ELSEVIER

Journal of Econometrics 68 (1995) 133–151

JOURNAL OF
Econometrics

Testing AR(1) against MA(1) disturbances in an error component model

Badi H. Baltagi*,[a], Qi Li[b]

[a]Department of Economics, Texas A&M University, College Station, TX 77843-4228, USA
[b]Department of Economics, University of Guelph, Guelph, Ont. N1G 2W1, Canada

Abstract

This paper derives three LM statistics for an error component model with first-order serially correlated errors. The first LM statistic jointly tests for zero first-order serial correlation and random individual effects, the second LM statistic tests for zero first-order serial correlation assuming fixed individual effects, and the third LM statistic tests for zero first-order serial correlation assuming random individual effects. In *all* three cases, the corresponding LM statistic is the *same* whether the alternative is AR(1) or MA(1). This paper also derives two extensions of the Burke, Godfrey, and Termayne (1990) test from the time-series to the panel data literature. The first tests the null of AR(1) disturbances against MA(1) disturbances, and the second tests the null of MA(1) disturbances against AR(1) disturbances in an error component model. These tests are computationally simple requiring only OLS or within residuals. The small sample performance of these tests are studied using Monte Carlo experiments.

Key words: Error components; Serial correlation; Lagrange multiplier; Random and fixed effects
JEL classification: C23

* Corresponding author.

The authors would like to thank Richard Blundell and two anonymous referees for their helpful comments and suggestions. Baltagi was funded by the Advanced Research Progam, Texas Higher Education Board, while Li was funded by the Social Sciences of Humanities Research Council of Canada.

SSDI 0304407694016 46 H

1. Introduction

A random or fixed effects error component model is by far the most widely used specification in panel data econometrics.[1] The error component model was extended to take into account first-order serial correlation in the remainder disturbances by Lillard and Willis (1978) for the random effects model and by Bhargava, Franzini, and Narendranathan (1982) for the fixed effects model. Both studies considered the AR(1) specification on the remainder disturbances. However, the MA(1) model is a viable alternative; see Nicholls, Pagan, and Terrell (1975). In fact, in the time-series literature, it is well known that the same Lagrange Multiplier test statistic is obtained for testing the null hypothesis of zero first-order serial correlation whether the alternative is AR(1) or MA(1); see Breusch and Godfrey (1981). This paper demonstrates that this result extends to the error component model with serial correlation. In Section 2, we derive three LM test statistics:[2] the first LM test statistic *jointly* tests for zero first-order serial correlation and random individual effects. The second LM test statistic tests for zero first-order serial correlation assuming a fixed effects model, and the third LM test statistic tests for zero first-order serial correlation assuming a random effects model. In *all* three cases, the corresponding LM statistic is the *same* whether the alternative is AR(1) or MA(1).

Section 3 deals with the problem of testing AR(1) against MA(1) in an error component model. This problem has been extensively studied in the time-series framework; see Walker (1967), King (1983), and more recently King and McAleer (1987), King (1988), and Burke, Godfrey, and Termayne (1990). For panel data, this paper proposes two extensions of the Burke, Godfrey, and Termayne (1990) test to the error component model: The first test assumes that the null is AR(1) while the alternative is MA(1). It is computationally simple using only the within residuals and performs well when T is large. However, for typical labor panels, N is large and T is small. Hence, we propose an alternative BGT-type test which relies on N large to achieve its asymptotic distribution. This test switches the hypotheses so that the null is MA(1) while the alternative is AR(1). It is also computationally simple, requiring only ordinary least squares residuals. Section 4 gives the Monte Carlo design along with the results of the experiments, while Section 5 gives our summary and conclusion.

2. Testing for first-order serial correlation in an error component model

Consider the following panel data regression:

$$y_{it} = x'_{it}\beta + u_{it}, \quad i = 1, 2, \dots, N, \quad t = 1, 2, \dots, T, \tag{1}$$

[1] See Hsiao (1986) and the references cited there.

[2] See Breusch and Pagan (1980), Engle (1984), and the recent monograph by Godfrey (1989) for the advantages of the LM test and its application to several econometric examples.

where β is a $K \times 1$ vector of regression coefficients including the intercept, i denoting individuals and t denoting time-periods. x_{it} is a $K \times 1$ vector of observations on K strictly exogenous explanatory variables. The composite disturbance is

$$u_{it} = \mu_i + v_{it}, \tag{2}$$

where $\mu_i \sim \mathrm{IIN}(0, \sigma_\mu^2)$ and the remainder disturbances follow a stationary AR(1) process: $v_{it} = \rho v_{i,t-1} + \varepsilon_{it}$ with $|\rho| < 1$, or a MA(1) process: $v_{it} = \varepsilon_{it} + \lambda \varepsilon_{i,t-1}$ with $|\lambda| < 1$. In both cases, we assume $\varepsilon_{it} \sim \mathrm{IIN}(0, \sigma_\varepsilon^2)$. In the next section, we will show that the joint LM test statistic for $\mathrm{H}_1^{\mathrm{a}}: \sigma_\mu^2 = 0; \lambda = 0$ is the same as the joint LM test statistic for $\mathrm{H}_1^{\mathrm{b}}$: $\sigma_\mu^2 = 0$; $\rho = 0$.

2.1. A joint LM test for serial correlation and random individual effects

Let us consider the LM test for the error component model where the remainder disturbances follow a MA(1) process. Writing the error term in vector form,

$$u = (I_N \otimes e_N)\mu + v, \tag{3}$$

where $\mu' = (\mu_1, \ldots, \mu_N)$ and $v' = (v_{11}, \ldots, v_{1T}, \ldots, v_{N1}, \ldots, v_{NT})$. e_T is a vector of ones of dimension T and I_N is an identity matrix of dimension N. The variance–covariance matrix for this model is

$$\Sigma = \mathrm{E}(uu') = \sigma_\mu^2 I_N \otimes J_T + \sigma_\varepsilon^2 I_N \otimes V_\lambda, \tag{4}$$

where

$$J_T = e_T e_T' \quad \text{and} \quad V_\lambda = \begin{bmatrix} 1+\lambda^2 & \lambda & 0 & \cdots & 0 & 0 \\ \lambda & 1+\lambda^2 & \lambda & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & \lambda & 1+\lambda^2 \end{bmatrix}.$$

The log-likelihood function is

$$l(\beta, \theta) = \text{const.} - \tfrac{1}{2}\log|\Sigma| - \tfrac{1}{2}u'\Sigma^{-1}u, \tag{5}$$

where $\theta = (\lambda, \sigma_\mu^2, \sigma_\varepsilon^2)'$. In order to construct the LM test statistic for $\mathrm{H}_1^{\mathrm{a}}: \sigma_\mu^2 = 0$; $\lambda = 0$, one needs $D(\theta) = \partial l(\theta)/\partial\theta$ and the information matrix $J(\theta) = \mathrm{E}[\partial^2 l(\theta)/\partial\theta\,\partial\theta']$ evaluated at the restricted maximum likelihood estimator $\hat{\theta}$. Following Hartley and Rao (1967) and Harville (1977), one can derive the following LM statistic for the null hypothesis $\mathrm{H}_1^{\mathrm{a}}$:

$$\mathrm{LM}_1 = \hat{D}'\hat{J}^{-1}\hat{D} = \frac{NT^2}{2(T-1)(T-2)}[A^2 - 4AB + 2TB^2], \tag{6}$$

where $A = [\hat{u}(I_N \otimes J_T)\hat{u}/(\hat{u}'\hat{u}) - 1]$, $B = (\hat{u}'\hat{u}_{-1}/\hat{u}'\hat{u})$, and $\hat{u}$ denotes the OLS residuals. This is asymptotically distributed (for large N) as χ_2^2 under H_1^a. LM_1 is exactly the same as the joint test statistic derived by Baltagi and Li (1991) for AR(1) residual disturbances and random individual effects. In fact, one can repeat the above derivation for H_1^b: $\sigma_\mu^2 = 0$; $\rho = 0$, by replacing the V_λ matrix by its AR(1) counterpart

$$V_\rho = \frac{1}{1-\rho^2}\begin{bmatrix} 1 & \rho & \cdots & \rho^{T-1} \\ \rho & 1 & \cdots & \rho^{T-2} \\ \vdots & \vdots & \ddots & \vdots \\ \rho^{T-1} & \rho^{T-2} & \cdots & 1 \end{bmatrix}.$$

Note that under the null hypothesis, we have $(V_\rho)_{\rho=0} = I_T = (V_\lambda)_{\lambda=0}$ and $(\partial V_\rho/\partial\rho)_{\rho=0} = G = (\partial V_\lambda/\partial\lambda)_{\lambda=0}$, where G is the bi-diagonal matrix with bi-diagonal elements all equal to one. Then it is easy to show that the same joint LM test statistic results whether the residual disturbances follow an AR(1) or a MA(1) process.

Note that the A^2 term in (6) is the basis for the LM test statistic for H_2: $\sigma_\mu^2 = 0$ assuming there is no serial correlation (see Breusch and Pagan, 1980). In fact, $LM_2 = \sqrt{NT/2(T-1)}A$ is asymptotically distributed (for large N) as N(0,1) under H_2 against the one-sided alternative H_2': $\sigma_\mu^2 > 0$. Also, the B^2 term in (6) is the basis for the LM test statistic for H_3: $\rho(\lambda = 0)$ assuming there are no individual effects (see Breusch and Godfrey, 1981). In fact, $LM_3 = \sqrt{NT^2/(T-1)}B$ is asymptotically distributed (for large N) as N(0,1) under H_3 against the one-sided alternative H_3': $\rho(\lambda) > 0$. The presence of an interaction term in the joint LM test statistic emphasizes the importance of the joint test when both serial correlation and random individual effects are suspected. However, when T is large the interaction term becomes negligible.

Note that all the LM tests considered assume that the underlying null hypothesis is that of white noise disturbances. However, in panel data applications, especially with large labor panels, one is concerned with individual effects and is guaranteed their existence. In this case, it is inappropriate to test for serial correlation assuming no individual effects as is done in H_3. In fact, if one uses LM_3 to test for serial correlation, one is very likely to reject the null hypothesis of H_3 even if the null is true. This is because the μ_i's are correlated for the same individual across time and this will contribute to rejecting the null of no serial correlation. Section 2.2 derives the appropriate LM test statistic in case of random individual effects, while Section 2.3 derives the appropriate LM test statistic in case of fixed individual effects.

2.2. An LM test for first-order serial correlation in a random effects model

Case 1: Let us start with the AR(1) model. The null hypothesis is H_4^b: $\rho = 0$ (given $\sigma_\mu^2 > 0$) vs. $H_4^{b'}$: $\rho \neq 0$ (given $\sigma_\mu^2 > 0$). The variance–covariance matrix (under the alternative) is

$$\Sigma_1 = \sigma_\mu^2(I_N \otimes J_T) + \sigma_\nu^2(I_N \otimes V_\rho).$$

Under the null hypothesis H_4^b, we have

$$\begin{aligned} D_\rho &= (\partial l/\partial \rho)|_{\rho=0} \\ &= [N(T-1)/T]\frac{\sigma_1^2 - \sigma_\varepsilon^2}{\sigma_1^2} \\ &\quad + (\sigma_\varepsilon^2/2)u'\{I_N \otimes [(\bar{J}_T/\sigma_1^2 + E_T/\sigma_\varepsilon^2)G(\bar{J}_T/\sigma_1^2 + E_T/\sigma_\varepsilon^2)]\}u, \end{aligned}$$

where $\bar{J}_T = e_T e_T'/T$, $E_T = I_T - \bar{J}_T$, G is the bi-diagonal matrix defined above, and $\sigma_1^2 = T\sigma_\mu^2 + \sigma_\varepsilon^2$. The information matrix,[3] when evaluated under the null hypothesis ($\rho = 0$), is

$$\hat{J} = \begin{bmatrix} \hat{J}_{\rho\rho} & N(T-1)\hat{\sigma}_\varepsilon^2/\hat{\sigma}_1^4 & \dfrac{N(T-1)}{T}\hat{\sigma}_\varepsilon^2[1/\hat{\sigma}_1^4 - 1/\hat{\sigma}_\varepsilon^4] \\ & (NT^2/2\hat{\sigma}_1^4) & NT/2\hat{\sigma}_1^4 \\ & & \dfrac{N}{2}\left[\dfrac{1}{\hat{\sigma}_1^4} + \dfrac{T-1}{\hat{\sigma}_\varepsilon^4}\right] \end{bmatrix},$$

where $\hat{J}_{\rho\rho} = N[2a^2(T-1)^2 + 2a(2T-3) + (T-1)]$ and $a = [(\hat{\sigma}_\varepsilon^2 - \hat{\sigma}_1^2)/T\hat{\sigma}_1^2]$. Thus the LM test statistic is

$$LM = \hat{D}'\hat{J}^{-1}D = (\hat{D}_\rho)^2\hat{J}^{11}, \tag{7}$$

where $\hat{J}^{11} = N^2T^2(T-1)/\det(\hat{J})4\hat{\sigma}_1^4\hat{\sigma}_\varepsilon^4$. Under the null hypothesis, LM is asymptotically distributed (for large N) as χ_1^2. $\hat{\sigma}_\varepsilon^2 = \hat{u}'(I_N \otimes E_T)\hat{u}/N(T-1)$ and $\hat{\sigma}_1^2 = \hat{u}'(I_N \otimes \bar{J}_T)\hat{u}/N$, where $\hat{u}$ denotes the maximum likelihood residuals under the null hypothesis. The one-sided LM test for the hypothesis H_4^b (corresponding to the alternative $\rho > 0$) is

$$LM_4 = \hat{D}_\rho\sqrt{\hat{J}^{11}}, \tag{8}$$

and this is asymptotically distributed (for large N) as N(0, 1).

Case 2: MA(1) model; H_4^a: $\lambda = 0$ (given $\sigma_\mu^2 > 0$) vs. $H_4^{a'}$: $\lambda \neq 0$ (given $\sigma_\mu^2 > 0$). The variance–covariance matrix is

$$\Sigma_2 = \sigma_\mu^2(I_N \otimes J_T) + \sigma_\varepsilon^2(I_N \otimes V_\lambda).$$

[3] Detailed derivation of the information matrix J is available upon request from the authors.

Note that under the null hypothesis H_4^a,

$$(\Sigma_2^{-1})|_{\lambda=0} = (\Sigma_1^{-1})|_{\rho=0},$$

$$(\partial\Sigma_2/\partial\lambda)|_{\lambda=0} = (\partial\Sigma_1/\partial\rho)|_{\rho=0},$$

$$(\partial\Sigma_2/\partial\sigma_\mu^2)|_{\lambda=0} = (\partial\Sigma_1/\partial\sigma_\mu^2)|_{\rho=0},$$

$$(\partial\Sigma_2/\partial\sigma_\varepsilon^2)|_{\lambda=0} = (\partial\Sigma_1/\partial\sigma_\varepsilon^2)|_{\rho=0}.$$

Therefore, it is easy to show that the LM test statistic for H_4^a is the same as that given in (7).

2.3. An LM test for first-order serial correlation in a fixed effects model

The model is the same as (1), and the null hypothesis is H_5^b: $\rho = 0$, given that the μ_i's are fixed parameters.[4] Writing each individual's variables in a $T \times 1$ vector form, we have

$$y_i = x_i\beta + \mu_i e_T + v_i, \tag{9}$$

where $y_i = (y_{i1}, y_{i2}, \ldots, y_{iT})'$, x_i is $T \times K$, and v_i is $T \times 1$. $v_i \sim \mathrm{N}(0, \Omega_\rho)$, where $\Omega_\rho = \sigma_\varepsilon^2 V_\rho$ for the AR(1) disturbances. The log-likelihood function is

$$l(\beta, \rho, \mu, \sigma_\varepsilon^2) = \text{const.} - \tfrac{1}{2}\log|\Sigma| - \frac{1}{2\sigma_\varepsilon^2}\sum_{i=1}^{N} [(y_i - x_i\beta - \mu_i e_T)' V_\rho^{-1}(y_i - x_i\beta - \mu_i e_T)], \tag{10}$$

where $\Sigma = I_N \otimes \Omega_\rho$ is the variance–covariance matrix of $v = (v_1', \ldots, v_N')'$. One can easily verify that the LM test statistic for testing H_5^b is

$$LM = [NT^2/(T-1)](\tilde{v}'\tilde{v}_{-1}/\tilde{v}'\tilde{v})^2, \tag{11}$$

where $\tilde{v}$ denote the usual within residuals. This LM statistic is asymptotically distributed (for large T) as χ_1^2 under the null hypothesis H_5^b. The one-sided test for H_5^b (corresponding to the alternative $\rho > 0$) is

$$LM_5 = \sqrt{NT^2/(T-1)}(\tilde{v}'\tilde{v}_{-1}/\tilde{v}'\tilde{v}), \tag{12}$$

which is asymptotically distributed (for large T) as N(0, 1). By a similar argument, one can show that the LM test statistic for H_5^a: $\lambda = 0$, in a fixed effects model with MA(1) residual disturbances is identical to LM_5. Note also that LM_5 differs from LM_3 only by the fact that the within residuals $\tilde{v}$ (in LM_5)

[4] It is important to note that Bhargava, Franzini, and Narendranathan (1982), hereafter BFN, modified the Durbin–Watson statistic to test for serial correlation of the AR(1) type when the individual effects are assumed fixed. BFN provide tables for the lower and upper bounds of their modified Durbin–Watson statistic for $N \geqslant 50$ and $T \leqslant 20$.

replace the OLS residuals $\hat{u}$ (in LM_3). Since the within transformation wipes out the individual effects whether fixed or random, one can also use (12) to test for serial correlation in the random effects models.

3. Testing AR(1) against MA(1) in an error component model

We now turn to the discussion of testing AR(1) against MA(1) in the random effects model.[5] Testing AR(1) against MA(1) in a time-series framework has been extensively studied; see King and McAleer (1987) for a Monte Carlo comparison of nonnested, approximate point optimal, as well as Lagrange Multiplier tests. In fact, King and McAleer (1987) found that the nonnested tests perform poorly in small samples, while King's (1983) point optimal test performs the best. Also, Godfrey's (1978a, 1978b) LM test for testing AR(1) against ARMA(1,1) is quite powerful even when testing AR(1) against MA(1). For the error component model, one can derive a corresponding LM test for testing AR(1) against ARMA(1,1). However, such a test is computationally burdensome. In this section, we propose two extensions of the Burke, Godfrey, and Termayne (1990) test (hereafter BGT) to the error component model. These tests are simple to implement requiring within or OLS residuals.

The basic idea of the BGT test is as follows: Under the null hypothesis of an AR(1) process, the remainder error term v_{it} satisfies

$$\operatorname{corr}(v_{it}, v_{i,t-\tau}) = \rho_\tau = (\rho_1)^\tau, \quad \tau = 1, 2, \ldots, \tag{13}$$

and, therefore, under the null hypothesis

$$\rho_2 - (\rho_1)^2 = 0. \tag{14}$$

Under the alternative hypothesis of an MA(1) process on v_{it}, $\rho_2 = 0$ and hence $\rho_2 - (\rho_1)^2 < 0$. Therefore, BGT recommends a test statistic based on (14) using estimates of ρ's obtained from OLS residuals. One problem remains. King (1983) suggest that any 'good' test should have a size which tends to zero, asymptotically, for $\rho_1 > 0.5$. The test based on (14) does not guarantee this property. To remedy this, BGT proposed supplementing (14) with the decision to accept the null hypothesis of AR(1) if $\hat{\rho}_1 > \frac{1}{2} + 1/\sqrt{T}$.

In an error component model, the within transformation wipes out the individual effects, and one can use the within residuals $\tilde{u}_{it}(= \tilde{v}_{it})$ instead of OLS

[5] A test that helps distinguish between the AR(1) and MA(1) process will be useful in obtaining a pre-test estimator that may be more efficient than the 'pure' estimators; see Griffiths and Beesley (1984) for a Monte Carlo study that compares pre-test and pure estimators in the time-series literature.

residuals $\hat{u}_{it}$ to construct the BGT test. Let

$$(\tilde{\rho}_1)_i = \sum_{t=2}^{T} \tilde{u}_{it}\tilde{u}_{i,t-1} \Big/ \sum_{t=1}^{T} \tilde{u}_{it}^2, \quad (\tilde{\rho}_2)_i = \sum_{t=3}^{T} \tilde{u}_{it}\tilde{u}_{i,t-2} \Big/ \sum_{t=1}^{T} \tilde{u}_{it}^2$$

$$\text{for} \quad i = 1, \dots, N.$$

The following test statistic, based on (14),

$$\tilde{\gamma}_i = \sqrt{T}((\tilde{\rho}_2)_i - (\tilde{\rho}_1^2)_i)/(1 - (\tilde{\rho}_2)_i), \tag{15}$$

is asymptotically distributed (for large T) as N(0,1) under the null hypothesis of an AR(1). Using the data on all N individuals, we can construct a generalized BGT test statistic for the error component model,

$$\tilde{\gamma} = \sqrt{N}\left(\sum_{i=1}^{N} \tilde{\gamma}_i/N\right) = \sqrt{NT} \sum_{i=1}^{N} \left[\frac{(\tilde{\rho}_2)_i - (\tilde{\rho}_1^2)_i}{1 - (\tilde{\rho}_2)_i}\right] \Big/ N. \tag{16}$$

$\tilde{\gamma}_i$'s are independent for different i since the $\tilde{u}_i$'s are independent. Hence $\tilde{\gamma}$ is also asymptotically distributed (for large T) as N(0, 1) under the null hypothesis of an AR(1) process. The test statistic (16) is supplemented by

$$\tilde{\rho}_1 = \sum_{i=1}^{N} (\tilde{\rho}_1)_i/N \equiv \frac{1}{N} \sum_{i=1}^{N} \left[\sum_{t=2}^{T} \tilde{u}_{it}\tilde{u}_{i,t-1} \Big/ \sum_{t=1}^{T} \tilde{u}_{it}^2\right], \tag{17}$$

and our proposed BGT_1 test can be summarized as follows:

1) Use the within residuals $\tilde{u}_{it}$ to calculate $\tilde{\gamma}$ and $\tilde{\rho}_1$ from (16) and (17).

2) Accept the AR(1) model if $\tilde{\gamma} > c_\alpha$ or $\tilde{\rho}_1 > \frac{1}{2} + 1/\sqrt{T}$, where Prob $[\mathrm{N}(0, 1) \leqslant c_\alpha] = \alpha$.

The bias in estimating ρ_s ($s = 1, 2$) by using within residuals is of $O(1/T)$ as $N \to \infty$ (see Nickell, 1981). Therefore, BGT_1 may not perform well for small T. Since for typical labor panels, N is large and T is small, we turn next to an alternative extension of the BGT test which performs well for large N rather than large T.

An alternative BGT-type test for testing AR(1) vs. MA(1)

Let the null hypothesis be H_7: $v_{it} = \varepsilon_{it} + \lambda\varepsilon_{i,t-1}$ and the alternative be H_7': $v_{it} = \rho v_{i,t-1} + \varepsilon_{it}$, where $\varepsilon_{it} \sim \mathrm{N}(0, \sigma_\varepsilon^2)$. Note that this test differs from the BGT_1 test in that the null hypothesis is MA(1) rather than AR(1).[6] The alternative

[6] In the time-series literature, Walker (1967) derived nonnested tests for the AR(1) against the MA(1) as well as for the case where the null and alternative are interchanged.

BGT-type test uses autocorrelation estimates derived from OLS residuals and can be motivated as follows:

Let $Q_0 = \sum\sum u_{it}^2/NT = u'u/NT$ and $Q_s = \sum\sum u_{it}u_{i,t-s}/N(T-s) = u'(I_N \otimes G_s)u/N(T-s)$ for all $s = 1, \dots, S$, where $G_s = \frac{1}{2}\,\mathrm{Toeplitz}\,(\iota_s)$, ι_s is a vector of zeros with the $(s+1)$th element being 1. $s = 1, \dots, S$ with $S \leqslant (T-1)$ and S is finite.[7] Given the true residuals (the u's), and assuming $[u'Au/n - \mathrm{E}(u'Au/n)] \xrightarrow{\mathrm{P}} 0$, where $n = NT$ and A is an arbitrary symmetric matrix, it is straightforward to show the following:[8]

(i) For the MA(1) model,

$$\operatorname{plim} Q_0 = \sigma_\mu^2 + \sigma_\nu^2 + \mathrm{o}_p(1) = \sigma_\mu^2 + \sigma_\varepsilon^2(1+\lambda^2) + \mathrm{o}_p(1),$$

$$\operatorname{plim} Q_1 = \sigma_\mu^2 + \lambda\sigma_\varepsilon^2 + \mathrm{o}_p(1),$$

$$\operatorname{plim} Q_s = \sigma_\mu^2 + \mathrm{o}_p(1) \quad \text{for} \quad s = 2, \dots, S.$$

(ii) For the AR(1) model,

$$\operatorname{plim} Q_0 = \sigma_\mu^2 + \sigma_\nu^2 + \mathrm{o}_p(1),$$

$$\operatorname{plim} Q_s = \sigma_\mu^2 + \rho^s\sigma_\nu^2 + \mathrm{o}_p(1) \quad \text{for} \quad s = 1, \dots, S.$$

We only give the proof for the MA(1) case since the AR(1) case can be similarly shown.

Proof. For the MA(1) case, we have $\operatorname{plim} Q_0 = \lim \mathrm{E}(u'u)/(NT) = \lim \mathrm{tr}(\Sigma)/(NT) = (\sigma_\mu^2 + \sigma_\nu^2)$. Similarly, we have

$$\begin{aligned}\operatorname{plim} Q_1 &= \lim \mathrm{E}(u'(I_N \otimes G_1)u)/N(T-1) = \lim \mathrm{tr}\,[\Sigma(I_N \otimes G_1)]/N(T-1)\\ &= \mathrm{tr}\,[\sigma_\mu^2 J_T G_1 + \sigma_\varepsilon^2 V_\lambda G_1]/(T-1)\\ &= [\sigma_\mu^2(T-1) + \sigma_\varepsilon^2\lambda(T-1)]/(T-1)\\ &= (\sigma_\mu^2 + \sigma_\varepsilon^2\lambda).\end{aligned}$$

The third equality uses the fact that $\mathrm{tr}(J_T G_1) = (T-1)$ and $\mathrm{tr}(V_\lambda G_1) = \lambda(T-1)$.

$$\begin{aligned}\operatorname{plim} Q_s &= \lim \mathrm{E}(u'(I_N \otimes G_s)u)/N(T-s) = \lim \mathrm{tr}[\Sigma(I_N \otimes G_s)]/N(T-s)\\ &= \mathrm{tr}[\sigma_\mu^2 J_T G_s + \sigma_\varepsilon^2 V_\lambda G_s]/(T-s) = [\sigma_\mu^2(T-s)]/(T-s) = \sigma_\mu^2\\ &\quad \text{for} \quad s = 2, \dots, S.\end{aligned}$$

[7] Let $a = (a_1, a_2, \dots, a_n)'$ denote an arbitrary $n \times 1$ vector, then Toeplitz(a) is an $n \times n$ symmetric matrix generated from the $n \times 1$ vector a with the diagonal elements all equal to a_1, second diagonal elements equal to a_2, etc. For example, Toeplitz(ι_1) is the G matrix defined in Section 2.1.

[8] Unless otherwise stated, plim is taken as $N \to \infty$.

The third equality uses the fact that $\mathrm{tr}(J_T G_s) = (T - s)$ and $\mathrm{tr}(V_\lambda G_s) = 0$ for $s = 2, \ldots, S$.

The inclusion of the individual effects μ_i makes the composite error term $u_{it} = \mu_i + v_{it}$ correlated across time, even for more than one period apart from the MA(1) model. Therefore, a direct comparison of the $\mathrm{corr}(u_{it}, u_{i,t-s})$ cannot distinguish the AR(1) process from the MA(1) process. However, if we take differences of $Q_s - Q_{s+l}$ for $s \geqslant 2$ and $l \geqslant 1$, we can see that $\mathrm{plim}(Q_s - Q_{s+l}) = 0$ for the MA(1) process and $\mathrm{plim}(Q_s - Q_{s+l}) = \sigma_v^2 \rho^s (1 - \rho^l) > 0$ for the AR(1) process. Hence for N very large we should be able to distinguish these two processes based on the information obtained from $Q_s - Q_{s+l}$.

Take $s = 2$ and $l = 1$, an asymptotic test of H_7 against H'_7 could, therefore, be based upon $(Q_2 - Q_3)$ with significantly small positive values leading to the rejection of H_7. Under the standard assumptions, an appropriate test statistic is

$$\gamma = \sqrt{N/V}\,(Q_2 - Q_3), \tag{18}$$

where $V = 2\mathrm{tr}\{[(\sigma_\mu^2 J_T + \sigma_\varepsilon^2 V_\lambda)(G_2/(T-2) - G_3/(T-3))]^2\}$. It can be shown, under some regularity condition, that γ is asymptotically distributed (for large N) as N(0, 1) under the null hypothesis of an MA(1) process.[9] Obviously, there are many different ways to construct such a test. For example, we can use $Q_2 + Q_3 - 2Q_4$ instead of $Q_2 - Q_3$ to define the γ test. In this case, $V = 2\mathrm{tr}\{[(\sigma_\mu^2 J_T + \sigma_\varepsilon^2 V_\lambda)(G_2/(T-2) + G_3/(T-3) - 2G_4/(T-4))]^2\}$. In order to calculate V, we need to know σ_μ^2, σ_ε^2, and λ. We make the following simple but important observation which enables us to compute the test statistic γ using only OLS residuals. Under the null hypothesis, v_{it} follows an MA(1) process. By noting the fact that $\sigma_\varepsilon^2(1 + \lambda^2) = \sigma_v^2$, we can write $\sigma_\varepsilon^2 V_\lambda = \sigma_v^2 I_T + (\sigma_\varepsilon^2 \lambda) G$. Thus we do not need to estimate λ in order to compute the test statistic γ, all we need to get are some consistent estimators for σ_v^2, $(\lambda\sigma_\varepsilon^2)$, and σ_μ^2. These are obtained as follows:

$$\hat\sigma_v^2 = \hat Q_0 - \hat Q_2, \quad (\lambda\hat\sigma_\varepsilon^2) = \hat Q_1 - \hat Q_2, \quad \hat\sigma_\mu^2 = \hat Q_2,$$

where $\hat Q_s$ are obtained from Q_s by replacing u_{it} by the OLS residuals $\hat u_{it}$.

It is easy to show that $\hat Q_s - Q_s = \mathrm{O}_p(1/N)$, hence $\hat Q_s$ and Q_s are asymptotically equivalent estimators for the test statistic γ. Substituting these consistent estimators into V we get $\hat V$, and the test statistic γ becomes

$$\hat\gamma = \sqrt{N/\hat V}\,(\hat Q_2 - \hat Q_3), \tag{19}$$

[9] A derivation of this result is available upon request from the authors.

where

$$(\hat{Q}_2 - \hat{Q}_3) = \sum_{i=1}^{N} \sum_{t=3}^{T} \hat{u}_{it}\hat{u}_{i,t-2}/N(T-2) - \sum_{i=1}^{N} \sum_{t=4}^{T} \hat{u}_{it}\hat{u}_{i,t-3}/N(T-3)$$

and

$$\hat{V} = 2\mathrm{tr}\{[(\hat{\sigma}_\mu^2 J_T + \hat{\sigma}_v^2 I_T + \sigma_\varepsilon^2 \hat{\lambda} G)(G_2/(T-2) + G_3/(T-3))]^2\}.$$

$\hat{\gamma}$ is asymptotically distributed (for large N) as N(0, 1) under the null hypothesis H_7, and is referred to as the BGT_2 test.

4. Monte Carlo results

The model is set as follows:

$$Y_{it} = 5 + 0.5X_{it} + \mu_i + v_{it}, \qquad i = 1, \dots, N, \quad t = 1, \dots, T, \tag{20}$$

where $X_{it} = 0.1t + 0.5X_{i,t-1} + \omega_{it}$, with ω_{it} uniformly distributed on the interval $[-0.5, 0.5]$. X_{i0} is chosen as $(5 + 10\omega_{i0})$; see Nerlove (1971). $\mu_i \sim \mathrm{IIN}(0, \sigma_\mu^2)$. $\sigma^2 = \sigma_\mu^2 + \sigma_v^2$ is fixed at 20 and $\omega = \sigma_\mu^2/\sigma^2$ takes five different values, $\omega = (0.2, 0.4, 0.6, 0.8, 0.9)$.[10] The v_{it}'s follow an AR(1) process, or an MA(1) process, see (2), where both ρ and λ take five different values (0.2, 0.4, 0.6, 0.8, 0.9). We use three different combinations of $(N, T) = (10, 20), (20, 20), (50, 10)$ for all the tests considered in this paper (see Table 1 for a summary). In order to study the asymptotic behavior of the BGT_1 and BGT_2 test statistics, we conduct additional Monte Carlo experiments for large $T = 30, 60, 100$ for the BGT_1 test with $N = 10, 20, 30$. For the BGT_2 test, we choose large $N = 50, 100, 150$ with $T = 10, 20, 30$.

For the various null hypotheses listed in Table 1, Table 2 gives the corresponding number of rejections of that null in 1000 replications for $(N, T) = (10, 20)$.[11] The main results can be summarized as follows:

(i) The joint LM_1 test performs well in testing the null of H_1: $\rho = \sigma_\mu^2 = 0$. Its estimated size is not statistically different from its nominal size. In the presence of large individual effects ($\omega > 0.2$) or high serial correlation $\rho(\lambda) > 0.2$, this test has high power rejecting the null in 99 to 100% of the cases. It only has low power when $\omega = 0$ and $\rho(\lambda) = 0.2$ or $\omega = 0.2$ and $\rho(\lambda) = 0$.

(ii) The test statistic LM_2 for testing H_2: $\sigma_\mu^2 = 0$ implicitly assumes that $\rho(\lambda) = 0$. When ρ is indeed equal to zero, this test performs well. However, as

---

[10] The Monte Carlo results are invariant to the choice of α, β, and σ^2 (see Breusch, 1980).

[11] Results for $(N, T) = (20, 20)$ and $(50, 10)$ are similar and are available upon request from the authors.

Table 1
Summary of the tests considered

	Null hypothesis H_0	Alternative hypothesis H_A	Test statistic	Asymptotic distribution under H_0
(1a)	H_1^a: $\sigma_\mu^2 = 0$; $\lambda = 0$	σ_μ^2 or $\lambda \neq 0$	LM_1	χ_2^2
(1b)	H_1^b: $\sigma_\mu^2 = 0$; $\rho = 0$	σ_μ^2 or $\rho \neq 0$	LM_1	χ_2^2
(2)	H_2: $\sigma_\mu^2 = 0$	$\sigma_\mu^2 > 0$	LM_2	N(0, 1)
(3a)	H_3^a: $\lambda = 0$	$\lambda > 0$	LM_3	N(0, 1)
(3b)	H_3^b: $\rho = 0$	$\rho > 0$	LM_3	N(0, 1)
(4a)	H_4^a: $\lambda = 0$ ($\sigma_\mu^2 > 0$)	$\lambda > 0$ ($\sigma_\mu^2 > 0$)	LM_4	N(0, 1)
(4b)	H_4^b: $\rho = 0$ ($\sigma_\mu^2 > 0$)	$\rho > 0$ ($\sigma_\mu^2 > 0$)	LM_4	N(0, 1)
(5a)	H_5^a: $\lambda = 0$ (μ_i's fixed)	$\lambda > 0$ (μ_i's fixed)	LM_5	N(0, 1)
(5b)	H_5^b: $\rho = 0$ (μ_i's fixed)	$\rho > 0$ (μ_i's fixed)	LM_5	N(0, 1)
(6)	H_6: AR(1)	MA(1)	BGT_1	N(0, 1)
(7)	H_7: MA(1)	AR(1)	BGT_2	N(0, 1)

Row (5b): For H_5^b we also consider the modified Durbin–Watson test considered by Bhargava, Franzini, and Narendranathan (1982).

ρ moves away from zero and increases, this test tends to be biased in favor of rejecting the null. This is because a large serial correlation coefficient (i.e., large ρ) contributes to a large correlation among the individuals in the sample, even though $\sigma_\mu^2 = 0$. The first block corresponds to $\sigma_\mu^2 = 0$, i.e., the null is true. In this case, the number of rejections of the null are monotonically increasing as ρ increases, and for $\rho = 0.9$, LM_2 rejects in 100% of the cases. Similar results are obtained in case v_{it} follows an MA(1) process. The other blocks correspond to $\omega > 0$, i.e., the null is false, and LM_2 performs well when $\rho(\lambda) = 0$. As $\rho(\lambda)$ increases, this helps LM_2 even more in rejecting the null hypothesis of $\sigma_\mu^2 = 0$. In general, the presence of positive serial correlation tend to bias the case in favor of finding nonzero individual effects.

(iii) Similarly, the LM_3 test for testing H_3: $\rho = 0$ implicitly assumes $\sigma_\mu^2 = 0$. This test performs well when $\sigma_\mu^2 = 0$. However, as σ_μ^2 increases, the performance of this test deteriorates when the null ($\rho = 0$) is true. For example, when $\omega = 0.9$ and $\rho = 0$, it rejects the null hypothesis, when it is true, 100% of the time. The large correlation among the μ_i's contribute to the rejection of null hypothesis of no serial correlation. These results strongly indicate that one should not ignore the individual effects when testing for serial correlation.

Table 2
Testing for serial correlation and/or individual effects; number of rejections in 1000 replications; $N = 10$, $T = 20$

Null hypothesis		H_1	H_2	H_3	H_4	H_5	H_6	H_7
ω	$\rho(\lambda)$	LM_1	LM_2	LM_3	LM_4	LM_5	BGT_1	BGT_2
0	$\rho = \lambda = 0$	39[c]	45	37	34	10	223[a]	54[b]
	$\rho = 0.2$	657	207	825	790	539	217	116
	0.4	1000	530	1000	1000	996	225	350
	0.6	1000	876	1000	1000	1000	231	780
	0.8	1000	988	1000	1000	1000	290	951
	0.9	1000	1000	1000	1000	1000	288	984
	$\lambda = 0.2$	643	279	802	770	523	367	65
	0.4	995	293	1000	997	992	809	52
	0.6	1000	358	1000	1000	1000	987	60
	0.8	1000	392	1000	1000	1000	999	56
	0.9	1000	405	1000	1000	1000	1000	45
0.2	$\rho = \lambda = 0$	898	941	639	47	10	223	42
	$\rho = 0.2$	983	968	983	790	539	217	108
	0.4	1000	977	1000	999	996	225	306
	0.6	1000	993	1000	1000	1000	231	770
	0.8	1000	1000	1000	1000	1000	290	961
	0.9	1000	1000	1000	1000	1000	288	975
	$\lambda = 0.2$	985	971	986	807	523	367	61
	0.4	1000	981	1000	1000	992	809	45
	0.6	1000	969	1000	1000	1000	987	50
	0.8	1000	973	1000	1000	1000	999	42
	0.9	1000	975	1000	1000	1000	1000	38
0.4	$\rho = \lambda = 0$	997	999	947	36	10	223	49
	$\rho = 0.2$	1000	999	999	796	539	217	104
	0.4	1000	996	1000	999	996	225	314
	0.6	1000	999	1000	1000	1000	231	745
	0.8	1000	1000	1000	1000	1000	290	946
	0.9	1000	1000	1000	1000	1000	288	973
	$\lambda = 0.2$	997[c]	995	999	811	523	367	55
	0.4	1000	1000	1000	1000	992	809	44
	0.6	1000	999	1000	1000	1000	987	45
	0.8	1000	999	1000	1000	1000	999	48
	0.9	1000	999	1000	1000	1000	1000	40
0.6	$\rho = \lambda = 0$	999	999	996	33	10	223	41
	$\rho = 0.2$	1000	1000	1000	817	539	217	95
	0.4	1000	1000	1000	999	996	225	309
	0.6	1000	1000	1000	1000	1000	231	686
	0.8	1000	1000	1000	1000	1000	290	895
	0.9	1000	1000	1000	1000	1000	288	911

Table 2 (continued)

Null hypothesis		H_1	H_2	H_3	H_4	H_5	H_6	H_7
ω	$\rho(\lambda)$	LM_1	LM_2	LM_3	LM_4	LM_5	BGT_1	BGT_2
0.6	$\lambda = 0.2$	1000	1000	1000	796	523	367	50
	0.4	1000	1000	1000	999	992	809	55
	0.6	1000	1000	1000	1000	1000	987	48
	0.8	1000	1000	1000	1000	1000	999	45
	0.9	1000	1000	1000	1000	1000	1000	65
0.8	$\rho = \lambda = 0$	999	1000	999	52	10	223	59
	$\rho = 0.2$	1000	1000	1000	812	539	217	113
	0.4	1000	1000	1000	1000	996	225	260
	0.6	1000	1000	1000	1000	1000	231	627
	0.8	1000	1000	1000	1000	1000	290	770
	0.9	1000	1000	1000	1000	1000	288	784
	$\lambda = 0.2$	1000	1000	1000	817	523	367	63
	0.4	1000	1000	1000	1000	992	809	57
	0.6	1000	1000	1000	1000	1000	987	60
	0.8	1000	1000	1000	1000	1000	999	61
	0.9	1000	1000	1000	1000	1000	1000	54
0.9	$\rho = \lambda = 0$	1000	1000	1000	39	10	223	52
	$\rho = 0.2$	1000	1000	1000	808	539	217	106
	0.4	1000	1000	1000	1000	996	225	244
	0.6	1000	1000	1000	1000	1000	231	481
	0.8	1000	1000	1000	1000	1000	290	632
	0.9	1000	1000	1000	1000	1000	288	712
	$\lambda = 0.2$	1000	1000	1000	816	523	367	68
	0.4	1000	1000	1000	999	992	809	62
	0.6	1000	1000	1000	1000	1000	987	59
	0.8	1000	1000	1000	1000	1000	999	61
	0.9	1000	1000	1000	1000	1000	1000	61

[a]The null hypothesis is AR(1).
[b]The null hypothesis, is MA(1).
[c]Counts between 37 and 63 are not statistically different from 50 at the 0.05 level.

(iv) In contrast to LM_3, both LM_4 and LM_5 take into account the presence of individual effects. Recall that LM_5 tests H_5: $\rho = 0$ (assuming the μ_i's are fixed effects). The performance of this test is independent of the values of σ_μ^2 (i.e., independent of the value of ω). This is because the within transformation wipes out the individual effects even if they are random. For large values of ρ or λ (ρ or λ greater than 0.4), both LM_4 and LM_5 have high power, rejecting the null more than 99% of the time. However, the estimated size of LM_4 is closer to the 5% nominal value than that of LM_5. In addition, for $N = 50$ and $T = 10$, our results show that the Bhargava, Franzini, and Narendranathan (1982) modified Durbin–Watson statistic has similar performance to that of LM_4 and is closer to its nominal size than LM_5.[12]

Now let us turn to the results of the tests of AR(1) against MA(1) specification in an error component model. Recall that for BGT_1, the null hypothesis is an AR(1) model, while for the BGT_2 test, the null hypothesis is an MA(1) model. From Table 2, the performance of BGT_1 is not satisfactory when the true model is AR(1) or MA(1) with small λ, while the performance of BGT_2 is not satisfactory when the true model is AR(1) with small ρ. These results suggest that $T = 20$ and $N = 10$ are not large enough for the BGT_1 and BGT_2 tests to achieve their asymptotic distribution. We therefore conduct additional Monte Carlo experiments with larger values of T for BGT_1 and larger values of N for BGT_2 (see Tables 3 and 4 below).

Table 3 reports the estimated power and size functions for BGT_1 for three values of $T = 30, 60, 100$. Note that this test is independent of the value of ω since the within transformation wipes out the individual effects. Overall, the BGT_1 test performs well if $T \geqslant 60$ and $T > N$. However, when T is small, or T is of moderate size but N is large, this test will tend to overreject the AR(1) process and therefore is not recommended in these cases.

Table 4 reports the estimated power and size function for BGT_2 for $N = 50$, 100, 150 and $T = 10, 20, 30$. Recall that for BGT_2 the null hypothesis is that of

Table 3
Estimated size and power functions of BGT_1 test; number of rejections of H_6 in 1000 replications

	$T = 30$			$T = 60$			$T = 100$		
$\rho(\lambda)$	$N = 10$	$N = 20$	$N = 30$	$N = 10$	$N = 20$	$N = 30$	$N = 10$	$N = 20$	$N = 30$
$\rho = \lambda = 0$	169	231	315	132	151	193	92	127	152
$\rho = 0.2$	165	244	310	135	142	193	103	127	149
0.4	181	255	316	125	155	205	102	122	155
0.6	185	276	354	117	166	205	87	119	152
0.8	134	214	300	1	0	0	0	0	0
0.9	22	11	3	0	0	0	0	0	0
$\lambda = 0.2$	363	523	583	376	596	745	435	670	821
0.4	849	984	1000	1000	1000	1000	1000	1000	1000
0.6	999	1000	1000	1000	1000	1000	1000	1000	1000
0.8	1000	1000	1000	1000	1000	1000	1000	1000	1000
0.9	1000	1000	1000	1000	1000	1000	1000	1000	1000

[12] $N = 50$ is the smallest N for which the upper and lower limits of the modified Durbin–Watson statistic are reported by Bhargava, Franzini, and Narendranathan (1982).

Table 4
Estimated size and power functions of BGT_2 test; number of rejections of H_7 in 1000 replications

		$T = 10$			$T = 20$			$T = 30$		
ω	$\rho(\lambda)$	$N = 50$	$N = 100$	$N = 150$	$N = 50$	$N = 100$	$N = 150$	$N = 50$	$N = 100$	$N = 150$
0.2	$\rho = 0.2$	134	182	258	228	331	306	272	417	548
	0.4	613	828	951	913	993	1000	982	1000	1000
	0.6	987	1000	1000	1000	1000	1000	1000	10000	1000
	0.8	1000	1000	1000	1000	1000	1000	1000	10000	1000
	0.9	1000	1000	1000	1000	1000	1000	1000	10000	1000
	$\lambda = 0.2$	54	53	52	52	47	51	49	50	49
	0.4	65	44	42	43	49	49	46	47	64
	0.6	52	41	42	57	52	50	40	52	49
	0.8	46	59	47	43	53	46	58	64	45
	0.9	56	53	45	43	40	52	60	46	54
0.4	$\rho = 0.2$	122	181	225	218	331	404	262	426	550
	0.4	582	828	935	890	989	999	979	1000	1000
	0.6	973	1000	1000	1000	1000	1000	1000	10000	1000
	0.8	1000	1000	1000	1000	1000	1000	1000	10000	1000
	0.9	1000	1000	1000	1000	1000	1000	1000	10000	1000
	$\lambda = 0.2$	40	49	48	69	54	57	46	55	47
	0.4	41	44	55	53	44	64	46	44	49
	0.6	49	52	48	56	48	49	39	56	57
	0.8	43	44	40	46	57	48	47	45	40
	0.9	53	57	45	40	47	44	52	69	55
0.6	$\rho = 0.2$	135	181	202	203	301	394	234	425	542
	0.4	553	780	923	869	988	999	977	1000	1000
	0.6	952	999	1000	1000	1000	1000	1000	10000	1000
	0.8	997	1000	1000	1000	1000	1000	1000	10000	1000
	0.9	1000	1000	1000	1000	1000	1000	1000	10000	1000
	$\lambda = 0.2$	50	49	50	62	50	48	37	54	51
	0.4	57	48	64	51	66	56	52	48	54
	0.6	41	41	50	56	46	51	54	60	55
	0.8	55	47	44	39	51	47	54	40	48
	0.9	51	57	62	48	45	45	63	45	41
0.8	$\rho = 0.2$	107	131	180	185	257	355	242	352	506
	0.4	464	711	834	824	979	994	961	998	1000
	0.6	886	993	1000	999	1000	1000	1000	10000	1000
	0.8	976	1000	1000	1000	1000	1000	1000	10000	1000
	0.9	984	1000	1000	1000	1000	1000	1000	10000	1000
	$\lambda = 0.2$	46	53	47	52	56	39	48	40	48
	0.4	49	57	50	45	53	61	56	45	61
	0.6	57	51	62	61	44	54	51	63	60
	0.8	50	48	47	56	47	53	43	45	53
	0.9	69	44	59	51	46	55	51	34	47

Table 4 (continued)

		$T = 10$			$T = 20$			$T = 30$		
ω	$\rho(\lambda)$	$N = 50$	$N = 100$	$N = 150$	$N = 50$	$N = 100$	$N = 150$	$N = 50$	$N = 100$	$N = 150$
0.9	$\rho = 0.2$	105	164	152	167	253	297	230	336	428
	0.4	389	559	663	750	920	988	929	988	1000
	0.6	743	935	986	986	1000	1000	1000	10000	1000
	0.8	884	985	998	996	1000	1000	1000	10000	1000
	0.9	870	993	999	991	1000	1000	1000	10000	1000
	$\lambda = 0.2$	60	65	53	55	58	43	59	44	49
	0.4	68	50	48	49	52	46	61	62	53
	0.6	46	63	63	60	51	53	61	53	53
	0.8	57	52	57	67	60	53	41	36	62
	0.9	49	64	65	56	49	37	56	43	45

an MA(1) process and the alternative is an AR(1) process. If the true model is MA(1), the performance of the BGT_2 is satisfactory, with its size close to 5% for all cases considered. When the true model is an AR(1) process, we observe the following: (i) For each fixed T, the performance of BGT_2 improves as N increases. (ii) For each fixed N, the performance of BGT_2 improves as T increases. This is in sharp contrast to the BGT_1 test where, when AR(1) is the true model, the performance of the test deteriorates as N increases for fixed T. Overall, BGT_2's performance is good except when the true model is AR(1) with small ρ.

5. Summary and conclusions

This paper recommends testing jointly for serial correlation and random individual effects in panel data applications. Rejecting the null, one can test for individual effects assuming serial correlation. Most likely, in labor panels, the presence of individual effects will not be rejected. In addition, one can test for serial correlation assuming fixed or random individual effects. This paper focuses on the latter tests, and shows that the resulting LM statistics are invariant to the form of the first-order serial correlation whether AR(1) or MA(1). This extends the Breusch and Godfrey (1981) result from time-series regressions to panel data regressions with individual effects. The Monte Carlo results show that tests for first-order serial correlation, assuming individual effects (whether fixed or random) perform better than those ignoring them. Of these tests, LM_4 and BFN give better estimated size than LM_5 and are therefore recommended.

This paper also derives two test statistics that extend the Burke, Godfrey, and Termayne (1990) test for the AR(1) against MA(1) disturbances from the time-series regressions to panel data regressions with individual effects. The first statistic, named the BGT_1 test, uses within residuals and tests the null of an AR(1) against the alternative of an MA(1). This test performs well if $T \geqslant 60$ and $T > N$. However, when T is small, or T is of moderate size but N is large, BGT_1 will tend to overreject the null hypothesis. Therefore BGT_1 is not recommended for these cases. For typical labor panels, N is large and T is small. For these cases, this paper recommends the BGT_2 test, which uses OLS residuals and tests the null of an MA(1) against the alternative of an AR(1). This test performs well when N is large and does not rely on T to achieve its asymptotic distribution. Our Monte Carlo results show that BGT_2's performance improves as either N or T increases.

References

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

Bhargava, A., L. Franzini, and W. Narendranathan, 1982, Serial correlation and the fixed effects model, Review of Economic Studies 49, 533–549.

Breusch, T.S., 1980, Useful invariance results for generalized regression models, Journal of Econometrics 13, 327–340.

Breusch, T.S. and L.G. Godfrey, 1981, A review of recent work on testing for autocorrelation in dynamic simultaneous models, in: D.A. Currie, R. Nobay, and D. Peel, eds., Macroeconomic analysis: Essays in macroeconomics and economics (Croom Helm, London).

Breusch, T.S. and A.R. Pagan, 1980, The Lagrange multiplier test and its applications to model specification, Review of Economic Studies 47, 225–238.

Burke, S.P., L.G. Godfrey, and A.R. Termayne, 1990, Testing AR(1) against MA(1) disturbances in the linear regression model: An alternative procedure, Review of Economic Studies 57, 135–145.

Engle, R.F., 1984, Wald, likelihood ratio, and Lagrange multiplier tests in econometrics, in: Z. Griliches and M.D. Intriligator, eds., Handbook of econometrics (North–Holland, Amsterdam).

Godfrey, L.G., 1978a, Testing against general autoregressive and moving average error models when the regressors include lagged dependent variables, Econometrica 46, 1293–1302.

Godfrey, L.G., 1978b, Testing for higher order serial correlation in regression equations when the regressors include lagged dependent variables, Econometrica 46, 1303–1310.

Godfrey, L.G., 1987, Discriminating between autocorrelation and misspecification in regression analysis: An alternative strategy, Review of Economics and Statistics 69, 128–134.

Godfrey, L.G., 1989, Misspecification tests in econometrics (Cambridge University Press, Cambridge).

Griffiths, W.E. and P.A.A. Beesley, 1984, The small-sample properties of some preliminary test estimators in a linear model with autocorrelated errors, Journal of Econometrics 25, 49–61.

Hartley, H.O. and J.N.K. Rao, 1967, Maximum likelihood estimation for the mixed analysis of variance model, Biometrika 54, 93–108.

Harville, D.A., 1977, Maximum likelihood approaches to variance component estimation and to related problems, Journal of the American Statistical Association 72, 320–338.

Hsiao, C., 1986, Analysis of panel data (Cambridge University Press, Cambridge).

King, M.L., 1983, Testing for autoregressive against moving average errors in the linear regression model, Journal of Econometrics 21, 35–51.

King, M.L., 1988, Towards a theory of point optimal testing, Econometric Reviews 6, 169–218.

King, M.L. and M. McAleer, 1987, Further results on testing AR(1) against MA(1) disturbances in the linear regression model, Review of Economic Studies 54, 649–636.

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

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

Nicholls, D.F., A.R. Pagan, and R.D. Terrell, 1975, The estimation and use of models with moving average disturbance terms: A survey, International Economic Review 16, 113–134.

Nickell, S., 1981, Biases in dynamic models with fixed effects, Econometrica 49, 1417–1426.

Walker, A.M., 1967, Some tests of separate families of hypotheses in time series analysis, Biometrika 54, 39–68.