found for  $LM_{SEC}$ . The two traditional tests show like results as found in the earlier studies by Anselin and Rey (1991) and Anselin and Florax (1995). Moran’s  $I$  is within the acceptable range for all but one case (regular  $n = 121$ ), and  $LM_{ERR}$  is within the range for all regular lattices, but under-rejects in all but the largest irregular situations.

A lognormal error distribution (Table 3) affects these properties significantly. Only

<sup>16</sup>The data used are not the ones reported in the book, but the ones available from <http://www.spacestat.com>. There are slight differences in the ordering of the observations between the two that may cause confusion.

<sup>17</sup>A complete set of results is available from the authors. To conserve space, only selected tables are included here.

in a few instances do the rejection levels fall within the proper range, but without a clear pattern.<sup>18</sup> Even the three tests that do not rely on normality ( $KR_W$ ,  $KR_{WW}$  and  $KR_{GMM}$ ) do not have the proper size for the sample sizes considered. Both  $KR_W$  and  $KR_{WW}$  tend to over-reject and  $KR_{GMM}$  shows rather erratic results, not suggesting any type of convergence with growing sample sizes. Again, in accordance with earlier findings, both Moran's  $I$  and  $LM_{ERR}$  seriously under-reject in this case.

A mixture error distribution (Table 4) similarly affects the size of the four tests against spatial error components. The Moran's  $I$  and  $LM_{ERR}$  are less influenced, rejecting within the acceptable range for the larger data sets, especially for the regular lattices.  $LM_{SEC}$  is clearly unreliable under the null in this case, yielding a high degree of over-rejection, particularly for the irregular lattices, and even in the large samples. Suprisingly, the three "robust" tests are similarly affected, with considerable over-rejection by  $KR_{GMM}$ , but some weak evidence of a convergence towards 0.05 for the two KR tests.

A more complete picture of the extent to which the test statistics follow the asymptotic distribution under the null in the simulated samples is provided by the recently introduced *P value plots* and *P value discrepancy plots* of Davidson and MacKinnon (1998). Figure 1 shows the empirical distribution function of the  $p$ -values for the all six tests plotted against their nominal sizes, for the smallest sample size ( $n = 46$ ) and under normality. In order to focus attention on the Type I errors typically used in practice, we limited the x-axis to  $p \leq 0.12$ .<sup>19</sup> One would want this plot to follow a 45 degree line, since the  $p$ -values are distributed uniformly between 0 and 1. Discrepancies from the 45 degree line suggest an empirical distribution that differs from the theoretical (asymptotic) one used to establish the critical values. Figure 1 illustrates the worst case for the normal error distribution, with the smallest sample and densest spatial weights. Clearly, Moran's  $I$  is considerably above the 45 degree line and  $KR_{WW}$  is considerably below. Neither of these test statistics seems to follow the "theoretical" null distribution well. This is not surprising for Moran's  $I$ , since previous work has shown that this statistic requires medium to large samples before it starts to approximate its asymptotic properties. The other statistics in Figure 1 are more or less parallel to the 45 degree line, although both  $KR_{GMM}$  and  $LM_{SEC}$  cross this line (going from over-rejection to under-rejection, and vice versa).<sup>20</sup>

We take a closer look at the differences between the theoretical distribution of  $p$ -values and its empirical counterpart for the four SEC test statistics in the P value discrepancy plots in Figures 2 to 5. The P value discrepancy plots highlight the differences between the empirical distribution function and the 45 degree line in the P value plot and are particularly useful to compare test statistics that perform similarly. The comparison is over the range of relevant critical values, rather than for a single value, as in Tables 2 to 4, although our primary interest centers on the behavior near the origin (for small  $p$ -values). We therefore truncated the graphs at  $p = 0.12$ .

Figures 2 and 3 compare the test statistics for the two largest sample sizes (irregular and regular) under normality. For all statistics, the discrepancy is smaller for the

<sup>18</sup>For example,  $LM_{SEC}$  is acceptable in the regular lattices for  $n = 81, 121$  and  $1024$ , but only for  $n = 46$  in the irregular lattice, where the rejection frequency increases with sample size.

<sup>19</sup>Graphs with the full range of  $p$ -values for the x-axis are available from the authors

<sup>20</sup> $LM_{SEC}$  crosses the line for values of  $p$  outside the range reported here.

regular lattice than for the irregular one, again illustrating the importance of spatial characteristics of the data beyond simple sample size. The  $KR_{GMM}$  test does not show any indication of a parallel pattern, although slightly less so for  $n = 1024$ . For  $LM_{SEC}$ , the situation is slightly better, although for the irregular case there is still considerable discrepancy. The two KR tests are more or less parallel to the horizontal line for this sample size. When the error term is not normal, the results are much less attractive. As illustrated in Figure 4 for the lognormal distribution in an irregular lattice with  $n = 413$ , both  $LM_{SEC}$  and  $KR_{GMM}$  show considerable discrepancies, with the latter yielding an erratic pattern. For the corresponding regular lattice (not shown here), the erratic pattern for  $KR_{GMM}$  remains, but  $LM_{SEC}$  has become more or less parallel. While the KR tests also show some discrepancy from the asymptotic distribution, this is much less the case than for the other statistics. Finally, in Figure 5, we illustrate how the strange pattern for  $KR_{GMM}$  persists in the mixed normal case as well, while the other statistics are more or less parallel to the horizontal line, with relatively minor discrepancies. Overall, the figures would suggest that the stated asymptotic distribution under the null is not achieved in the absence of normality, or for the sample sizes considered here. While this is somewhat disturbing, it is much less of a problem for the smallest  $p$ -values, which are of interest in applied work. Here, some size adjustment may be necessary to reflect the correct Type I error.

## 5.2 Power of the Tests

The power of the six test statistics is compared for a nominal size of 0.05 and for three values of the variance ratio  $\theta$ , as illustrated in Tables 5 to 7.<sup>21</sup> A number of interesting patterns can be distinguished. First, and somewhat surprisingly, the  $LM_{SEC}$  statistic dominates all others and achieves high power even in moderately sized samples, especially in the regular lattices. For example, it is the only statistic that rejects in more than 80% of the cases for all values of  $\theta$  when  $n \geq 400$ .<sup>22</sup> The closest competitor is  $KR_{WW}$ , which also achieves 90% rejection levels in moderately sized samples, for alternatives with  $\theta \geq 4$ . The surprising element is that this performance is virtually unaffected by the error distribution and  $LM_{SEC}$  retains its high power even in the lognormal and mixed normal cases.

A second finding is the weak power of the  $KR_{GMM}$  approach. It consistently rates lowest and never rejects much more than 25% of the cases for lognormal and mixed errors, even when  $\theta = 8$ . In the normal case, performance is slightly more acceptable, but only for large  $\theta$  and in the two largest irregular data sets. It is possible that much larger sample sizes may be required before the asymptotic properties of this test are obtained, (especially in the non-normal case), but the evidence in the samples considered here would suggest that this statistic is not appropriate for small to medium-sized data

<sup>21</sup>The rejection frequencies are based on 1,000 replications. For some cases, we experimented with increasing the number of replications to 10,000. However, this increased precision did not affect the qualitative results and relative ranking of the test statistics. In order to conserve computing time, we performed the simulations for 1,000 replications.

<sup>22</sup>For comparison purposes, it may be useful to repeat the results given in Anselin (2001) for  $n = 4000$  and values of  $\theta$  less than 1. With the values of  $\theta$  in parentheses, the rejection frequencies were: 0.0477 (0.01), 0.0739 (0.05), 0.0982 (0.10), 0.2372 (0.25), 0.5188 (0.50), and 0.7587 (0.75).

sets.

A third characteristic is the superiority of the  $KR_{WW}$  over the  $KR_W$  form of the statistic, across the board. This contrasts with what we found under the null, where  $KR_{WW}$  tended to under-reject, suggesting possibly the need for an adjustment to its nominal size. It indicates that, at least in terms of power, there is a positive payoff when the “proper” range of interaction for the alternative is taken into account.

Finally, the “classical” Moran’s  $I$  and  $LM_{ERR}$  statistics show some power against the spatial error components alternative, although this is rather weak and only appears for the very large irregular lattices and with  $\theta \geq 4$ . Their performance is very similar and slightly dominates that of  $KR_W$ .

We close with a brief investigation of the complex interrelationship between test power, the characteristics of the spatial layout and the properties of the error term by means of a simple response surface regression. In Table 8 we show the results of a regression of the rejection frequencies on the size of the sample ( $n$ ), the percent non-zero elements (DENSE) to capture the effect of sparseness of the layout, the value of  $\theta$ , indicator variables for the lognormal error distribution (LOG) and for the mixed normal error distribution (MIX), as well as an indicator for irregular lattices (IRREG).<sup>23</sup> The results are remarkably consistent across the six test statistics. As suggested by the interpretation of the Tables 6 and 7, there is no significant effect of the error distribution.<sup>24</sup> In terms of the relative values of the coefficients, it is interesting to note that the largest changes in rejection with  $\theta$  are obtained for  $LM_{SEC}$ , followed by  $KR_{WW}$ , but clearly dominating the others. Similarly, the relative effect of sparseness is greater for the  $LM_{SEC}$  than for  $KR_{WW}$ , with a considerable margin. This suggests that the denser the spatial weights matrix (higher % non-zero elements), the lower is the power of the tests, except for  $KR_{GMM}$ , where this variable is not significant. The roles are reversed in terms of the effect of sample size, where the largest coefficient is obtained by  $KR_{WW}$ . The role of lattice structure is interesting as well. All but  $LM_{SEC}$  obtain a strongly significant positive coefficient, pointing to higher power in irregular lattices, but for the latter the effect is only marginally significant. Overall, the results of the surface response regression confirm the superiority of the  $LM_{SEC}$  and  $KR_{WW}$  statistics in terms of the way in which the associated rejection frequencies change with the characteristics of interest.

## 6 Conclusion

In this paper, we reviewed and compared six test statistics against alternatives of the spatial error components form. This represents the first extensive study of this particular spatial process model. Overall, we found that the new  $LM_{SEC}$  statistic performed remarkably well, especially in terms of power and even in situations not covered by its normality assumption. Similarly, the  $KR_{WW}$  variant of the Kelejian-Robinson statistic, which was suggested in this paper to account for second order neighbors, performed

<sup>23</sup>The results are for an OLS regression with heteroskedastic-robust standard errors. The rejection frequencies were subject to a logistic transformation to ensure that predicted values were in the proper range.

<sup>24</sup>Confirming our earlier interpretation of the tables, there is only such an effect for  $KR_{GMM}$ .

well. Both these statistics suggest themselves as useful diagnostics for empirical practice. A less satisfactory finding was that the easy to implement  $KR_{GMM}$  procedure performed very poorly in the samples considered here, and may indeed require “very large” samples before it can be used reliably. The Moran’s  $I$  and  $LM_{ERR}$  statistics were shown to have some limited power against spatial error component alternatives, but not sufficiently to warrant their use instead of the more specialized diagnostics.

A few cautionary remarks are in order as well, however. The results under the null hypothesis illustrate how the asymptotic critical values may be a poor guide in empirical practice when no spatial error components problem is present. In particular the  $KR_{GMM}$  test likely requires very large samples (larger than the ones considered here) before it approaches its theoretical size, while the  $LM_{SEC}$  performs very poorly for lognormal errors (and to a lesser degree for mixed normal errors). Of the six test statistics considered, the traditional Moran’s  $I$  and  $LM_{ERR}$  remain the most reliable when it comes to rejecting the null hypothesis when it should not be rejected. Interestingly, the two original KR tests do not show substantially greater robustness to non-normality than the other statistics.

Finally, the degree of generality of the results reported here is of course limited by the designs taken into account in the simulations. While other parameter combinations and more complex error structures could be considered, the selection used nevertheless provides a useful insight into the tradeoffs involved in the application of these diagnostics and into their relative power. In particular, it clearly reveals that test statistics with attractive theoretical but asymptotic properties do not necessarily perform acceptably in realistic samples. Theory alone is insufficient to guide us in this respect and much remains to be done to acquire further experimental and empirical evidence.

## References

- Anselin, L. (1988). *Spatial Econometrics: Methods and Models*. Kluwer Academic, Dordrecht.
- Anselin, L. (2001). Rao’s score test in spatial econometrics. *Journal of Statistical Planning and Inference*. in press.
- Anselin, L. and Bera, A. (1998). Spatial dependence in linear regression models with an introduction to spatial econometrics. In Ullah, A. and Giles, D. E., editors, *Handbook of Applied Economic Statistics*, pp. 237–289. Marcel Dekker, New York.
- Anselin, L. and Florax, R. J. (1995). Small sample properties of tests for spatial dependence in regression models: some further results. In Anselin, L. and Florax, R. J., editors, *New Directions in Spatial Econometrics*, pp. 21–74. Springer-Verlag, Berlin.
- Anselin, L. and Rey, S. J. (1991). Properties of tests for spatial dependence in linear regression models. *Geographical Analysis*, 23:112–131.
- Cliff, A. and Ord, J. K. (1981). *Spatial Processes: Models and Applications*. Pion, London.

- Davidson, R. and MacKinnon, J. G. (1998). Graphical methods for investigating the size and power of hypothesis tests. *The Manchester School*, 66:1–26.
- Kelejian, H. H. and Prucha, I. (1998). A generalized spatial two stage least squares procedures for estimating a spatial autoregressive model with autoregressive disturbances. *Journal of Real Estate Finance and Economics*, 17:99–121.
- Kelejian, H. H. and Prucha, I. (1999). A generalized moments estimator for the autoregressive parameter in a spatial model. *International Economic Review*, 40:509–533.
- Kelejian, H. H. and Robinson, D. P. (1992). Spatial autocorrelation: A new computationally simple test with an application to per capita country police expenditures. *Regional Science and Urban Economics*, 22:317–333.
- Kelejian, H. H. and Robinson, D. P. (1993). A suggested method of estimation for spatial interdependent models with autocorrelated errors, and an application to a county expenditure model. *Papers in Regional Science*, 72:297–312.
- Kelejian, H. H. and Robinson, D. P. (1995). Spatial correlation: A suggested alternative to the autoregressive model. In Anselin, L. and Florax, R. J., editors, *New Directions in Spatial Econometrics*, pp. 75–95. Springer-Verlag, Berlin.
- Kelejian, H. H. and Robinson, D. P. (1997). Infrastructure productivity estimation and its underlying econometric specifications: A sensitivity analysis. *Papers in Regional Science*, 76:115–131.
- Kelejian, H. H. and Robinson, D. P. (1998). A suggested test for spatial autocorrelation and/or heteroskedasticity and corresponding Monte Carlo results. *Regional Science and Urban Economics*, 28:389–417.

Table 1: Connectedness Characteristics of Spatial Weights

Regular Lattices			Irregular Lattices		
$n$	Avg. Links <sup>a</sup>	Sparseness <sup>b</sup>	$n$	Avg. Links <sup>a</sup>	Sparseness <sup>b</sup>
49	3.43	7.14	46	5.00	11.11
81	3.56	4.44	80	5.08	6.42
121	3.64	3.03	124	5.29	4.30
256	3.75	1.47	264	5.53	2.10
400	3.80	0.95	413	5.56	1.35
1024	3.88	0.38	1023	5.52	0.54

<sup>a</sup> Average number of neighbors for each location.<sup>b</sup> Percent non-zero links.Table 2: Empirical Rejection Frequencies under  $H_0$ , Normal<sup>a</sup>

$n$	LM <sub>SEC</sub>	KR <sub>W</sub>	KR <sub>WW</sub>	KR <sub>GMM</sub>	$I$	LM <sub>ERR</sub>
49	0.0366	0.0685	0.0390	0.0613	0.0504	0.0494
81	0.0425	0.0595	0.0407	0.0662	0.0521	0.0508
121	0.0413	0.0578	0.0391	0.0614	0.0540	0.0515
256	0.0444	0.0536	0.0445	0.0646	0.0523	0.0508
400	0.0499	0.0504	0.0489	0.0604	0.0484	0.0472
1024	0.0484	0.0509	0.0479	0.0613	0.0492	0.0492
46	0.0247	0.0482	0.0254	0.0701	0.0498	0.0382
80	0.0317	0.0484	0.0339	0.0622	0.0485	0.0460
124	0.0367	0.0469	0.0341	0.0607	0.0484	0.0438
264	0.0466	0.0476	0.0399	0.0619	0.0511	0.0488
413	0.0452	0.0465	0.0446	0.0602	0.0470	0.0460
1023	0.0480	0.0496	0.0506	0.0553	0.0487	0.0486

<sup>a</sup>  $\alpha = 0.05$ ; 10,000 replications



Table 3: Empirical Rejection Frequencies under  $H_0$ , Lognormal<sup>a</sup>

$n$	$LM_{SEC}$	$KR_W$	$KR_{WW}$	$KR_{GMM}$	$I$	$LM_{ERR}$
49	0.0441	0.0892	0.0784	0.0788	0.0348	0.0287
81	0.0472	0.0823	0.0763	0.0659	0.0363	0.0313
121	0.0502	0.0834	0.0721	0.0570	0.0365	0.0336
256	0.0592	0.0767	0.0686	0.0606	0.0402	0.0373
400	0.0610	0.0704	0.0625	0.0660	0.0399	0.0383
1024	0.0472	0.0847	0.0517	0.0812	0.0439	0.0425
46	0.0477	0.0727	0.0620	0.0900	0.0321	0.0238
80	0.0718	0.0817	0.0642	0.0700	0.0350	0.0278
124	0.0923	0.0744	0.0612	0.0600	0.0349	0.0281
264	0.1361	0.0734	0.0707	0.0674	0.0429	0.0430
413	0.1852	0.0689	0.0632	0.0550	0.0402	0.0381
1023	0.1561	0.0731	0.0572	0.0454	0.0412	0.0414

<sup>a</sup>  $\alpha = 0.05$ ; 10,000 replicationsTable 4: Empirical Rejection Frequencies under  $H_0$ , Mixture<sup>a</sup>

$n$	$LM_{SEC}$	$KR_W$	$KR_{WW}$	$KR_{GMM}$	$I$	$LM_{ERR}$
49	0.0538	0.0737	0.0649	0.0826	0.0275	0.0264
81	0.0626	0.0628	0.0625	0.0631	0.0340	0.0327
121	0.0711	0.0579	0.0602	0.0569	0.0389	0.0376
256	0.0769	0.0575	0.0595	0.0652	0.0484	0.0476
400	0.0761	0.0512	0.0518	0.0636	0.0498	0.0486
1024	0.0721	0.0541	0.0524	0.0701	0.0514	0.0507
46	0.0614	0.0694	0.0604	0.0906	0.0273	0.0228
80	0.0989	0.0614	0.0574	0.0684	0.0314	0.0284
124	0.1308	0.0601	0.0534	0.0636	0.0351	0.0336
264	0.1615	0.0554	0.0571	0.0641	0.0434	0.0423
413	0.2210	0.0516	0.0526	0.0600	0.0443	0.0445
1023	0.1873	0.0394	0.0432	0.0685	0.0494	0.0499

<sup>a</sup>  $\alpha = 0.05$ ; 10,000 replications

Table 5: Empirical Rejection Frequencies under  $H_1$ , Normal Errors <sup>a</sup>

$n$	$\theta$	$LM_{SEC}$	$KR_W$	$KR_{WW}$	$KR_{GMM}$	$I$	$LM_{ERR}$
49	1	0.172	0.134	0.093	0.107	0.110	0.107
	4	0.599	0.236	0.256	0.171	0.225	0.210
	8	0.828	0.298	0.407	0.241	0.261	0.262
81	1	0.265	0.123	0.135	0.108	0.127	0.119
	4	0.825	0.233	0.421	0.203	0.245	0.240
	8	0.967	0.291	0.607	0.242	0.293	0.289
121	1	0.373	0.119	0.168	0.108	0.125	0.129
	4	0.963	0.255	0.632	0.243	0.222	0.231
	8	0.995	0.312	0.817	0.280	0.303	0.301
256	1	0.663	0.136	0.324	0.108	0.127	0.127
	4	1.000	0.238	0.934	0.280	0.244	0.238
	8	1.000	0.329	0.991	0.364	0.318	0.318
400	1	0.822	0.108	0.504	0.135	0.105	0.107
	4	1.000	0.267	0.990	0.305	0.254	0.250
	8	1.000	0.329	0.997	0.414	0.312	0.306
1024	1	0.996	0.121	0.877	0.158	0.148	0.145
	4	1.000	0.265	1.000	0.425	0.256	0.252
	8	1.000	0.320	1.000	0.518	0.325	0.327
46	1	0.141	0.095	0.076	0.131	0.116	0.090
	4	0.448	0.249	0.287	0.278	0.287	0.240
	8	0.608	0.353	0.433	0.367	0.421	0.360
80	1	0.195	0.114	0.119	0.140	0.142	0.104
	4	0.664	0.377	0.483	0.322	0.429	0.377
	8	0.878	0.531	0.701	0.488	0.603	0.546
124	1	0.280	0.174	0.177	0.141	0.200	0.162
	4	0.849	0.482	0.656	0.352	0.604	0.549
	8	0.972	0.699	0.856	0.506	0.775	0.734
264	1	0.526	0.223	0.337	0.190	0.307	0.268
	4	0.992	0.765	0.922	0.519	0.842	0.826
	8	1.000	0.925	0.993	0.742	0.950	0.946
413	1	0.762	0.290	0.486	0.373	0.381	0.354
	4	1.000	0.884	0.993	0.790	0.914	0.905
	8	1.000	0.988	1.000	0.906	0.990	0.987
1023	1	0.976	0.523	0.842	0.335	0.657	0.644
	4	1.000	0.998	1.000	0.848	0.999	0.999
	8	1.000	1.000	1.000	0.956	1.000	1.000

<sup>a</sup>  $\alpha = 0.05$ ; 1,000 replications

Table 6: Empirical Rejection Frequencies under  $H_1$ , Lognormal Errors <sup>a</sup>

$n$	$\theta$	$LM_{SEC}$	$KR_W$	$KR_{WW}$	$KR_{GMM}$	$I$	$LM_{ERR}$
49	1	0.164	0.138	0.086	0.077	0.075	0.076
	4	0.563	0.248	0.243	0.104	0.189	0.190
	8	0.799	0.333	0.362	0.160	0.230	0.224
81	1	0.248	0.157	0.092	0.078	0.090	0.085
	4	0.821	0.257	0.356	0.092	0.203	0.202
	8	0.968	0.318	0.569	0.129	0.280	0.282
121	1	0.324	0.143	0.162	0.063	0.090	0.094
	4	0.969	0.277	0.575	0.094	0.211	0.210
	8	0.998	0.329	0.782	0.112	0.249	0.256
256	1	0.620	0.159	0.297	0.069	0.103	0.114
	4	0.999	0.266	0.905	0.114	0.214	0.212
	8	1.000	0.361	0.976	0.142	0.294	0.294
400	1	0.835	0.135	0.465	0.075	0.107	0.112
	4	1.000	0.268	0.980	0.132	0.214	0.211
	8	1.000	0.341	0.998	0.173	0.292	0.291
1024	1	0.998	0.152	0.909	0.095	0.117	0.114
	4	1.000	0.260	1.000	0.161	0.228	0.230
	8	1.000	0.350	1.000	0.172	0.307	0.306
46	1	0.129	0.105	0.098	0.104	0.084	0.061
	4	0.418	0.262	0.296	0.156	0.268	0.218
	8	0.553	0.364	0.414	0.203	0.416	0.357
80	1	0.196	0.135	0.137	0.079	0.130	0.092
	4	0.621	0.352	0.496	0.136	0.403	0.345
	8	0.806	0.506	0.666	0.189	0.542	0.483
124	1	0.285	0.142	0.170	0.078	0.174	0.138
	4	0.796	0.498	0.670	0.131	0.570	0.520
	8	0.914	0.663	0.841	0.188	0.718	0.667
264	1	0.497	0.196	0.365	0.091	0.282	0.258
	4	0.950	0.767	0.918	0.140	0.812	0.786
	8	0.994	0.897	0.979	0.218	0.935	0.924
413	1	0.615	0.268	0.543	0.088	0.362	0.331
	4	0.988	0.895	0.993	0.192	0.910	0.905
	8	1.000	0.973	0.996	0.270	0.973	0.971
1023	1	0.915	0.554	0.902	0.091	0.627	0.604
	4	1.000	0.998	1.000	0.197	0.999	0.997
	8	1.000	1.000	1.000	0.275	1.000	1.000

<sup>a</sup>  $\alpha = 0.05$ ; 1,000 replications

Table 7: Empirical Rejection Frequencies under  $H_1$ , Mixed Normal Errors <sup>a</sup>

$n$	$\theta$	$LM_{SEC}$	$KR_W$	$KR_{WW}$	$KR_{GMM}$	$I$	$LM_{ERR}$
49	1	0.152	0.134	0.109	0.074	0.087	0.077
	4	0.575	0.248	0.245	0.131	0.137	0.143
	8	0.798	0.298	0.321	0.169	0.197	0.192
81	1	0.269	0.115	0.157	0.070	0.077	0.076
	4	0.814	0.235	0.418	0.082	0.156	0.157
	8	0.956	0.277	0.525	0.099	0.199	0.202
121	1	0.352	0.139	0.209	0.064	0.078	0.078
	4	0.934	0.250	0.566	0.088	0.176	0.175
	8	0.988	0.313	0.748	0.098	0.225	0.223
256	1	0.625	0.125	0.396	0.067	0.102	0.098
	4	0.994	0.262	0.858	0.117	0.191	0.194
	8	1.000	0.331	0.952	0.106	0.231	0.235
400	1	0.817	0.113	0.542	0.086	0.105	0.102
	4	0.999	0.269	0.954	0.112	0.229	0.224
	8	1.000	0.360	0.985	0.142	0.282	0.273
1024	1	0.995	0.134	0.867	0.083	0.105	0.104
	4	1.000	0.259	1.000	0.137	0.218	0.220
	8	1.000	0.350	1.000	0.173	0.295	0.294
46	1	0.147	0.136	0.126	0.109	0.086	0.057
	4	0.391	0.295	0.330	0.165	0.283	0.218
	8	0.551	0.393	0.453	0.181	0.397	0.345
80	1	0.229	0.144	0.193	0.097	0.132	0.090
	4	0.580	0.391	0.507	0.108	0.404	0.353
	8	0.744	0.515	0.640	0.163	0.543	0.487
124	1	0.298	0.187	0.254	0.081	0.177	0.144
	4	0.759	0.559	0.689	0.128	0.592	0.542
	8	0.896	0.676	0.802	0.156	0.715	0.682
264	1	0.479	0.265	0.410	0.092	0.266	0.234
	4	0.940	0.765	0.889	0.128	0.817	0.800
	8	0.988	0.905	0.966	0.160	0.932	0.927
413	1	0.598	0.383	0.571	0.105	0.365	0.338
	4	0.986	0.878	0.970	0.151	0.899	0.888
	8	1.000	0.946	0.998	0.211	0.960	0.956
1023	1	0.890	0.574	0.838	0.099	0.668	0.645
	4	1.000	0.979	1.000	0.177	0.994	0.994
	8	1.000	0.998	1.000	0.253	1.000	1.000

<sup>a</sup>  $\alpha = 0.05$ ; 1,000 replications

Table 8: Response Surface Regressions <sup>a</sup>

VARIABLE <sup>b</sup>	LM <sub>SEC</sub>	KR <sub>W</sub>	KR <sub>WW</sub>	KR <sub>GMM</sub>	I	LM <sub>ERR</sub>
CONSTANT	1.329 (0.506)	-2.650 (0.047)	-1.502 (0.324)	-1.812 (0.164)	-2.794 (0.437)	-2.768 (0.442)
<i>n</i>	0.009 0.00337 (0.00054)	0.000 0.00165 (0.00072)	0.000 0.00498 (0.00060)	0.000 0.00069 (0.00025)	0.000 0.00184 (0.00074)	0.000 0.00182 (0.00074)
DENSE	0.000 -0.452 (0.063)	0.022 -0.165 (0.039)	0.000 -0.284 (0.048)	0.005 -0.022 (0.021)	0.013 -0.190 (0.039)	0.014 -0.206 (0.039)
$\theta$	0.000 0.628 (0.041)	0.000 0.319 (0.043)	0.000 0.529 (0.038)	0.310 0.159 (0.018)	0.000 0.330 (0.042)	0.000 0.335 (0.043)
LOG	0.000 -0.355 (0.329)	0.000 0.029 (0.3113)	0.000 -0.155 (0.277)	0.000 -1.269 (0.137)	0.000 -0.192 (0.301)	0.000 -0.207 (0.302)
MIX	0.281 -0.581 (0.315)	0.926 -0.079 (0.274)	0.576 -0.247 (0.275)	0.000 -1.349 (0.141)	0.524 -0.312 (0.287)	0.493 -0.313 (0.294)
IRREG	0.065 -0.516 (0.253)	0.772 2.171 (0.251)	0.368 0.631 (0.237)	0.000 0.639 (0.109)	0.276 2.724 (0.254)	0.287 2.557 (0.257)
<i>R</i> <sup>2</sup>	0.041	0.000	0.008	0.000	0.000	0.000
<i>R</i> <sup>2</sup>	0.844	0.667	0.867	0.755	0.722	0.712

<sup>a</sup>  $\alpha = 0.05$ ; 1,000 replications; OLS regression with heteroskedasticity-robust standard errors.<sup>b</sup> Estimated standard errors in parentheses; probability of null coefficient.

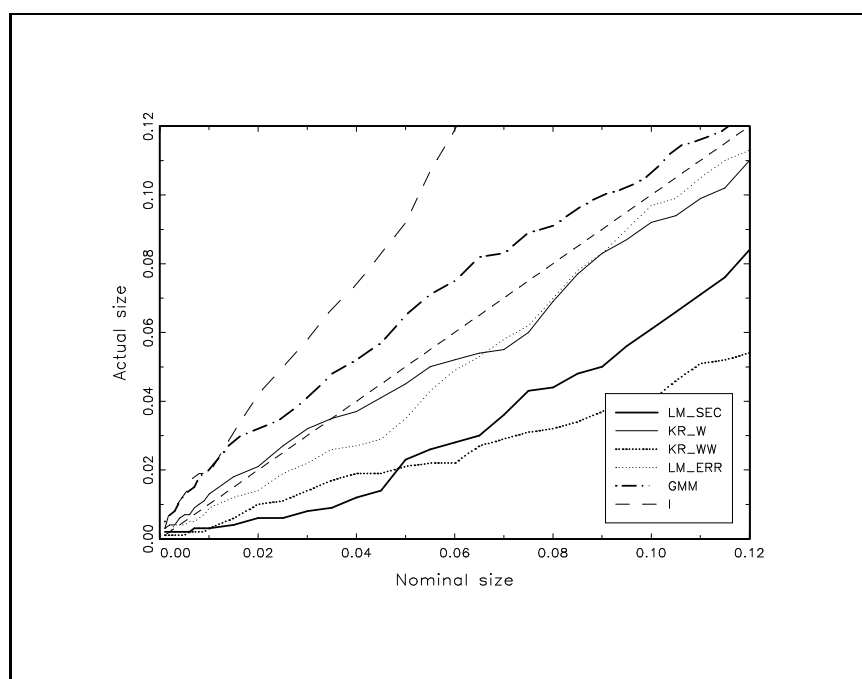
Figure 1: P Value Plot, Normal,  $n = 46$ 

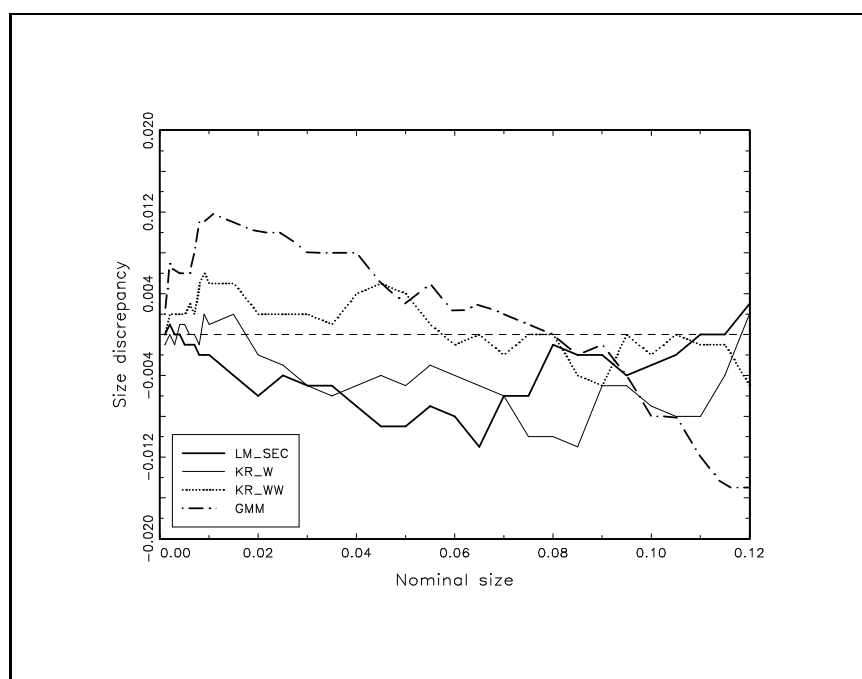
Figure 2: P Value Discrepancy Plot, Normal,  $n = 1023$ 

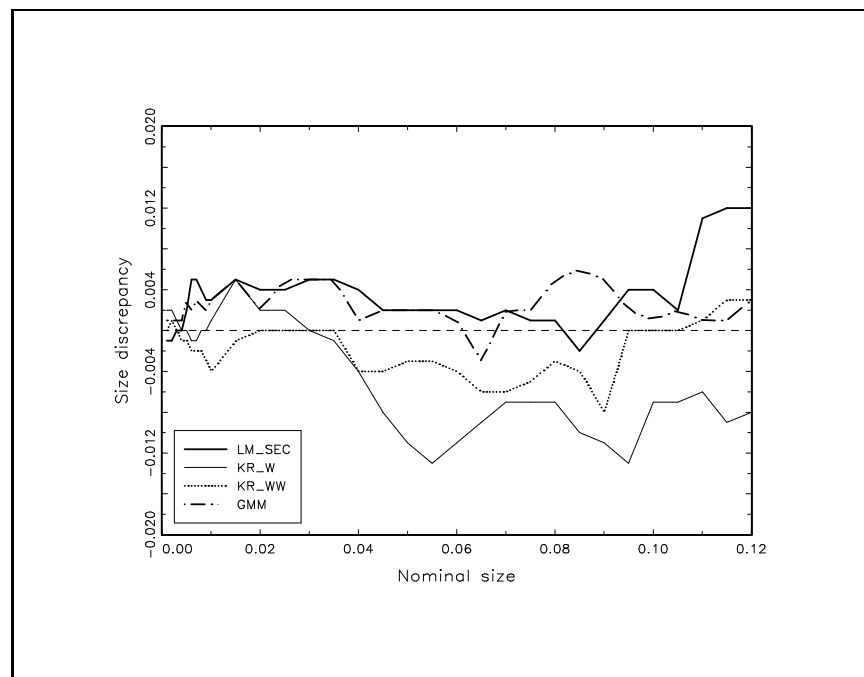
Figure 3: P Value Discrepancy Plot, Normal,  $n = 1024$ 



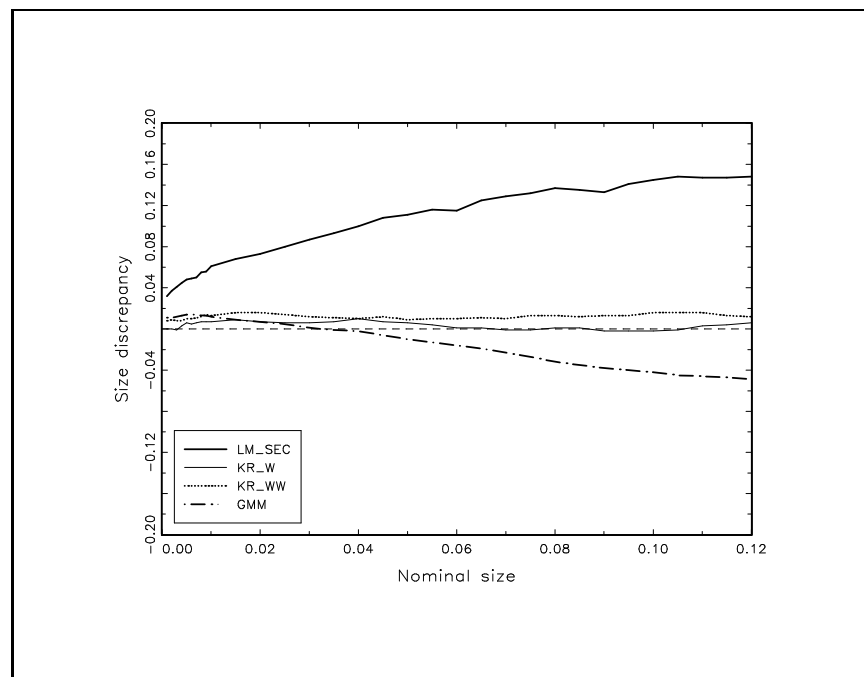
Figure 4: P Value Discrepancy Plot, Lognormal,  $n = 413$ 

Figure 5: P Value Discrepancy Plot, Mixed Normal,  $n = 81$ 