Second Revision for Empirical Economics MS2012519-2
Robust Tests for Time-invariant Individual Heterogeneity vs. Dynamic State Dependence

Abstract

We derive tests for persistent effects in a general linear dynamic panel data context. Two sources of persistent behavior are considered: time-invariant unobserved factors (captured by an individual random effect) and dynamic persistence or 'state dependence' (captured by autoregressive behavior). We will use a maximum likelihood framework to derive a family of tests that help researchers learn whether persistence is due to individual heterogeneities, dynamic effect, or both. The proposed tests have power only in the direction they are designed to perform, that is, they are locally robust to the presence of alternative sources of persistence, and consequently, are able to identify which source of persistence is active. A Monte Carlo experiment is implemented to explore the finite sample performance of the proposed procedures. The tests are applied to a panel data series of real GDP growth for the period 1960-2005.

JEL Classification: C12, C23.
Keywords: Dynamic panel, local misspecification, random effects, testing.

1 Introduction

One of the most important advantages of panel models is to distinguish among alternative sources of persistent behavior. After controlling for observed factors, two sources of persistence are relevant. First, time-invariant unobserved factors that reflects individual heterogeneity (captured by an individual fixed or random effects) induce persistence. A second source is dynamic persistence through serial correlation, in the error term or as a lagged dependent variable. Distinguishing among sources of persistence is a much relevant issue for policy purposes at the microeconomic and macroeconomic level. Unobserved heterogeneities call for interventions to remedy individual factors that keep individuals or countries persistently in poverty, like improving education. Dynamic persistence, on the other hand, may better be handled through helping households or countries to cope with the persistent effects of negative shocks, like insurance programs. (See Lillard and Willis', 1978, classic article on earnings persistence, and Sosa-Escudero, Marchionni and Arias, 2011, for a recent application in rural El Salvador.) However, as argued by Angrist and Pischke (2009, p.245), models for each persistence source "are not nested, which means that we cannot hope to estimate one and get the other as a special case if need be."

The purpose of this paper is to derive tests for persistent effects in a general linear dynamic panel data context. We derive a family of tests that help researchers learn whether persistence is due to individual heterogeneities, dynamic effects, or both.

Baltagi and Li (1995) derive a test for serial correlation, when random effects are present and controlled for in an error components model, based on a maximum-likelihood context. Similarly Holtz-Eakin (1988) proposed a test for random individual effets, in a dynamic panel structure estimated by GMM. These two proposals can be seen as 'conditional', in the sense that they test for a particular source of persistence, controlling for (estimating) the other one. Bera, Sosa Escudero and Yoon (2001) show that standard 'unconditional' tests for random effects (Breusch and Pagan, 1980) or serial correlation (Baltagi and Li, 1991), are of limited use for these purposes because each of them implicitly assumes that the other source of persistence is absent. For example, the classical test by Breusch and Pagan (1980) is shown to reject its null not only when random effects are present, but also due to the presence positive serial correlation. A similar and symmetric concern affects the test by Baltagi and Li (1991), which confounds serial correlation with random effects. Bera et al. (2001) circumvent this problem by deriving modified tests for each source of persistence, that are insensitive to the local presence of the other one, i.e., a test for random effects (serial correlation) that is insensitive to the local presence of serial correlation (random effects). The local nature of the solution might seem restrictive, but a comprehensive Monte Carlo experiment by these authors shows that the proposed tests perform well, even in non-local contexts and small samples. A major advantage of this strategy, as compared to a conditional approach as implicit in Baltagi and Li (1995) or Holtz-Eakin (1988) is that tests can be based on simple pooled-OLS estimation of a model, under the joint null of neither random effects nor serial correlation.

As stressed by Hendry and Mizon (1978) and Hendry (1995), serial correlation is only a particular form of dynamic misspecification, which does not necessarily capture more general dynamic persistence patterns. An autoregressive specification is thus preferred as a more general model to analyze dynamic behavior. This is the underlying idea behind the 'general-to-specific' approach advocated by Hendry (1995). Consequently, in our panel data context, first order serial correlation is only one possible specification of a more general dynamic panel model, where a 'common factor' restriction holds.

We construct tests for persistent effects in a general linear dynamic panel data context. Unobserved individual heterogeneity is captured by random individual effects and dynamic persistence is handled through the presence of a lagged dependent variable, closer to the idea of 'state dependence' in the applied literature. Our testing strategy is based on pooled OLS estimation of a model without persistence. Hence the proposed testing strategy can help researchers decide whether a truly dynamic model is required, or whether simpler random effects model would suffice to capture persistent behavior.

A Monte Carlo experiment is implemented to explore the finite sample performance of our tests. They are shown to have power only in the direction they are designed to perform, that is, they are robust to the presence of alternative sources of persistence, and consequently, are able to identify which source of persistence is active. Moreover, the tests have correct size and power for alterantive distributional assumptions in the error components.

The paper is organized as follows. Section 2 presents the model and the assumptions.

Section 3 derives the test statistics. Section 4 studies the small sample behavior of the proposed tests. Section 5 contains an application of the proposed tests to the study of real GDP per capita growth in a panel data of countries. Section 6 concludes.

2 Model and Assumptions

Consider a first order dynamic panel data model with random individual effects:

$$\begin{aligned} y_{it} &= \gamma y_{i,t-1} + x'_{it}\beta + u_{it}, \\ u_{it} &= \mu_i + \varepsilon_{it}, \end{aligned}$$

where $i = 1, 2, \ldots, N$ and $t = 1, 2, \ldots, T$. In this model, y_{it} is the dependent variable, x_{it} is a $(k \times 1)$ vector of exogenous variables, μ_i is the random effect component, and ε_{it} is the general disturbance term. β is a $(k \times 1)$ vector of coefficients and γ is a scalar parameter. In this context, dynamic effects or state dependence relates to the relevance of $y_{i,t-1}$ as a determinant of current values of the dependent variable. The time persistent presence of the term μ_i induces an alternative source of persistence, usually referred as unobserved time-invariante individual heterogeneity or random effects.

To derive the asymptotic properties of our tests, we impose the following regularity assumptions. Define $\tilde{x}_i = (x_{i,1}, ..., x_{i,T})$ and $\tilde{\varepsilon}_i = (\varepsilon_{i,1}, ..., \varepsilon_{i,T})$ as random matrices of dimension $(k \times T)$ and $(1 \times T)$, respectively.

Assumptions. $\{(y_{i0}, \tilde{x}_i, \mu_i, \tilde{\varepsilon}_i) : i = 1, ..., N\}$ *are independent and identically distributed random vectors that satisfy the following requirements:* $\tilde{x}_i$ *is independent of* $(\mu_i, \tilde{\varepsilon}_i)$*;* $E(\tilde{x}_i\tilde{x}'_i)$ *is finite and nonsingular;* y_{i0} *is stochastic and independent of* $(\tilde{x}_i, \mu_i, \tilde{\varepsilon}_i)$*;* μ_i *and* $\tilde{\varepsilon}_i$ *are unobservable and independent of each other with* $\tilde{\varepsilon}_i \sim N(0, \sigma^2_\varepsilon I_T)$ *and* $\mu_i \sim N(0, \omega\sigma^2_\varepsilon)$*. Moreover,* $(\beta', \sigma^2_\varepsilon, \gamma, \omega)'$ *belongs to a compact subset of* $\mathbb{R}^k \times \mathbb{R}_{>0} \times [0, 1) \times \mathbb{R}_{\geq 0}$.

The asymptotic results will be derived assuming that N grows to infinity and T is fixed. This serves for the most common case where the T dimension is short while the number of individuals N is large. Given that T is fixed, our assumptions imposes mild conditions on the time series properties of $\tilde{x}_i$. More specifically, we are just requiring $E(\tilde{x}_i\tilde{x}'_i)$ to be finite and nonsingular, and among other properties, x_{it} may have a unit root. In addition, our assumptions allow us to include a constant term as a component of x_{it}, and hence, the individual effect μ_i may be interpreted as deviation from a common mean. We remark that such a constant term cannot be identified without the restriction $E(\mu_i) = 0$.

The log-likelihood function for our model is given by

$$L(\theta) = -\frac{NT}{2}\ln(\sigma_\varepsilon^2) - \frac{N}{2}\ln(1+T\omega) - \frac{u'u}{2\sigma_\varepsilon^2} + \left(\frac{\omega}{1+T\omega}\right)\left(\frac{u'H_{NT}u}{2\sigma_\varepsilon^2}\right), \tag{1}$$

where $\theta = (\beta', \sigma_\varepsilon^2, \gamma, \omega)'$, $u = (u_{11}, u_{12}, \ldots, u_{it}, \ldots, u_{N(T-1)}, u_{NT})'$, $H_{NT} = I_N \otimes e_T e_T'$, e_T denotes a $(T \times 1)$ vector of ones, and $\otimes$ stands for the Kronecker product. Bhargava and Sargan (1983, pp. 1641) and Hsiao (2003, ch. 4) present a derivation of this function. We refer to these references for further details, in particular for consistency and asymptotic normality of all the parameter estimates in θ, and remark that its functional form depends on the normality of $\tilde{\varepsilon}_i$ and μ_i. In the next section, our proposed tests will be based on the log-likelihood function (1), and therefore, this function will serve for our purposes of establishing the sources of persistence.

3 Tests for Persistent Effects

In our model, a test for the presence of dynamic effects or state dependence corresponds to evaluating $H_0^\gamma : \gamma = 0$. A test for random effects or time-invariant individual heterogeneity involves checking $H_0^\omega : \omega = 0$. And a joint test for the presence of both types of persistence corresponds to evaluating $H_0^{\gamma\omega} : (\gamma, \omega) = (0, 0)$.

To derive Lagrange multiplier (LM) tests, we require the score functions and the Fisher information matrix of the log-likelihood model (1). Denote $d_a(\theta) = \partial L(\theta)/\partial a$ as the score function of $L(\theta)$ with respect to a, where a can be any sub-vector of $(\beta', \sigma_\varepsilon^2, \gamma, \omega)'$. Denote the elements of the Fisher information matrix as $J_{ab}(\theta) = -(NT)^{-1}E[\partial^2 L(\theta)/\partial a \partial b']$, where a, b, c can be any sub-vectors of $(\beta', \sigma_\varepsilon^2, \gamma, \omega)'$. In addition, let define $J_{a,b}(\theta) = J_{aa}(\theta) - J_{ab}(\theta)J_{bb}^{-1}(\theta)J_{ba}(\theta)$ and $J_{ac,b}(\theta) = J_{ac}(\theta) - J_{ab}(\theta)J_{bb}^{-1}(\theta)J_{bc}(\theta) = J'_{ca,b}(\theta)$. Note that these terms involve unknown expectations, thus in order to derive feasible tests, they will be replaced by the corresponding sample analogues evaluated at the parameter values estimated under the null hypothesis. Explicit formulas for $d_a(\theta)$ and $J_{ab}(\theta)$ are provided in Appendices A.1 and A.2, respectively, while details about the construction of the statistics below are given in Appendix A.3. Assuming that N grows to infinity and T is fixed, hereafter, we follow closely Bera and Yoon (1993, Sections 2 and 3) to derive the asymptotic distributions of our tests.

A first approach consists in deriving marginal tests for H_0^{γ}, H_0^{ω}, and a joint test for $H_0^{\gamma\omega}$. By 'marginal' we mean a test for dynamic effects (random effects) assuming no random effects (dynamic effects). The proposed tests can be expressed as

$$LM_a = \frac{1}{NT} d_a(\hat{\theta})' J_{a,b}^{-1}(\hat{\theta}) d_a(\hat{\theta}), \tag{2}$$

where $a = \gamma$, ω, and $(\gamma, \omega)'$, respectively, $b = (\beta', \sigma_\varepsilon^2)'$, and $\hat{\theta} = (\hat{\beta}', \hat{\sigma}_\varepsilon^2)'$ is the maximum likelihood estimate of b under joint null of no persistence, $H_0^{\gamma\omega} : (\gamma, \omega) = (0, 0)$. The formulas for the three LM statistics are

$$\begin{aligned} LM_\gamma &= (NT)\frac{B^2}{C}, & (3)\\ LM_\omega &= (NT)\frac{A^2}{2(T-1)}, \text{ and} & (4)\\ LM_{\gamma\omega} &= (NT)\frac{[B + (A/T)]^2}{C - 2(T-1)/T^2} + (NT)\frac{A^2}{2(T-1)}, & (5) \end{aligned}$$

where $A = 1 - (\hat{u}' H_{NT} \hat{u})/(\hat{u}'\hat{u})$, $B = (y_{-1}'\hat{u})/(\hat{u}'\hat{u})$, $C = (\hat{e}'\hat{e})/(\hat{u}'\hat{u}) + (T-1)/T$, $\hat{u} = Qy$, $Q = I_{NT} - X(X'X)^{-1}X'$, X is a $(NT \times k)$ matrix of regressors, and y is a $(NT \times 1)$ vector of dependent variables. Moreover, $\hat{e} = Q\hat{y}_{-1}$, where $\hat{y}_{-1}$ is a $(TN \times 1)$ vector obtained from the vectorization of the $(T \times N)$ matrix $[\hat{y}_{-1,it}]_{t,i}$ with $\hat{y}_{-1,i1} = y_{i0}$ and $\hat{y}_{-1,it} = x_{i,t-1}'\hat{\beta}$ for $t \geq 2$.[1]

Under the joint null hypothesis of no persistence, the marginal statistics, LM_γ and LM_ω, converge in distribution to $\chi_1^2(0)$ while $LM_{\gamma\omega}$ to $\chi_2^2(0)$, where $\chi_m^2(0)$ denotes a central chi-square distribution with m degrees of freedom. These asymptotic results can be derived from a well-known property of the LM statistics (see for example, Bera and Yoon, 1993, pp. 651): under the null $a = 0$ and when the alternative is correctly specified, the asymptotic distribution of LM_a is $\chi_{\dim(a)}^2(0)$.

The fact that $J_{\gamma\omega,b}(\theta_0) = (T-1)/T \neq 0$, where $\theta_0 = (\beta', \sigma_\varepsilon^2, 0, 0)'$, implies that marginal tests, though useful to determine the falseness of the joint null of no persistence, are

[1] The matrices X and y are ordered following the usual approach in the literature; see e.g., Bera et al (2001) and the vector u defined in Section 2. The notation for the matrix $[\hat{y}_{-1,it}]_{t,i}$ is as follows: the $(t \times i)$-th element of the matrix $[\hat{y}_{-1,it}]_{t,i}$ is $\hat{y}_{-1,it}$. Naturally, LM_ω is the classic Breusch-Pagan (1980) test for random effects.

of limited use for the goal of identifying the source of persistence once the joint null is determined to be false. As established by Davidson and MacKinnon (1987) and Saikkonen (1989), a marginal LM test for one parameter is affected by the other one being incorrectly set to zero. More concretely, Saikkonen (1989)'s results imply that when the source of persistence not tested for is locally misspecified, the marginal LM for the other source will converge to a non-central asymptotic chi-square variable, hence leading to spurious rejections due to the misspecified nuisance parameter and not to the falseness of the null hypothesis.

The precedent argument can be formalized as follows. Consider $a \in \{\gamma, \omega\}$ and $c \in \{\gamma, \omega\} \backslash \{a\}$. Under $H_0^a : a = 0$ but $c = \delta_c / \sqrt{NT}$ with $\delta_c \neq 0$, the marginal statistic LM_a converges in distribution to $\chi_1^2(\lambda_{a(c)})$ with $\lambda_{a(c)} = \delta_c^2 J_{ac,b}^2(\theta_0)/J_{a,b}(\theta_0)$, where $\chi_1^2(\lambda)$ denotes a chi-square distribution with one degree of freedom and non-centrality parameter λ; see for example, Bera and Yoon (1993, eq. 2.2).

After applying this argument to our marginal tests, we obtain the following results. Under $H_0^\gamma : \gamma = 0$ but when $\omega = \delta_\omega / \sqrt{NT}$ with $\delta_\omega > 0$, LM_γ converges in distribution to $\chi_1^2(\lambda_{\gamma(\omega)})$ where the non-centrality parameter $\lambda_{\gamma(\omega)} = [\delta_\omega (T-1)]^2 / [J_{\gamma,b}(\theta_0) T^2]$, where $J_{\gamma,b}(\theta_0)$ is defined in eq. (A.4) of Appendix A.3.1. This means that when testing for state dependence, the presence of time-invariant individual heterogeneity makes the test to wrongly reject the null because of misspecification of the alternative hypothesis. In a similar vein, under $H_0^\omega : \omega = 0$ but when $\gamma = \delta_\gamma / \sqrt{NT}$ with $\delta_\gamma > 0$, LM_ω converges in distribution to $\chi_1^2(\lambda_{\omega(\gamma)})$ with $\lambda_{\omega(\gamma)} = \delta_\gamma^2 2(T-1)/T^2$. Consequently, and in an analogous way to the problem found by Bera et al. (2001), the classic test for random effects by Breusch and Pagan (1980) will reject its null not only due to the presence of unobserved heterogeneity but also due to the presence of state dependence.[2]

In words, when marginal tests reject, they suggest the presence of some persistence without clear indication about which source (if not both) are relevant. Marginal tests do

[2] Although this paper considers $N \to \infty$ and T fixed, we briefly discusses what happens when T also grows to infinity. Observe first that $\lambda_{\omega(\gamma)} \to 0$ as $T \to \infty$, which implies that a local misspecification of the form $\gamma = \delta_\gamma / \sqrt{NT}$ vanishes in large panels. Therefore, LM_ω converges in distribution to a central chi-square with one degree of freedom and this marginal test can be used in the presence of local state dependence. In contrast, since $\lambda_{\gamma(\omega)}$ does not necessarily converges to zero when $T \to \infty$, LM_γ is not robust to a local misspecification in the variance of the random effect component. Whether or not $\lambda_{\gamma(\omega)}$ converges to zero depends on the time series properties of x_{it}, which are beyond the scope of this paper.

not add much information besides the one already provided by the test for the joint null of no persistence. Hereafter, to identify the source of departure away from the joint null of no persistence, $H_0^{\gamma\omega}$, we will follow two strategies.

First, we construct conditional LM tests for H_0^{γ} and H_0^{ω}, where ω and γ, respectively, are estimated by maximum likelihood. By 'conditional' we mean a test for dynamic effects (random effects) considering the presence of non-local random effects (dynamic effects). The two conditional LM statistics are denoted by $LM_{\gamma/\omega}$ and $LM_{\omega/\gamma}$.

A conditional test for the presence of dynamic effects, H_0^{γ}, requires the implementation of a random effects GLS estimator for (β', ω) under $\gamma = 0$. Denoting such an estimator by $(\hat{\beta}'_{\gamma}, \hat{\omega}_{\gamma})$, the formula for $LM_{\gamma/\omega}$ is

$$LM_{\gamma/\omega} = (NT)\frac{B_{\gamma}^2}{C_{\gamma} - 2(T-1)/T^2}, \tag{6}$$

where

$$\begin{aligned}
B_{\gamma} &= (1 + T\hat{\omega}_{\gamma})\frac{y'_{-1}R\hat{u}^{\gamma}}{\hat{u}^{\gamma\prime}H_{NT}\hat{u}^{\gamma}}, \\
C_{\gamma} &= \frac{1 + T\hat{\omega}_{\gamma}}{\hat{u}^{\gamma\prime}H_{NT}\hat{u}^{\gamma}}\{\hat{y}^{\gamma\prime}_{-1}[R - RX(X'RX)^{-1}X'R]\hat{y}^{\gamma}_{-1}\} + \frac{T-1}{T}\left(1 + \frac{\hat{\omega}_{\gamma}^2}{1 + T\hat{\omega}_{\gamma}}\right), \\
R &= I_{NT} - \frac{\hat{\omega}_{\gamma}}{1 + T\hat{\omega}_{\gamma}}H_{NT},
\end{aligned}$$

$\hat{u}^{\gamma} = y - X\hat{\beta}_{\gamma}$, and $\hat{y}^{\gamma}_{-1}$ is a $(TN \times 1)$ vector obtained from the vectorization of the $(T \times N)$ matrix $[\hat{y}^{\gamma}_{-1,it}]_{t,i}$ with $\hat{y}^{\gamma}_{-1,i1} = y_{i0}$ and $\hat{y}^{\gamma}_{-1,it} = x'_{i,t-1}\hat{\beta}_{\gamma}$ for $t \geq 2$. Observe that $[\hat{y}^{\gamma}_{-1,it}]_{t,i}$ is defined in an analogous way to $[\hat{y}_{-1,it}]_{t,i}$. Under the null H_0^{γ}, the asymptotic distribution of $LM_{\gamma/\omega}$ is $\chi_1^2(0)$ regardless of the presence of random effects; see Appendix A.3.2. This test is similar to that of Holtz-Eakin (1988).

A similar conditional test for H_0^{ω} involves a simple OLS estimator of (β', γ). Denoting such estimator by $(\hat{\beta}'_{\omega}, \hat{\gamma}_{\omega})$, the corresponding LM statistic becomes

$$LM_{\omega/\gamma} = (NT)\frac{A_{\omega}^2}{2(T-1) - D_{\omega}}, \tag{7}$$

where $A_{\omega} = 1 - (\hat{u}^{\omega\prime}H_{TN}\hat{u}^{\omega})(\hat{u}^{\omega\prime}\hat{u}^{\omega})$,

$$D_{\omega} = \left[\frac{(T-1) + \sum_{t=2}^{T}(T-t)\hat{\gamma}_{\omega}^{t-1}}{T}\right]^2 \frac{\hat{u}^{\omega\prime}\hat{u}^{\omega}}{\hat{u}'_{-1}\hat{u}_{-1}},$$

$\hat{u}^{\omega} = y - y_{-1}\hat{\gamma}_{\omega} - X\hat{\beta}_{\omega}$, and $\hat{u}_{-1} = Qy_{-1}$. Under the null H_0^{ω}, the asymptotic distribution of $LM_{\omega/\gamma}$ is $\chi_1^2(0)$ regardless of the presence of state dependence; see Appendix A.3.3.

A second strategy avoids estimating the nuisance parameters (that is, still based on the joint null of no persistence) and consists in adjusting the original LM statistic using the robustification procedure of Bera and Yoon (1993, Section 3). This approach allows the construction of a test for a particular source of persistence that does not require the estimation of the parameters of the other one, provided that departures from zero in the nuisance parameters are small. In particular, it is based on assuming local departures in the nuisance parameter and the validity of the tests for non-local departures need to be studied for each case. These tests are referred as robust tests. These tests are useful procedures to evaluate specific departures from a joint null hypothesis. More specifically, for either $a = \gamma$ or $a = \omega$, consider a test for $H_0^a : a = 0$ that is robust to *local* misspecification in the parameter $c \in \{\gamma, \omega\} \backslash \{a\}$ with $c = \delta_c/\sqrt{NT}$. Observe that only the parameters $b = (\beta', \sigma_{\varepsilon}^2)'$ are estimated. The modified Bera-Yoon statistic is given by

$$LM_{a/c}^{*} = \frac{1}{NT} d_{a/c,b}(\hat{\theta})' J_{a/c,b}^{-1}(\hat{\theta}) d_{a/c,b}(\hat{\theta}), \tag{8}$$

where $d_{a/c,b}(\theta) = d_a(\theta) - J_{ac,b}(\theta) J_{c,b}^{-1}(\theta) d_c(\theta)$ and $J_{a/c,b}^{-1}(\theta) = J_{a,b}(\theta) - J_{ac,b}(\theta) J_{c,b}^{-1}(\theta) J_{ca,b}(\theta)'$.

The main result in Bera, Montes-Rojas, and Sosa-Escudero (2009) implies that the modified locally robust Bera-Yoon statistics can be constructed in a simple way, once marginal and joint tests have been derived. Specifically, we have that $LM_{\gamma/\omega}^{*} = LM_{\gamma\omega} - LM_{\omega}$ and $LM_{\omega/\gamma}^{*} = LM_{\gamma\omega} - LM_{\gamma}$; Appendix A.3.1 also provides an alternative derivation based on formula (8). As established in Bera and Yoon (1993, eq. 3.10), the robust tests converge in distribution to $\chi_1^2(0)$ under the corresponding null and in the presence of local misspecification in the unconsidered parameter. That is, modified tests are locally robust to unconsidered sources of persistence.

Naturally, when nuisance parameters are indeed zero, the marginal LM tests are locally optimal implying a sort of 'robustification cost.' That is, a power loss for unnecessarily estimating an additional nuisance parameter that was indeed zero (in the case of conditional tests), or for robustifying a test statistic when the marginal one would have sufficed (in the case of the Bera-Yoon tests). These 'robustification' costs can be quantified using the results of Bera and Yoon (1993, Section 3). To do so, consider the local alternative

$H_1^a : a = \delta_a/\sqrt{NT}$ with $c = 0$, where $a \in \{\gamma, \omega\}$, $c \in \{\gamma, \omega\}\backslash\{a\}$, and $\delta_a \neq 0$. Under this alternative, LM_a and LM_a^* converge in distribution to non-central chi-squares $\chi_1^2(\lambda_a)$ and $\chi_1^2(\lambda_{a/c}^*)$, respectively, where the non-centrality parameters are $\lambda_a = \delta_a^2 J_{a,b}(\theta_0)$ and $\lambda_{a/c}^* = \delta_a^2[J_{a,b}(\theta_0) - J_{ac,b}^2(\theta_0)/J_{c,b}(\theta_0)]$. The presence of a robustification is due to the fact that $\lambda_a \geq \lambda_{a/c}^*$.

The previous argument can be applied to our tests in a straightforward way. When $\gamma = \delta_\gamma/\sqrt{NT}$, $\delta_\gamma > 0$, and $\omega = 0$, LM_γ and $LM_{\gamma/\omega}^*$ converge in distribution to $\chi_1^2(\lambda_\gamma)$ and $\chi_1^2(\lambda_{\gamma/\omega}^*)$, respectively. Under $\omega = \delta_\omega/\sqrt{NT}$, $\delta_\omega > 0$, and $\gamma = 0$, LM_ω and $LM_{\omega/\gamma}^*$ converge in distribution to $\chi_1^2(\lambda_\omega)$ and $\chi_1^2(\lambda_{\omega/\gamma}^*)$, respectively. The non-centrality parameters are given by

$$\begin{aligned}
\lambda_\gamma &= \delta_\gamma^2 J_{\gamma,b}(\theta_0), \\
\lambda_{\gamma/\omega}^* &= \delta_\gamma^2 \left[J_{\gamma,b}(\theta_0) - \frac{2(T-1)}{T^2} \right], \\
\lambda_\omega &= \delta_\omega^2 \frac{(T-1)}{2}, \text{ and} \\
\lambda_{\omega/\gamma}^* &= \delta_\omega^2 \left[\frac{(T-1)}{2} - \frac{(T-1)^2}{J_{\gamma,b}(\theta_0) T^2} \right].
\end{aligned}$$

As can be noted, since $\lambda_\gamma \geq \lambda_{\gamma/\omega}^*$ and $\lambda_\omega \geq \lambda_{\omega/\gamma}^*$, the asymptotic power of the robust statistics is less (or equal) than that of the marginal statistics when there is no misspecification. Due to the shape of the Fisher information matrix, particularly $J_{\beta\sigma_\varepsilon^2}(\theta_0) = J_{\beta\omega}(\theta_0) = 0_{[k\times 1]}$ and $J_{\gamma\sigma_\varepsilon^2}(\theta_0) = 0$, the statistics $LM_{\gamma/\omega}$ and $LM_{\omega/\gamma}$ also converge in distribution to $\chi_1^2(\lambda_{\gamma/\omega}^*)$ and $\chi_1^2(\lambda_{\omega/\gamma}^*)$, respectively. This result implies that both conditional and Bera-Yoon robust tests have the same asymptotic power. This is important in practice, since this result implies that when local misspecifications are small, there are no power gains of estimating nuisance parameters when the goal is to detect whether a particular source of persistence is active.

The performance of the robust test in a non-local context, and the importance of the robustification and conditioning costs in small samples is an empirical question that will be studied through the extensive Monte Carlo experiment of the following section.

4 Monte Carlo Experiments

The results of the previous section suggest three testing strategies to detect persistence and to identify which source (unobserved heterogeneity, state dependence, or both) is active. The first strategy is based on estimating the model under the joint null hypothesis of no persistence effects. This leads to two marginal tests and a joint test. A second strategy derives conditional tests, that is, tests for one source after having estimated the relevant parameters that handle the other one. The final strategy, based on the Bera-Yoon (1993) principle, produces robustified marginal tests, that are still based on the joint null, and hence avoids estimating nuisance parameters.

There are several concerns that deserve to be explored empirically. First, as mentioned before, the use of robustified or conditional tests may imply a power loss when marginal tests would have sufficed, i.e., when the source not being considered is indeed inactive. Second, modified tests are meant to be resistant to misspecified alternatives in a *local* sense (that is, for small deviations from zero in the nuisance parameter), so its performance in a non-local context is a matter of concern. Third, the likelihood framework involves a strict normality assumption whose relevance must be assessed. Finally, and for all testing procedures, the adequacy of asymptotic approximations for sample sizes similar to those used in practice, is a much relevant issue. The purpose of this section is to study these issues empirically, through a Monte Carlo experiment.

To facilitate comparison, we use a design similar to the one used in previous work on the subject: Bera et al. (2001) and Baltagi, Chang, and Li (1992). We refer to these papers for further details. We consider different values of (γ, ω) and (N, T). The data generating process (DGP) is:

$$\begin{aligned} y_{it} &= \gamma y_{it-1} + \alpha + x_{it}\beta + u_{it}, \\ u_{it} &= \mu_i + \varepsilon_{it}, \end{aligned}$$

where $(\alpha, \beta) = (5, 0.5)$, $\mu_i \sim N(0, 20\omega)$, and $\varepsilon_{it} \sim N(0, 20)$. The independent variable x_{it} is generated following Nerlove (1971), i.e., $x_{it} = 0.1t + 0.5x_{it-1} + w_{it}$, where w_{it} is uniformly distributed on the interval $[-0.5, 0.5]$ and $x_{i0} \sim 5 + U(-5, 5)$. The initial value y_{i0} is taken from the uniform distribution in $[-1, 1]$. The number of replications is 5000 and the nominal size is 5%.

In order to study the empirical size, Table 1 reports rejection rates for alternative sample sizes with $N \in \{50, 100\}$ and $T \in \{2, 5, 10\}$, while the parameters are set at the joint null $H_0^{\gamma\omega} : \gamma = \omega = 0$. As can be noted, the empirical size is in general below 5%, that is, they are undersized. In all cases, except for $LM_{\gamma/\omega}$, empirical size gets closer to the nominal size as T increases. For $LM_{\gamma/\omega}$, however, a large T reduces the rejection rates.

In order to explore further the asymptotic and small sample properties of the test, Table 2 reports simulations with $N \in \{10, 20, 50, 100, 200, 1000\}$ and $T = 5$. The top rows study the empirical size of the test under $H_0^{\gamma\omega}$. The results show that the empirical size remains undersized for small and large N for most of the tests. However, the simulations show that the tests have good size properties for small N.

Table 2 also studies the power properties of the tests for $N \in \{10, 20, 50, 100, 200, 1000\}$ and $T = 5$. Simulations for $(\gamma = 0.1, \omega = 0)$ show that tests for detecting dynamic persistence, $H_0^{\gamma} : \gamma = 0$, that is, LM_{γ}, $LM_{\gamma/\omega}$ and $LM^*_{\gamma/\omega}$, are all consistent as $N \to \infty$. As expected the marginal test LM_{γ} has the greatest power, followed by the robust $LM^*_{\gamma/\omega}$, and finally the conditional $LM_{\gamma/\omega}$.[3] In this case, a value of $\gamma \neq 0$ affects the marginal test for unobserved heterogeneity, LM_{ω}, making it to wrongly reject its null hypothesis of $H_0^{\omega} : \omega = 0$. However, the conditional test $LM_{\omega/\gamma}$, which estimates γ, corrects the rejection rates and makes them similar to the top rows. The Bera-Yoon robustification procedure $LM^*_{\omega/\gamma}$ partially corrects the rejection rates, which achieve a value of 0.141 with the largest $N = 1000$.

Simulations for $(\gamma = 0, \omega = 0.1)$ show that tests for detecting unobserved heterogeneity, $H_0^{\omega} : \omega = 0$, that is, LM_{ω}, $LM_{\omega/\gamma}$ and $LM^*_{\omega/\gamma}$, are also consistent as $N \to \infty$. As in the previous case, the greatest power is achieved by the marginal test, followed by the robust and the conditional tests. Moreover, a value of $\omega \neq 0$ affects the marginal test for dynamic persistence, LM_{γ}, making it to wrongly reject its null hypothesis of $H_0^{\gamma} : \gamma = 0$. However, the conditional test $LM_{\gamma/\omega}$, which estimates γ, reduces the rejection rates but make them very undersized (empirical size goes to 0 as $N \to \infty$). The Bera-Yoon robust test $LM^*_{\gamma/\omega}$ fully corrects the rejection rates, with similar values to those achieved in the top rows.

Table 3 explores power for different values of ω and γ. We report the case $(N, T) =$

[3]However, as noted by an anonymous referee power comparisons require size-correction. Given the difficulty of doing these corrections in empirical work we do not pursue this strategy here and all power comparisons are evaluated using the actual rejection rates.

$(50, 5)$. Results for alternative sizes only reinforce those of this case, and are omitted to save space, and available from the authors by request. Consider first the case when the only source of persistence is due to individual random effects, that is, ω is allowed to vary while keeping $\gamma = 0$. First, and as predicted by the theory, the marginal LM test (LM_γ) is negatively affected by model misspecification ($\omega \neq 0$), that is, it spuriously rejects the null of no dynamic effects due to the relevance of random effects. Interestingly, the conditional LM and the Bera-Yoon robust tests have decreasing size as ω increases. Power is increasing for all tests specifically designed to react to random individual effects, i.e., LM_ω, $LM_{\omega/\gamma}$, $LM^*_{\omega/\gamma}$, and for the joint tests $LM_{\gamma\omega}$. As expected the highest power is achieved by the optimal marginal test (LM_γ), followed by the Bera-Yoon robust test, and then by the conditional LM test. A very important result is that the Bera-Yoon robustifed procedure has a smaller cost (in terms of power) than that of the conditional LM test, where the additional ω parameter is estimated by maximum likelihood. In addition, the robustification cost is very small.

The case where only dynamic effects induce persistence shows comparable results. The Breusch-Pagan marginal LM test, LM_ω, is negatively affected by the presence of dynamic effects, whereas the conditional test, $LM_{\omega/\gamma}$, has correct size. The Bera-Yoon test, $LM^*_{\omega/\gamma}$, is affected by misspecification ($\gamma \neq 0$) but its rejection rates are much better than those of the marginal LM test. This emphasizes the fact that Bera-Yoon robustification procedure works when local departures from the joint null are considered. The highest power is again that of the optimal marginal test, followed by the Bera-Yoon robust test, and then by the conditional LM test. The comparison of the latter two show that the Bera-Yoon robustification procedure has a smaller cost in terms of power than that of the conditional LM test, where the additional parameter γ is estimated by maximum likelihood.

Finally, Table 4 evaluates the performance of the test statistics under non Gaussian DGP's. Specifically, both μ and ε follow either a t-Student with 4 degrees of freedom or a χ^2 with 1 degree of freedom. In this case, we repeat the same specification as in Table 3 for different values of ω and γ when $(N, T) = (50, 5)$. The table shows that our tests (derived under normality) still have correct empirical size and excellent power even when other DGP's are used.

5 Empirical Application: Income Growth

As an application of the proposed tests, we study the source of persistent behavior in the series of real GDP per capita growth in a panel data set of countries. Understanding the behavior of this series contributes to the long-standing debate about convergence rates. Both sources of persistence, in the form of unobserved heterogeneity and state-dependence, are recurrently cited in the empirical literature on economic growth. First, country-specific unobserved effects can be interpreted as differences in the countries' technology parameters in Solow-Swan production function regressions that correspond to differences in country-specific variables, e.g., institutions, natural resources, etc.[4] This raises the suspicion that underdevelopment is a state of equilibrium and that there are forces at work that tend to restore the equilibrium every time there are small improvements in living conditions. Second, dynamic persistence can be associated with the effect of past shocks or economic decisions on the countries' growth. For instance, Rosenstein-Rodan's (1943) 'big-push' theory stated that countries needed a large inflow of capital to break the vicious cycle of poverty. In this case, income shocks (natural disasters, wars) have enduring consequences on the country's income growth. Understanding the specific source of persistence (if any or both) helps to understand differences between poor and rich countries and the nature of economic development.

To explore these alternative persistence patterns, we consider the model

$$\begin{aligned} g_{it} &= \gamma g_{i,t-1} + \beta_1 + \beta_2 t + u_{it}, \\ u_{it} &= \mu_i + \varepsilon_{it}, \end{aligned}$$

where $i = 1, 2, \ldots, N$, $t = 1, 2, \ldots, T$, g_{it} is real GDP per capita growth, μ_i is the country-specific effect component, and ε_{it} is the general disturbance term. We consider models

[4] For instance, Graham and Temple (2006) find that multiple equilibria are associated to differences in aggregate total factor productivity. It is also reasonable to assume that these country-specific effects are themselves functions of the capital stock, as in Romer (1986) and Azariadis and Drazen (1990), or that they depend on the initial conditions of the endogenous variables in the presence of historical self-reinforcement (Mookherjee and Ray, 2001). The theory of different 'convergence clubs' (Baumol, 1986; DeLong, 1988; and Quah, 1993,1996,1997) relates to the existence of an exclusionary mechanism that keeps members of one group or club facing a lower level equilibrium from moving to another group or club with a higher level equilibrium. Moreover, this gives the idea of a vicious circle of poverty as a 'constellation of forces tending to act and react upon one another in such a way as to keep a poor country in a state of poverty' (Nurkse, 1953, pp. 4).

with and without a time trend.

We use data on real GDP per capita growth, calculated as the difference of log real GDP per capita, from the Penn World Tables (series `rgdpl`, PPP GDP per capita at 2005 constant prices). Our dataset is a balanced panel of 109 countries over the period 1960 to 2005, containing five year periods. Thus, we have $T = 8$ with growth periods 1960-1965, 1965-1970,. . . ,2000-2005; and $N = 109$.[5] The average logarithmic growth rate is 0.092 with a standard deviation of 0.178.

The tests developed in this paper appear in Table 5. The joint test, $LM_{\gamma\omega}$, indicates strong persistence in the panel data in both models, with and without a time trend. Also both marginal tests, LM_{γ} and LM_{ω}, indicate that both sources of persistence are present. Nevertheless, as discussed in the Sections 3 and 4, this conclusion can be misleading because marginal tests are not useful to detect the source of persistence. Moreover, the parameter estimates differ considerably depending on the estimated model. The parameter estimate for γ assuming H_0^{ω} is $\hat{\gamma} = 0.244$ and $\hat{\gamma} = 0.249$ for the model with and without time trend, respectively. The parameter estimate for ω assuming H_0^{γ} is $\hat{\omega} = 0.149$ and $\hat{\omega} = 0.148$ for the model with and without time trend, respectively. These figures suggest that models with and without time trend mostly coincide. However, when estimating the full model with maximum likelihood we obtain ($\hat{\gamma} = 0.185, \hat{\omega} = 0.055$) for the model with time trend and ($\hat{\gamma} = 0.249, \hat{\omega} = 0$; the binding constraint $\omega = 0$ is reached) for the model with no time trend. These estimates indicate that only dynamic persistence is present but individual heterogeneity is not and that testing for the presence of only one persistence is necessary in order to avoid the unnecessary inclusion of country-specific heterogeneity.

The conditional tests, $LM_{\gamma/\omega}$ and $LM_{\omega/\gamma}$, indeed suggest that only dynamic persis-

[5]The countries included in the sample are Argentina, Australia, Austria, Burundi, Belgium, Benin, Burkina Faso, Bangladesh, Bolivia, Brazil, Barbados, Botswana, Central African Republic, Canada, China, Switzerland, Chile, Cote d'Ivoire, Cameroon, Congo, Republic of, Colombia, Comoros, Cape Verde, Costa Rica, Cyprus, Denmark, Dominican Republic, Algeria, Ecuador, Egypt, Spain, Ethiopia, Finland, Fiji, France, Gabon, United Kingdom, Ghana, Guinea, Gambia, Guinea-Bissau, Equatorial Guinea, Greece, Guatemala, Hong Kong, Honduras, Haiti, Indonesia, India, Ireland, Iran, Iceland, Israel, Italy, Jamaica, Jordan, Japan, Kenya, Korea, Republic of, Sri Lanka, Lesotho, Luxembourg, Morocco, Madagascar, Mexico, Mali, Mozambique, Mauritania, Mauritius, Malawi, Malaysia, Namibia, Niger, Nigeria, Nicaragua, Netherlands, Norway, Nepal, New Zealand, Pakistan, Panama, Peru, Philippines, Papua New Guinea, Puerto Rico, Portugal, Paraguay, Romania, Rwanda, Senegal, Singapore, El Salvador, Sweden, Seychelles, Syria, Chad, Togo, Thailand, Trinidad and Tobago, Turkey, Taiwan, Tanzania, Uganda, Uruguay, United States, Venezuela, South Africa, Congo, Dem. Rep., Zambia, Zimbabwe.

tence is observed. In fact, the conditional test for unobserved time-invariant heterogeneity with a time trend, $LM_{\omega/\gamma}$ accepts the null hypothesis H_0^ω at the 5% significance level. The Bera-Yoon robust test $LM^*_{\gamma/\omega}$ indicates that the null hypothesis of absence of dynamic persistence is rejected but also the test $LM^*_{\omega/\gamma}$ indicates rejection. In order to evaluate the validity of the Bera-Yoon tests we consider that deviations from $\gamma = 0$, the nuisance parameter for testing H_0^ω, are large in all considered models and therefore, it might not be valid to take those γ values as local/small departures as analyzed in the Monte Carlo simulations. Overall these results show that dynamic persistence appears to explain countries' growth differences with an autoregressive parameter of 0.25, and that country-specific heterogeneity does not explain growth persistence once dynamic persistence is taken into account. Thus, countries' growth performance is path-dependent and is not conditioned by the countries' specific characteristics.

6 Conclusion

This paper derives simple tests for persistent effects in a dynamic linear panel data model with unobserved individual effects. It improves upon the previous literature by handling state persistence through a truly dynamic model, instead of relegating it to first order serial correlation in the error term, which is seen as just one particular restriction that arises from imposing a common factor restriction on the general specification. This is in line with the classical literature on dynamic econometrics, that strongly emphasizes general dynamic structures. The classic test by Breusch-Pagan (1980) for random effect is found to be negatively affected by dynamic misspecification, that is, when it rejects its null it is due to unobserved heterogeneity and/or dynamic misspecification, along the results previously found by Bera et al. (2001).

We suggest two alternatives to identify the sources of persistence. The first 'conditional' strategy involves estimating the parameters handling the source of persistence not tested for. The second 'robust' strategy is based on the Bera-Yoon (1993) principle. A main advantage of the latter is that is does not require previous estimation of nuisance parameters, and hence can be implemented after estimating a pooled panel model with no persistence.

A Monte Carlo study shows that the conditional and robust tests perform well in small samples. Specifically, they do not suffer from the oversize of marginal tests, and also, they have power only in the direction designed even in non-local contexts. When the alternative hypothesis is correctly specified, the power loss with respect to the optimal marginal test is very small, although it is not optimal. Furthermore, they still perform well in non-gaussian contexts.

An important advantage of our tests is that they can be implemented after pooled OLS estimation of a static model with no random or dynamic effects. This is relevant in practice, in light of the well known concerns affecting instrumental variables /GMM strategies, aimed at dealing with biases induced by the presence of lagged dependent variables in a linear panel model (see Bond, 2002, for a useful review of advantages and disadvantages or linear dynanic panel specifications). Our proposed tests have the ability of distinguishing which source of persistence is active (random individual or dynamic effects) without requiring the estimation of a dynamic structure, based on simple OLS estimation. Hence the results of our tests should be useful to decide whether it is relevant to involve a truly dynamic model (when persistences are due to dynamic misspecification) or whether a simpler, random effects structure would suffice (when random effects are the sole source of persistence). From a practical perspective, the results in this paper suggest to start with a joint test for both sources of persistence, and then if the null hypothesis of no persistence is rejected, conditional and robust tests should be used to evaluate which source of persistence is present, while marginal tests can be misleading.

References

Anderson, T. W., Hsiao, C. (1982). Formulation and estimation of dynamic models using panel data. *Journal of Econometrics*. 18:47-82.

Angrist, J.D., Pischke, J.S. (2009). *Mostly Harmless Econometrics. An Empiricist's Companion*. Princeton: Princeton University Press.

Arellano, M., Bond, S. (1991). Some tests of specification for panel data: Monte Carlo evidence and an application to employment equations. *Review of Economic Studies*. 58:277-297.

Azariadis, C., and A. Drazen (1990). Threshold Externalities in Economic Development. *Quarterly Journal of Economics*. 105:501-526.

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

Baltagi, B. H., Li, Q. (1995). Testing for AR(1) against MA(1) disturbances in and eror component model. *Journal of Econometrics*. 68.

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

Baumol, W. (1986). Productivity Growth, Convergence, and Welfare: What the Long-Run Data Show. *American Economic Review*. 76:1072-1085.

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

Bera, A., Montes-Rojas, G. and Sosa-Escudero, W. (2010) General specification testing with locally misspecified models. *Econometric Theory*. 26:1838-1845.

Bera, A. K., Yoon, M. (1993). Specification testing with locally misspecified alternatives. *Econometric Theory*. 9:649-658.

Bera, A. K., Sosa-Escudero, W., Yoon, M. (2001). Test for the error component model in the presence of local misspecification. *Journal of Econometrics*. 101:1-23.

Bhargava, A., Sargan, J. D. (1983). Estimating dynamic random effects models from panel data covering short time periods. *Econometrica*. 51:1635-1659.

Bond, S. (2002). Dynamic panel data models: a guide to micro data methods and practice. CENMMAP working paper CWP09/02

Breusch, T. S., Pagan, A. R. (1980). The lagrange multiplier test and its application to model specification in econometrics. *Review of Economic Studies*. 47:239-253.

Davidson, J., MacKinnon, J. (1987). Implicit alternatives and the local power of test statistics. *Econometrica*. 55:1305-1329.

DeLong, J. B. (1988). Productivity growth, convergence, and welfare: Comment. *American Economic Review*. 78:1138-1154.

Graham, B. and Temple, J. (2006). Rich nations poor nations: how much can multiple equilibria explain?. *Journal of Economic Growth*. 11:5-41.

Hendry, D. F. (1995). *Dynamic Econometrics*. New York: Oxford University Press.

Hendry, D. F. and Mizon, G.E. (1978). Serial correlation as a convenient simplification, not a nuisance: A comment on a study of the demand for money by the Bank of England. *Economic Journal*. 88:549-563.

Holtz-Eakin, D. (1988). Testing for individual effects in autorregresive models. *Journal of Econometrics*. 39:297-307.

Hsiao, C. (2003). *Analysis of Panel Data*. New York: Cambridge University Press.

Judson, R. A. and Owen, A. L. (1999). Estimating dynamic panel data models: a guide for macroeconomists. *Economic Letters*. 65:9-15.

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

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

Mookherjee, D. and Ray, D. (2001). *Readings in the theory of economic development*. New York: Blackwell Publishing.

Nurkse, R. (1953). *Problems of capital formation in underdeveloped countries*. New York: Oxford University Press.

Quah, D. (1993). Empirical cross-section dynamics in economic growth. *European Economic Review*. 37:426-434.

Quah, D. (1996). Convergence empirics across economies with (some) capital mobility. *Journal of Economic Growth*. 1:95-124.

Quah, D. (1997). Empirics for Growth and Distribution: Stratification, Polarization, and Convergence Clubs. *Journal of Economic Growth*. 2:27-59.

Romer, P. (1986). Increasing returns and long-run growth. *Journal of Political Economy*. 94:1002-1037.

Rosenstein-Rodan, P. (1943). The problem of industrialization of eastern and south-eastern Europe. *Economics Journal*. 53:202-211.

Saikkonen, P. (1989). Asymptotic relative efficiency of the classical tests under misspecification. *Journal of Econometrics*. 42:351-369.

Sosa-Escudero, W., Marchionni, M. and Arias, O. (2011). Sources of income persistence: evidence from rural El Salvador. *Journal of Income Distribution*. 20:3-28.

Wooldridge, J. M. (2002). *Econometric Analysis of Cross Section and Panel Data*. Massachusetts: MIT Press.

A Appendix: Derivation of the Test Statistics

A.1 First and Second Derivatives of the Log-Likelihood Function

The below results are helpful to obtain the Fisher information matrix. The first partial derivatives of the function L with respect to θ are:

$$
\begin{aligned}
\frac{\partial L(\theta)}{\partial \beta} &= \frac{X'u}{\sigma_\varepsilon^2} - \frac{\omega}{\sigma_\varepsilon^2(1+T\omega)}(X'H_{NT}u), \\
\frac{\partial L(\theta)}{\partial \sigma_\varepsilon^2} &= -\frac{NT}{2\sigma_\varepsilon^2} + \frac{u'u}{2(\sigma_\varepsilon^2)^2} - \frac{\omega}{2(\sigma_\varepsilon^2)^2(1+T\omega)}(u'H_{NT}u), \\
\frac{\partial L(\theta)}{\partial \gamma} &= \frac{y_{-1}'u}{\sigma_\varepsilon^2} - \frac{\omega}{\sigma_\varepsilon^2(1+T\omega)}(y_{-1}'H_{NT}u), \text{ and} \\
\frac{\partial L(\theta)}{\partial \omega} &= -\frac{NT}{2(1+T\omega)} + \frac{u'H_{NT}u}{2(\sigma_\varepsilon^2)(1+T\omega)^2}.
\end{aligned}
$$

From these expressions, the second derivatives can be written as

$$
\begin{aligned}
\frac{\partial^2 L(\theta)}{\partial \beta^2} &= -\frac{X'X}{\sigma_\varepsilon^2} + \frac{\omega}{\sigma_\varepsilon^2(1+T\omega)}(X'H_{NT}X), \\
\frac{\partial^2 L(\theta)}{\partial \beta \partial \sigma_\varepsilon^2} &= -\frac{X'u}{(\sigma_\varepsilon^2)^2} + \frac{\omega}{(\sigma_\varepsilon^2)^2(1+T\omega)}(X'H_{NT}u), \\
\frac{\partial^2 L(\theta)}{\partial \beta \partial \gamma} &= -\frac{X'y_{-1}}{\sigma_\varepsilon^2} + \frac{\omega}{\sigma_\varepsilon^2(1+T\omega)}(X'H_{NT}y_{-1}), \\
\frac{\partial^2 L(\theta)}{\partial \beta \partial \omega} &= -\frac{1}{\sigma_\varepsilon^2(1+T\omega)^2}(X'H_{NT}u), \\
\frac{\partial^2 L(\theta)}{(\partial \sigma_\varepsilon^2)^2} &= \frac{NT}{2(\sigma_\varepsilon^2)^2} - \frac{u'u}{(\sigma_\varepsilon^2)^3} + \frac{\omega}{(\sigma_\varepsilon^2)^3(1+T\omega)}(u'H_{NT}u), \\
\frac{\partial^2 L(\theta)}{\partial \sigma_\varepsilon^2 \partial \gamma} &= -\frac{y_{-1}'u}{(\sigma_\varepsilon^2)^2} + \frac{\omega}{(\sigma_\varepsilon^2)^2(1+T\omega)}(y_{-1}'H_{NT}u), \\
\frac{\partial^2 L(\theta)}{\partial \sigma_\varepsilon^2 \partial \omega} &= -\frac{u'H_{NT}u}{2(\sigma_\varepsilon^2)^2(1+T\omega)^2}, \\
\frac{\partial^2 L(\theta)}{\partial \gamma^2} &= -\frac{y_{-1}'y_{-1}}{\sigma_\varepsilon^2} + \frac{\omega}{\sigma_\varepsilon^2(1+T\omega)}(y_{-1}'H_{NT}y_{-1}), \\
\frac{\partial^2 L(\theta)}{\partial \gamma \partial \omega} &= -\frac{y_{-1}'H_{NT}u}{\sigma_\varepsilon^2(1+T\omega)^2}, \text{ and} \\
\frac{\partial^2 L(\theta)}{\partial \omega^2} &= \frac{NT^2}{2(1+T\omega)^2} - \frac{T}{(\sigma_\varepsilon^2)(1+T\omega)^3}(u'H_{NT}u).
\end{aligned}
$$

A.2 Fisher Information Matrix

Define now $J(\theta) = -(NT)^{-1}E[\partial^2 L(\theta)/\partial\theta\partial\theta']$ as the Fisher information evaluated at the true parameters $\theta = (\beta', \sigma_\varepsilon^2, \gamma, \omega)'$. Using previous results, it is easy to show that

$$J(\theta) = \begin{pmatrix} J_{\beta\beta} & J_{\beta\varepsilon} & J_{\beta\gamma} & J_{\beta\omega} \\ J'_{\beta\varepsilon} & J_{\varepsilon\varepsilon} & J_{\varepsilon\gamma} & J_{\varepsilon\omega} \\ J'_{\beta\gamma} & J_{\varepsilon\gamma} & J_{\gamma\gamma} & J_{\gamma\omega} \\ J'_{\beta\omega} & J_{\varepsilon\omega} & J_{\gamma\omega} & J_{\omega\omega} \end{pmatrix}, \tag{A.1}$$

with

$$\begin{aligned}
J_{\beta\beta} &= \frac{1}{T\sigma_\varepsilon^2}\left[E(\tilde{x}_1\tilde{x}_1') - \frac{\omega}{(1+T\omega)}E(\tilde{x}_1 e_T e_T' \tilde{x}_1')\right], \\
J_{\beta\varepsilon} &= 0_{[k\times 1]}, \\
J_{\beta\gamma} &= \frac{1}{T\sigma_\varepsilon^2}\left[E(\tilde{x}_1\tilde{y}_{1,-1}') - \frac{\omega}{(1+T\omega)}E(\tilde{x}_1 e_T e_T' \tilde{y}_{1,-1}')\right], \\
J_{\beta\omega} &= 0_{[k\times 1]}, \\
J_{\varepsilon\varepsilon} &= \frac{1}{2(\sigma_\varepsilon^2)^2}, \\
J_{\varepsilon\gamma} &= 0_{[1\times 1]}, \\
J_{\varepsilon\omega} &= \frac{1}{2\sigma_\varepsilon^2(1+T\omega)}, \\
J_{\gamma\gamma} &= \frac{1}{T\sigma_\varepsilon^2}\left[E(\tilde{y}_{1,-1}\tilde{y}_{1,-1}') - \frac{\omega}{(1+T\omega)}E(\tilde{y}_{1,-1} e_T e_T' \tilde{y}_{1,-1}')\right], \\
J_{\gamma\omega} &= E\left[\frac{\tilde{y}_{1,-1} e_T e_T' \tilde{u}_1'}{T\sigma_\varepsilon^2(1+T\omega)^2}\right], \text{ and} \\
J_{\omega\omega} &= \frac{T}{2(1+T\omega)^2},
\end{aligned}$$

where $\tilde{u}_i = (u_{i1}, \ldots, u_{iT})$ and $\tilde{y}_{i,-1} = (y_{i0}, y_{i1}, \ldots, y_{i(T-1)})$ are $(1 \times T)$ vectors whose t-th components are u_{it} and $y_{i,t-1}$, respectively. To derive above expressions, recall that the expectation of the score is zero at the true parameters (see for example, $J_{\varepsilon\varepsilon}$ and $J_{\varepsilon\gamma}$), and also that $(y_{i0}, x_{i1}', \ldots, x_{iT}')$ are identically distributed across i, so the expectations have been written in terms of $i = 1$.

A.3 Construction of Statistics

Before proceeding, the next expressions are helpful to construct the statistics. Under our assumptions and when $\gamma = 0$, for any $(\beta', \sigma_\varepsilon^2, \omega) \in \mathbb{R}^k \times \mathbb{R}_{>0} \times \mathbb{R}_{\geq 0}$, we have that

$$\begin{aligned}
E(\tilde{x}_1 \tilde{y}'_{1,-1}) &= E(\tilde{x}_1 \tilde{y}_1^{e\prime}), \\
E(\tilde{x}_1 e_T e'_T \tilde{y}'_{1,-1}) &= E(\tilde{x}_1 e_T e'_T \tilde{y}_1^{e\prime}), \\
E(\tilde{y}_{1,-1} \tilde{y}'_{1,-1}) &= E(\tilde{y}_1^e \tilde{y}_1^{e\prime}) + (T-1)(1+\omega)\sigma_\varepsilon^2, \text{ and} \\
E(\tilde{y}_{1,-1} e_T e'_T \tilde{y}'_{1,-1}) &= E(\tilde{y}_1^e e_T e_T \tilde{y}_1^{e\prime}) + (T-1)[1+(T-1)\omega]\sigma_\varepsilon^2,
\end{aligned}$$

where $\tilde{y}_1^e = (y_{10}, x'_{11}\beta, \ldots, x'_{1(T-1)}\beta)$ is a $(1 \times T)$ vector whose t-th component is the conditional expectation $E[y_{1(t-1)}|y_{i0}, \tilde{x}_1]$ under $\gamma = 0$. In addition, the following results hold:

1. When $\gamma = 0$, for any $(\beta', \sigma_\varepsilon^2, \omega) \in \mathbb{R}^k \times \mathbb{R}_{>0} \times \mathbb{R}_{\geq 0}$, we have that

$$E\left[\frac{\tilde{y}_{1,-1} e_T e'_T \tilde{u}_1}{T\sigma_\varepsilon^2 (1+T\omega)^2}\right] = \frac{T-1}{T(1+T\omega)}.$$

2. When $\omega = 0$, for any $(\beta', \sigma_\varepsilon^2, \gamma) \in \mathbb{R}^k \times \mathbb{R}_{>0} \times (-1, 1)$, we have that

$$E\left[\frac{\tilde{y}_{1,-1} e_T e'_T \tilde{u}_1}{T\sigma_\varepsilon^2 (1+T\omega)^2}\right] = \frac{1}{T}\left[(T-1) + \sum_{t=2}^{T} (T-t)\gamma^{t-1}\right]. \tag{A.2}$$

   This expression is obtained by induction on T.

In the next subsections, we build the test statistics exploiting the formulas LM_a and $LM^*_{a/c}$ detailed in Section 3.

A.3.1 Construction of LM_γ, LM_ω, $LM^*_{\gamma/\omega}$, $LM^*_{\omega/\gamma}$, $LM_{\gamma\omega}$

These statistics can be computed by estimating just the restricted model, that is, by estimating $(\beta', \sigma_\varepsilon^2)$ under $(\gamma, \omega) = (0, 0)$. Denote such estimate by $\hat{\theta}_0 = (\hat{\beta}', \hat{\sigma}_\varepsilon^2, 0, 0)'$ and recall that we have defined $\theta_0 = (\beta', \sigma_\varepsilon^2, 0, 0)'$, as well as, $b = (\beta', \sigma_\varepsilon^2)'$. From the general definition of $J_{a,b}(\theta)$ in Section 3, we can write

$$\begin{aligned}
J_{\gamma,b}(\theta_0) &= J_{\gamma\gamma}(\theta_0) - J_{\gamma b}(\theta_0)J_{bb}^{-1}(\theta_0)J_{b\gamma}(\theta_0), \\
J_{\omega,b}(\theta_0) &= J_{\omega\omega}(\theta_0) - J_{\omega b}(\theta_0)J_{bb}^{-1}(\theta_0)J_{b\omega}(\theta_0), \\
J_{\gamma\omega,b}(\theta_0) &= J_{\gamma\omega}(\theta_0) - J_{\gamma b}(\theta_0)J_{bb}^{-1}(\theta_0)J_{b\omega}(\theta_0), \text{ and} \\
J_{(\gamma,\omega)',b}(\theta_0) &= J_{(\gamma,\omega)'(\gamma,\omega)'}(\theta_0) - J_{(\gamma,\omega)'b}(\theta_0)J_{bb}^{-1}(\theta_0)J_{b(\gamma,\omega)'}(\theta_0).
\end{aligned}$$

Note that $J_{\gamma\omega,b}(\theta_0)$ is different from $J_{(\gamma,\omega)',b}(\theta_0)$. After combining the above terms with the expressions of eq. (A.1), we obtain

$$\begin{aligned}
J_{\gamma,b}(\theta_0) &= \frac{1}{T\sigma_\varepsilon^2}\left\{E(\tilde{y}_{1,-1}\tilde{y}_{1,-1}') - E(\tilde{y}_{1,-1}\tilde{x}_1')[E(\tilde{x}_1\tilde{x}_1')]^{-1}E(\tilde{x}_1\tilde{y}_{1,-1}')\right\}, \\
J_{\omega,b}(\theta_0) &= \frac{T-1}{2}, \\
J_{\gamma\omega,b}(\theta_0) &= \frac{T-1}{T}, \text{ and} \\
J_{(\gamma,\omega)',b}(\theta_0) &= \begin{pmatrix} J_{\gamma,b}(\theta_0) & J_{\gamma\omega,b}(\theta_0) \\ J_{\gamma\omega,b}(\theta_0) & J_{\omega,b}(\theta_0) \end{pmatrix}. \qquad \text{(A.3)}
\end{aligned}$$

From our assumptions and since $(\gamma, \omega) = (0, 0)$, it can be shown that

$$J_{\gamma,b}(\theta_0) = \frac{1}{T\sigma_\varepsilon^2}\left\{E(\tilde{y}_1^e\tilde{y}_1^{e\prime}) - E(\tilde{y}_1^e\tilde{x}_1')[E(\tilde{x}_1\tilde{x}_1')]^{-1}E(\tilde{x}_1\tilde{y}_1^{e\prime})\right\} + \frac{T-1}{T}, \qquad \text{(A.4)}$$

where $\tilde{y}_i^e = (y_{i0}, x_{i1}'\beta, \ldots, x_{i(T-1)}'\beta)$ is a $(1 \times T)$ vector whose t-th component is the conditional expectation $E[y_{i(t-1)}|y_{i0}, \tilde{x}_i]$ under $\gamma = 0$.

In order to build feasible test statistics, which can be computed from a random sample, we replace the unknown expectation of expression (A.3) with the corresponding sample analogues; for example, $E(\tilde{x}_1\tilde{x}_1')$ is replaced by $(1/N)\sum_{i=1}^N \tilde{x}_i\tilde{x}_i'$. After doing so,

$$\hat{J}_{\gamma,b}(\hat{\theta}_0) = \frac{\hat{y}_{-1}'Q\hat{y}_{-1}}{(NT)\hat{\sigma}_\varepsilon^2} + \frac{T-1}{T} = \frac{\hat{e}'\hat{e}}{\hat{u}'\hat{u}} + \frac{T-1}{T},$$

where second equality follows from the fact that Q is idempotent. Trivially, we have that $\hat{J}_{\omega,b}(\hat{\theta}_0) = (T-1)/2$ and $\hat{J}_{\gamma\omega,b}(\hat{\theta}_0) = (T-1)/T$. Then, it is straightforward to construct $\hat{J}_{(\gamma,\omega)',b}(\hat{\theta}_0)$.

After plugging-in the above terms in eq. (2), the marginal and joint test statistics become

$$LM_\gamma = d_\gamma(\hat\theta_0)' J^{-1}_{\gamma,b}(\hat\theta_0) d_\gamma(\hat\theta_0)/(NT) = (NT)\frac{B^2}{C},$$
$$LM_\omega = d_\omega(\hat\theta_0)' J^{-1}_{\omega,b}(\hat\theta_0) d_\omega(\hat\theta_0)/(NT) = (NT)\frac{A^2}{2(T-1)}, \text{ and}$$
$$LM_{\gamma\omega} = d_{(\gamma,\omega)'}(\hat\theta_0)' J^{-1}_{(\gamma,\omega)',b}(\hat\theta_0) d_{(\gamma,\omega)'}(\hat\theta_0)/(NT) = (NT)\left\{\frac{[B+(A/T)]^2}{C-2(T-1)/T^2} + \frac{A^2}{2(T-1)}\right\}.$$

Note that the scores $d_\gamma(\hat\theta_0)$, $d_\omega(\hat\theta_0)$, and $d_{(\gamma,\omega)'}(\hat\theta_0)$ can be obtained from Appendix A.1

Proceeding in a similar manner, we build the robust statistics. After plugging in the formulas of $\hat J_{\gamma,b}(\hat\theta_0)$, $\hat J_{\omega,b}(\hat\theta_0)$, $\hat J_{\gamma\omega,b}(\hat\theta_0)$, and $\hat J_{(\gamma,\omega)',b}(\hat\theta_0)$ in eq. (8), we obtain that

$$\begin{aligned} LM^*_{\gamma/\omega} &= (NT)\frac{[B+(A/T)]^2}{C-2(T-1)/T^2} \text{ and} \\ LM^*_{\omega/\gamma} &= (NT)\frac{[A/2+(T-1)B/(TC)]^2}{(T-1)/2-(T-1)^2/(T^2C)}. \end{aligned}$$

Alternatively, and as it was stated in Section 3, $LM^*_{\gamma/\omega}$ and $LM^*_{\omega/\gamma}$ can be obtained from Bera et al (2009)'s results. Observe that the above expressions are to equal $LM_{\gamma\omega} - LM_\omega$ and $LM_{\gamma\omega} - LM_\gamma$.

A.3.2 Construction of $LM_{\gamma/\omega}$

From eq. (2), the formula for $LM_{\gamma/\omega}$ becomes

$$LM_{\gamma/\omega} = \frac{1}{NT} d_\gamma(\hat\theta_\gamma)' J^{-1}_{\gamma,b_\gamma}(\hat\theta_\gamma) d_\gamma(\hat\theta_\gamma), \tag{A.5}$$

where $b_\gamma = (\beta', \sigma^2_\varepsilon, \omega)'$ and $\hat\theta_\gamma = (\hat\beta'_\gamma, \hat\sigma^2_{\varepsilon,\gamma}, 0, \hat\omega_\gamma)'$ denotes the maximum likelihood estimator of $(\beta', \sigma^2_\varepsilon, \gamma, \omega)'$ under the restriction $\gamma = 0$. Under the null $\gamma = 0$, $LM_{\gamma/\omega}$ converges in distribution to $\chi^2_1(0)$; see for example, Bera and Yoon (1993, pp. 651).

The score $d_\gamma(\hat\theta_\gamma)$ can be easily obtained from Appendix A.1. To construct $J_{\gamma,b_\gamma}(\hat\theta_\gamma)$, note first that

$$
\begin{aligned}
J_{\gamma\gamma}(\theta_\gamma) &= \frac{1}{T\sigma_\varepsilon^2}\left[E(\tilde{y}_{1,-1}\tilde{y}'_{1,-1}) - \frac{\omega}{(1+T\omega)}E(\tilde{y}_{1,-1}e_T e'_T\tilde{y}'_{1,-1})\right], \\
J_{b_\gamma\gamma}(\theta_\gamma) &= \begin{pmatrix} \frac{1}{T\sigma_\varepsilon^2}\left[E(\tilde{x}_1\tilde{y}'_{1,-1}) - \frac{\omega}{(1+T\omega)}E(\tilde{x}_1 e_T e'_T\tilde{y}'_{1,-1})\right] \\ 0_{[1\times 1]} \\ \frac{T-1}{T(1+T\omega)} \end{pmatrix}, \text{ and} \\
J_{b_\gamma b_\gamma}(\theta_\gamma) &= \begin{pmatrix} \frac{1}{T\sigma_\varepsilon^2}\left[E(\tilde{x}_1\tilde{x}'_1) - \frac{\omega}{(1+T\omega)}E(\tilde{x}_1 e_T e'_T\tilde{x}'_1)\right] & 0_{[k\times 1]} & 0_{[k\times 1]} \\ 0_{[1\times k]} & \frac{1}{2(\sigma_\varepsilon^2)^2} & \frac{1}{2\sigma_\varepsilon^2(1+T\omega)} \\ 0_{[1\times k]} & \frac{1}{2\sigma_\varepsilon^2(1+T\omega)} & \frac{T}{2(1+T\omega)^2} \end{pmatrix},
\end{aligned}
$$

where $\theta_\gamma = (\beta', \sigma_\varepsilon^2, 0, \omega)'$; see expression (A.1).

Since $J_{\gamma,b_\gamma}(\hat{\theta}_\gamma) = J_{\gamma\gamma}(\hat{\theta}_\gamma) - J_{\gamma b_\gamma}(\hat{\theta}_\gamma)J^{-1}_{b_\gamma b_\gamma}(\hat{\theta}_\gamma)J_{b_\gamma\gamma}(\hat{\theta}_\gamma)$, we obtain $J_{\gamma,b_\gamma}(\hat{\theta}_\gamma)$ by using elementary algebra and noting $J_{\gamma b_\gamma}(\theta_\gamma) = J'_{b_\gamma\gamma}(\theta_\gamma)$. Finally, after replacing the expectations by the sample analogues, we obtain the desired result:

$$
LM_{\gamma/\omega} = (NT)\frac{B_\gamma^2}{C_\gamma - 2(T-1)/T^2}.
$$

A.3.3 Construction of $LM_{\omega/\gamma}$

Again, from eq. (2), the formula for $LM_{\omega/\gamma}$ becomes

$$
LM_{\omega/\gamma} = \frac{1}{NT}d_\omega(\hat{\theta}_\omega)'J^{-1}_{\omega,b_\omega}(\hat{\theta}_\omega)d_\omega(\hat{\theta}_\omega), \tag{A.6}
$$

where $b_\omega = (\beta', \sigma_\varepsilon^2, \gamma)'$ and $\hat{\theta}_\omega = (\hat{\beta}'_\omega, \hat{\sigma}^2_{\varepsilon,\omega}, \hat{\gamma}_\omega, 0)'$ denotes the maximum likelihood estimator of $(\beta', \sigma_\varepsilon^2, \gamma, \omega)$ under the restriction $\omega = 0$. Under the null $\omega = 0$, $LM_{\omega/\gamma}$ converges in distribution to $\chi_1^2(0)$. To construct $J_{\omega,b_\omega}(\hat{\theta}_\omega)$, observe that

$$
\begin{aligned}
J_{\omega\omega}(\theta_\omega) &= \frac{T}{2}, \\
J_{b_\omega\omega}(\theta_\omega) &= \begin{pmatrix} 0_{[k\times 1]} \\ \frac{1}{2\sigma_\varepsilon^2} \\ \frac{1}{T}\left[(T-1) + \sum_{t=2}^{T}(T-t)\gamma^{t-1}\right] \end{pmatrix}, \text{ and} \\
J_{b_\omega b_\omega}(\theta_\omega) &= \begin{pmatrix} \frac{1}{T\sigma_\varepsilon^2}E(\tilde{x}_1\tilde{x}'_1) & 0_{[k\times 1]} & \frac{1}{T\sigma_\varepsilon^2}E(\tilde{x}_1\tilde{y}'_{1,-1}) \\ 0_{[1\times k]} & \frac{1}{2\sigma_\varepsilon^4} & 0_{[1\times 1]} \\ \frac{1}{T\sigma_\varepsilon^2}E(\tilde{y}_{1,-1}\tilde{x}'_1) & 0_{[1\times 1]} & \frac{1}{T\sigma_\varepsilon^2}E(\tilde{y}_{1,-1}\tilde{y}'_{1,-1}) \end{pmatrix}.
\end{aligned}
$$

Finally, following similar arguments to that of the previous subsection, we obtain the formula of eq. (7).

Tables

Table 1: Empirical size

N	T	LM_γ	$LM_{\gamma/\omega}$	$LM^*_{\gamma/\omega}$	LM_ω	$LM_{\omega/\gamma}$	$LM^*_{\omega/\gamma}$	$LM_{\gamma\omega}$
50	2	0.013	0.052	0.003	0.047	0.009	0.027	0.018
50	5	0.031	0.027	0.024	0.048	0.021	0.042	0.032
50	10	0.044	0.022	0.041	0.042	0.030	0.045	0.040
100	2	0.014	0.048	0.004	0.053	0.010	0.032	0.020
100	5	0.032	0.025	0.026	0.052	0.022	0.045	0.034
100	10	0.039	0.013	0.035	0.045	0.034	0.049	0.040

Notes: Monte Carlo simulations based on 5000 replications. Theoretical size 5%. $(\gamma, \omega) = (0, 0)$. Panel data models with (N, T) = (50, 5).

Table 2: Empirical size: Consistency

N	T	LM_γ	$LM_{\gamma/\omega}$	$LM^*_{\gamma/\omega}$	LM_ω	$LM_{\omega/\gamma}$	$LM^*_{\omega/\gamma}$	$LM_{\gamma\omega}$
					$\gamma = 0, \omega = 0$			
10	5	0.027	0.072	0.021	0.032	0.013	0.035	0.025
20	5	0.030	0.036	0.028	0.039	0.017	0.036	0.026
50	5	0.027	0.023	0.024	0.039	0.023	0.035	0.030
100	5	0.028	0.019	0.022	0.047	0.022	0.037	0.028
200	5	0.040	0.023	0.029	0.047	0.022	0.044	0.031
500	5	0.032	0.011	0.024	0.039	0.021	0.031	0.031
1000	5	0.032	0.010	0.024	0.055	0.023	0.046	0.035
				$\gamma = 0.1, \omega = 0$				
10	5	0.048	0.029	0.025	0.059	0.014	0.053	0.043
20	5	0.118	0.013	0.077	0.114	0.021	0.061	0.105
50	5	0.256	0.057	0.161	0.183	0.014	0.073	0.226
100	5	0.547	0.158	0.358	0.262	0.016	0.071	0.436
200	5	0.818	0.390	0.604	0.440	0.022	0.101	0.731
500	5	0.998	0.904	0.971	0.822	0.018	0.135	0.995
1000	5	1.000	0.958	0.990	0.897	0.020	0.141	1.000
				$\gamma = 0, \omega = 0.1$				
10	5	0.038	0.033	0.017	0.116	0.029	0.091	0.082
20	5	0.065	0.009	0.010	0.231	0.066	0.188	0.154
50	5	0.152	0.003	0.019	0.455	0.165	0.349	0.333
100	5	0.302	0.002	0.018	0.726	0.393	0.615	0.627
200	5	0.546	0.002	0.017	0.953	0.725	0.892	0.917
500	5	0.913	0.006	0.018	1.000	0.986	0.999	0.998
1000	5	0.996	0.006	0.020	1.000	1.000	1.000	1.000

Notes: Monte Carlo simulations based on 5000 replications. Theoretical size 5%. Panel data models with (N, T) = (50, 5).

Table 3: Empirical size: Power and robustness

γ	ω	LM_{γ}	$LM_{\gamma/\omega}$	$LM^{*}_{\gamma/\omega}$	LM_{ω}	$LM_{\omega/\gamma}$	$LM^{*}_{\omega/\gamma}$	$LM_{\gamma\omega}$
0.000	0.000	0.028	0.023	0.022	0.041	0.020	0.036	0.025
0.000	0.050	0.063	0.009	0.026	0.182	0.062	0.143	0.124
0.000	0.100	0.151	0.005	0.019	0.456	0.179	0.370	0.358
0.000	0.150	0.298	0.005	0.019	0.718	0.342	0.625	0.624
0.000	0.200	0.458	0.004	0.018	0.868	0.509	0.793	0.807
0.000	0.250	0.601	0.003	0.014	0.950	0.647	0.906	0.915
0.000	0.300	0.733	0.003	0.013	0.975	0.746	0.952	0.959
0.000	0.350	0.832	0.002	0.012	0.992	0.828	0.980	0.984
0.000	0.400	0.895	0.003	0.012	0.997	0.875	0.993	0.995
0.050	0.000	0.078	0.014	0.053	0.070	0.020	0.053	0.070
0.100	0.000	0.258	0.050	0.163	0.161	0.019	0.071	0.212
0.150	0.000	0.565	0.192	0.363	0.293	0.018	0.099	0.476
0.200	0.000	0.849	0.420	0.620	0.484	0.014	0.144	0.769
0.250	0.000	0.963	0.666	0.805	0.657	0.014	0.193	0.929
0.300	0.000	0.996	0.871	0.936	0.816	0.013	0.278	0.988
0.350	0.000	1.000	0.964	0.985	0.920	0.012	0.384	0.999
0.400	0.000	1.000	0.995	0.998	0.968	0.014	0.512	1.000

Notes: Monte Carlo simulations based on 5000 replications. Theoretical size 5%. Panel data models with (N, T) = (50, 5).

Table 4: Empirical size: Different DGP

γ	ω	LM_{γ}	$LM_{\gamma/\omega}$	$LM^*_{\gamma/\omega}$	LM_{ω}	$LM_{\omega/\gamma}$	$LM^*_{\omega/\gamma}$	$LM_{\gamma\omega}$
				$DGP: t_4$ *Student*				
0.000	0.000	0.035	0.033	0.027	0.044	0.020	0.039	0.033
0.000	0.050	0.066	0.007	0.023	0.197	0.076	0.159	0.137
0.000	0.100	0.159	0.003	0.014	0.454	0.190	0.371	0.357
0.000	0.150	0.285	0.005	0.019	0.667	0.349	0.588	0.579
0.000	0.200	0.444	0.004	0.013	0.811	0.490	0.741	0.743
0.000	0.250	0.579	0.003	0.012	0.896	0.618	0.844	0.851
0.000	0.300	0.682	0.006	0.012	0.942	0.696	0.904	0.917
0.000	0.350	0.768	0.004	0.010	0.964	0.767	0.942	0.944
0.000	0.400	0.823	0.004	0.011	0.982	0.819	0.964	0.973
0.050	0.000	0.073	0.015	0.055	0.073	0.022	0.053	0.069
0.100	0.000	0.261	0.056	0.161	0.158	0.020	0.069	0.209
0.150	0.000	0.571	0.196	0.369	0.294	0.018	0.093	0.475
0.200	0.000	0.837	0.417	0.599	0.487	0.017	0.139	0.761
0.250	0.000	0.969	0.683	0.819	0.667	0.013	0.201	0.938
0.300	0.000	0.994	0.876	0.932	0.809	0.012	0.278	0.988
0.350	0.000	0.999	0.960	0.981	0.914	0.011	0.387	0.997
0.400	0.000	1.000	0.993	0.996	0.964	0.013	0.505	0.999
				$DGP: \chi^2_1$				
0.000	0.000	0.030	0.018	0.021	0.041	0.019	0.035	0.030
0.000	0.050	0.076	0.008	0.021	0.195	0.083	0.156	0.147
0.000	0.100	0.173	0.006	0.019	0.461	0.213	0.384	0.372
0.000	0.150	0.324	0.006	0.020	0.642	0.353	0.571	0.569
0.000	0.200	0.436	0.006	0.018	0.770	0.462	0.707	0.703
0.000	0.250	0.560	0.004	0.012	0.853	0.574	0.795	0.799
0.000	0.300	0.650	0.006	0.011	0.903	0.664	0.860	0.866
0.000	0.350	0.734	0.004	0.011	0.939	0.737	0.908	0.909
0.000	0.400	0.791	0.005	0.013	0.962	0.774	0.933	0.937
0.050	0.000	0.079	0.017	0.056	0.074	0.018	0.045	0.073
0.100	0.000	0.255	0.061	0.158	0.157	0.022	0.071	0.208
0.150	0.000	0.565	0.182	0.361	0.296	0.021	0.102	0.463
0.200	0.000	0.848	0.402	0.602	0.472	0.017	0.139	0.755
0.250	0.000	0.973	0.674	0.798	0.667	0.015	0.201	0.938
0.300	0.000	0.998	0.866	0.927	0.830	0.017	0.289	0.994
0.350	0.000	1.000	0.959	0.977	0.911	0.012	0.371	0.999
0.400	0.000	1.000	0.992	0.995	0.969	0.011	0.482	1.000

Notes: Monte Carlo simulations based on 5000 replications. Theoretical size 5%. Panel data models with (N, T) = (50, 5).

Table 5: Empirical application: real GDP per capita growth

	Dynamic persistence $H_0^{\gamma}: \gamma = 0$			Time-invariant persistence $H_0^{\omega}: \omega = 0$			Both $H_0^{\gamma\omega}: \gamma = \omega = 0$
	marginal LM_{γ}	conditional $LM_{\gamma/\omega}$	robust $LM^*_{\gamma/\omega}$	marginal LM_{ω}	conditional $LM_{\omega/\gamma}$	robust $LM^*_{\omega/\gamma}$	joint $LM_{\gamma\omega}$
	with time trend						
statistic	42.75	8.52	14.17	51.87	2.51	23.28	66.03
decision	reject	reject	reject	reject	accept	reject	reject
	without time trend						
statistic	46.37	10.12	16.79	50.90	2.09	21.32	67.69
decision	reject	reject	reject	reject	accept	reject	reject

Notes: See text for details. Tests based on a 5% significance level.