View metadata, citation and similar papers at core.ac.uk
brought to you by CORE
provided by SEDICI - Repositorio de la UNLP

Universidad Nacional de La Plata

Departamento de Economía
Facultad de Ciencias Económicas
Universidad Nacional de La Plata

Tests for the Error Component Model in the Presence of Local Misspecification

Anil K. Bera, Walter Sosa Escudero and Mann Yoon

Documento de Trabajo Nro. 22
Marzo (2000)

Tests for the Error Component Model in the Presence of Local Misspecification

Anil K. Bera*
Department of Economics
University of Illinois at Urbana-Champaign

Walter Sosa Escudero
Department of Economics
National University of La Plata, Argentina

Mann Yoon
Department of Economics
California State University at Los Angeles

This Version: January 2000

Abstract

It is well known that most of the standard specification tests are not valid when the alternative hypothesis is misspecified. This is particularly true in the error component model, when one tests for either random effects or serial correlation without taking account of the presence of the other effect. In this paper we study the size and power of the standard Rao's score tests analytically and by simulation when the data is contaminated by local misspecification. These tests are adversely affected under misspecification. We suggest simple procedures to test for random effects (or serial correlation) in the presence of local serial correlation (or random effects), and these tests require ordinary least squares residuals only. Our Monte Carlo results demonstrate that the suggested tests have good finite sample properties for *local* misspecification, and in some cases even for far distant misspecification. Our tests are also capable of detecting the right direction of the departure from the null hypothesis. We also provide some empirical illustrations to highlight the usefulness of our tests.

*Corresponding author: Anil Bera, Department of Economics, University of Illinois, Urbana-Champaign, 1206 S. 6th Street, Champaign, IL 61820, Phone: (217) 333-4596; Fax: (217) 244-6678; e-mail: anil@fisher.econ.uiuc.edu

1 Introduction

The random error component model introduced by Balestra and Nerlove (1966) was extended by Lillard and Willis (1978) to include serial correlation in the remainder disturbance term. Such an extension, however, raises questions about the validity of the existing specification tests such as the Rao's (1948) score (RS) test for random effects assuming no serial correlation as derived in Breusch and Pagan (1980). In a similar way doubts could be raised about tests for serial correlation derived assuming no random effects. Baltagi and Li (1991) proposed a RS test that jointly tests for serial correlation and random effects. One problem with the joint test is that, if the null hypothesis is rejected, it is not possible to infer whether the misspecification is due to serial correlation or to random effects. Also, as we will discuss later, because of higher degrees of freedom the joint test will not be optimal if the departure from the null occurs only in *one* direction. More recently, Baltagi and Li (1995) derived RS statistics for testing serial correlation assuming fixed/individual effects. These tests require maximum likelihood estimation of individual effects parameters.

For a long time econometricians have been aware of the problems that arise when the alternative hypothesis used to construct a test deviates from the data generating process (DGP). As emphasized by Haavelmo (1944, pp. 65-66), in testing any economic relations, specification of a given fixed set of possible alternatives, called the priori admissible hypothesis, - ^{0}, is of fundamental importance. Misspecification of the priori admissible hypotheses was termed as type-III error by Bera and Yoon (1993). Welsh (1996, p. 119) also pointed out a similar concept in the statistics literature. Typically, the alternative hypothesis may be misspecified in three different ways. In the first one, which we shall call "complete misspecification," the set of assumed alternatives, - ^{0}, and the DGP, - $'$, say, are mutually exclusive. This happens, for instance, if one tests for serial independence when the DGP has random individual effects but no serial dependence. The second case occurs when the alternative is underspecified in that it is a subset of a more general model representing the DGP, i.e., - $^0 \subset$ - $'$. This happens, for example, when both serial correlation and individual effects are present, but are tested separately (one at a time). The last case is "overtesting" which results from overspecification, that is, when - $^0 \supset$ - $'$. This can happen when, say, Baltagi and Li (1991) joint test for serial correlation and random individual effects is used when only one effect is present. [For a detailed discussion of the concepts of undertesting and overtesting, see Bera and Jarque (1982)]. In this paper, we study analytically the asymptotic effects of misspecifications on the one-directional and joint tests for serial dependence and random individual effects. These results compliment the simulation results of Baltagi and Li (1995). Then, applying the modified RS test developed by Bera and Yoon (1993), we derive a test for random effects (serial correlation) in the presence of serial correlation (random effects). Our tests can be easily implemented using ordinary least squares (OLS) residuals from the standard linear model for panel data. Our testing strategy is close to that of Hillier (1991) in the sense that we try to partition an overall rejection region to obtain evidence about the direction (or directions) in which the model needs revision.

The plan of the paper is as follows. In the next section we review a general theory of the distribution and adjustment of the standard RS statistic in the presence of local misspecification. In Section 3, the general results are specialized to the error component model. In Section 4, we present two empirical illustrations. Section 5 reports the results of an extensive Monte Carlo study. These results, along with the empirical examples, clearly demonstrate the inappropriateness of one-directional tests in identifying the specific source of misspecification(s), and highlight the usefulness of our adjusted tests. Section 6 provides some concluding remarks.

2 Effects of misspecification and a general approach to testing in the presence of a nuisance parameter

Consider a general statistical model represented by the log-likelihood $L(\gamma, \psi, \phi)$. Here, the parameters ψ and ϕ are taken as scalars to conform with our error component model, but in general they could be vectors. Suppose an investigator sets $\phi = \phi_0$ and tests $H_0 : \psi = \psi_0$ using the log-likelihood function $L_1(\gamma, \psi) = L(\gamma, \psi, \phi_0)$, where ϕ_0 and ψ_0 are known values. The RS statistic for testing H_0 in $L_1(\gamma, \psi)$ will be denoted by RS_ψ. Let us also denote $\theta = (\gamma', \psi', \phi')'$ and $\widetilde{\theta} = (\widetilde{\gamma}', \psi_0', \phi_0')'$, where $\widetilde{\gamma}$ is the maximum likelihood estimator (MLE) of γ when $\psi = \psi_0$ and $\phi = \phi_0$. The score vector and the information matrix are defined, respectively, as

$$d_a(\theta) = \frac{\partial L(\theta)}{\partial a} \quad \text{for} \quad a = \gamma, \psi, \phi$$

and

$$J(\theta) = -E\left[\frac{1}{n}\frac{\partial^2 L(\theta)}{\partial\theta\partial\theta'}\right] = \begin{bmatrix} J_\gamma & J_{\gamma\psi} & J_{\gamma\phi} \\ J_{\psi\gamma} & J_\psi & J_{\psi\phi} \\ J_{\phi\gamma} & J_{\phi\psi} & J_\phi \end{bmatrix},$$

where n denotes the sample size. If $L_1(\gamma,\psi)$ were the true model, then it is well known that under $H_0 : \psi = \psi_0$,

$$RS_\psi = \frac{1}{n} d_\psi(\tilde{\theta})' J_{\psi\cdot\gamma}^{-1}(\tilde{\theta}) d_\psi(\tilde{\theta}) \xrightarrow{D} \chi_1^2(0),$$

where $\xrightarrow{D}$ denotes convergence in distribution and $J_{\psi\cdot\gamma}(\theta) \equiv J_{\psi\cdot\gamma} = J_\psi - J_{\psi\gamma}J_\gamma^{-1}J_{\gamma\psi}$. And under $H_1 : \psi = \psi_0 + \xi/\sqrt{n}$,

$$RS_\psi \xrightarrow{D} \chi_1^2(\lambda_1), \tag{1}$$

where the noncentrality parameter λ_1 is given by $\lambda_1 \equiv \lambda_1(\xi) = \xi' J_{\psi\cdot\gamma}\xi$. Given this set-up, asymptotically the test will have correct size and will be locally optimal. Now suppose that the true log-likelihood function is $L_2(\gamma,\phi)$ so that the alternative $L_1(\gamma,\psi)$ becomes *completely* misspecified. Using a sequence of local values $\phi = \phi_0 + \delta/\sqrt{n}$, Davidson and MacKinnon (1987) and Saikkonen (1989) obtained the asymptotic distribution of RS_ψ under $L_2(\gamma,\phi)$ as

$$RS_\psi \xrightarrow{D} \chi_1^2(\lambda_2), \tag{2}$$

where the non-centrality parameter λ_2 is given by $\lambda_2 \equiv \lambda_2(\delta) = \delta' J_{\phi\psi\cdot\gamma} J_{\psi\cdot\gamma}^{-1} J_{\psi\phi\cdot\gamma}\delta$ with $J_{\psi\phi\cdot\gamma} = J_{\psi\phi} - J_{\psi\gamma}J_\gamma^{-1}J_{\gamma\phi}$. Due to this non-centrality parameter, RS_ψ will have power in the model $L(\gamma,\psi,\phi)$ even when $\psi = \psi_0$; and, therefore, the test will have incorrect size. Notice that the crucial quantity is $J_{\psi\phi\cdot\gamma}$ which can be interpreted as the partial covariance between d_ψ and d_ϕ after eliminating the effect of d_γ on d_ψ and d_ϕ. If $J_{\psi\phi\cdot\gamma} = 0$, then the local presence of the parameter ϕ has no effect on RS_ψ.

Turning now to the case of *underspecification*, let the true model be represented by the log-likelihood $L(\gamma,\psi,\phi)$. The alternative $L_1(\gamma,\psi)$ is now underspecified with respect to the nuisance parameter ϕ, leading to the problem of undertesting. In order to derive the asymptotic distribution of RS_ψ under the true model $L(\gamma,\psi,\phi)$, we again consider the local departures $\phi = \phi_0 + \delta/\sqrt{n}$ together with $\psi = \psi_0 + \xi/\sqrt{n}$. It can be shown that [see Bera and Yoon (1991)]

$$RS_{\psi} \xrightarrow{D} \chi_1^2(\lambda_3), \tag{3}$$

where

$$\begin{aligned} \lambda_3 \equiv \lambda_3(\xi, \delta) &= (\delta' J_{\phi\psi\cdot\gamma} + \xi' J_{\psi\cdot\gamma}) J_{\psi\cdot\gamma}^{-1} (J_{\psi\phi\cdot\gamma}\delta + J_{\psi\cdot\gamma}\xi) \\ &= \lambda_1(\xi) + \lambda_2(\delta) + 2\xi' J_{\psi\phi\cdot\gamma}\delta. \end{aligned}$$

Using this result, we can compare the asymptotic local power of the underspecified test with that of the optimal test. It turns out that the contaminated non-centrality parameter $\lambda_3(\xi, \delta)$ may actually increase or decrease the power depending on the configuration of the term $\xi' J_{\psi\phi\cdot\gamma}\delta$.

The problem of overtesting occurs when multi-directional joint tests are applied based on an overstated alternative model. Suppose we apply a joint test for testing hypothesis of the form $H_0 : \psi = \psi_0$ and $\phi = \phi_0$ using the alternative model $L(\gamma, \psi, \phi)$. Let $RS_{\psi\phi}$ be the joint RS test statistic for H_0. To find the asymptotic distribution of $RS_{\psi\phi}$ under overspecification, i.e., when the DGP is represented by the log-likelihood either $L_1(\gamma, \psi)$ or $L_2(\gamma, \phi)$, let us consider the following result, which could be obtained from (1) by replacing ψ with $[\psi', \phi']'$. Assuming correct specification, i.e., under the true model represented by $L(\gamma, \psi, \phi)$ with $\psi = \psi_0 + \xi/\sqrt{n}$ and $\phi = \phi_0 + \delta/\sqrt{n}$,

$$RS_{\psi\phi} \xrightarrow{D} \chi_2^2(\lambda_4), \tag{4}$$

where

$$\lambda_4 \equiv \lambda_4(\xi, \delta) = [\xi' \quad \delta'] \begin{bmatrix} J_{\psi\cdot\gamma} & J_{\psi\phi\cdot\gamma} \\ J_{\phi\psi\cdot\gamma} & J_{\phi\cdot\gamma} \end{bmatrix} \begin{bmatrix} \xi \\ \delta \end{bmatrix}.$$

Using this fact, we can easily find the asymptotic distribution of the overspecified test. Consider testing $H_0 : \psi = \psi_0$ and $\phi = \phi_0$ in $L(\gamma, \psi, \phi)$ where $L_1(\gamma, \psi)$ represents the true model. Under $L_1(\gamma, \psi)$ with $\psi = \psi_0 + \xi/\sqrt{n}$, we obtain by setting $\delta = 0$ in (4)

$$RS_{\psi\phi} \xrightarrow{D} \chi_2^2(\lambda_5), \tag{5}$$

where $\lambda_5 \equiv \lambda_5(\xi) = \xi' J_{\psi\cdot\gamma}\xi$.

Note that the non-centrality parameter $\lambda_5(\xi)$ of the overspecified test $RS_{\psi\phi}$ is identical to $\lambda_1(\xi)$ of the optimal test RS_ψ in (1). Although $\lambda_5 = \lambda_1$, some loss of power is to be expected, as shown in Das Gupta and Perlman (1974), due to the higher degrees of freedom of the joint test $RS_{\psi\phi}$.

Using the result (2), Bera and Yoon (1993) suggested a modification to RS_ψ so that the resulting test is valid in the *local* presence of ϕ. The modified statistic is given by

$$
\begin{aligned}
RS^*_\psi \;\; = \;\; & \frac{1}{n}[d_\psi(\tilde{\theta}) - J_{\psi\phi\cdot\gamma}(\tilde{\theta})J^{-1}_{\phi\cdot\gamma}(\tilde{\theta})d_\phi(\tilde{\theta})]' \\
& [J_{\psi\cdot\gamma}(\tilde{\theta}) - J_{\psi\phi\cdot\gamma}(\tilde{\theta})J^{-1}_{\phi\cdot\gamma}(\tilde{\theta})J_{\phi\psi\cdot\gamma}(\tilde{\theta})]^{-1} \\
& [d_\psi(\tilde{\theta}) - J_{\psi\phi\cdot\gamma}(\tilde{\theta})J^{-1}_{\phi\cdot\gamma}(\tilde{\theta})d_\phi(\tilde{\theta})]. \qquad (6)
\end{aligned}
$$

This new test essentially adjusts the mean and variance of the standard RS_ψ. Bera and Yoon (1993) proved that under $\psi = \psi_0$ and $\phi = \phi_0 + \delta/\sqrt{n}$ RS^*_ψ has a *central* χ^2_1 distribution. Thus, RS^*_ψ has the same asymptotic null distribution as that of RS_ψ based on the correct specification, thereby producing an asymptotically correct size test under locally misspecified model. Bera and Yoon (1993) further showed that for local misspecification the adjusted test is asymptotically equivalent to Neyman's $C(\alpha)$ test and, therefore, shares the optimality properties of the $C(\alpha)$ test. There is, however, a price to be paid for all these benefits. Under the local alternatives $\psi = \psi_0 + \xi/\sqrt{n}$

$$
RS^*_\psi \xrightarrow{D} \chi^2_1(\lambda_6), \qquad (7)
$$

where $\lambda_6 \equiv \lambda_6(\xi) = \xi'(J_{\psi\cdot\gamma} - J_{\psi\phi\cdot\gamma}J^{-1}_{\phi\cdot\gamma}J_{\phi\psi\cdot\gamma})\xi$. Note that $\lambda_1 - \lambda_6 \geq 0$, where λ_1 is given in (1). Result (7) is valid both in the presence or absence of the local misspecification $\phi = \phi_0 + \delta/\sqrt{n}$, since the asymptotic distribution of RS^*_ψ is unaffected by the local departure of ϕ from ϕ_0. Therefore, RS^*_ψ will be less powerful than RS_ψ when there is no misspecification. The quantity

$$
\lambda_7 = \lambda_1 - \lambda_6 = \xi' J_{\psi\phi\cdot\gamma}J^{-1}_{\phi\cdot\gamma}J_{\phi\psi\cdot\gamma}\xi \qquad (8)
$$

can be regarded as the premium we pay for the validity of RS^*_ψ under local misspecification. Two other observations regarding RS^*_ψ are also worth noting. First, RS^*_ψ requires estimation only under the joint null, namely $\psi = \psi_0$ and $\phi = \phi_0$. Given the full specification of the model $L(\gamma, \psi, \phi)$ it is, of course, possible to derive a RS test for $\psi = \psi_0$ in the presence of ϕ. However, that requires MLE of ϕ which could be difficult to obtain in some cases. Second, when $J_{\psi\phi\cdot\gamma} = 0$, $RS^*_\psi = RS_\psi$. In practice this is a very simple condition to check. As mentioned earlier, if this condition is true, RS_ψ is an asymptotically valid test in the local presence of ϕ.

3 Tests for error component model

We consider the following one-way error component model introduced by Lillard and Willis (1978), which combines random individual effects and first order autocorrelation in the disturbance term:

$$\begin{aligned} y_{it} &= x'_{it}\beta + u_{it}, \quad i = 1, 2, \ldots, N, \ t = 1, 2, \ldots, T, \\ u_{it} &= \mu_i + \nu_{it}, \\ \nu_{it} &= \rho\nu_{i,t-1} + \epsilon_{it}, \quad |\rho| < 1, \end{aligned} \tag{9}$$

where β is a $(k \times 1)$ vector of parameters including the intercept, $\mu_i \sim IIDN(0, \sigma^2_\mu)$ is a random individual component, and $\epsilon_{it} \sim IIDN(0, \sigma^2_\epsilon)$. The μ_i and ν_{it} are assumed to be independent of each other with $\nu_{i,0} \sim N(0, \sigma^2_\epsilon/(1 - \rho^2))$. N and T denote the number of individual units and the number of time periods, respectively. For the validity of the tests discussed here, we need to assume that the regularity conditions of Anderson and Hsiao (1982) are satisfied. Also, testing for σ^2_μ involves the issue of the parameter being at the boundary. Although for the nonregular problem of testing at the boundary value, both the likelihood ratio and Wald test statistics do not have their usual asymptotic chi-squared distribution, the RS test statistic does [see, e.g., Bera, Ra and Sarkar (1998)].

Let us set $\theta = (\gamma, \psi, \phi)' = (\sigma^2_\epsilon, \sigma^2_\mu, \rho)'$. Consider the problem of testing for the existence of the random effects $(H_0 : \psi = 0)$ in the presence of serial correlation $(\phi \neq 0)$. To derive our RS^*_ψ, which will now be denoted as RS^*_μ, we note that it is sufficient to consider the scores and the information matrix evaluated at $\theta_0 = (\gamma_0, \psi_0, \phi_0)' = (\sigma^2_\epsilon, 0, 0)'$ because of the block-diagonality of the information matrix involving the β and θ parameters. These quantities have been derived in Baltagi and Li (1991):

$$\frac{\partial L}{\partial \sigma^2_\epsilon} = d_\gamma = -\frac{NT}{2\sigma^2_\epsilon} + \frac{u'u}{2\sigma^4_\epsilon},$$

$$
\begin{aligned}
\frac{\partial L}{\partial \sigma_\mu^2} = d_\mu \equiv d_\psi &= -\frac{NT}{2\sigma_\epsilon^2}\left[1 - \frac{u'(I_N \quad e_T e_T')u}{u'u}\right], \\
\frac{\partial \mu}{\partial \rho} = d_\rho \equiv d_\phi &= NT\left(\frac{u'u_{-1}}{u'u}\right),
\end{aligned} \tag{10}
$$

where I_N is an identity matrix of dimension N, e_T is a vector of ones of dimension T, $u' = (u_{11}, \ldots, u_{1T}, \ldots, u_{N1}, \ldots, u_{NT})$ and u_{-1} is an $(NT \times 1)$ vector containing $u_{i,t-1}$. To simplify notation, here the score for the parameter σ_μ^2 is denoted as d_μ. We will continue to follow this convention for the elements of the information matrix and for expressing our test statistics. Denoting $J = (NT)^{-1}E(-\partial^2 L/\partial\theta\partial\theta')$ evaluated at θ_0, we have

$$
J = \frac{1}{2\sigma_\epsilon^4}\begin{bmatrix} 1 & 1 & 0 \\ 1 & T & \frac{2(T-1)\sigma_\epsilon^2}{T} \\ 0 & \frac{2(T-1)\sigma_\epsilon^2}{T} & \frac{2(T-1)\sigma_\epsilon^4}{T} \end{bmatrix}.
$$

This implies that

$$
\begin{aligned}
J_{\mu\rho\cdot\gamma} = J_{\psi\phi\cdot\gamma} &= \frac{T-1}{T\sigma_\epsilon^2}, \\
J_{\mu\cdot\gamma} = J_{\psi\cdot\gamma} &= \frac{T-1}{2\sigma_\epsilon^4}, \\
J_{\rho\cdot\gamma} = J_{\phi\cdot\gamma} &= \frac{T-1}{T},
\end{aligned} \tag{11}
$$

where γ stands for the parameter σ_ϵ^2. Since $J_{\mu\rho\cdot\gamma} > 0$, indicating the asymptotic positive correlation between the scores d_μ and d_ρ, the one-directional test for the random effects reported in Breusch and Pagan (1980) is not valid asymptotically in the presence of serial correlation. For this case our RS_μ^* can be easily constructed, from equation (6), as

$$
RS_\mu^* = \frac{NT(A+2B)^2}{2(T-1)(1-\frac{2}{T})}, \tag{12}
$$

where A and B denote, as in Baltagi and Li (1991),

$$
A = 1 - \frac{\tilde{u}'(I_N \quad e_T e_T')\tilde{u}}{\tilde{u}'\tilde{u}},
$$

and

$$
B = \frac{\tilde{u}'\tilde{u}_{-1}}{\tilde{u}'\tilde{u}}.
$$

Note that $\tilde{u}$ are the OLS residuals from the standard linear model $y_{it} = x_{it}'\beta + u_{it}$ without the random effects and serial correlation. Also notice that A and B are closely related to the estimates of the scores d_μ and d_ρ, respectively. It is easy to see that the RS^*_μ adjusts the conventional RS statistic given in Breusch and Pagan (1980), i.e.,

$$RS_\mu = \frac{NTA^2}{2(T-1)}, \tag{13}$$

by correcting the mean and variance of the score d_μ for its asymptotic correlation with d_ρ.

To see the behavior of RS_μ let us first consider the case of complete misspecification, i.e., $\sigma^2_\mu = 0$ but $\rho \neq 0$. Using (2) and (11), the noncentrality parameter of RS_μ for this case is:

$$\lambda_2(\rho) = \delta' J_{\rho\mu\cdot\gamma} J^{-1}_{\mu\cdot\gamma} J_{\mu\rho\cdot\gamma}\delta = 2\rho^2 \frac{T-1}{T^2}, \tag{14}$$

where for simplicity we use ρ in place of δ. In this case, the use or RS_μ will lead to rejection of the null hypothesis $\sigma^2_\mu = 0$ too often. For local departures RS^*_μ will not have this drawback when $\rho \neq 0$ since under $\sigma^2_\mu = 0$, RS^*_μ will have a *central* χ^2 distribution. Let us now consider the underspecification situation i.e., when we have both $\sigma^2_\mu > 0$ and $\rho \neq 0$, and we use RS_μ to test $H_o : \sigma^2_\mu = 0$. From (1), (3) and (11), we see that the change in the noncentrality parameter of RS_μ due to nonzero ρ is given by

$$\begin{aligned} \lambda_3(\xi, \delta) - \lambda_1(\xi) &= \lambda_2(\rho) + 2\xi' J_{\mu\rho\cdot\gamma}\delta \\ &= \rho^2 \frac{2(T-1)}{T^2} + 2\sigma^2_\mu \rho \frac{T-1}{T\sigma^2_\epsilon} \\ &= \frac{2(T-1)}{T}\left[\frac{\rho^2}{T} + \frac{\sigma^2_\mu \rho}{\sigma^2_\epsilon}\right], \end{aligned} \tag{15}$$

where we use σ^2_μ in place of ξ. From (15), it is easy to see that when $\rho > 0$, the presence of autocorrelation will add power to RS_μ; but when $\rho < 0$ it can loose power if the individual effect is very high and σ^2_ϵ is low. In this situation, the noncentrality parameter of RS^*_μ is not affected. From (7) and (11), the noncentrality parameter of RS^*_μ under $\sigma^2_\mu > 0$ and $\rho \neq 0$, can be written as

$$\begin{aligned} \lambda_6 \equiv \lambda_6(\sigma^2_\mu) &= \sigma^4_\mu (J_{\mu\cdot\gamma} - J_{\mu\rho\cdot\gamma} J^{-1}_{\rho\cdot\gamma} J_{\rho\mu\cdot\gamma}) \\ &= \sigma^4_\mu \left[\frac{T-1}{2\sigma^4_\epsilon} - \frac{(T-1)^2}{T^2\sigma^4_\epsilon}\frac{T}{T-1}\right] \\ &= \frac{\sigma^4_\mu}{\sigma^4_\epsilon}(T-1)(\frac{1}{2} - \frac{1}{T}), \end{aligned} \tag{16}$$

which does not depend on ρ. There is, however, a cost in applying RS_μ^* when ρ is indeed zero. From (8) the cost is

$$\lambda_7 \equiv \lambda_7(\sigma_\mu^2) = \sigma_\mu^4 J_{\mu\rho\cdot\gamma} J_{\rho\cdot\gamma}^{-1} J_{\rho\mu\cdot\gamma} = \frac{\sigma_\mu^4}{\sigma_\epsilon^4}\frac{T-1}{T}. \tag{17}$$

Note that this cost is present only under $\sigma_\mu^2 > 0$. That is, there is a cost only in terms of the power of RS_μ^*; the size is unaffected. Later we will provide an interesting interpretation of this cost of RS_μ^* in terms of the behavior of the unadjusted test RS_ρ under $\sigma_\mu^2 > 0$.

As mentioned before, Baltagi and Li (1995) derived a RS test for serial correlation in the presence of random individual effects. Naturally, the test requires MLE of σ_μ^2. Our procedure gives a simple test for serial correlation in the random effects model. In this situation RS_ρ^* is obtained simply by switching σ_μ^2 and ρ to yield

$$RS_\rho^* = \frac{NT^2(B + \frac{A}{T})^2}{(T-1)(1-\frac{2}{T})}. \tag{18}$$

If we assume that the random effects are absent throughout, then RS_ρ^* in (18) reduces to

$$RS_\rho = \frac{NT^2B^2}{T-1}. \tag{19}$$

This conventional RS statistic (19) is also given in Baltagi and Li (1991).

As we have done for RS_μ, we can also study the performance of RS_ρ under various misspecifications. When there is complete misspecification, i.e., when $\rho = 0$ but $\sigma_\mu^2 > 0$, the noncentrality parameter of RS_ρ is

$$\lambda_2(\sigma_\mu^2) = \xi' J_{\mu\rho\cdot\gamma} J_{\rho\cdot\gamma}^{-1} J_{\rho\mu\cdot\gamma}\xi = \frac{\sigma_\mu^4}{\sigma_\epsilon^4}\frac{T-1}{T}, \tag{20}$$

where we have used σ_μ^2 in place of ξ. Therefore, RS_ρ will reject $H_0 : \rho = 0$ too often when $\sigma_\mu^2 > 0$. Similarly, when there is underspecification, i.e, $\rho \neq 0$ with $\sigma_\mu^2 > 0$, the change in the noncentrality parameter due to the presence of the random effect, is

$$\begin{aligned} \lambda_3(\xi,\delta) - \lambda_1(\delta) &= \lambda_2(\sigma_\mu^2) + 2\delta' J_{\rho\mu\cdot\gamma}\xi \\ &= \frac{T-1}{T}\frac{\sigma_\mu^2}{\sigma_\epsilon^2}\left[\frac{\sigma_\mu^2}{\sigma_\epsilon^2} + 2\rho\right]. \end{aligned} \tag{21}$$

Therefore, we have an increase in (or a possible loss of) power when $\rho > 0$ (or $\rho < 0$). The noncentrality parameter of RS_ρ^* will not be affected at all under $\sigma_\mu^2 > 0$. On the other hand, we do, however, pay a penalty when $\sigma_\mu^2 = 0$ and we use the adjusted test RS_ρ^*. The penalty is

$$\lambda_7(\rho) = \rho^2 J_{\rho\mu\cdot\gamma} J_{\mu\cdot\gamma}^{-1} J_{\mu\rho\cdot\gamma} = 2\rho^2 \frac{T-1}{T^2}. \tag{22}$$

Due to this factor the power of RS_ρ^* will be somewhat less than that of RS_ρ when σ_μ^2 is indeed zero; the size of RS_ρ^*, however, remains unaffected. It is very interesting to note that

$$\lambda_7(\rho) = \lambda_2(\rho) \tag{23}$$

given in (14). Similarly, from (17) and (20)

$$\lambda_7(\sigma_\mu^2) = \lambda_2(\sigma_\mu^2). \tag{24}$$

An implication of (23) is that the cost of using RS_ρ^* when $\sigma_\mu^2 = 0$ is the same as the cost of using RS_μ when $\rho \neq 0$. Similarly, (24) implies that the loss in the noncentrality parameter of RS_μ^* when $\rho = 0$ is equal to the unwanted gain in the noncentrality parameter of RS_ρ when $\sigma_\mu^2 > 0$. We will explain these seemingly unintuitive phenomena after we find a relationship among the four statistics, RS_μ^*, RS_μ, RS_ρ^*, and RS_ρ. It should be noted that the equalities of equations (23) and (24) are not specific for the error component model, and they hold in general. This can be seen by comparing $\lambda_2(\delta)$ below (2) with λ_7 in equation (8), where ψ swaps position with ϕ and ξ is replaced by δ.

Baltagi and Li (1991, 1995) derived a joint RS test for serial correlation and random individual effects which is given by

$$RS_{\mu\rho} = \frac{NT^2}{2(T-1)(T-2)}[A^2 + 4AB + 2TB^2]. \tag{25}$$

Under the joint null $\sigma_\mu^2 = \rho = 0$, $RS_{\mu\rho}$ is asymptotically distributed as χ_2^2. Use of this will result in a loss of power compared with the proper one-directional tests when only one of the two forms of misspecification is present, as we noted while discussing (5). For example, when $\rho = 0$ and $\sigma_\mu^2 > 0$, the noncentrality parameter of both RS_μ and $RS_{\mu\rho}$ is [see (1) and (5)]

$$\lambda_1(\sigma_\mu^2) = \sigma_\mu^4 J_{\mu\cdot\gamma} = \frac{\sigma_\mu^4}{\sigma_\epsilon^4}\frac{T-1}{2}. \tag{26}$$

Since for RS_μ and $RS_{\mu\rho}$ we will use respectively χ_1^2 and χ_2^2 critical values, $RS_{\mu\rho}$ will be less powerful. An interesting result follows from (12), (13), (18), (19) and (25), namely,

$$RS_{\mu\rho} = RS_\mu^* + RS_\rho = RS_\mu + RS_\rho^*, \tag{27}$$

i.e., the two directional RS test for σ_μ^2 and ρ can be decomposed into the sum of the adjusted one-directional test of one type of alternative and the unadjusted form for the other one. Using (27) we can easily explain some of our earlier observations. First, consider the identities in (23) and (24). From (27), we have

$$RS_\rho - RS_\rho^* = RS_\mu - RS_\mu^*. \tag{28}$$

Let us consider the case of $\sigma_\mu^2 = 0$ and $\rho \neq 0$. Then the left-hand side of (28) represents the "penalty" of using RS_ρ^* (instead of RS_ρ) while the right-hand side amounts to the "cost" of using RS_μ. (28) implies that these penalty and cost should be the same, as noted in (23). A reverse argument explains (24). Secondly, the local presence of ρ (or σ_μ^2) has no effect on RS_μ^* (or RS_ρ^*); therefore, from (5) and (27), we can clearly see why the noncentrality parameter of $RS_{\mu\rho}$ will be equal to that of RS_ρ (or RS_μ) when $\sigma_\mu^2 = 0$ (or $\rho = 0$).

So far we have considered only two sided-tests for $H_0 : \sigma_\mu^2 = 0$. Since $\sigma_\mu^2 \geq 0$, it is natural to consider one-sided tests, and it is expected that will lead to more powerful tests. Within our framework, it is easy to construct appropriate one-sided tests by taking the *signed* square root of our earlier two-sided statistics, RS_μ and RS_μ^*. We will denote these one-sided test statistics as RSO_μ and RSO_μ^*, and they are given by

$$RSO_\mu = -\sqrt{\frac{NT}{2(T-1)}}A$$

and

$$RSO_\mu^* = -\sqrt{\frac{NT}{2(T-1)(1-\frac{2}{T})}}(A-2B)$$

The negative sign is due to the fact that $\partial L/\partial\sigma_\mu^2 = -(NT/2\sigma_\varepsilon^2)A$ and the one-sided tests are based on this score function or its adjustment. Under $H_0 : \sigma_\mu^2 = 0$, the adjusted test RSO_μ^* will be asymptotically distributed as $N(0,1)$. The unadjusted RSO_μ will be asymptotically normal but with a nonzero mean $\sqrt{2(T-1)/T^2}\rho$ when $\rho \neq 0$ as can be seen from (14). The statistic RSO_μ^* was first suggested by Honda (1985) and its finite sample properties have been investigated by Baltagi, Chang and Li (1992). Similar one-sided versions for RS_ρ and RS_ρ^* can also be used. However, in practice, the direction of serial correlation is rarely known for sure for the one-directional tests to be more powerful. It is easy to see that the one directional tests will not satisfy the equality in (28). In our empirical illustrations below and in the Monte Carlo study we also use these one-sided tests and study their comparative finite sample performance.

4 Empirical illustrations

In this section we present two empirical examples that illustrate the usefulness of the proposed tests. The first is based on a data set used by Greene (1983, 2000). The equation to be estimated is a simple, log-linear cost function:

$$\ln C_{it} = \beta_0 + \beta_1 \ln R_{it} + u_{it},$$

where R_{it} is measured as output of firm i in year t in millions of kilowatt-hours, and C_{it} is the total generation cost in millions of dollars, $i = 1, 2, \ldots, 6$, and $t = 1, 2, 3, 4$. The second example is based on the well-known Grunfeld (1958) investment data set for five US manufacturing firms measured over 20 years which is frequently used to illustrate panel issues. It has been used in the illustration of misspecification tests in the error-component model in Baltagi et al. (1992), and in recent books such as those by Baltagi (1995, p.20) and Greene (2000, p.592). The equation to be estimated is a panel model of firm investment using the real value of the firm and the real value of capital stock as explanatory variables:

$$I_{it} = \beta_0 + \beta_1 F_{it} + \beta_2 C_{it} + u_{it},$$

where I_{it} denotes real gross investment for firm i in period t, F_{it} is the real value of the firm and C_{it} is the real value of the capital stock, $i = 1, 2, \ldots, 5$, and $t = 1, 2, \ldots, 20$.

We estimated the parameters of both models by OLS and implemented the following seven tests based on OLS residuals: the Breusch-Pagan test for random effects (RS_μ), the proposed modified version (RS_μ^*), the LM serial correlation test (RS_ρ), the corresponding modified version (RS_ρ^*), the joint test for serial correlation and random effects ($RS_{\mu\rho}$), and the two one-sided tests for random effects (RSO_μ and RSO^*_μ). The test statistics for both examples are presented in Table 1; the p-values are given in parentheses.

All of the test statistics were computed individually, and the equality in (27) is satisfied for both data sets. In the example based on Greene's data the unmodified tests for serial correlation (RS_ρ) and for random effects (RS_μ to some extent, and RSO_μ quite strongly) reject the respective null hypothesis of no serial correlation and no random effects, and the omnibus test rejects the joint null. But our modified tests suggest that in this example the problem seems to be serial correlation rather the presence of both effects. For Grunfeld's data, applications of our modified tests point to the presence of the other effect. The unmodified tests soundly reject their corresponding null hypotheses. The modified versions of the random effect tests (RS^*_μ and RSO^*_μ) also reject the null but the modified serial correlation test (RS^*_ρ) barely rejects the null at the 5% significance level. It is interesting to note the substantial reduction of the autocorrelation test statistic, from 73.351 to 3.712. So in this example the misspecification can be thought to come from the presence of random effects rather than serial correlation. As expected, the joint test statistic is highly significant.

In spite of the small sample size of the data sets, these examples seem to illustrate clearly the main points of the paper: the proposed modified versions of the test are more informative than a test for serial correlation or random effect that ignores the presence of the other effect. In the first case, serial correlation spuriously induces rejection of the no-random effects hypothesis, and in the second case the opposite happens: the presence of a random effect induces rejection of the no-serial correlation hypothesis. The joint test $RS_{\mu\rho}$ rejects the joint null but is not informative about the direction of the misspecification.

$RS_{\mu\rho}$ provides a correct measure of the joint effects of individual component and serial correlation. The main problem is how to decompose this measure to get an idea about the true departure(s). From a practical standpoint if $RS_{\mu\rho} = RS_\mu + RS_\rho$ does not hold, that should be an indication of the presence of an interaction between random effects and serial correlation; and the unadjusted statistics RS_μ and RS_ρ will be contaminated by the presence of other departures. For example, for the Grunfeld data

$$RS_\mu + RS_\rho - RS_{\mu\rho} = RS_\mu - RS^*_\mu = RS_\rho - RS^*_\rho = 69.638.$$

This provides a measure of the interaction between σ^2_μ and ρ, and is also equal to the correction needed for each unadjusted test.

It is important to emphasize that the implementation of the modified tests is based solely on OLS residuals. It could be argued that a more efficient test procedure could be based on the estimation of a general model that allows for both serial correlation and random effects, and then the tests of the hypotheses of no-serial correlation and no-random effects as restrictions on this general model (either jointly or individually) could be carried out. But this would require the maximization of a likelihood function whose computational tractability is substantially more involved than computing simple OLS residuals. Hsiao (1986, p.55) commented that the "computation of the MLE is very complicated." For more on the estimation issues of the error component model with serial correlation see Baltagi (1995, pp. 18-19), Majumder and King (1999) and Phillips (1999).

5 Monte Carlo results

In this section we present the results of a Monte Carlo study to investigate the finite sample behavior of the tests. To facilitate comparison with existing results we follow a structure similar to the one adopted by Baltagi, et al. (1992) and Baltagi and Li (1995).

The model was set as a special case of (9):

$$
\begin{aligned}
y_{it} &= \alpha + \beta x_{it} + u_{it}, \quad i = 1, 2, \ldots, N, \ t = 1, 2, \ldots, T \\
u_{it} &= \mu_i + v_{it}, \\
v_{it} &= \rho v_{i,t-1} + \varepsilon_{it}, \qquad |\rho| < 1,
\end{aligned}
$$

where $\alpha = 5$ and $\beta = 0.5$. The independent variable x_{it} was generated following Nerlove (1971):

$$x_{it} = 0.1t + 0.5x_{i,t-1} + \omega_{it},$$

where ω_{it} has the uniform distribution on $[-0.5, 0.5]$. Initial values were chosen as in Baltagi, et al. (1992). Let $\sigma^2, \sigma^2_\mu, \sigma^2_v$ and σ^2_ε represent the variances of u_{it}, μ_i, v_{it} and ε_{it}, respectively, and let $\tau = \sigma^2_\mu/\sigma^2$, which represents the "strength" of the random effects. Here, $\sigma^2 = \sigma^2_\mu + \sigma^2_v$, and we set $\sigma^2 = 20$. τ and ρ were allowed to take seven different values $(0, 0.05, 0.1, 0.2, 0.4, 0.6, 0.8)$, and three different sample sizes (N, T) were considered: $(25, 10), (25, 20)$ and $(50, 10)$. Since for each i, v_{it} follows an AR(1) process, $\sigma^2_v = \sigma^2_\varepsilon/(1-\rho^2)$.

Then, according to this structure, the random effect term and the innovation were generated as:

$$\begin{aligned} \mu_i &\sim IIDN(0, 20(1-\tau)) \\ \varepsilon_{it} &\sim IIDN(0, 20(1-\tau)(1-\rho^2)). \end{aligned}$$

For each sample size the model described above was generated 1,000 times under different parameter settings. Therefore, the maximum standard errors of the estimates of the size and powers would be $\sqrt{0.5(1-0.5)/1000} \simeq 0.015$. In each replication the parameters of the model were estimated using OLS, and seven test statistics, namely, RS_μ, RS_μ^*, RS_ρ, RS_ρ^*, $RS_{\mu\rho}$, RSO_μ and RSO_μ^* were computed. The tables and graphs are based on the nominal size of 0.05. Our simulation study was quite extensive; we carried out experiments for all possible parameter combinations for the three sample sizes. We present here only a portion of our extensive tables and graphs; the rest is available from the authors upon request.

Calculated statistics under $\tau = \rho = 0$ were used to estimate the empirical sizes of the tests and to study the closeness of their distributions to χ^2 through $Q-Q$ plots and the Kolmogorov-Smirnov test. From Table 2 we note that both RS_μ and RS_μ^* have similar empirical sizes, but these are below the nominal size 0.05 for $N = 25, T = 10$ and $N = 50, T = 10$. The results for $RS_\rho, RS_\rho^*, RS_{\mu\rho}$ are not good. All of them reject the null too frequently, but the empirical sizes improve as we increase N or T. Comparing the performances of RS_ρ and RS_ρ^*, we notice that RS_ρ^* has somewhat better size properties. As expected, the one-sided tests RSO_μ and RSO_μ^* have larger empirical sizes than their two-sided counterparts. Overall, except for a couple of cases, the size performance of all tests are within one standard errors of the nominal size 0.05.

The results of Table 2 are consistent with the Q-Q plots in Figure 1 for $N = 25$, $T = 10$. To save space figures for the other two combinations of (N, T) are not included. We also do not present the figures for the joint and one-sided tests, since they resemble those reported for the other tests. From the plots note that the empirical distributions of the test statistics diverge from that of the χ_1^2 at the right tail parts. For RS_μ and RS_μ^* the points are below the 45^o line, particularly for the high values, and that leads to sizes being *below* 0.05 as we just noted from Table 2. However, the number of points (out of 1,000) that are far away from the 45^o line at the tail parts are not many. For RS_ρ and RS_ρ^* we observe a higher degree of departure from the 45^o line in the opposite direction, and this leads to much higher sizes of the tests. Results from the Kolmogorov-Smirnov test, not reported here, accept the null hypothesis of the *overall* distribution being the same as χ^2 for the first five, and standard normal for the last two statistics. For the true sizes of the tests, however, it is only the tail part, not the overall distribution, that matters.

Let us now turn into the performance of tests in terms of power. For $N = 25$ and $T = 10$, the estimated rejection probabilities of the tests are reported in Table 3, and are also illustrated in Figures (2a)-(2d). The results for $\tau = \rho = 0.08$ are not reported since in most cases the rejection probabilities were one or very close to one. Moreover, our adjusted tests are designed for locally misspecified alternatives clsoe to $\tau = \rho = 0.0$, and the main objective of our Monte Carlo study is to investigate the performance of our suggested tests in the neighborhood of $\tau = \rho = 0.0$. Let us first concentrate on RS_μ, RS_μ^*, RSO_μ and RSO_μ^* which are designed to test the null hypothesis $H_0 : \sigma_\mu^2 = 0$. When $\rho = 0$, RS_μ and RSO_μ are, respectively, the two- and one-sided optimal tests. This is clearly evident looking at all the rows in Table 3 with $\rho = 0$; RSO_μ has the highest powers among all the tests and RS_μ just trails behind it. The power of RS_μ^* is less than that of RS_μ when $\rho = 0$. The losses in power are, however, not very large, as can also be seen from Figure 2(a). When τ exceeds 0.2 (or σ_μ^2 exceeds 4, since we set $\sigma_\mu^2 = 20\tau$) both tests have power equal to 1. The amount of loss in using RS_μ^* when $\rho = 0$ was characterized by (17) in terms of the decrease in the noncentrality parameter. That loss increases with $\tau(\sigma_\mu^2)$. However, the overall power of RS_μ^* is guided by the noncentrality parameter in (16):

$$\lambda_6(\sigma_\mu^2) = \frac{\sigma_\mu^4}{2\sigma_\epsilon^4}(T-1) - \frac{\sigma_\mu^4}{\sigma_\epsilon^4}\frac{T-1}{T},$$

where the second term is the amount of penalty in using RS_μ^* when $\rho = 0$, and it is given in (17). Since the first term dominates, the relative value of the loss is negligible. While RS_μ^* and RSO_μ^* do not sustain much loss in power when $\rho = 0$, we notice some problems in RS_μ and RSO_μ when $\sigma_\mu^2 = 0$ but $\rho \neq 0$. RS_μ and RSO_μ reject $H_0 : \sigma_\mu^2 = 0$ too frequently. For example, when $\tau = 0$ (i.e., $\sigma_\mu^2 = 0$) and $\rho = 0.4$, RS_μ and RSO_μ have rejection probabilities 0.847 and 0.888, respectively. For other values of ρ the proportion of rejections of $\sigma_\mu^2 = 0$ (when it is true) for RS_μ can be seen in Figure 2(b). As we discussed in Section 3, this unwanted rejection probabilities is due to the noncentrality parameter $\lambda_2(\rho)$ in (14), which is "purely" a funtion of the degree of departure of ρ from zero. RS_μ^*

and RSO_μ^* also have some unwanted rejection probabilities but the problem is less severe. For the above case of $\tau = 0$ and $\rho = 0.4$ the rejection probabilities for RS_μ^* and RSO_μ^* are, respectively, 0.325 and 0.354. Figure 2(b) gives the power of RS_μ^* when $\tau = 0$ for different values of ρ. As we mentioned earlier, RS_μ^* and RSO_μ^* are designed to be robust only under local misspecification, i.e, for low values of ρ. From that point of view, they do a very good job—their performances deteriorate only when ρ takes high values. Now by directly comparing the one- and two-sided tests for $H_0 : \sigma_\mu^2 = 0$, we note that the former has higher rejection probabilities except for a few cases when $\tau = 0.0$. For these cases the score $\partial L/\partial\sigma_\mu^2$ takes large negative values and that leads to acceptance of H_0 when one-sided tests are used and rejection of H_0 when we use the two-sided test. Note that $RSO_\mu = sign\sqrt{RS_\mu}$ rejects H_0 if $RSO_\mu > 1.645$ while using RS_μ rejection occurs if $RS_\mu > 3.84$ which exceeds 1.645^2.

From Table 3 and Figure 2(c), we note that when $\tau > 0$, an increase in $\rho(> 0)$ enhances the rejection probabilities of RS_μ. For example, when $\tau = 0.05$ the rejection probabilities of RS_μ for $\rho = 0.0$ and 0.2 are, respectively, 0.344 and 0.734. This can be explained using the expression (15), which gives the changes in the noncentrality parameter of RS_μ due to ρ. From (16) we see that the noncentrality parameter of RS_μ^* does not depend on ρ. This result is, of course, valid only asymptotically and for local departures of ρ from zero. Figure 2(d) shows that there is some uniform gain in rejection probabilities of RS_μ^* only when $\rho = 0.4$. For smaller values of ρ, the rejection probabilities sometimes even decrease but are always close to values for the case $\rho = 0$.

As we indicated earlier there could be some *loss* of power of RS_μ when $\rho < 0$. We performed a small-scale experiment for this case, results of which are reported in Table 4. First note that when $\tau = 0$, an increase in the absolute value of ρ leads to an increase in the size of RS_μ. For example, when $N = 25$, $T = 10$ and $\tau = 0$, the rejection frequencies for $\rho = 0$ and $\rho = -0.4$ are, respectively, 0.047 and 0.573. This is due to the noncentrality parameter (14) which is a function of ρ^2. When $\tau > 0$ ($\sigma_\mu^2 > 0$), the changes in the noncentrality parameter could be negative, and there could be a substantial loss in power of RS_μ. For instance, for the above (25,10) sample size combinations, and $\tau = 0.05$, the powers of RS_μ, for $\rho = 0.0$ and -0.4 are, respectively, 0.344 and 0.039. RS_μ^* does not suffer from these detrimental effects as we see from Table 4. Its size remains small for all $\rho < 0$, and power even increases as the absolute value of ρ becomes larger.

In a similar way, we can explain the behavior of RS_ρ and RS_ρ^* using Table 3 and Figures

3(a)-3(d). From Table 3 we note that, as expected, when $\sigma_\mu^2 = 0$, RS_ρ has the highest powers among all the tests. The powers of RS_ρ^* are very close to those of RS_ρ. Therefore, the premium we pay for the wider validity of RS_ρ^* is minimal.

The real benefit of RS_ρ^* is noticed when $\rho = 0$ but $\tau > 0$; the performance of RS_ρ^* is quite remarkable, as can be seen from Figure 3(b). RS_ρ rejects $H_0 : \rho = 0$ too often, whereas, quite correctly, RS_ρ^* does not reject H_0 so often. For example, when $\tau = 0.2$ and $\rho = 0$, the rejection proportions for RS_ρ and RS_ρ^* are 0.802 and 0.042, respectively. Even when we increase τ to 0.6, the rejection proportion for RS_ρ^* is only 0.045, whereas RS_ρ rejects 100% of the time. In a way, RS_ρ^* is doing more than it is designed to do, that is, not rejecting $\rho = 0$ when ρ is indeed zero even for *large* values of τ.

From Figure 3(c), we observe that the power of RS_ρ is strongly affected by the presence of random effects, while there is virtually no effect on the power of RS_ρ^* as seen from Figure 3(d) even for large values of τ. This performance of RS_ρ^* is exceptionally good. For negative values of ρ in Table 4, we see that the presence of τ has a less detrimental effect on RS_ρ^*. For example, when $\rho = -0.10$, the rejection probabilities of RS_ρ are 0.396 and 0.184 for $\tau = 0.0$ and 0.05, respectively; for the same situations, the powers of RS_ρ^* are, respectively, 0.346 and 0.314.

Comparing the performance of RS_ρ^* and RS_μ^*, we see that the former is even more "robust" in the presence of τ, both in terms of size and power, than is the latter in the presence of serial correlation. To see this from a theoretical point of view, let us consider (17) and (22), which are, respectively, the penalties of using RS_μ^* and RS_ρ^*. From (17), $\frac{\sigma_\mu^4}{\sigma_\epsilon^4}\frac{T-1}{T}$, the penalty in using RS_μ^*, also depends on ρ through $\sigma_\epsilon^2 = 20(1-\tau)(1-\rho^2)$, while (22), $2\rho^2(T-1)/T^2$, is a function of ρ only and is of smaller magnitude in terms of T.

Finally, we discuss briefly the performance of the joint statistic $RS_{\mu\rho}$ in the light of our results (4) and (5). This test is optimal when $\sigma_\mu^2 > 0$ and $\rho \neq 0$. As we can see from Table 3, in this situation $RS_{\mu\rho}$ has the highest power most of the time. However, when the departure from $\sigma_\mu^2 = 0, \rho = 0$ is one-directional (say, $\sigma_\mu^2 > 0, \rho = 0$), RS_μ and $RS_{\mu\rho}$ have the same non-centrality parameter [see (26)]. Since $RS_{\mu\rho}$ and RS_ρ use the χ_2^2 and χ_1^2 tests, respectively, there will be a loss of power in using $RS_{\mu\rho}$. For example, when $\tau = 0.10$ and $\rho = 0$, the powers for RS_μ and $RS_{\mu\rho}$ are 0.752 and 0.702, respectively. Similarly, when $\tau = 0, \rho = 0.2$, the power of RS_ρ and $RS_{\mu\rho}$ are respectively, 0.869 and 0.818. These results are consistent with those of Baltagi and Li (1995). Although $RS_{\mu\rho}$ has overall good power, it cannot help to identify the exact source of misspecification when there is only a one-directional departure.

The qualitative performance of all the tests do not change when we increase the sample sizes to $N = 25, T = 20$, and $N = 50, T = 10$ and they further illustrate the usefulness of our modified tests. These results are not presented but are available from the authors upon request.

6 Conclusions

In this paper we have proposed some simple tests, based on OLS residuals for random effects in the presence of serial correlation, and for serial correlation allowing for the presence of random effects. These tests are obtained by adjusting the existing test procedures. We have investigated the finite sample size and power performance of these and some of the available tests through a Monte Carlo study. We have also provided some empirical examples. The Monte Carlo study, along with the examples, clearly show the usefulness of our procedures to identify the exact source(s) of misspecification. One drawback of our methodology is that we allow for only *local* misspecification. For non-local departures, efficient tests could be obtained after estimating full model(s) by maximum likelihood; that, however, will loose the simplicity of our tests using only OLS residuals.

Acknowledgements

We would like to thank the co-editor A. Ronald Gallant, an associate editor and two anonymous referees for many pertinent comments that helped us to improve the paper. Thanks are also due to Miki Naoko for her help in preparing the manuscript. An earlier version of this paper was presented at Texas A&M University, March 1996; the Midwest Econometric Group Meeting, the University of Wisconsin at Madison, November 1996; the Economics seminar at University of San Andres, Argentina, November 1997; and the Annual Meeting of the Argentine Association of Political Economy, Bahia Blanca, Argentina. We wish to thank the participants and Badi Baltagi for helpful comments and discussion. However, we retain responsibility for any remaining errors.

References

Andereson, T.W and C. Hsiao, 1982, Formulation and estimation of dynamic models using panel data, Journal of Econometrics, 18, 47-82.

Balestra, P. and M. Nerlove, 1966, Pooling cross-section and time-series data in the estimation of a dynamic model: The demand for natural gas, Econometrica, 34, 585–612.

Baltagi, B., 1995, Econometric Analysis of Panel Data, New York: Wiley & Sons.

Baltagi, B., Y. Chang and Q. Li, 1992, Monte Carlo results on several new and existing tests for the error component model, Journal of Econometrics, 54, 95-120.

Baltagi, B. and Q. Li, 1991, A joint test for serial correlation and random individual effects, Statistics & Probability Letters, 11, 277–280.

Baltagi, B. and Q. Li, 1995, Testing AR(1) against MA(1) disturbances in an error component model, Journal of Econometrics, 68, 133-151.

Bera, A.K. and C. M. Jarque, 1982, Model specification tests: A simultaneous approach, Journal of Econometrics, 20, 59-82.

Bera, A.K. and M.J. Yoon, 1991, Specification testing with misspecified alternatives, Bureau of Economic and Business Research Faculty Working Paper 91-0123, University of Illinois, and presented at the Econometric Society Winter Meeting, Washington DC, December 1990.

Bera, A.K.and M.J. Yoon, 1993, Specification testing with locally misspecified alternatives, Econometric Theory, 9, 649-658.

Bera, A.K., S.S. Ra and N. Sarkar, 1998, Hypothesis testing for some nonregular cases in econometrics, In S. Chakravorty, D. Coondoo and R. Mukherjee, Ed., Econometrics: Theory and Practice, New Delhi: Allied Publishers, pp. 319-351.

Breusch, T.S. and A.R. Pagan, 1980, The Lagrange multiplier test and its applications to model specification in econometrics, Review of Economic Studies, 47, 239–253.

Das Gupta, S. and M.P. Perlman (1974), Power of the noncentral F-test: Effect of additional variate on Hotteling's T^2-test, Journal of the American Statistical Association, 69, 174-180.

Davidson, R. and J.G. MacKinnon, 1987, Implicit alternatives and local power of test statistics, Econometrica, 55, 1305-1329.

Greene, W., 1983, Simultaneous estimation of factor substitution, economies of scale and non-neutral technical change, In A. Dogramaci, ed., Econometric Analyses of Productivity, Boston: Kluwer-Nijoff, pp. 121-144.

Greene, W., 2000, Econometric Analysis, 4th Edition, New Jersey: Prentice Hall.

Grunfeld, Y., 1958, The Determinants of Corporate Investment, Unpublished Ph.D thesis, Department of Economics, University of Chicago.

Grunfeld, Y. and Z. Griliches, 1960, Is aggregation necessarily bad?, Review of Economics and Statistics, 42, 1-13.

Haavelmo, T., 1944, The probability approach in econoemtrics, Supplement to Econometrica, 12.

Hillier, G.H., 1991, On multiple diagnostic procedures for the linear model, Journal of Econometrics, 47, 47-66.

Honda, Y., 1985, Testing the error component model with non-normal disturbances, Review of Economic Studies, 52, 681-690.

Hsiao, C., 1986, Analysis of Panel Data, Cambridge: Cambridge University Press.

Lillard, L.A. and R.J. Willis, 1978, Dynamic aspects of earning mobility, Econometrica, 46, 985-1012.

Majunder, A.K., and M.L. King, 1999, Estimation and testing of a regression with a serially correlated error component, Working Paper, Department of Econometrics and Business Statistics, Monash University, Australia.

Nerlove, M., 1971, Further evidence of the estimation of dynamic economic relations from a time-series of cross-sections, Econometrica, 39, 359-382.

Phillips, R.F., 1999, Estimation of mixed-effects and error components models with an AR(1) component: Some new likelihood maximization procedures and Monte Carlo evidence, Working Paper, Department of Economics, George Washington University.

Rao, C.R., 1948, Large sample tests of statistical hypothesis concerning several parameters with applications to problems of estimation, Proceedings of the Cambridge Philosophical Society, 44, 50-57.

Saikkonen, P., 1989, Asymptotic relative efficiency of the classical test statistics under misspecification, Journal of Econometrics, 42, 351-369.

Welsh, A. H., 1996, Aspects of Statistical Inference, New York: John Wiley & Sons.

Table 1

Empirical illustration

Tests for random effects and serial correlation

Data	RS_μ	RS_μ^*	RS_ρ	RS_ρ^*	$RS_{\mu\rho}$	RSO_μ	RSO_μ^*
Greene	5.872	0.269	15.569	9.966	15.838	2.423	0.518
	(0.015)	(0.604)	(0.000)	(0.002)	(0.000)	(0.007)	(0.3020)
Grunfeld	453.822	384.183	73.351	3.712	457.535	21.303	19.605
	(0.000)	(0.000)	(0.000)	(0.054)	(0.000)	(0.000)	(0.000)

Note: p-values are given in parenthesis.

Table 2

Empirical size of tests

(nominal size=0.05)

	Tests						
(N,T)	RS_μ	RS_μ^*	RS_ρ	RS_ρ^*	$RS_{\mu\rho}$	RSO_μ	RSO_μ^*
(25,10)	0.047	0.048	0.087	0.072	0.062	0.045	0.051
(25,20)	0.050	0.051	0.060	0.056	0.057	0.052	0.058
(50,10)	0.043	0.040	0.065	0.062	0.059	0.046	0.053

Table 3: Estimated rejection probabilities of different tests
Sample size: N=25; T=10

τ	ρ	RS_μ	RS_μ^*	RS_ρ	RS_ρ^*	$RS_{\mu\rho}$	RSO_μ	RSO_μ^*
0.00	0.00	0.047	0.048	0.087	0.072	0.062	0.045	0.051
0.00	0.05	0.053	0.050	0.143	0.141	0.122	0.085	0.039
0.00	0.10	0.123	0.080	0.381	0.333	0.342	0.187	0.061
0.00	0.20	0.322	0.158	0.869	0.788	0.818	0.416	0.128
0.00	0.40	0.847	0.325	1.000	0.999	1.000	0.888	0.354
0.00	0.60	0.998	0.776	1.000	1.000	1.000	0.998	0.804
0.05	0.00	0.344	0.298	0.153	0.072	0.308	0.435	0.373
0.05	0.05	0.442	0.301	0.351	0.118	0.423	0.530	0.402
0.05	0.10	0.514	0.296	0.598	0.326	0.605	0.591	0.359
0.05	0.20	0.734	0.364	0.949	0.789	0.932	0.776	0.428
0.05	0.40	0.955	0.576	1.000	1.000	1.000	0.971	0.641
0.05	0.60	0.998	0.867	1.000	1.000	1.000	1.000	0.890
0.10	0.00	0.752	0.691	0.371	0.047	0.702	0.808	0.760
0.10	0.05	0.759	0.630	0.563	0.123	0.728	0.818	0.707
0.10	0.10	0.830	0.644	0.792	0.301	0.852	0.876	0.723
0.10	0.20	0.907	0.648	0.990	0.794	0.980	0.937	0.710
0.10	0.40	0.988	0.790	1.000	0.999	1.000	0.991	0.830
0.10	0.60	1.000	0.933	1.000	1.000	1.000	1.000	0.949
0.20	0.00	0.983	0.968	0.802	0.042	0.977	0.988	0.982
0.20	0.05	0.977	0.962	0.906	0.139	0.981	0.984	0.973
0.20	0.10	0.987	0.967	0.966	0.300	0.988	0.992	0.975
0.20	0.20	0.991	0.942	0.997	0.785	0.998	0.994	0.958
0.20	0.40	0.999	0.954	1.000	0.999	1.000	0.999	0.964
0.20	0.60	1.000	0.990	1.000	1.000	1.000	1.000	0.992
0.40	0.00	1.000	1.000	0.995	0.045	1.000	1.000	1.000
0.40	0.05	0.999	0.999	0.999	0.125	0.999	1.000	0.999
0.40	0.10	1.000	1.000	0.999	0.321	1.000	1.000	1.000
0.40	0.20	1.000	1.000	1.000	0.774	1.000	1.000	1.000
0.40	0.40	1.000	1.000	1.000	0.998	1.000	1.000	1.000
0.40	0.60	1.000	1.000	1.000	1.000	1.000	1.000	1.000
0.60	0.00	1.000	1.000	1.000	0.045	1.000	1.000	1.000
0.60	0.05	1.000	1.000	1.000	0.156	1.000	1.000	1.000
0.60	0.10	1.000	1.000	1.000	0.311	1.000	1.000	1.000
0.60	0.20	1.000	1.000	1.000	0.739	1.000	1.000	1.000
0.60	0.40	1.000	1.000	1.000	0.998	1.000	1.000	1.000
0.60	0.60	1.000	1.000	1.000	1.000	1.000	1.000	1.000

Table 4
Estimated rejection probabilities of different tests for negative ρ

τ	ρ	RS_μ	RS^*_μ	RS_ρ	RS^*_ρ	$RS_{\rho,\mu}$
Sample size: N = 25; T = 10						
0.00	-0.05	0.039	0.031	0.173	0.170	0.118
0.00	-0.10	0.044	0.019	0.396	0.346	0.285
0.00	-0.20	0.162	0.016	0.902	0.857	0.833
0.00	-0.40	0.573	0.048	1.000	1.000	1.000
0.05	-0.05	0.254	0.289	0.097	0.130	0.269
0.05	-0.10	0.202	0.340	0.184	0.314	0.365
0.05	-0.20	0.097	0.369	0.680	0.830	0.770
0.05	-0.40	0.039	0.679	0.997	1.000	1.000
Sample size: N = 25; T = 20						
0.00	-0.05	0.041	0.025	0.247	0.217	0.168
0.00	-0.10	0.049	0.025	0.640	0.600	0.520
0.00	-0.20	0.136	0.010	0.999	0.999	0.992
0.00	-0.40	0.610	0.018	1.000	1.000	1.000
0.05	-0.05	0.652	0.707	0.090	0.200	0.665
0.05	-0.10	0.613	0.758	0.244	0.557	0.806
0.05	-0.20	0.507	0.829	0.882	0.987	0.992
0.05	-0.40	0.303	0.963	1.000	1.000	1.000

Figure 1: Q-Q plots. Sample size (25,10)

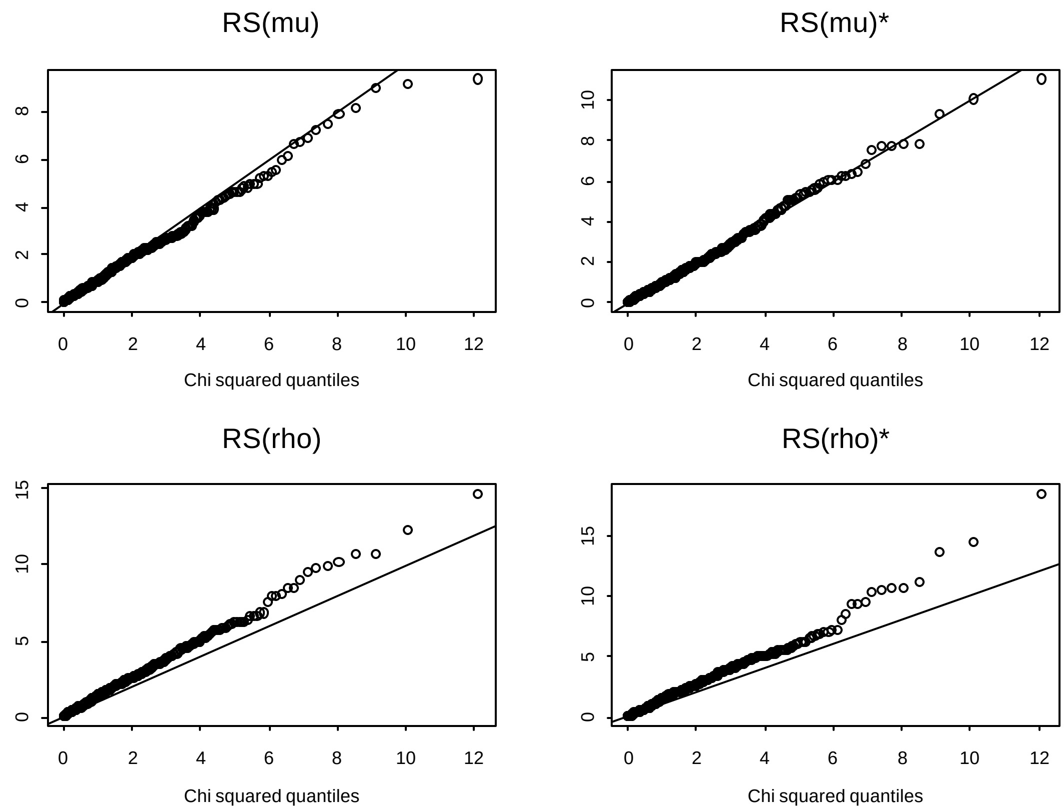

Figure 2: Tests for random effects. Sample Size (25,10)

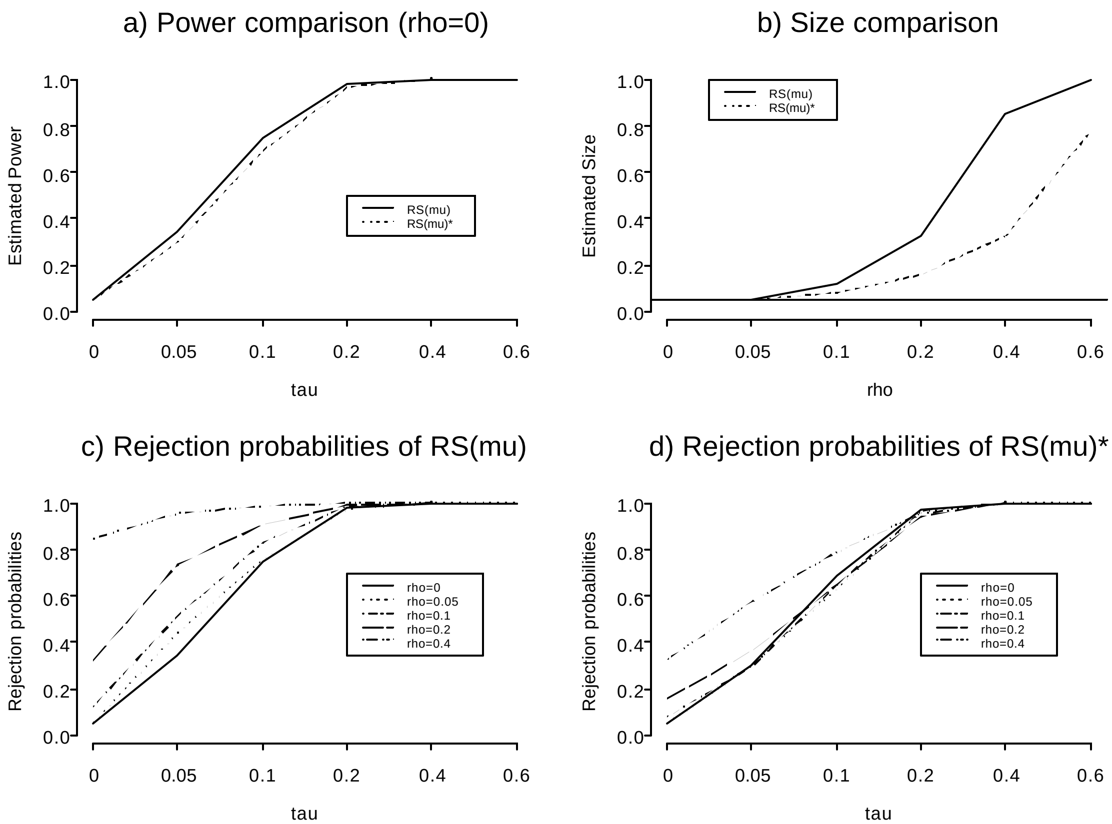

Figure 3: Tests for serial correlation. Sample Size (25,10)