Spatial Economic Analysis

ISSN: (Print) (Online) Journal homepage: www.tandfonline.com/journals/rsea20

To use, or not to use the spatial Durbin model? – that is the question

Malabika Koley & Anil K. Bera

To cite this article: Malabika Koley & Anil K. Bera (2024) To use, or not to use the spatial Durbin model? – that is the question, Spatial Economic Analysis, 19:1, 30-56, DOI: 10.1080/17421772.2023.2256810

To link to this article: https://doi.org/10.1080/17421772.2023.2256810

SPATIAL ECONOMIC ANALYSIS
2024, VOL. 19, NO. 1, 30–56
https://doi.org/10.1080/17421772.2023.2256810

Routledge
Taylor & Francis Group

Regional Studies Association

Check for updates

To use, or not to use the spatial Durbin model? – that is the question

Malabika Koley [a] and Anil K. Bera [a]

ABSTRACT

The spatial Durbin model (SDM) is one of the most widely used models in spatial econometrics. It originated as a generalisation of the spatial error model (SEM) under a *non-linear* parametric restriction (see Anselin (1988, pp. 110–111)). This restriction should be tested to select an appropriate model between SDM and SEM. Perhaps, due to the complexity of executing a test for a *non-linear* hypothesis, this restriction is rarely tested in practice, though see Burridge (1981), Mur and Angulo (2006) and LeSage and Pace (2009, p. 164). This paper considers an alternative *linear* hypothesis to test the suitability of the SDM. To achieve this, we first use Rao's score (RS) testing principle and then Bera and Yoon (1993)'s methodology to robustify the original RS tests. The robust tests that require only ordinary least squares (OLS) estimation are able to identify the specific source(s) of departure(s) from the baseline linear regression model. An extensive Monte Carlo study provides evidence that our suggested tests possess excellent finite sample properties, both in terms of size and power. Our empirical illustrations, with two real data sets, attest that the tests developed in this paper could be very useful in judging the suitability of the SDM for the spatial data in hand.

KEYWORDS

SDM, common factor restriction, specification testing, Rao's score (RS) tests, parametric misspecification, robust RS tests

JEL C12, C21

HISTORY Received 7 October 2022; in revised form 20 July 2023

1. INTRODUCTION

Spatial econometrics was developed as a confluence of econometrics and spatial analysis in the late 19^{th} century. It includes a variety of techniques of estimation and testing of economic models to take account of the cross-sectional dependence arising due to the unique features of spatial data. These techniques are increasingly being used in an array of different fields of economics such as real estate and urban economics, economics of education, social networks, and international trade and finance see, for instance, Villar (1999), Anselin et al. (2000), Elhorst (2014) and Wang and Guan (2017). A basic spatial econometric model is the spatial autoregressive (SAR) model which was first proposed by Whittle (1954). The SAR model includes a (endogenous) spatially lagged dependent variable which captures the direct spatial dependence. For an early review of the genesis of spatial econometric models see Anselin and Bera (1998). For a recent review, see, LeSage and Pace (2009). A second type of spatial econometric model is the spatial error model (SEM) which takes account of the spatial dependence of the errors, i.e., when

CONTACT Malabika Koley mkoley2@illinois.edu
[a] Department of Economics, University of Illinois at Urbana Champaign, IL, USA
Supplemental data for this article can be accessed online at https://doi.org/10.1080/17421772.2023.2256810

the error of one spatial unit is dependent on the errors of the neighbouring units. This model, essentially, is the spatial counterpart of the time series serial correlation model. Despite the prevalent acclamation for SAR and SEM, the model that is most commonly used in empirical research is the spatial Durbin model (SDM). An online search reveals that there are a few hundreds of empirical papers published during the last five years that used SDM, see, for example, Autant-Bernard and LeSage (2011), Yang et al. (2015), Tientao et al. (2016), Feng and Chen (2018), Han et al. (2018), Sabater and Graham (2019), Sun et al. (2019), Li and Li (2020), Khezri et al. (2021) and Chen et al. (2022).

The SAR, SEM and SDM models are fundamentally different in their ways of taking account of spatial dependence through different channels. The SAR and the SDM models both allow for direct effects through their own exogenous regressors as well as indirect effects through the exogenous regressors of the neighbouring locations. On the contrary, there is no neighbourhood impact for the SEM model. The indirect spillover effect produced by the SDM model is both local and global with no prior restrictions on the magnitudes of these effects (see Elhorst (2010) for a thorough discussion on local and global spillover effects). Additionally, in the case of the SDM model, the exogenous regressor of one location casts an effect on the dependent variable of the neighbouring locations through *feedback loops*. The *feedback loops* are a result of the Durbin term, due to which a location, say i, affects a neighbouring location, say j, which eventually has a consequence on i; $i, j = 1, 2, \ldots, n$, where n is the sample size (see LeSage and Pace (2009, pp. 34–39) for a detailed discussion on impact effects). The relative advantages of using the SDM model over the contending models, such as SAR and SEM, are manifold. Note that the Durbin term constituting the SDM model is nothing but the lag of the exogenous regressors that captures the dependence due to the characteristics of the neighbouring units (locations). Thus, it carries a very intuitive physical interpretation based on the underlying economic theory. Apart from this, as noted by Elhorst (2010, p. 14), 'A concomitant advantage of the spatial Durbin model is that it produces correct standard errors or $t-$values of the coefficient estimates also if the true data generating process is a spatial error model'. Therefore, in terms of estimation, the additional Durbin term in the SDM model does not pose any further econometric complexity.

The SDM model reduces to SEM under a *non-linear* common factor restriction. Thus, theoretically speaking, the choice between SDM and SEM can be settled by performing a statistical test for the common factor restriction. Tests for this restriction have been developed in the literature by Burridge (1981) who suggested a likelihood ratio (LR) test for the common factor hypothesis. Much later, Mur and Angulo (2006) developed the Lagrange multiplier (LM) test and compared their performances with the LR test of Burridge. However, such tests are rarely implemented in practice, with few exceptions, such as Ertur and Koch (2007)[1] who tested the *non-linear* common factor restriction(s) using the LR test in their spatially augmented Solow growth model. Using the Penn World Tables (PWT, version 6.1) from Alan et al. (2002), they rejected the non-linear restrictions and favoured the SDM specification, concluding the presence of physical capital externalities. More recently, Juhl (2021)[2] did a comprehensive study on the use of the Wald test of common factors in spatial models. The main problem with the Wald test in testing *non-linear* hypothesis is that it is sensitive to algebraically equivalent alternative formulations of the null hypothesis. Juhl (2021, p. 193)'s suggestion is that the,

> 'researchers should either base inferences on bootstrap critical values for the Wald statistic or use the likelihood ratio test which is invariant to such reparameterizations when deciding on the model specification that adequately reflects the spatial process generating the data'.

Use of the likelihood ratio test (LRT) is not attractive for testing a *non-linear* hypothesis, since LRT requires a maximum likelihood estimation both under the null and the alternative hypotheses. A vast majority of empirical papers such as Xu and Wang (2017), Feng and Chen (2018) and Li and Li (2020) among many others used a SDM model but without any formal test for the common factor hypothesis. We first consider a *linear* hypothesis to test the suitability of the SDM and then use Rao's score (RS) test principle, which is asymptotically equivalent to the LRT but computationally much simpler as we will discuss in detail later.

Elhorst (2010, p. 10) supported LeSage and Pace (2009, p. 156)'s assertion that, 'There is too much emphasis in the spatial econometrics literature on use of statistical testing procedures to infer the appropriate model specification, and much of this literature ignores the SDM model'. The primary objective of this paper is to bring SDM to the forefront of spatial econometrics testing practice and establish the relevance of SDM in empirical research. We first start with the Durbin specification and test the significance of the spatially lagged independent variable WX, say, where W is a spatial weight matrix and X is a set of purely exogenous independent variables. The performance of this test may be adversely affected by the possible presence of a spatially lagged dependent variable Wy, where y is the dependent variable of interest. To deal with this, we use Bera and Yoon (1993)'s (BY, henceforth) testing principle which needs estimation only under the joint null hypothesis to identify the specific sources of departure(s) from the basic regression model. As detailed in subsequent sections, none of our tests require the estimation of the full SDM and are purely based on only the ordinary least squares (OLS) estimation. Thus, our suggested tests are easy to implement in practice. Another attractive feature of our tests, as we will explain later, is the additivity property that comes in handy as an automatic check for algebraic derivations and numerical calculations of our tests.

An alternative testing approach would be to use the conditional Rao's score (CRS) tests. For instance, testing the significance of the spatially lagged independent variable WX after estimating the coefficient of Wy. This, however, will require the maximum likelihood estimation (MLE) of the SAR model. Bera et al. (2020) demonstrated that the finite sample performance of the adjusted RS statistics are very similar to those of the CRS tests both in terms of size and power. Thus, there is practically no loss in terms of finite sample size and power in using the adjusted RS tests and at the same time we can avoid the MLE needed for the CRS tests. Therefore, in this paper, we solely concentrate on the adjusted RS tests.

The plan for the rest of the paper is as follows. Section 2 gives a brief background of SDM, illustrating how the model was adopted in the spatial literature as an extension of the time series Durbin model. The underlying assumptions for the model(s) are also presented in this section. Section 3 gives an account of the RS test, that we use as our fundamental testing principle, under a general set-up. Section 4 provides details of the tests constructed for specification of the model in this chapter. Section 5 demonstrates the finite sample performance of our asymptotic tests with the help of an elaborate Monte Carlo simulation study. Section 6 illustrates the usefulness and practical applicability of our tests using two real data sets. We find supportive evidence of our theory, both in simulations and empirical applications. Finally, Section 7 concludes the paper.

2. BACKGROUND AND MODEL

James Durbin (1923–2012) first suggested a completely novel method of estimation, currently known as the estimating function (EF) approach in his seminal paper Durbin (1960). As an application of the EF technique he considered the following standard linear regression with a stationary autoregressive (AR) process of the error:

$$\left.\begin{array}{rcl} y_t &=& x_t'\beta + u_t \\ u_t &=& \lambda u_{t-1} + \epsilon_t \end{array}\right\}, \tag{2.1}$$

where y_t is the observation of the dependent variable at time t, x_t is a vector of exogenous non-stochastic regressors of order $k \times 1$ and $\epsilon_t \overset{iid}{\sim} (0, \sigma^2)$ is the random error, $t = 1, 2, \ldots, n$.

Durbin (1960, Section 7) combined the two equations in (2.1) into the following reduced form:

$$\begin{aligned} y_t &= \lambda y_{t-1} + x_t'\beta - x_{t-1}'\beta\lambda + \epsilon_t \\ \Rightarrow \quad y_t &= \lambda y_{t-1} + x_t'\beta + x_{t-1}'\gamma + \epsilon_t. \end{aligned} \tag{2.2}$$

The specification in (2.2) can be viewed as an unrestricted model, which reduces to (2.1) under the restriction $\gamma = -\beta\lambda$. Hence one should test the *non-linear* hypothesis $H_0^T{:}\gamma + \beta\lambda = 0$ in (2.2), where 'T' signifies the time series context, in order to decide between the models (2.1) and (2.2). Rejection of H_0^T signifies appropriateness of the unrestricted model (2.2) whereas acceptances denotes the suitability of (2.1).

Anselin (1988, pp. 110–111) adopted Durbin's time series approach in the spatial context by starting with the following spatial error model (SEM):

$$\left.\begin{aligned} y &= \alpha_0 \iota_n + X\beta + u \\ u &= \lambda W u + \epsilon \end{aligned}\right\}, \tag{2.3}$$

where ι_n is a $n \times 1$ vector of ones, y is a $n \times 1$ vector of observations of the dependent variable, X represents a $n \times (k-1)$ matrix of non-stochastic observations on $(k-1)$ independent variables, $W \equiv ((w_{ij}))$ with w_{ij} representing the degree of potential interactions (geographic or social or economic, etc.) between the i^{th} and the j^{th} locations or spatial units, $i, j = 1, 2, \ldots, n$.[3] In addition to $w_{ii} = 0$ for $i = 1, 2, \ldots, n$, we consider row-normalised W, i.e., $\sum_{j=1}^n w_{ij} = 1$, for $i = 1, 2, \ldots, n$. Thus, we have $W\iota_n = \iota_n$. The $n \times 1$ vector of errors, u, is generated using a spatial autoregressive model and ϵ is an $n \times 1$ random error distributed as $\epsilon \sim N(0, \sigma^2 I_n)$.

Combining the specifications in (2.3) and using $W\iota_n = \iota_n$, we obtain the reduced form model as,

$$\begin{aligned} y &= \alpha_0\iota_n - \lambda\alpha_0\iota_n + \lambda Wy + X\beta - WX\beta\lambda + \epsilon \\ &= \alpha\iota_n + \lambda Wy + X\beta - WX\beta\lambda + \epsilon \\ &= \alpha\iota_n + \lambda Wy + X\beta + WX\gamma + \epsilon, \quad \text{say,} \end{aligned} \tag{2.4}$$

Where $\alpha = \alpha_0(1-\lambda)$ is the intercept term and $\gamma = -\beta\lambda$. In the spatial literature, model (2.4) is called the spatial Durbin model (SDM) and WX is known as the Durbin term (see Lee and Yu (2016)).

To provide intuitive justification of the Durbin model in (2.4), let us first note that one purpose of spatial modelling is to take account of externalities, for instance, the neighbourhood effects. Anselin (1988), Anselin and Bera (1998), LeSage and Pace (2009) and Elhorst (2014) provide excellent reviews of how spatial externalities are taken into account with the help of spatial econometric modelling. This can be accomplished in a number of ways. The SAR part λWy in (2.4) captures the effect on say, y_i of its *own* neighbourhood values, through the term $\sum_{j=1}^n w_{ij}y_j$, $i = 1, 2, \ldots, n$. On the other hand, the term $WX\gamma$ takes care of the externality effect of neighbours' exogenous variables on y_i through the term $\sum_{j=1}^n w_{ij}X_{jl}$, $l = 1, 2, \ldots, k-1$. When both the effects are present, considering only one of them will lead to the incorrect statistical inference, both in terms of estimation and testing. In this paper we deal with the testing problem by proposing robust tests for λ and γ that are immune to the possible presence of the other parameter.

Note that the common factor hypothesis suggests that SDM should be used only if the hypothesis $H_0^S{:}\gamma + \beta\lambda = 0$, where "$S$" refers to spatial, is rejected. On the other hand, if H_0^S is accepted, one should use the spatial error model (SEM). However, such a test is rarely executed in practice, perhaps due to the analytical complexity involved in dealing with the *non-linear* hypothesis H_0^S. The standard practice in a larger part of the empirical spatial econometrics literature is to estimate a SDM without questioning its suitability. In an attempt to bridge this gap, in this paper we suggest studying the bigger question of whether or not one should include the Durbin specification WX in the model without referring to the SEM structure. To do so, we alternatively propose to test the significance of WX using the hypothesis $H_0^D{:}\gamma = 0$, where "D" signifies Durbin. Thus, it is reasonable to consider testing $H_0^D{:}\gamma = 0$ instead of H_0^S.

The objective of this paper is to construct tests that are easy to implement in practice without requiring complex estimations of the unrestricted model. First, we construct the joint test, $RS_{\lambda\gamma}$, for the hypothesis $H_{01}{:}\lambda = 0,\ \gamma = 0$. Naturally, this test will be based on the OLS estimation of the model under the joint null $\lambda = 0,\ \gamma = 0$, i.e.,

$$y \;=\; \alpha\iota_n + X\beta + \epsilon. \tag{2.5}$$

We then develop a test for $H_0^D{:}\gamma = 0$, setting $\lambda = 0$, in (2.5). This test, denoted by RS_γ, as we will discuss later, will have excessive size when $\lambda \neq 0$. Therefore, next we employ BY methodology to test H_0^D in the possible presence of λ, and develop a robust version of RS_γ, namely RS_γ^* that also requires only OLS estimation.

It is important to clarify that RS_γ^*, which will be robust in the local possible presence of λ, is not same as the common-factor tests developed in Burridge (1981) and Mur and Angulo (2006) but rather an alternative and simple way to test the validity of SDM. To illustrate the link between the two hypotheses H_0^D and H_0^S note that, rejection of both H_0^D and H_0^S indicate appropriateness of the specification as in (2.5). However, their acceptances have different implications: acceptance of H_0^S signifies the suitability of the SEM whereas acceptance of H_0^D denotes the suitability of the SAR model. In a similar way, we develop the adjusted and unadjusted tests given by RS_λ and RS_λ^*, respectively, for the hypothesis $H_0^\lambda{:}\lambda = 0$, while assuming $\gamma = 0$ denoted by RS_λ. RS_λ^* will be robust to the local presence of γ.

In the next section we will describe how to construct *robust* test statistics and derive their asymptotic distributions under a *general* set-up. Descriptions of these tests are also available elsewhere, such as in Koley and Bera (2022), Bera, Doğan, and Taşpınar (2019) and Bera, Doğan, Taşpınar, and Leiluo (2019b). However, we include these here to make this paper self-dependent. Proofs of the asymptotic results are based on a set of assumptions following Lee (2004) and Jenish and Prucha (2009, 2012) that are listed below.

Assumption 1. All observations are located on an infinitely countable lattice $D \subset \mathbb{R}^{d_0}$, $d_0 \geq 1$ and are at least $\rho^* > 0$ distance away from each other. Thus, if ρ_{ij} be the distance between the i^{th} and the j^{th} locations or individuals $l(i),\ l(j),\ l{:}\{1,\ 2,\ \cdots\ n\} \to D_n \subset D$, then $\rho_{ij} \geq \rho^*$.

This assumption specifies an increasing domain asymptotic instead of infill asymptotics for the spatial near epoch dependence (NED) structure. It also ensures that the agents have some kind of distance, for instance, physical, economic, social or political, between them.

We impose the following normalisation assumption on the structure of the non-stochastic spatial weight matrix $W \equiv ((w_{ij})),\ i,\ j = 1,\ 2,\ 3,\ \ldots,\ n$.

Assumption 2. For $i,\ j = 1,\ 2,\ 3,\ \ldots,\ n,\ n \in \mathbb{N},\ w_{ii} = 0$ and $w_{ij} \geq 0,\ i \neq j$. Furthermore, $sup_n \|W\|_\infty = c_w < \infty$, where $\|.\|_\infty$ denotes maximum absolute row sum norm.

Assumption 3. The matrix of exogenous regressors X is a $n \times (k-1)$ matrix of distinct columns. All the elements of X are non-stochastic and $\lim\limits_{n\to\infty} \dfrac{1}{n} X'X$ exists and is non-singular.

Assumption 4. The vector of random errors $\epsilon \sim N(0,\ \sigma^2 I)$.

The above normality assumption is used to obtain the score functions and the information matrix based on the normal log-likelihood function. However, using the results in Fang et al. (2014), it can be shown that the tests derived in this paper are also asymptotically valid under non-normality.

3. ROBUST RAO SCORE (RS) TEST UNDER PARAMETRIC MISSPECIFICATION

Let $\theta = (\alpha_1', \alpha_2', \alpha_3')'$, $\theta \in \mathbb{R}^k$, $\alpha_1 \in \mathbb{R}^p$, $\alpha_2 \in \mathbb{R}^q$, $\alpha_3 \in \mathbb{R}^r$, $k = p + q + r$, be the set of parameters and $l(\alpha_1, \alpha_2, \alpha_3)$ denote the log-likelihood function. The parameter sets α_2 and α_3 can be taken as any combination of the testing parameters and α_1 is the set of nuisance parameters. Suppose a researcher wants to test $H_0^{\alpha_2}:\alpha_{20} = \alpha_2^*$ while assuming $H_0^{\alpha_3}:\alpha_{30} = \alpha_3^*$ holds true, using the log-likelihood function $l_1(\alpha_1, \alpha_2) = l(\alpha_1, \alpha_2, \alpha_{30})$, where α_2^* and α_3^* are known, typically taken as zeros. The maximum likelihood estimator (MLE) of θ under $\alpha_{20} = \alpha_2^*$ and $\alpha_{30} = \alpha_3^*$ is given by $\tilde{\theta} = (\tilde{\alpha}_1, \alpha_2^*, \alpha_3^*)$, obtained by constrained maximisation of the likelihood under $H_0^{\alpha_2}$ and $H_0^{\alpha_3}$. RS_{α_2} denotes the RS statistic for testing $H_0^{\alpha_2}:\alpha_{20} = \alpha_2^*$.

We now introduce some notations. Define $d_a(\theta) = \dfrac{\partial l(\theta)}{\partial a}$ and $J_{ab}(\theta) = E\left[-\dfrac{1}{n}\dfrac{\partial^2 l(\theta)}{\partial a \partial b'}\right]$, a and $b \in \theta$. For notational simplicity, denote $d_a(\theta) \equiv d_a$ and $J_{ab}(\theta) \equiv J_{ab}$ at the true value of θ, i.e., θ_o. The information matrix can be partitioned as,

$$J(\theta) = E\left[-\frac{1}{n}\frac{\partial^2 l(\theta)}{\partial \theta \partial \theta'}\right] = \begin{pmatrix} J_{\alpha_1\alpha_1}(\theta) & J_{\alpha_1\alpha_2}(\theta) & J_{\alpha_1\alpha_3}(\theta) \\ J_{\alpha_2\alpha_1}(\theta) & J_{\alpha_2\alpha_2}(\theta) & J_{\alpha_2\alpha_3}(\theta) \\ J_{\alpha_3\alpha_1}(\theta) & J_{\alpha_3\alpha_2}(\theta) & J_{\alpha_3\alpha_3}(\theta) \end{pmatrix}. \tag{3.1}$$

The RS statistic for testing $H_o^{\alpha_2}:\alpha_{20} = \alpha_2^*$ is given by,

$$RS_{\alpha_2} = \frac{1}{n} d_{\alpha_2}'(\tilde{\theta}) J_{\alpha_2\cdot\alpha_1}^{-1}(\tilde{\theta}) d_{\alpha_2}(\tilde{\theta}), \tag{3.2}$$

where $J_{\alpha_2\cdot\alpha_1}(\tilde{\theta}) = J_{\alpha_2\alpha_2}(\tilde{\theta}) - J_{\alpha_2\alpha_1}(\tilde{\theta}) J_{\alpha_1\alpha_1}^{-1}(\tilde{\theta}) J_{\alpha_1\alpha_2}(\tilde{\theta})$. If $l_1(\alpha_1, \alpha_2)$ denotes the true log-likelihood function, then, following Davidson and MacKinnon (1987), we have,

1. under $H_o^{\alpha_2}:\alpha_{20} = \alpha_2^*$,

$$RS_{\alpha_2} \xrightarrow{d} \chi_q^2,$$

a *central* chi-squared distribution with q degrees of freedom.

2. Under a sequence of local alternatives, $H_A^{\alpha_2}:\alpha_{20} = \alpha_2^* + \dfrac{\tau_2}{\sqrt{n}}$, τ_2 being a non-stochastic bounded vector,

$$RS_{\alpha_2} \xrightarrow{d} \chi_q^2(\nu_1), \tag{3.3}$$

where ν_1 is the *non-centrality* parameter given by $\nu_1 \equiv \nu_1(\tau_2) = \tau_2' J_{\alpha_2\cdot\alpha_1} \tau_2$.

RS_{α_2} is optimal and has correct size and power properties when $l_1(\alpha_1, \alpha_2)$ is the true log-likelihood function. However, RS_{α_2} is no longer optimal when the true likelihood is different to $l_1(\alpha_1, \alpha_2)$. Suppose the true log-likelihood be given by $l_2(\alpha_1, \alpha_3) = l(\alpha_1, \alpha_2^*, \alpha_3)$ so that the assumed log-likelihood $l_1(\alpha_1, \alpha_2)$ is misspecified. Following Davidson and MacKinnon (1987) and Saikkonen (1989), it can be shown that under $H_0^{\alpha_2}:\alpha_{20} = \alpha_2^*$ and $H_A^{\alpha_3}:\alpha_{30} = \alpha_3^* + \dfrac{\tau_3}{\sqrt{n}}$, where τ_3 is a non-stochastic bounded vector,

$$RS_{\alpha_2} \xrightarrow{d} \chi_q^2(\nu_2),$$

where $\nu_2 \equiv \nu_2(\tau_3) = \tau_3' J_{\alpha_3\alpha_2\cdot\alpha_1} J^{-1}_{\alpha_2\cdot\alpha_1} J_{\alpha_2\alpha_3\cdot\alpha_1} \tau_3$ with $J_{\alpha_3\alpha_2\cdot\alpha_1} = J_{\alpha_3\alpha_2} - J_{\alpha_3\alpha_1} J^{-1}_{\alpha_1\alpha_1} J_{\alpha_1\alpha_2}$ and $J_{\alpha_2\alpha_3\cdot\alpha_1} = J'_{\alpha_3\alpha_2\cdot\alpha_1}$. Thus, when the assumed model is misspecified, the asymptotic distribution of RS_{α_2} shifts from *central* to *non-central* chi-squared distribution resulting in the excessive size of the test. Note that the misspecification that we allow here is somewhat limited. The amount of misspecification is local in nature, namely, $\tau_3/\sqrt{n}$. Therefore, our adjusted tests are asymptotically robust for *local misspecification* only. However, as our Monte Carlo results will later demonstrate, the tests could be robust for *non-local misspecification* as well.

For testing $H_0^{\alpha_2}:\alpha_{20} = \alpha_2^*$, BY suggested the adjusted (net) score $d^*_{\alpha_2}$, instead of the raw score d_{α_2}, which is given by,

$$\begin{aligned} d^*_{\alpha_2}(\tilde{\theta}) &= d_{\alpha_2}(\tilde{\theta}) - E[d_{\alpha_2}|d_{\alpha_3}]_{(\tilde{\theta})} \\ &= d_{\alpha_2}(\tilde{\theta}) - J_{\alpha_2\alpha_3\cdot\alpha_1}(\tilde{\theta}) J^{-1}_{\alpha_3\cdot\alpha_1}(\tilde{\theta}) d_{\alpha_3}(\tilde{\theta}). \end{aligned} \tag{3.4}$$

At this stage, we note *two* things. *First*, from Equation (3.4), it is clear that the adjusted score $d^*_{\alpha_2}$ is a function of the raw score of the misspecified parameter α_3, i.e., d_{α_3} which in turn captures part of the information about α_3 contained in the data. Therefore, without doing a full-fledged MLE of α_3, as needed by the CRS test, we do take care of α_3, albeit indirectly. *Second*, the crux of the adjustment in (3.4) is the term,

$$J_{\alpha_2\alpha_3\cdot\alpha_1}(\tilde{\theta}) = J_{\alpha_2\alpha_3}(\tilde{\theta}) - J_{\alpha_2\alpha_1}(\tilde{\theta}) J^{-1}_{\alpha_1\alpha_1}(\tilde{\theta}) J_{\alpha_1\alpha_3}(\tilde{\theta}), \tag{3.5}$$

which can be interpreted as the partial covariance between d_{α_2} and d_{α_3} after eliminating the linear effect of d_{α_1}. $RS^*_{\alpha_2}$ is given by,

$$RS^*_{\alpha_2} = \frac{1}{n} d^{*'}_{\alpha_2}(\tilde{\theta}) [J_{\alpha_2\cdot\alpha_1}(\tilde{\theta}) - J_{\alpha_2\alpha_3\cdot\alpha_1}(\tilde{\theta}) J^{-1}_{\alpha_3\cdot\alpha_1}(\tilde{\theta}) J_{\alpha_3\alpha_2\cdot\alpha_1}(\tilde{\theta})]^{-1} d^*_{\alpha_2}(\tilde{\theta}). \tag{3.6}$$

The distribution of $RS^*_{\alpha_2}$ is unaffected by the local presence of α_3, such as $\alpha_{30} = \alpha_3^* + \frac{\tau_3}{\sqrt{n}}$. We have the following proposition.

Proposition 1.

- Under $H_A^{\alpha_2}:\alpha_{20} = \alpha_2^* + \frac{\tau_2}{\sqrt{n}}$, and $H_A^{\alpha_3}:\alpha_{30} = \alpha_3^* + \frac{\tau_3}{\sqrt{n}}$, it can be shown that

$$RS_{\alpha_2} \xrightarrow{d} \chi^2_q(\nu_3),$$

where $\nu_3 \equiv \nu_3(\tau_2, \tau_3) = \tau_2' J_{\alpha_2\cdot\alpha_1} \tau_2 + 2\tau_2' J_{\alpha_2\alpha_3\cdot\alpha_1} \tau_3 + \tau_3' J'_{\alpha_2\alpha_3\cdot\alpha_1} J^{-1}_{\alpha_2\cdot\alpha_1} J_{\alpha_2\alpha_3\cdot\alpha_1} \tau_3$ is the non-centrality parameter.

- Under $H_o^{\alpha_2}:\alpha_{20} = \alpha_2^*$ and irrespective of the value of α_3 we have,

$$RS^*_{\alpha_2} \xrightarrow{d} \chi^2_q.$$

- Under $H_A^{\alpha_2}:\alpha_{20} = \alpha_2^* + \tau_2/\sqrt{n}$ we have,

$$RS^*_{\alpha_2} \xrightarrow{d} \chi^2_q(\nu_4),$$

where $\nu_4 \equiv \nu_4(\tau_2) = \tau_2' (J_{\alpha_2\cdot\alpha_1} - J_{\alpha_2\alpha_3\cdot\alpha_1} J^{-1}_{\alpha_3\cdot\alpha_1} J_{\alpha_3\alpha_2\cdot\alpha_1}) \tau_2$ is the non-centrality parameter.

Proof, see the Appendices in the supplemental data online.

Since the distribution of $RS^*_{\alpha_2}$ is free of the local departure $\tau_3/\sqrt{n}$ of α_3 from its hypothesised value α_3^*, $RS^*_{\alpha_2}$ is said to be robust under local misspecification. In a similar way we can construct the adjusted test statistic $RS^*_{\alpha_3}$ from RS_{α_3} for testing $H_0^{\alpha_3}:\alpha_{30} = \alpha_3^*$ in the possible local presence of α_2, such as $\alpha_{20} = \alpha_2^* + \frac{\tau_2}{\sqrt{n}}$. The advantage of using RS^* over RS is that it has an asymptotic *central* chi-squared distribution under the null irrespective of whether there is misspecification in the model or not.

Our suggestion would be to start the testing process by carrying out the *joint* test of $H_0:\alpha_{20} = \alpha_2^*,\ \alpha_{30} = \alpha_3^*$. To construct the joint test let us first partition the information matrix in (3.1) as,

$$J(\theta) = \left(\begin{array}{c|cc} J_{\alpha_1\alpha_1}(\theta) & J_{\alpha_1\alpha_2}(\theta) & J_{\alpha_1\alpha_3}(\theta) \\ \hline J_{\alpha_2\alpha_1}(\theta) & J_{\alpha_2\alpha_2}(\theta) & J_{\alpha_2\alpha_3}(\theta) \\ J_{\alpha_3\alpha_1}(\theta) & J_{\alpha_3\alpha_2}(\theta) & J_{\alpha_3\alpha_3}(\theta) \end{array}\right) = \left(\begin{array}{c|c} J_{11}(\theta) & J_{12}(\theta) \\ \hline J_{12}(\theta) & J_{22}(\theta) \end{array}\right), \quad \text{say.}$$

Therefore,

$$\begin{aligned} J^{-1}(\theta) &= \begin{pmatrix} J_{11}(\theta) & J_{12}(\theta) \\ J_{21}(\theta) & J_{22}(\theta) \end{pmatrix}^{-1} \\ &= \begin{pmatrix} J^{11}(\theta) & J^{12}(\theta) \\ J^{21}(\theta) & J^{22}(\theta) \end{pmatrix}, \end{aligned} \tag{3.7}$$

where, $J^{11} = (J_{11}^{-1} + J_{11}^{-1}J_{12}(J_{22} - J_{21}J_{11}^{-1}J_{12})^{-1}J_{21}J_{11}^{-1})$, $J^{12} = -J_{11}^{-1}J_{12}(J_{22} - J_{21}J_{11}^{-1}J_{12})^{-1}$, $J^{21} = (J^{12})'$ and $J^{22} = (J_{22} - J_{21}J_{11}^{-1}J_{12})^{-1}$.

The joint test statistic is,

$$RS_{\alpha_2\alpha_3} = \frac{1}{n}\left[\, d'_{\alpha_2}(\tilde{\theta}) \quad d'_{\alpha_3}(\tilde{\theta}) \,\right] J^{22}(\tilde{\theta}) \left[\, d'_{\alpha_2}(\tilde{\theta}) \quad d'_{\alpha_3}(\tilde{\theta}) \,\right]'. \tag{3.8}$$

Proposition 2.

- Under $H_A^{\alpha_2}:\alpha_{20} = \alpha_2^* + \frac{\tau_2}{\sqrt{n}}$ and $H_A^{\alpha_3}:\alpha_{30} = \alpha_3^* + \frac{\tau_3}{\sqrt{n}}$

$$RS_{\alpha_2\alpha_3} \xrightarrow{d} \chi^2_{q+r}(\nu_5),$$

where $\nu_5 \equiv \nu_5(\tau_2,\ \tau_3) = \begin{pmatrix} \tau_2' & \tau_3' \end{pmatrix} \begin{pmatrix} J_{\alpha_2\cdot\alpha_1} & J_{\alpha_2\alpha_3\cdot\alpha_1} \\ \cdot & J_{\alpha_3\cdot\alpha_1} \end{pmatrix} \begin{pmatrix} \tau_2 \\ \tau_3 \end{pmatrix}$.

- Under $H_A^{\alpha_2}:\alpha_{20} = \alpha_2^* + \frac{\tau_2}{\sqrt{n}}$ and $H_0^{\alpha_3}:\alpha_{30} = \alpha_3^*$

$$RS_{\alpha_2\alpha_3} \xrightarrow{d} \chi^2_{q+r}(\nu_6),$$

where $\nu_6 \equiv \nu_6(\tau) = \tau_2' J_{\alpha_2\cdot\alpha_1}\tau_2$ which can be easily obtained by substituting $\tau_3 = 0$ in ν_5.

$RS_{\alpha_2\alpha_3}$ tests the *simultaneous presence* of the effects of α_2 and α_3 in the model. Thus, the possibility of parametric misspecification does not arise. The joint test is conclusive only when H_0 is accepted. In the event of rejection of H_0, our prescription would be to use $RS^*_{\alpha_2}$ and $RS^*_{\alpha_3}$ separately to identify the exact source of departure from $H_0:\alpha_{20} = \alpha_2^*,\ \alpha_{30} = \alpha_3^*$. There is absolutely no need for the use of the unadjusted tests RS_{α_2} and RS_{α_3}. Therefore, our complete specification search will consist of at most these three tests: $RS_{\alpha_2\alpha_3}$, $RS^*_{\alpha_2}$ and $RS^*_{\alpha_3}$.

An attractive additivity result holds connecting all the five test statistics mentioned so far (for proof, see for instance, Bera et al. (2009), Bera et al. (2019a)),

$$RS_{\alpha_2\alpha_3} = RS_{\alpha_2} + RS^*_{\alpha_3} = RS^*_{\alpha_2} + RS_{\alpha_3}. \tag{3.9}$$

As is well known, the joint test statistic $RS_{\alpha_2\alpha_3}$ and the two unadjusted ones RS_{α_2} and RS_{α_3} are, in general, available in the existing computer packages. Therefore, the additivity property provides an easy way to obtain the complete battery of tests needed for any practical application.

4. CONSTRUCTION OF THE TESTS

In this section we develop specification tests for the SDM as given in (2.4):

$$y = \alpha\iota_n + \lambda Wy + X\beta + WX\gamma + \epsilon.$$

For convenience, let us re-write the model as,

$$y = \lambda Wy + X_1\beta_1 + WX\gamma + \epsilon, \tag{4.1}$$

where $X_1 = \begin{bmatrix} \iota_n & X \end{bmatrix}$ is the augmented matrix of order $n \times k$ and $\beta_1' = (\alpha, \beta')'$ is of order $k \times 1$. The hypotheses of interest are

1. The joint test for spatial autocorrelation and Durbin effect,

$$H_{01}:\lambda = 0, \ \gamma = 0.$$

2. The test for spatial autocorrelation,

$$H_0^{\lambda}:\lambda = 0.$$

3. Test the significance of the lagged independent regressors or the Durbin effect,

$$H_0^{D}:\gamma = 0.$$

The parameter set in model (4.1) can be partitioned into subvectors such as $\theta = (\beta_1', \sigma^2, \lambda, \gamma)' \equiv (\psi', \lambda, \gamma)'$ where λ and γ are the testing parameters and $\psi = (\beta_1', \sigma^2)'$ is the set of nuisance parameters. Comparing with the general notations of RS tests discussed in Section 3, here we have $\theta = (\alpha_1', \alpha_2', \alpha_3')'$, where $\alpha_1' = \psi'$ with $(\alpha_1', \alpha_2')'$ as any combination of the testing parameters $(\lambda, \gamma')'$. As noted earlier, the advantage of using RS^* is that it takes account of the possible misspecification(s) while only requiring OLS estimation of the model in (4.1) under the joint null H_{01}, i.e.,

$$Y = X_1\beta_1 + \epsilon. \tag{4.2}$$

The restricted estimate of θ under H_{01} is $\tilde{\theta} = (\tilde{\beta}_1', \tilde{\sigma}^2, 0, 0)'$, where,

$$\begin{aligned} \tilde{\beta}_1 &= (X_1'X_1)^{-1}X_1'y \\ \tilde{\epsilon} &= y - X_1\tilde{\beta}_1 \quad \text{and} \\ \tilde{\sigma}^2 &= (\tilde{\epsilon}'\tilde{\epsilon})/n. \end{aligned}$$

For the construction of the RS tests, we first need to define the log-likelihood function. Under our normality assumption on ϵ, the log-likelihood function of (4.1) is given by,

$$l(\theta) \;=\; -\frac{n}{2}log(2\pi) - \frac{n}{2}log(\sigma^2) - \frac{1}{2\sigma^2}log(\epsilon'\epsilon) + log(|S(\lambda)|), \tag{4.3}$$

where $\epsilon = (y - X_1\beta_1)$ and $S(\lambda) = (I - \lambda W)$.

The score function $d_a(\theta) = \frac{\partial l(\theta)}{\partial \theta}$ and the information matrix are defined as, $J(\theta) = ((J_{ab}(\theta))) = \left(\left(E\left[-\frac{1}{n}\frac{\partial^2 l(\theta)}{\partial a \partial b'}\right]\right)\right)$ with $a,\, b \in\, \theta$ being evaluated at the restricted estimate $\tilde{\theta}$ for the construction of the tests. The analytical expressions of the scores are given below (detailed derivations are provided in the Appendices in the supplemental data online).

$$d_{\beta_1}(\tilde{\theta}) = \frac{X_1'\tilde{\epsilon}}{\tilde{\sigma}^2}, \quad d_{\sigma^2}(\tilde{\theta}) = \frac{n}{2\tilde{\sigma}^2} + \frac{\tilde{\epsilon}'\tilde{\epsilon}}{2\tilde{\sigma}^4}, \quad d_\lambda(\tilde{\theta}) = \frac{y'W'\tilde{\epsilon}}{\tilde{\sigma}^2}, \quad \text{and} \quad d_\gamma = \frac{X'W'\tilde{\epsilon}}{\tilde{\sigma}^2}. \tag{4.4}$$

The information matrix $J(\theta)$ for our model (4.1) is represented as,

$$J(\theta) = \begin{pmatrix} J_{\beta_1\beta_1}(\theta) & J_{\beta_1\sigma^2}(\theta) & J_{\beta_1\lambda}(\theta) & J_{\beta_1\gamma}(\theta) \\ J_{\sigma^2\beta_1}(\theta) & J_{\sigma^2\sigma^2}(\theta) & J_{\sigma^2\lambda}(\theta) & J_{\sigma^2\gamma}(\theta) \\ J_{\lambda\beta_1}(\theta) & J_{\lambda\sigma^2}(\theta) & J_{\lambda\lambda}(\theta) & J_{\lambda\gamma}(\theta) \\ J_{\gamma\beta_1}(\theta) & J_{\gamma\sigma^2}(\theta) & J_{\gamma\lambda}(\theta) & J_{\gamma\gamma}(\theta) \end{pmatrix},$$

and when evaluated at $\theta = \tilde{\theta}$, it is,

$$J(\tilde{\theta}) = \frac{1}{n}\begin{pmatrix} \frac{X_1'X_1}{\tilde{\sigma}^2} & 0 & \frac{X_1'WX_1\tilde{\beta}_1}{\tilde{\sigma}^2} & \frac{X_1'WX}{\tilde{\sigma}^2} \\ 0 & \frac{n}{2\tilde{\sigma}^4} & 0 & 0 \\ \frac{\tilde{\beta}_1'X_1'W'X_1}{\tilde{\sigma}^2} & 0 & \frac{\tilde{\beta}_1'X_1'W'WX_1\tilde{\beta}_1' + T\tilde{\sigma}^2}{\tilde{\sigma}^2} & \frac{\tilde{\beta}_1'X_1'W'WX}{\tilde{\sigma}^2} \\ \frac{X'W'X_1}{\tilde{\sigma}^2} & 0 & \frac{X'W'WX_1\tilde{\beta}_1}{\tilde{\sigma}^2} & \frac{X'W'WX}{\tilde{\sigma}^2} \end{pmatrix} \tag{4.5}$$

In the following subsections, the scores and the information matrix in (4.4) and (4.5), respectively, will be used as the main ingredients in constructing our tests.

4.1. Testing the joint hypothesis $H_{01}\!:\!\lambda = 0,\ \gamma = 0$

For the construction of the joint test let us partition the information matrix in (4.5) as,

$$\begin{aligned} J(\tilde{\theta}) \;&=\; \frac{1}{n}\left(\begin{array}{cc|cc} \frac{X_1'X_1}{\tilde{\sigma}^2} & 0 & \frac{X_1'WX_1\tilde{\beta}_1}{\tilde{\sigma}^2} & \frac{X_1'WX}{\tilde{\sigma}^2} \\ 0 & \frac{n}{2\tilde{\sigma}^4} & 0 & 0 \\ \hline \frac{\tilde{\beta}_1'X_1'W'X_1}{\tilde{\sigma}^2} & 0 & \frac{\tilde{\beta}_1'X_1'W'WX_1\tilde{\beta}_1' + T\tilde{\sigma}^2}{\tilde{\sigma}^2} & \frac{\tilde{\beta}_1'X_1'W'WX}{\tilde{\sigma}^2} \\ \frac{X'W'X_1}{\tilde{\sigma}^2} & 0 & \frac{X'W'WX_1\tilde{\beta}_1}{\tilde{\sigma}^2} & \frac{X'W'WX}{\tilde{\sigma}^2} \end{array}\right) \\ &=\; \left(\begin{array}{c|c} J_{11}(\tilde{\theta}) & J_{12}(\tilde{\theta}) \\ \hline J_{21}(\tilde{\theta}) & J_{22}(\tilde{\theta}) \end{array}\right), \qquad \text{say.} \end{aligned} \tag{4.6}$$

Thus, substituting the scores $d_\lambda(\tilde{\theta})$ and $d_\gamma(\tilde{\theta})$ given in (4.4) in the expression of the joint test statistic in (3.8), we get the joint test as,

$$RS_{\lambda\gamma} = \frac{1}{n}\left[\, d_\lambda(\tilde{\theta}) \quad d_\gamma'(\tilde{\theta}) \,\right] J^{22}(\tilde{\theta}) \left[\, d_\lambda(\tilde{\theta}) \quad d_\gamma'(\tilde{\theta}) \,\right]', \tag{4.7}$$

where $J^{22} = (J_{22} - J_{21}J_{11}^{-1}J_{12})^{-1}$ as in (3.7).

Corollary 1.

- Under $H_A^\lambda{:}\lambda = \frac{\eta}{\sqrt{n}}$ and $H_A^D{:}\gamma = \frac{\zeta}{\sqrt{n}}$, where ζ is a vector of length $(k-1)$,

$$RS_{\lambda\gamma} \xrightarrow{d} \chi_k^2(\nu_7),$$

where $\nu_7 \equiv \nu_7(\eta, \zeta) = \frac{1}{n}\left[P\eta^2 + \frac{(\eta\beta_1'X_1' + \zeta'X')Q_{X_1}(\eta X_1\beta_1 + X\zeta)}{\sigma^2}\right]$, with $P = [tr(W'W) + tr(W^2)]$ and $Q_{X_1} = W'M_{X_1}W$, where $M_{X_1} = [I - X_1(X_1'X_1)^{-1}X_1']$.

- Under $H_A^\lambda{:}\lambda = \frac{\eta}{\sqrt{n}}$ and $H_0^D{:}\gamma = 0$

$$RS_{\lambda\gamma} \xrightarrow{d} \chi_k^2(\nu_8),$$

where $\nu_8 \equiv \nu_8(\eta) = \frac{1}{n}\left[P\eta^2 + \frac{\eta^2\beta_1'X_1'Q_{X_1}X_1\beta_1'}{\sigma^2}\right]$.

Now we derive the adjusted tests for testing $H_0^\lambda{:}\lambda = 0$ and $H_0^D{:}\gamma = 0$ that are robust to local misspecification in the alternative model. The properties of these tests under different scenarios are presented as corollaries.

4.2. Testing $H_0^\lambda{:}\lambda = 0$ assuming $\gamma \neq 0$

We first want to test for the presence of spatial autocorrelation in the dependent variable in the possible presence of the Durbin effect, WX. Comparing with the general notation in Section 3, we have $\alpha_2 \equiv \lambda$, $\alpha_3 \equiv \gamma$ and the nuisance parameters $\alpha_1 \equiv \psi = (\beta_1', \sigma^2)'$. Using the expression of the test statistic given in (3.2), we get the *unadjusted RS* test for $H_0^\lambda{:}\lambda = 0$ as follows (detailed derivation given in the Appendices in the supplemental data online).

$$RS_\lambda = \frac{1}{\tilde{\sigma}^2}(y'W'\tilde{\epsilon})'[P\tilde{\sigma}^2 + \tilde{\beta}_1'X_1'Q_{X_1}X_1\tilde{\beta}_1]^{-1}(y'W'\tilde{\epsilon}). \tag{4.8}$$

Next, we formulate the adjusted test RS_λ^* which is robust to the local presence of the Durbin parameter γ. For the adjusted tests, the crux of the adjustment as noted in Equation (3.4) is given by,

$$J_{\lambda\gamma\cdot\psi} = J_{\lambda\gamma} - J_{\lambda\psi}J_{\psi\psi}^{-1}J_{\psi\gamma}.$$

Note that this term is nothing but the partial covariance between the scores d_λ and d_γ after eliminating the linear effect of the nuisance parameters ψ, i.e.,

$$J_{\lambda\gamma\cdot\psi} = Cov(d_\lambda, d_\gamma | d_\psi).$$

If $J_{\lambda\gamma\cdot\psi} = 0$, it indicates that there is no effect of λ and γ on one another, and thus no need for adjustment in the RS tests. However, for our model we have,

$$J_{\lambda\gamma\cdot\psi} = \frac{\beta_1' X_1' Q_{X_1} X}{n\sigma^2},$$

which is not equal to zero as long as there is at least one non-zero β_1 coefficient. Thus, adjustments are required to take account of the possible presence of the parameter that is not tested. For developing the adjusted test we first obtain the adjusted score $d_\lambda^* = [d_\lambda - E(d_\lambda|d_\gamma)]$ which, by construction, is orthogonal to d_γ. From (3.4), we can see that, here,

$$E[d_\lambda|d_\gamma] = J_{\lambda\gamma\cdot\psi} J_{\gamma\cdot\psi}^{-1} d_\gamma,$$

where,

$$J_{\gamma\cdot\psi} = \frac{X' Q_{X_1} X}{n\sigma^2}. \tag{4.9}$$

Using the expression of $RS_{\alpha_2}^*$ in (3.5), the adjusted test statistic RS_λ^* can be expressed as,

$$RS_\lambda^* = \frac{[y' W' \epsilon - \beta_1' X_1' Q_{X_1} X (X' Q_{X_1} X)^{-1} X' W' \epsilon]^2}{\sigma^2 [\beta_1' X_1' Q_{X_1} X_1 \beta_1 + P\sigma^2 - \beta_1' X_1' Q_{X_1} X (X' Q_{X_1} X)^{-1} X' Q_{X_1} X_1 \beta_1]}. \tag{4.10}$$

The asymptotic distributions of the tests RS_λ and RS_λ^* under different scenarios are given in the following corollary.

Corollary 2. Under the assumptions 1–4, we have the following two results.

- Under $H_A^\lambda : \lambda = \frac{\eta}{\sqrt{n}}$ and $H_A^D : \gamma = \frac{\zeta}{\sqrt{n}}$, τ_3 is a $(k-1) \times 1$ vector.

$$RS_\lambda \xrightarrow{d} \chi_1^2(\nu_9),$$

where $\nu_9 \equiv \nu_9(\eta, \zeta) = \eta_2 J_{\lambda\cdot\psi} + 2\eta J_{\lambda\gamma\cdot\psi}\zeta + \zeta_3' J_{\gamma\lambda\cdot\psi} J_{\lambda\cdot\psi}^{-1} J_{\lambda\gamma\cdot\psi}\zeta$, where

$$J_{\lambda\cdot\psi} = \frac{\beta_1' X_1' Q_{X_1} X_1 \beta_1 + P\sigma^2}{n\sigma^2}. \tag{4.11}$$

- Under $H_A^\lambda : \lambda = \frac{\eta}{\sqrt{n}}$ we obtain,

$$RS_\lambda^* \xrightarrow{d} \chi_1^2(\nu_{10}), \tag{4.12}$$

  where $\nu_{10} \equiv \nu_{10}(\eta) = \eta^2 [J_{\lambda\cdot\psi} - J_{\lambda\gamma\cdot\psi} J_{\gamma\cdot\psi}^{-1} J_{\gamma\lambda\cdot\psi}]$.

Proof, see Appendices.

4.3. Testing $H_0^D : \gamma = 0$ assuming $\lambda \neq 0$

We now construct the unadjusted and adjusted RS tests for the significance of the Durbin term. The unadjusted test is given by,

$$RS_\gamma = \frac{(X' W' \tilde{\epsilon})' (X' Q_{X_1} X)^{-1} (X' W' \tilde{\epsilon})}{\tilde{\sigma}^2}, \tag{4.13}$$

and the adjusted test, constructed employing the adjusted score $d_\gamma^* = [d_\gamma - E(d_\gamma | d_\lambda)]$, is given by,

$$\begin{aligned} RS_\gamma^* &= \frac{1}{\tilde{\sigma}^2}[X'\, W'\, \tilde{\epsilon} - X'\, Q_{X_1} X_1\, \tilde{\beta}_1 (n\tilde{\sigma}^2 J_{\lambda.\psi}(\tilde{\theta}))^{-1}\, y'\, W'\, \tilde{\epsilon}]' \\ &\quad [X'\, Q_{X_1} X - X'\, Q_{X_1} X_1 \tilde{\beta}_1 (n\tilde{\sigma}^2 J_{\lambda.\psi}(\tilde{\theta}))^{-1}\, \tilde{\beta}_1'\, X_1'\, Q_{X_1}\, X]^{-1} \\ &\quad [X'\, W'\, \tilde{\epsilon} - X'\, Q_{X_1}\, X_1\, \tilde{\beta}_1 (n\tilde{\sigma}^2 J_{\lambda.\psi}(\tilde{\theta}))^{-1}\, y'\, W'\, \tilde{\epsilon}]. \end{aligned} \tag{4.14}$$

The following corollary provides the asymptotic distributions of these tests.

Corollary 3. Under assumptions 1–5, we have.

- Under $H_A^D{:}\gamma = \frac{\zeta}{\sqrt{n}}$ and $H_A^\lambda{:}\lambda = \frac{\eta}{\sqrt{n}}$

$$RS_\gamma \xrightarrow{d} \chi^2_{k-1}(\nu_{11}),$$

where $\nu_{11} \equiv \nu_{11}(\eta, \zeta) = \zeta' J_{\gamma\cdot\psi}\zeta + 2\eta\zeta' J_{\gamma\lambda\cdot\psi} + \eta^2 J_{\lambda\gamma\cdot\psi} J^{-1}_{\gamma\cdot\psi} J_{\gamma\lambda\cdot\psi}$.

- Under $H_A^D{:}\gamma = \frac{\tau_3}{\sqrt{n}}$ we have,

$$RS_\gamma^* \xrightarrow{d} \chi_1^2(\nu_{12}), \tag{4.15}$$

where $\nu_{12} \equiv \nu_{12}(\zeta) = \zeta'(J_{\gamma\cdot\psi} - J_{\gamma\lambda\cdot\psi} J^{-1}_{\lambda\cdot\psi} J_{\lambda\gamma\cdot\psi})\zeta$, $J_{\gamma\cdot\psi}$, $J_{\lambda\cdot\psi}$ and $J_{\lambda\gamma\cdot\psi}$ are as defined earlier.

Proof, see Appendices.

Finally, the additivity result in (3.9) can be expressed in the context of testing SDM as,

$$RS_{\lambda\gamma} = RS_\lambda^* + RS_\gamma = RS_\lambda + RS_\gamma^*. \tag{4.16}$$

From the implementation point of view this result is very useful to ensure correct computation of the test statistics.

4.4. Recommendation for practitioners

In pursuit of the appropriate model for the data in hand, our suggestion would be to start with the unrestricted SDM given in (2.4) and test the joint hypothesis $H_{01}{:}\lambda = \gamma = 0$ using $RS_{\lambda\gamma}$ in (4.7). If H_{01} is accepted, it means that there is neither lag dependence nor Durbin effect in the data. Then, no further testing is required and one should settle with the linear regression model in (2.5). If H_{01} is rejected, to detect the exact sources(s) of departure from the null, one needs to test $H_0^\lambda{:}\lambda = 0$ and $H_0^D{:}\gamma = 0$, separately using our adjusted tests RS_λ^* in (4.10) and RS_γ^* in (4.14), respectively. Thus, our model specification search requires conducting, at most, three tests, namely, $RS_{\lambda\gamma}$, RS_λ^* and RS_γ^*, all requiring only OLS estimation. Additionally, a practitioner can check their computations of the test statistics by calculating the values of $RS_{\lambda\gamma}$, RS_λ, RS_λ^*, RS_γ and RS_γ^* separately and then empirically verifying the additivity result (4.16).

5. MONTE CARLO SIMULATION

In this section, we investigate the finite sample performance of our suggested tests through an extensive simulation study. For the simulation experiments we generate data from the following model (as in Equation (2.4)) where, for simplicity, we have only one independent variable X.

$$y = \lambda W y + \alpha \iota_n + X\beta + WX\gamma + \epsilon.$$

For the data under the joint null $H_{01}{:}\lambda = 0,\ \gamma = 0$, we set $\alpha = 1,\ \beta = 1.5$, $\epsilon \sim N(0, 2)$ and the regressor X is generated from $N(0, 4)$. The testing parameters λ and γ are taken from the sets $\{0.00, 0.01, \cdots, 0.60\}$ and $\{0.00, 0.01, \cdots, 0.35\}$, respectively. The parameter values are chosen in such a way that the empirical goodness-of-fit R^2 varies around 0.5. For a representation of the neighbourhood effects, we use the weight matrix for the crime data of 49 districts in Columbus, Ohio (Anselin (1988, p. 160)).

The results presented here are for sample size $n = 49$ with the corresponding weight matrix W_{49}. To verify our results for a larger sample $n = 98$, we repeated the experiments using a weight matrix, $W_{98} = (W_{49} \otimes I_2)$. All the experiments are conducted with nominal size 0.05. We perform 5000 simulations of each experiment to ensure that the standard errors (s.e.) of the estimated sizes and powers of the tests are not more than $\sqrt{0.05(1 - 0.05)/5000} \approx 0.0071$. To save space, in the text-part of the paper we only present results for the sample size $n = 49$. Some of the results for $n = 98$ are provided in the Appendices (see Table T1).

We start with estimating the empirical sizes of the five tests namely $RS_{\lambda\gamma}$, RS_λ, RS^*_λ, RS_γ and RS^*_γ. Table 1 presents the empirical sizes both for sample size $n = 49$ and $n = 98$. For $n = 49$ the unadjusted tests RS_λ and RS_γ over-reject the null hypotheses while the adjusted tests RS^*_λ, RS^*_γ as well as the joint test $RS_{\lambda\gamma}$ under-reject. The empirical sizes of all the tests improve, i.e., move closer to 0.05 as we increase the sample to $n = 98$.

To gauge the overall behaviour at the tail parts of the distributions of different test statistics under the null, the kernel density plots for $n = 49$ are presented in Figures 1 and 2. For our purpose, portions of the plots that are relevant for us are on the right of the solid vertical lines intersecting the x-axis at $x = 3.841$ and $x = 5.991$, which are, respectively, the 5% critical values for χ^2_1 and χ^2_2 densities. To compare the empirical sizes of the tests with the nominal size 0.05, we examine the closeness of their kernel density plots with those of χ^2 in the tail regions. Let us begin with the χ^2_1 tests RS_λ, RS^*_λ, RS_γ and RS^*_γ. From Figure 1, it is clear that the plots for RS_λ and RS_γ given by the green and red curves, respectively, are slightly over the χ^2_1 plot whereas the grey and blue curves for the adjusted tests RS^*_λ and RS^*_γ are below. Thus, it is evident why the unadjusted tests are oversized while the adjusted are somewhat undersized. Similarly, comparing the kernel density of the joint test $RS_{\lambda\gamma}$ with χ^2_2 in Figure 2, we observe that the two plots overlap, and the plot for $RS_{\lambda\gamma}$ is by and large somewhat below that for χ^2_2 resulting in an empirical size slightly less than 0.05. Thus, our kernel density plots in Figures 1 and 2 conform to the numbers in Table 1.

As discussed in Section 4.4, the starting point of our model specification search is to use the joint test $RS_{\lambda\gamma}$. Thus, it is quite important for $RS_{\lambda\gamma}$ to have good power properties in a finite sample. This is demonstrated with the help of a $3 - d$ power-surface plot, displayed in Figure 3, for the combination of values of (λ, γ) from (0, 0) to (0.50, 1.50). From Figure 3 we observe that the power of $RS_{\lambda\gamma}$ ascends to 1.00 quite swiftly. This ensures the reliability of $RS_{\lambda\gamma}$ in detecting the departure from $H_{01}{:}\lambda = \gamma = 0.0$. Thus, our suggestion of using $RS_{\lambda\gamma}$ as the starting point of the testing exercise will lead to the correct inference with a high probability.

Table 1. Empirical sizes of the tests (nominal size = 0.05).

Sample size (n)	RS_λ	RS^*_λ	RS_γ	RS^*_γ	$RS_{\lambda\gamma}$
49	0.0516	0.0392	0.0544	0.0464	0.0462
98	0.0472	0.0472	0.0522	0.0458	0.0470

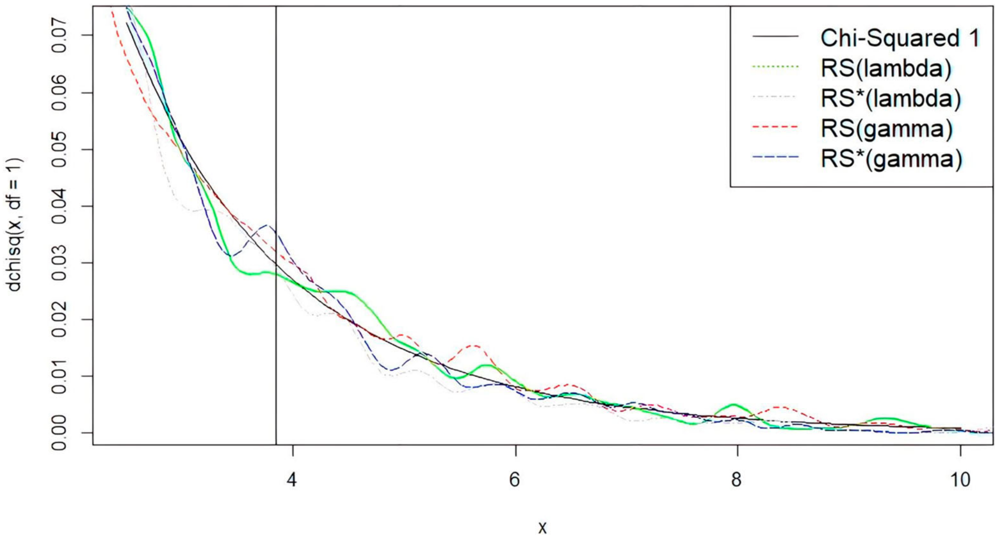

Figure 1. Kernel density plots for four test statistics in comparison with that of χ_1^2; ($n = 49$).
Note: Readers of the print issue can view the figures in colour online at https://doi.org/10.1080/17421772.2023.2256810.

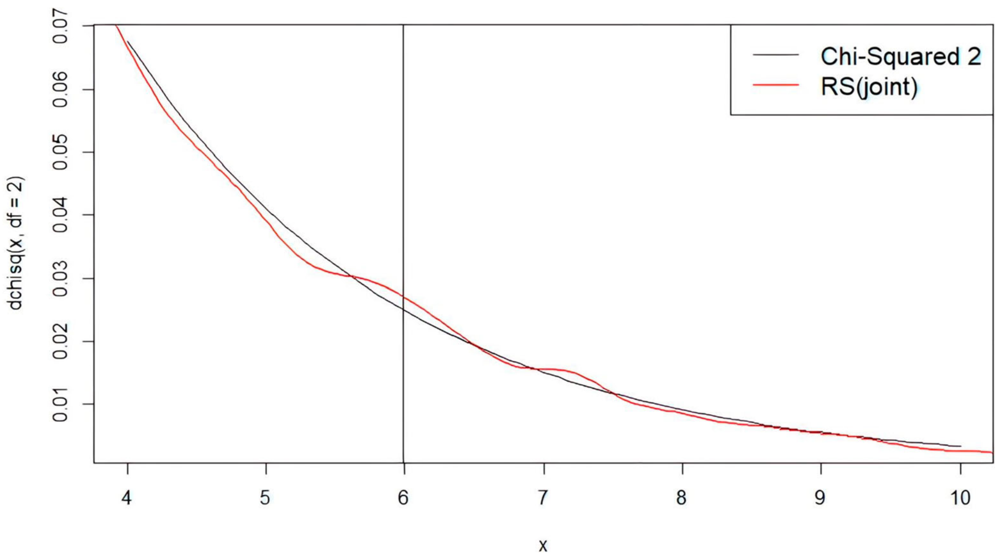

Figure 2. Kernel density plot for $RS_{\lambda\gamma}$ in comparison with that of χ_2^2; ($n = 49$).

Next we evaluate the comparative performance of all the five tests RS_λ, RS_λ^*, RS_γ, RS_γ^* and $RS_{\lambda\gamma}$ in terms of size and power. The numerical values of the rejection probabilities of the tests are given in Table 2 for the combinations of $\lambda = \{0.00, 0.10, 0.30, 0.50, 0.90\}$ and $\gamma = \{0.00, 0.05, 0.10, 0.30, 1.00\}$. There are five blocks according to the five distinct values of λ. Consider the numbers reported in the first row of the first block, i.e., for values $(\lambda, \gamma) = (0.00, 0.00)$; naturally, those coincide with the values given in Table 1 for $n = 49$. For the combination $(\lambda, \gamma) = (0.00, 0.30)$ the value for $RS_\lambda = 0.1882$ which is more than that of $RS_\lambda^* = 0.0366$. Similar differences of values are observed when we consider the rejection probabilities of RS_γ and RS_γ^* in the presence of λ. For instance, for $(\lambda, \gamma) = (0.30, 0.00)$, the

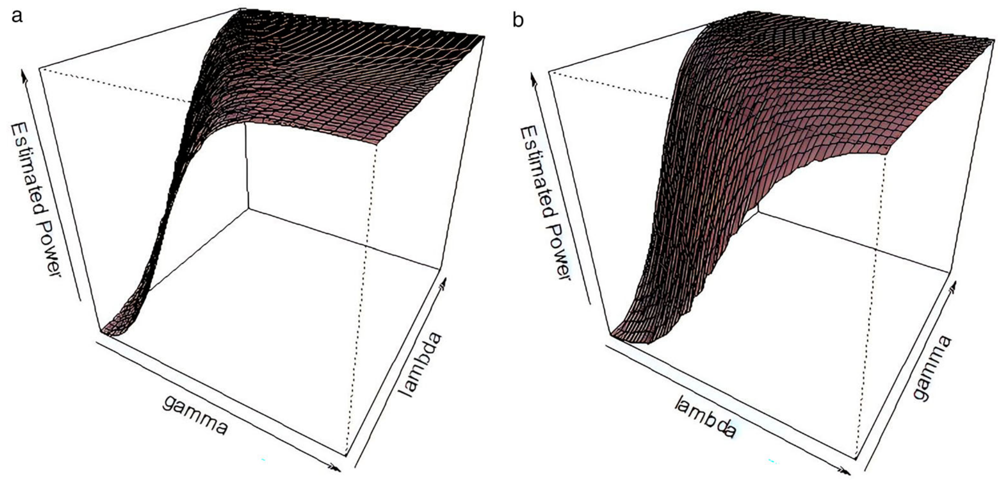

Figure 3. Power surface plots for $RS_{\lambda\gamma}$ under log-normal distribution, $n = 49$. Figures (a) and (b) are 90° rotations of each other.

Table 2. Estimated rejection probabilities of the tests, $n = 49$.

λ	γ	RS_λ	RS^*_λ	RS_γ	RS^*_γ	$RS_{\lambda\gamma}$
0.00	0.00	0.0516	0.0392	0.0544	0.0464	0.0462
0.00	0.05	0.0542	0.0414	0.0650	0.0444	0.0500
0.00	0.10	0.0626	0.0436	0.0762	0.0560	0.0606
0.00	0.30	0.1822	0.0336	0.2560	0.0944	0.1740
0.00	1.00	0.9490	0.1072	0.9936	0.3610	0.9768
0.10	0.00	0.1244	0.0568	0.1110	0.0462	0.0968
0.10	0.05	0.1858	0.0676	0.1724	0.0454	0.1468
0.10	0.10	0.2214	0.0630	0.2230	0.0534	0.1770
0.10	0.30	0.4902	0.0822	0.5464	0.0794	0.4394
0.10	1.00	0.9954	0.2996	0.9992	0.2676	0.9980
0.30	0.00	0.7372	0.3406	0.6136	0.0422	0.6490
0.30	0.05	0.7996	0.3716	0.7046	0.0366	0.7280
0.30	0.10	0.8446	0.3718	0.7742	0.0368	0.7906
0.30	0.30	0.9684	0.4908	0.9536	0.0352	0.9490
0.30	1.00	1.0000	0.8544	1.0000	0.1446	1.0000
0.50	0.00	0.9916	0.8560	0.9644	0.0322	0.9864
0.50	0.05	0.9964	0.8690	0.9810	0.0258	0.9926
0.50	0.10	0.9974	0.8866	0.9886	0.0190	0.9952
0.50	0.30	0.9998	0.9474	0.9994	0.0090	0.9992
0.50	1.00	1.0000	0.9988	1.0000	0.0916	1.0000
0.90	0.00	1.0000	1.0000	1.0000	0.0000	1.0000
0.90	0.05	1.0000	1.0000	1.0000	0.0014	1.0000
0.90	0.10	1.0000	1.0000	1.0000	0.0002	1.0000
0.90	0.30	1.0000	1.0000	1.0000	0.0048	1.0000
0.90	1.00	1.0000	1.0000	1.0000	0.1538	1.0000

rejection probability of $RS_\gamma = 0.6136$ is much higher than that of $RS^*_\gamma = 0.0422$. However, we will note later from Figure 6(d), the performance of RS^*_γ is quite good as we move away from the local alternative, i.e., for higher values of γ.

Finally, we look at the power of the joint test $RS_{\lambda\gamma}$ given in the last column of Table 2. We see that even when $\lambda = 0.0$, as we increase the value of γ, the power of $RS_{\lambda\gamma}$ keeps increasing and finally, reaches 0.9768 at $\gamma = 1.0$.. This indicates that $RS_{\lambda\gamma}$, a *two* degree of freedom test, has good power even for *one* directional departure (in the direction of γ), from the joint null $H_{01}:\lambda = \gamma = 0.0$. When $\lambda = 0.0$ and γ varies from 0.00 to 1.00, the optimal test is RS_γ while $RS_{\lambda\gamma}$ over-tests. Numbers in Table 2 imply that although $RS_{\lambda\gamma}$ is not the optimal test, in these cases it has rejection probabilities close to that of the optimal test. A similar pattern is observed when we compare RS_λ and $RS_{\lambda\gamma}$ for one-directional departure, when $\gamma = 0.0$ and λ takes values in {0.10, 0.30, 0.50, 0.90}.

Four plots are presented in Figure 4 to obtain a clear picture of the relative performance of RS_λ and RS^*_λ. From Figure 4(a), we observe that, as the misspecification due to γ increases, the size of RS_λ increases very rapidly. On the other hand, the size of the adjusted test RS^*_λ remains close to the nominal value 0.05 to a certain extent and then increases, however, the rate of increase is much less than that of RS_λ. One of the reviewers wanted us to investigate the threshold up to where RS^* works well and especially pointed out the performance of RS^*_λ when $\lambda = 0$ and γ diverge from zero. As we see from Table 2, when $\gamma = 1.00$, the empirical size of RS^*_λ = 0.1072 which is much higher than the nominal size, 0.05. Note that, by construction, RS^*_λ is designed to deal with the problem of *local* departure of γ from zero in the alternative model. In response to the reviewer's query, we zoom in on part of Figure 4(a) (as indicated by the rectangle) and present it in Figure 5. The solid black line is for $\alpha = 0.05$, the nominal size of the test. The empirical size of RS^*_λ starts below this line and crosses it around $\gamma = 0.62$ and steadily increases thereafter. However, even at $\gamma = 0.90$ the estimated size remains below 0.10. On the other hand, the estimated size of the unadjusted test RS_λ, as is very clear from Figure 4(a), reaches close to 1.0. Thus, our suggested test RS^*_λ performs quite well even when the misspecification is *not local*.

The power curves of the two tests are given in Figure 4(b) where there is no misspecification, i.e., $\gamma = 0$; in other words, when no adjustment is necessary. We see that the power of RS^*_λ given by the red curve trails below the power curve of RS_λ given by the black curve. Nevertheless, it is reassuring to see that the gain in size which can be quantified by the area between the two size curves in Figure 4(a) substantially outweighs the loss in power given by the area between the two power curves in Figure 4(b). The rejection probabilities of RS_λ and RS^*_λ as we deviate from $H_0^\lambda:\lambda = 0.0$ are given under different levels of misspecifications of γ starting at $\gamma = 0.0$ and then for $\gamma = 0.05,\ 0.10,\ 0.50$ and 1.0, in Figure 4(c and d). Note that for $\gamma = 0.0$, the rejection probabilities are already depicted in Figure 4(b). As misspecification due to γ increases RS^*_λ rejects the H_0^λ less frequently than RS_λ. This is also reflected in the rejection probability values in Table 2 as discussed earlier. Rejection probabilities for RS^*_λ are always less compared to RS_λ for any value of γ. However, once we balance that with the huge size problem of RS_λ, this seemingly 'less power of RS^*_λ' is really not a drawback of the adjusted test compared to its unadjusted counterpart. Also, an attractive feature of RS^*_λ is that, as γ is varied, its rejection probabilities are more stable compared to that of RS_λ. Thus, we can conclude that overall, RS^*_λ is fairly immune to the values of γ, as it has been designed to be.

Finally, we look at the size and power properties of RS_γ and RS^*_γ using Figure 6. From Figure 6(a) we see that, as we increase misspecification due to λ, the size of RS_γ steadily increases and reaches to the level 1.0 when $\lambda = 0.60$. In contrast, the empirical size of RS^*_γ seems to be too good to be true. RS^*_γ has a flat size, equal to that of the nominal size 0.05 even for large values of λ. Thus, there is an enormous gain in terms of size. Of course, there is a price (premium) we need to pay for this gain. As is evident from Figure 6(b) the power of RS^*_γ is less than that of RS_γ. By comparing the respective areas of gain and loss, we can see that the net gain is substantial.

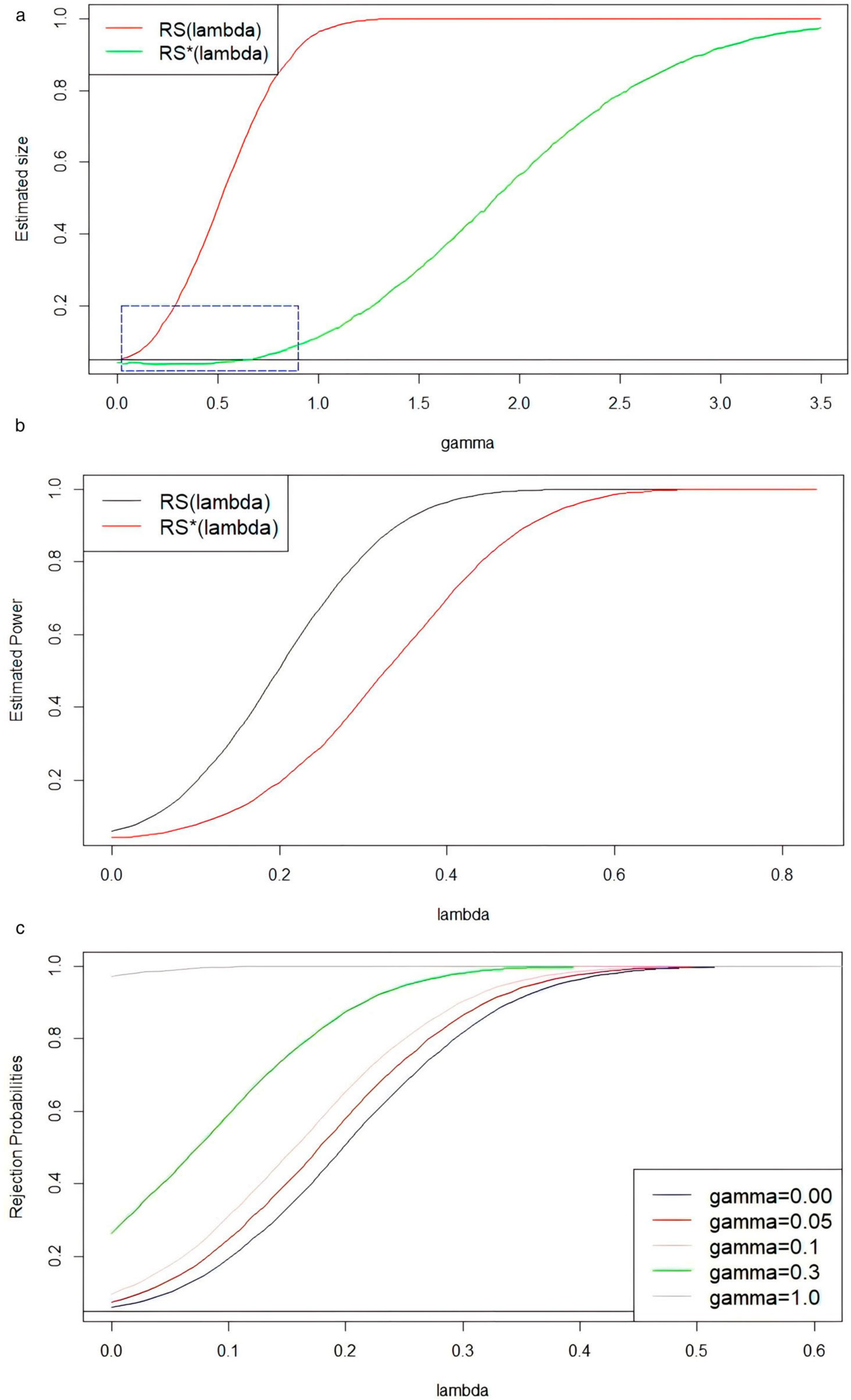

Figure 4. Test for spatial autocorrelation H_0^{λ}:$\lambda = 0$; $n = 49$.

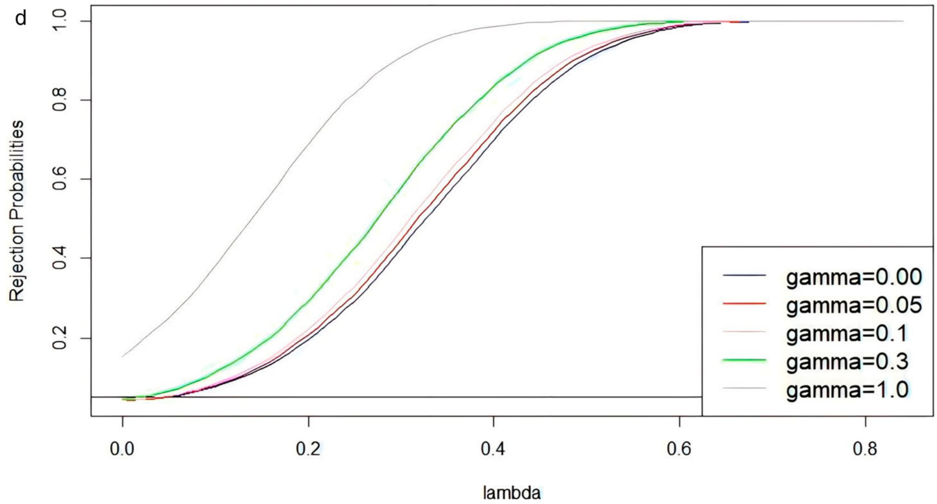

Figure 4. *Continued*

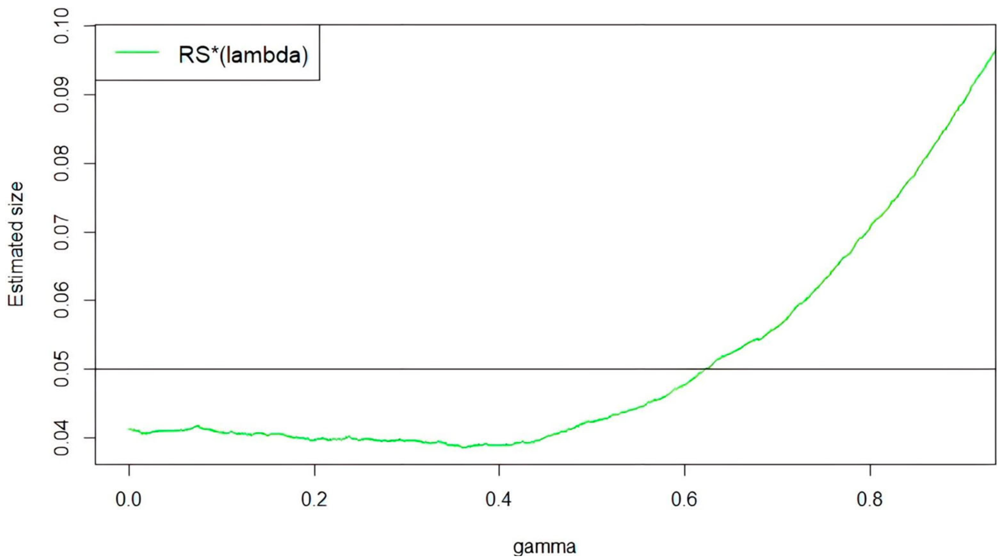

Figure 5. Size of RS^*_λ crossing the line of nominal size at 0.05 under misspecification ($\gamma > 0$).

Finally, from Figure 6(c and d), we notice more stability in the rejection probabilities of RS^*_γ compared to that of RS_γ as the amount of misspecification (values of λ) increases.

6. EMPIRICAL ILLUSTRATION

To gain insight into the properties and usefulness of our suggested tests for real data, we employ them in the context of two models used in literature. The first one is a spatial model that captures the relationship between crime and income and housing value for 49 neighbourhoods in Columbus, Ohio. The data are listed in Table 12.1, p.189 of Anselin (1988) and have been used in a

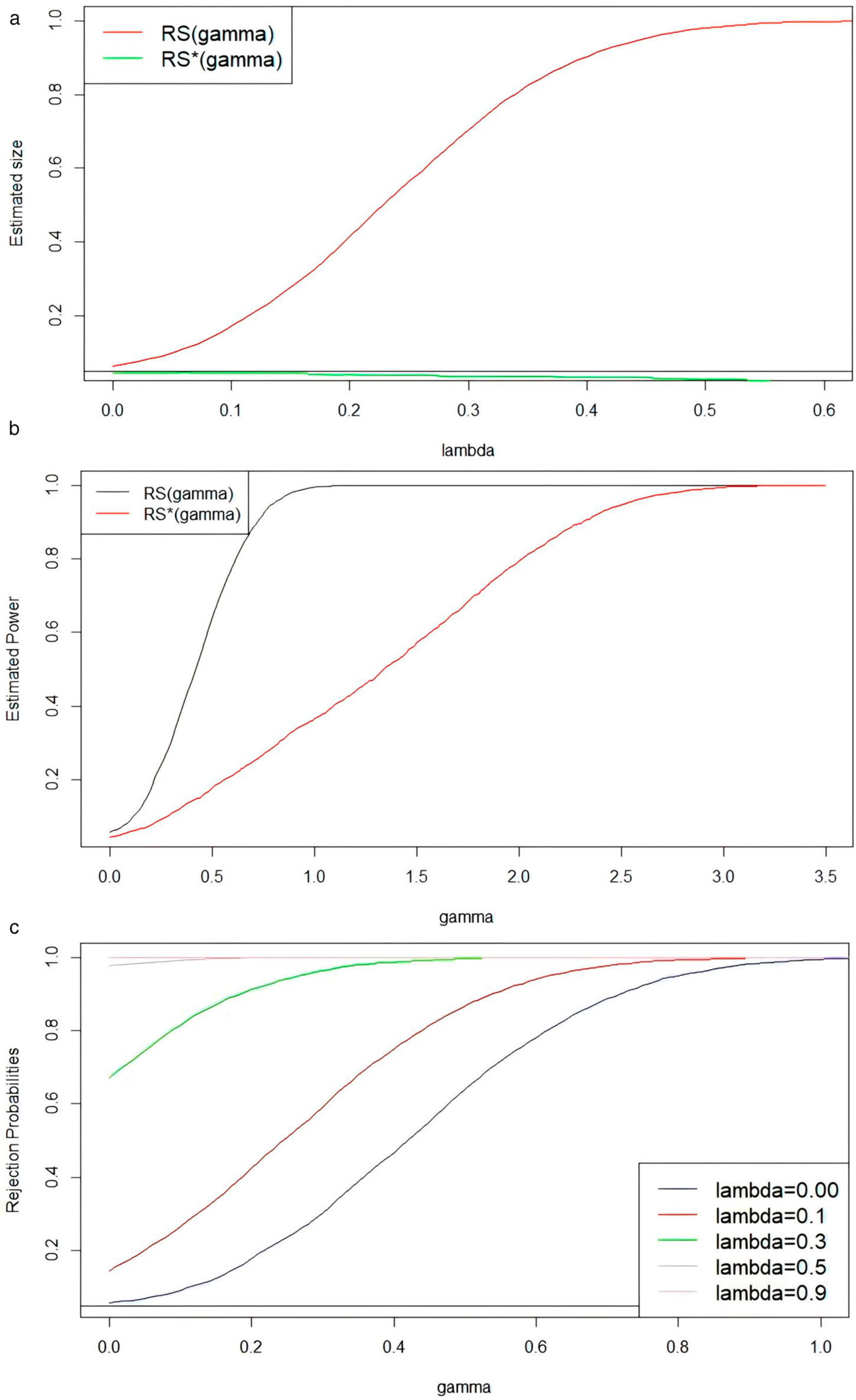

Figure 6. Test for the Durbin effect $H_0^D\colon \gamma = 0$; $n = 49$.

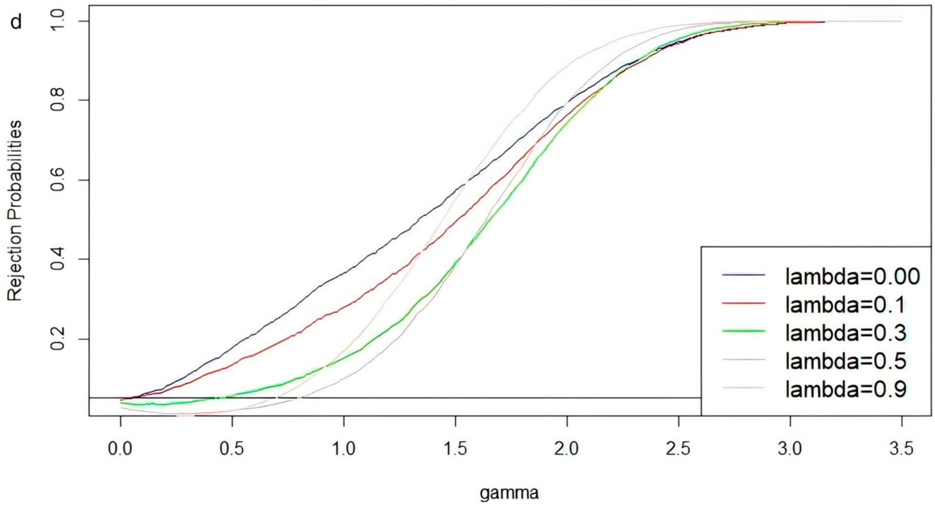

Figure 6. *Continued*

number of papers to benchmark different estimators and specification tests, see, for instance, Getis (1990), McMillen (1992), Anselin et al. (1996), LeSage (1997), Griffith (2000) and Elhorst (2014, pp. 27–32).

The test statistics based on the OLS residuals (of model (4.1) under the joint null $H_{01}:\lambda = 0,\ \gamma = 0$) are presented in Table 3 below. All the test statistics were computed separately. From Table 3, we note that the equalities in (4.16) are satisfied, confirming the correctness of our computation.

The joint test statistic, $RS_{\lambda\gamma} = 10.7487$ rejects the joint null hypothesis H_{01} when compared to χ^2_3 critical values up to significance level 0.0132. In the previous section we demonstrated good finite sample properties of $RS_{\lambda\gamma}$ both in terms of size and power. The joint test is, however, not informative about the specific deviation(s) of misspecifications. The two unadjusted statistics RS_λ and RS_γ reject the respective null hypotheses. If an investigator takes these rejections at their face values, then they would incorporate both SAR and Durbin components into the final model. Specifications of λ and γ can only be evaluated correctly by considering our two adjusted tests, namely, RS^*_λ and RS^*_γ. RS^*_λ still rejects the null hypothesis $H^\lambda_0:\lambda = 0$ while RS^*_γ turns out to be insignificant. Although insignificant, let us look at the decomposition of RS^*_γ. We have $RS^*_{\gamma_1} = 0.2637$, $RS^*_{\gamma_2} = 2.099$ and the interaction term between $RS^*_{\gamma_1}$ and $RS^*_{\gamma_2}$, say, $RS^*_{\gamma_1\gamma_2} = 0.531$. Thus, only $RS^*_{\gamma_2}$ has some closeness to significance. We will later present estimation of γ_1 and γ_2 to illustrate this further.

It is interesting to note how values of the test statistics reduce after adjustment (robustification). For instance, the reduction of $RS_\gamma = 6.1376$ to $RS^*_\gamma = 2.8931$ is very noticeable. From the analytical results of the previous section it is clear that the value 6.1376 is not due to the presence of the Durbin part only. The presence of the SAR part, which seems to be much stronger for this data set, also contributes to this value. Thus, the misspecification of the basic linear regression

Table 3. Test statistics for Columbus crime data (*p*-values are in parentheses).

$RS_{\lambda\gamma}$	RS_λ	RS_γ	RS^*_λ	RS^*_γ
10.7487	7.8556	6.1376	4.6111	2.8931
(0.0132)	(0.0051)	(0.0465)	(0.0318)	(0.2354)

model can be thought of as coming from the spatial lag dependence of the dependent variable rather than that of the independent variables (i.e., the Durbin part).

We further demonstrate how much extra information is contained in our adjusted tests by estimating the unrestricted model incorporating both SAR and Durbin components and the model with the SAR part only. A detailed interpretation of the results in Table 4 is not necessary for the purpose of our paper. The OLS results are given just as a benchmark. As can be foreseen from the values of our RS^*_γ, the coefficients of lag of income and lag of housing value are not significant in the SDM. Again, recall the values of $RS^*_{\gamma_1}$ and $RS^*_{\gamma_2}$ based on OLS. These values strikingly corroborate the p-values for γ_1 and γ_2 based on full-fledged MLE of SDM. The third column of the table represents the estimates from a pure SAR model, i.e., without the Durbin components. The estimates or the coefficients of income and housing value are close to those of SDM. However, this is more parsimonious and a better model, as evident from the low Akaike information criteria (AIC) value. Of course, both the models give better results relative to OLS. In conclusion we can say that a pure SAR model, as indicated by our adjusted tests based only on an OLS estimation, is better than SDM. Thus, reliable estimates for the coefficients of income and housing value can be obtained by filtering out the (direct) spatial dependence in the crime variable.

Our second application considers the spatial impact of house and neighbourhood characteristics and crime on housing prices in the City of Chicago based on Lu and Hewings (2022). The independent variables include number of bathrooms, floor area, total amount of crime and distance to the central business district (CBD). The spatial unit is a census tract, chosen due to the homogeneity of socioeconomic and demographic characteristics. Residential sales data from Illinois realtors is used for the dependent variable: the house value in each census tract. Crime data are obtained from the Chicago Police Department. Data used is from the year 2017 for 801 census tracts of the City of Chicago. For further details on the data, see Lu and Hewings (2022).

The test statistic values are provided in Table 5. We first look at the value of the joint test $RS_{\lambda\gamma} = 971.9565$ which rejects the null hypothesis $H_{01}{:}\lambda_0 = 0,\ \gamma_0 = 0$. The robust test

Table 4. Estimation of different models for the Columbus crime data (p-values in parentheses).

	OLS	SDM	SAR
Intercept	68.62***	45.59***	46.85***
	$(< 2e^{-16})$	(0.0005)	$(1.503e^{-10})$
Income	−1.60***	−0.94**	−1.07***
	$(1.83e^{-05})$	(0.0055)	(0.0006)
Housing value	−0.27*	−0.30***	−0.27**
	(0.0109)	(0.0009)	(0.0027)
Lag of income		−0.62	
		(0.2839)	
Lag of housing value		0.27	
		(0.1473)	
Spatial lag (λ)		0.38*	0.40***
		(0.0413)	(0.0037)
Number of parameters	4	7	5
Log likelihood	−187.3772	−182.02	−183.17
AIC	382.7545	378.03	376.34

***$p < 0.001$; **$p < 0.01$; *$p < 0.05$.

Table 5. Test statistics for the Chicago housing data (p-values are in parentheses).

$RS_{\lambda\gamma}$	RS_{λ}	RS_{γ}	RS^*_{λ}	RS^*_{γ}
971.9565	863.0719	140.3908	831.5657	108.8846
(< .00001)	(< .00001)	(< .00001)	(< .00001)	(< .00001)

statistics RS^*_{λ} and RS^*_{γ} indicate significance of spatial autocorrelation and the spatial Durbin part, respectively. As for the Columbus crime data, to get a precise idea about the significance of the neighbourhood effect of each independent variable, number of bathrooms, floor area, total crime and distance to CBD, we decompose RS^*_{γ} into $RS^*_{\gamma_1} = 123.788$, $RS^*_{\gamma_2} = 71.4650$, $RS^*_{\gamma_3} = 17.0841$ and $RS^*_{\gamma_4} = 0.0372$ plus some interaction terms. Comparing these values with the 5 per cent critical value of $\chi^2_1 = 3.841$, we see that $RS^*_{\gamma_1}$ and $RS^*_{\gamma_2}$ strongly reject the null hypotheses of no significance whereas $RS^*_{\gamma_3}$ and $RS^*_{\gamma_4}$ denote moderate to strong insignificance.

As we will illustrate now, these findings concur with those from the estimation of the full-fledged model provided in Table 6. We start with OLS estimation of a simple linear regression

Table 6. Estimation of different models for the Chicago housing value data.

	OLS	**SDM 1**	**SDM 2**
Intercept	6.29***	1.27***	1.04***
	$(< 2e^{-16})$	$(6.907e^{-07})$	$(1.261e^{-05})$
Number of bathrooms	0.34***	0.22***	0.23***
	$(< 2e^{-16})$	$(< 2.2e^{-16})$	$(< 2.2e^{-16})$
Floor area	−0.13	0.21***	0.19***
	(0.0872)	$(2.722e^{-07})$	$(1.867e^{-06})$
Total crime	−0.38***	−0.03	
	$(< 2e^{-16})$	(0.1450)	
Distance to CBD	−0.05***	−0.01	
	$(< 2e^{-16})$	(0.6095)	
Lag of number of bathrooms		−0.12***	−0.07*
		(0.0007)	(0.0272)
Lag of floor area		−0.31***	−0.40***
		$(7.382e^{-05})$	$(2.481e^{-08})$
Lag of total crime		−0.08*	
		(0.0209)	
Lag of distance to CBD		0.01	
		(0.7751)	
Spatial lag (λ)		0.84***	0.87***
		$(< 2.22e^{-16})$	$(< 2.22e^{-16})$
Number of parameters	6	11	7
Log likelihood	−115.5142	312.79	305.48
AIC	243.0284	−603.58	−596.97

***$p < 0.001$; **$p < 0.01$; *$p < 0.05$.

model disregarding the spatial dependence in the data. The OLS model is not a good fit as is evident from the high AIC score. Next, we estimate the spatial Durbin model including the entire set of independent variables (see Lu and Hewings (2022)). The improvement in the AIC value and significance of the spatial lag λ points towards the fact that this is a *better* model since it takes account of the spatial dependence in the data. However, not all the independent variables are significant. Total crime and distance to CBD are insignificant as is noted from their high p-values. Quite interestingly, the neighbourhood effects of these two variables given by lag of total crime and lag of distance to CBD are insignificant as well. This implies that total crime and distance to CBD have no significant effect on the house value for a census tract. Moreover, they are insignificant in explaining house value of the neighbouring census tracts as well. Finally, we estimate the model, given by SDM 2 in Table 6, using the subset of significant variables, i.e., number of bathrooms and floor area allowing for their lags as well. Quite obviously, this model is better than the all inclusive model SDM 1, as is evident from the lower AIC value. Hence we see that a decent idea about the estimation of the full-fledged model can be obtained just by looking at the robust test statistics values calculated using simple OLS estimations.

7. CONCLUSION

In this paper we consider the SDM which is one of the most popular models used in spatial econometrics. It reduces to the SEM under a *non-linear* parametric restriction. However, in practice, a test for such a non-linear restriction is rarely conducted. We suggest an alternative approach of testing simple *linear* hypothesis using Bera and Yoon (1993) methodology. Our tests, which are robust to the local misspecification of the alternative model, are analytically simpler and very easy to implement in practice. These tests exhibit excellent finite sample performance in terms of size and power. We further illustrate the usefulness of our proposed tests using Anselin's Columbus crime and Chicago City housing value data. In both the cases our tests prove to be useful in identifying the most appropriate models. This has been demonstrated by empirically establishing that the MLE results of the full models, i.e., the unrestricted SDM, can almost be anticipated by our robust tests. Finally, to conclude, this paper fills in the gap in the spatial econometrics testing literature that 'ignores the SDM model', as pointed out by LeSage and Pace (2009) and Elhorst (2010). Therefore, we hope our tests will be useful for empirical researchers in their search for the appropriate model specification.

ACKNOWLEDGEMENTS

We are most grateful to the Editor and the two anonymous Reviewers for their pertinent comments and helpful suggestions that greatly helped in improving the content and exposition of the paper. An earlier version of the paper was presented at the XVI World Conference Spatial Econometrics Association (SEA 2022), Warsaw, Poland, June 23-24, 2022. We are thankful to the participants of that conference for their comments, especially to the discussant, Professor Roman Minguez for careful reading of the paper and offering his valuable feedback. This paper was also presented at the Economic Research Unit (ERU), Indian Statistical Institute (ISI), 42nd Annual Seminar of the Centre for Urban Economic Studies (CUES), University of Calcutta and at the Department of Economics, Jadavpur University. We would like to thank the organizing committees for giving us the opportunity to present our paper, and the attendees of these seminars for their constructive feedback that further helped in producing an improved version of the paper. Finally, we are very grateful to Professor Geoffrey Hewings and Dr. Chang Lu for supplying us with their Illinois REALTORS data. Dr. Lu also assisted in defining the variables that we used in our model. Of course, we retain the responsibility for any remaining errors and omissions.

DISCLOSURE STATEMENT

No potential conflict of interest was reported by the author(s).

NOTES

[1] We appreciate one of the referees for bringing the empirical results of this paper to our attention.
[2] We are thankful to the editor for drawing our attention to this paper.
[3] It is important to emphasise that this spatial interaction w_{ij} may be exogenous such as geographic distance between two locations or endogenous such as social and economic relationships among regions, unobserved characteristics in social network models or strategic tax interactions among local governments (see, for instance, Case et al. (1993), Hsieh and Lee (2016) and Delgado et al. (2018)). The construction and estimation of a spatial econometric model is subject to whether the weight matrix that best represents the spatial characteristics of the data is exogenous or endogenous in nature. In this paper, we focus on purely exogenous spatial weight matrices. Endogenous W poses some additional complexities, both in terms of estimation and testing, for instance see Qu and Lee (2015), Qu et al. (2017) and Bera, Doğan, and Taşpınar (2019).

ORCID

Malabika Koley http://orcid.org/0000-0002-1485-9604
Anil K. Bera http://orcid.org/0009-0000-2206-8143

REFERENCES

Alan, H., Summers, R., & Aten, B. (2002). Penn world tables, version 6.1, Center for International Comparisons at the University of Pennsylvania.

Anselin, L. (1988). *Spatial econometrics: Methods and models*. Springer.

Anselin, L., & Bera, A. K. (1998). Spatial dependence in linear regression models with an introduction to spatial econometrics. *Handbook of Applied Economic Statistics*, *155*(5), 237–290.

Anselin, L., Bera, A. K., Florax, R., & Yoon, M. J. (1996). Simple diagnostic tests for spatial dependence. *Regional Science and Urban Economics*, *26*(1), 77–104.

Anselin, L., Cohen, J., Cook, D., Gorr, W., & Tita, G. (2000). Spatial analyses of crime. *Criminal Justice*, *4*(2), 213–262.

Autant-Bernard, C., & LeSage, J. P. (2011). Quantifying knowledge spillovers using spatial econometric models. *Journal of Regional Science*, *51*(3), 471–496. https://doi.org/10.1111/j.1467-9787.2010.00705.x

Bera, A. K., Bilias, Y., Yoon, M. J.Taşpınar, S.. (2019a). Adjustments of Rao's score test for distributional and local parametric misspecifications. *Journal of Econometric Methods*, *9*(1), 20170022.

Bera, A. K., Doğan, O., & Taşpınar, S. (2019). Testing spatial dependence in spatial models with endogenous weights matrices. *Journal of Econometric Methods*, *8*(1), 1–34.

Bera, A. K., Doğan, O., Taşpınar, S., & Leiluo, Y. (2019b). Robust LM tests for spatial dynamic panel data models. *Regional Science and Urban Economics*, *76*, 47–66. https://doi.org/10.1016/j.regsciurbeco.2018.08.001

Bera, A. K., Doğan, O., Taşpınar, S., & Sen, M. (2020). Specification tests for spatial panel data models. *Journal of Spatial Econometrics*, *1*(1), 1–39. https://doi.org/10.1007/s43071-020-0001-4

Bera, A. K., Montes-Rojas, G., & Sosa-Escudero, W. (2009). Testing under local misspecification and artificial regressions. *Economics Letters*, *104*(2), 66–68.

Bera, A. K., & Yoon, M. J. (1993). Specification testing with locally misspecified alternatives. *Econometric Theory*, *9*(4), 649–658. https://doi.org/10.1017/S0266466600008021

Burridge, P. (1981). Testing for a common factor in a spatial autoregression model. *Environment and Planning A*, *13*(7), 795–800. https://doi.org/10.1068/a130795

Case, A. C., Rosen, H. S., & HinesJrJ. R. (1993). Budget spillovers and fiscal policy interdependence: Evidence from the states. *Journal of Public Economics*, *52*(3), 285–307. https://doi.org/10.1016/0047-2727(93)90036-S

Chen, Y., Shao, S., Fan, M., Tian, Z., & Yang, L. (2022). One man's loss is another's gain: Does clean energy development reduce co2 emissions in China? Evidence based on the spatial durbin model. *Energy Economics*, *107*, 1–25.

Davidson, R., & MacKinnon, J. G. (1987). Implicit alternatives and the local power of test statistics. *Econometrica*, *55*(6), 1305–1329. https://doi.org/10.2307/1913558

Delgado, F. J., Lago-Peñas, S., & Mayor, M. (2018). Local tax interaction and endogenous spatial weights based on quality of life. *Spatial Economic Analysis*, *13*(3), 296–318. https://doi.org/10.1080/17421772.2018.1420213

Durbin, J. (1960). Estimation of parameters in time-series regression models. *Journal of the Royal Statistical Society: Series B*, *22*(1), 139–153.

Elhorst, J. P. (2010). Applied spatial econometrics: Raising the bar. *Spatial Economic Analysis*, *5*(1), 9–28. https://doi.org/10.1080/17421770903541772

Elhorst, J. P. (2014). *Spatial econometrics from cross-sectional data to spatial panels*. Springer.

Ertur, C., & Koch, W. (2007). Growth, technological interdependence and spatial externalities: Theory and evidence. *Journal of Applied Econometrics*, *22*(6), 1033–1062. https://doi.org/10.1002/jae.963

Fang, Y., Park, S. Y., & Zhang, J. (2014). A simple spatial dependence test robust to local and distributional misspecifications. *Economics Letters*, *124*(2), 203–206. https://doi.org/10.1016/j.econlet.2014.05.015

Feng, Z., & Chen, W. (2018). Environmental regulation, green innovation, and industrial green development: An empirical analysis based on the spatial Durbin model. *Sustainability*, *10*(2), 1–22. https://doi.org/10.3390/su10020001

Getis, A. (1990). Screening for spatial dependence in regression analysis. *Papers of the Regional Science Association*, *69*(1), 69–81. https://doi.org/10.1007/BF01933897

Griffith, D. A. (2000). A linear regression solution to the spatial autocorrelation problem. *Journal of Geographical Systems*, *2*(2), 141–156. https://doi.org/10.1007/PL00011451

Han, F., Xie, R., & Lai, M. (2018). Traffic density, congestion externalities, and urbanization in China. *Spatial Economic Analysis*, *13*(4), 400–421. https://doi.org/10.1080/17421772.2018.1459045

Hsieh, C.-S., & Lee, L. F. (2016). A social interactions model with endogenous friendship formation and selectivity. *Journal of Applied Econometrics*, *31*(2), 301–319. https://doi.org/10.1002/jae.2426

Jenish, N., & Prucha, I. R. (2009). Central limit theorems and uniform laws of large numbers for arrays of random fields. *Journal of Econometrics*, *150*(1), 86–98. https://doi.org/10.1016/j.jeconom.2009.02.009

Jenish, N., & Prucha, I. R. (2012). On spatial processes and asymptotic inference under near-epoch dependence. *Journal of Econometrics*, *170*(1), 178–190. https://doi.org/10.1016/j.jeconom.2012.05.022

Juhl, S. (2021). The wald test of common factors in spatial model specification search strategies. *Political Analysis*, *29*(2), 193–211. https://doi.org/10.1017/pan.2020.23

Khezri, M., Karimi, M. S., Khan, Y., & Abbas, S. (2021). The spillover of financial development on co2 emission: A spatial econometric analysis of Asia-pacific countries. *Renewable and Sustainable Energy Reviews*, *145*, 1–11. https://doi.org/10.1016/j.rser.2021.111110

Koley, M., & Bera, A. K. (2022). Testing for spatial dependence in a spatial autoregressive (SAR) model in the presence of endogenous regressors. *Journal of Spatial Econometrics*, *3*(1), 11.

Lee, L. F. (2004). Asymptotic distributions of quasi-maximum likelihood estimators for spatial autoregressive models. *Econometrica*, *72*(6), 1899–1925. https://doi.org/10.1111/j.1468-0262.2004.00558.x

Lee, L. F., & Yu, J. (2016). Identification of spatial Durbin panel models. *Journal of Applied Econometrics*, *31*(1), 133–162. https://doi.org/10.1002/jae.2450

LeSage, J., & Pace, R. K. (2009). *Introduction to spatial econometrics*. Chapman and Hall/CRC.

LeSage, J. P. (1997). Bayesian estimation of spatial autoregressive models. *International Regional Science Review*, *20*(1-2), 113–129. https://doi.org/10.1177/016001769702000107

Li, J., & Li, S. (2020). Energy investment, economic growth and carbon emissions in China—Empirical analysis based on spatial Durbin model. *Energy Policy*, *140*, 1–11.

Lu, C., & Hewings, G. (2022). Modeling the spatial impact of crime on housing prices: Evidences from Chicago city. *Manuscript*, 1–15.

McMillen, D. P. (1992). Probit with spatial autocorrelation. *Journal of Regional Science*, *32*(3), 335–348. https://doi.org/10.1111/j.1467-9787.1992.tb00190.x

Mur, J., & Angulo, A. (2006). The spatial Durbin model and the common factor tests. *Spatial Economic Analysis*, *1* (2), 207–226. https://doi.org/10.1080/17421770601009841

Qu, X., & Lee, L. F. (2015). Estimating a spatial autoregressive model with an endogenous spatial weight matrix. *Journal of Econometrics*, *184*(2), 209–232. https://doi.org/10.1016/j.jeconom.2014.08.008

Qu, X., Lee, L. F., & Yu, J. (2017). QML estimation of spatial dynamic panel data models with endogenous time varying spatial weights matrices. *Journal of Econometrics*, *197*(2), 173–201. https://doi.org/10.1016/j.jeconom.2016.11.004

Sabater, A., & Graham, E. (2019). International migration and fertility variation in Spain during the economic recession: A spatial Durbin approach. *Applied Spatial Analysis and Policy*, *12*(3), 515–546. https://doi.org/10.1007/s12061-018-9255-9

Saikkonen, P. (1989). Asymptotic relative efficiency of the classical test statistics under misspecification. *Journal of Econometrics*, *42*(3), 351–369. https://doi.org/10.1016/0304-4076(89)90058-4

Sun, J., Wang, J., Wang, T., & Zhang, T. (2019). Urbanization, economic growth, and environmental pollution: Partial differential analysis based on the spatial Durbin model. *Management of Environmental Quality*, *30*(2), 483–494. https://doi.org/10.1108/MEQ-05-2018-0101

Tientao, A., Legros, D., & Pichery, M. C. (2016). Technology spillover and TFP growth: A spatial Durbin model. *International Economics*, *145*, 21–31. https://doi.org/10.1016/j.inteco.2015.04.004

Villar, O. A. (1999). Spatial distribution of production and international trade: A note. *Regional Science and Urban Economics*, *29*(3), 371–380. https://doi.org/10.1016/S0166-0462(98)00041-6

Wang, X., & Guan, J. (2017). Financial inclusion: Measurement, spatial effects and influencing factors. *Applied Economics*, *49*(18), 1751–1762. https://doi.org/10.1080/00036846.2016.1226488

Whittle, P. (1954). On stationary processes in the plane. *Biometrika*, *41*(3-4), 434–449. https://doi.org/10.1093/biomet/41.3-4.434

Xu, X., & Wang, Y. (2017). Study on spatial spillover effects of logistics industry development for economic growth in the Yangtze river delta city cluster based on spatial durbin model. *International Journal of Environmental Research and Public Health*, *14*(12), 1508. https://doi.org/10.3390/ijerph14121508

Yang, T. C., Noah, A. J., & Shoff, C. (2015). Exploring geographic variation in US mortality rates using a spatial Durbin approach. *Population, Space and Place*, *21*(1), 18–37. https://doi.org/10.1002/psp.1809