Computational Statistics and Data Analysis ▮ (▮▮▮▮) ▮▮▮–▮▮▮

Contents lists available at ScienceDirect

Computational Statistics and Data Analysis

journal homepage: www.elsevier.com/locate/csda

Testing for serial independence of panel errors

Zaichao Du *

Research Institute of Economics and Management, Southwestern University of Finance and Economics, 55 GuangHuaCun St., Chengdu, 610074, China

ARTICLE INFO

Article history:
Received 9 November 2012
Received in revised form 21 July 2013
Accepted 21 July 2013
Available online xxxx

Keywords:
Empirical characteristic function
Panel data
Parameter estimation uncertainty
Permutation test
Serial dependence
Unobservable errors

ABSTRACT

A test for the serial independence of errors in panel data models is proposed. The test is based on the difference between the joint empirical characteristic function of residuals at different lags and the product of their marginal empirical characteristic functions. The test is nuisance-parameter-free and powerful against any type of pairwise dependence at all lags. A simple random permutation procedure is used to approximate the limit distribution of the test. A Monte Carlo experiment illustrates the finite sample performance of the test, and supports that the test statistic based on the estimated residuals has the same asymptotic distribution as the corresponding statistic based on the unobservable true errors.

1. Introduction

The serial independence of unobservable errors is a key assumption for the validity of many statistical inferences and asymptotic results. It is also a useful identification assumption in many econometric models for consumer choice, treatment effects and binary choice, to give a few examples. Brown and Wegkamp (2002) propose a minimum distance from independence estimator, where the independence between the unobserved errors and the exogenous variables is the crucial condition for identification. When one applies Efron's (1979) bootstrap to residuals, one often assumes the independence of errors for the bootstrap to be valid, see e.g. Singh (1981). Peretti (2005) shows that testing the significance of the departures from utility maximization boils down to testing the independent and identically distributed (iid) assumption of some unobserved errors. For some other cases where the independence of errors is crucial, see e.g. Diebold et al. (1998), Clement and Smith (2000) and Brown et al. (2007).

Despite its importance, there are only a few studies on testing independence of errors. Box and Pierce (1970) and Ljung and Box (1978), BLP hereafter, propose a test for the autocorrelation of errors in ARMA models. Brock et al. (1991, 1996) construct a test, BDS henceforth, for the independence of errors using chaos theory. Their test is nuisance-parameter-free only under conditional mean models, and it has no power against certain types of dependence. Hong and Lee (2003) overcome the aforementioned problems of BDS test, but they use kernel smoothing techniques and their test is affected by the choice of bandwidth. Du and Escanciano (in press) propose a distribution-free test based on the Hoeffding–Blum–Kiefer–Rosenblatt-type empirical process (HBKR hereafter, see Hoeffding, 1948; Blum et al., 1961; Delgado, 1999) applied to residuals at different lags.

There are no existing tests for serial independence of panel errors, to the best of our knowledge. Instead, there are only some tests for serial correlation of panel errors. Many of them are just for first-order serial correlation, see Breusch and

* Tel.: +86 28 87353797; fax: +86 28 87356958.
E-mail addresses: duzc@swufe.edu.cn, smilingchao@hotmail.com.

0167-9473/$ – see front matter © 2013 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.csda.2013.07.031

Please cite this article in press as: Du, Z., Testing for serial independence of panel errors. Computational Statistics and Data Analysis (2013), http://dx.doi.org/10.1016/j.csda.2013.07.031

Pagan (1980) and Wooldridge (2002, p. 275) for example. Arellano and Bond (1991) propose a test for lack of second-order serial correlation in the first-differenced errors. There are also some portmanteau tests for panel errors, such as Hong and Kao's (2004) wavelet-based test, Inoue and Solon's (2006) Lagrange multiplier test and Okui (2009), which is an extension of BLP test to panels. All these tests assume the cross-sectional independence of the errors, which is problematic sometimes, see e.g. Ng (2006), Baltagi et al. (2009, 2012) and Shin et al. (2009).

In this paper we propose a test for serial independence of panel errors. Our test is nuisance-parameter-free and powerful against any type of pairwise dependence at all lags. It does not involve bandwidth selection. Moreover, unlike the previous tests, it is well defined even under the presence of cross-sectional dependence in the errors.

We consider a balanced linear panel data model with individual specific effects

$$y_{it} = X_{it}'\theta_0 + \alpha_i + u_{it}, \quad i = 1, \ldots, N;\ t = 1, \ldots, T, \tag{1}$$

where y_{it} is the dependent variable of interest, X_{it} is a vector of covariates, α_i is the unobserved individual effect that could be correlated with X_{it}, u_{it} is the error term, and θ_0 is some unknown parameter in a compact set $\Theta \subset \mathbb{R}^p$.

The null hypothesis that we test in this paper is

$$H_0 : \{u_t(\theta_0)\}_{t\in\mathbb{Z}} \text{ is iid for some} \theta_0 \in \Theta \subset \mathbb{R}^p, \tag{2}$$

where henceforth

$$u_t(\theta) = (u_{1t}(\theta), u_{2t}(\theta), \ldots, u_{Nt}(\theta))' \in \mathbb{R}^N, \quad \theta \in \Theta,$$

with

$$u_{it}(\theta) = y_{it} - \overline{y}_i - (X_{it} - \overline{X}_i)'\theta,$$

$$\overline{y}_i = \frac{1}{T}\sum_{t=1}^{T} y_{it}; \qquad \overline{X}_i = \frac{1}{T}\sum_{t=1}^{T} X_{it}.$$

The alternative hypothesis is the negation of the null (2).

Our test statistic is based on the difference between the joint characteristic function of $(u_t(\theta_0), u_{t-j}(\theta_0))$ and the product of their marginal characteristic functions. As θ_0 is unknown, we replace it with a suitable estimate $\widehat{\theta}$ and construct residuals $\widehat{u}_t = u_t(\widehat{\theta})$. In contrast to the general theory on empirical processes with estimated parameters, see e.g. Durbin (1973), our test enjoys the "nuisance-parameter-free" property. We actually show that our test statistic based on the residuals has the same asymptotic distribution as the corresponding statistic based on the true errors. By choosing a proper weighting function in our test statistic, we avoid the multivariate numerical integration in $2N$ dimensions. The limit distribution of our test statistic is approximated directly by a simple random permutation procedure, see e.g. Delgado (1996).

Our test here extends (Du and Escanciano, in press) to panel data models. Their test based on the HBKR process suffers from the curse of dimensionality as the indicators $\{I(u_t \le x)\}_{t=1}^T$ are essentially all 0 when N gets large in panels, where hereinafter the indicator function $I(\cdot) = 1$, if the statement in the parentheses is true, and $I(\cdot) = 0$ otherwise. Our use of the characteristic function alleviates this problem.

To assess the finite sample performance of our proposed test, we do some Monte Carlo studies. We compare our test with Wooldridge (2002, p. 275) and Okui (2009). Generally, our test has good size and power performance. We also find that our test based on residuals has sizes and powers very close to its counterpart based on the true errors.

We illustrate our method under large T small N setup. We then discuss how one can modify our test to large N large T case if one is ready to assume iid across individuals.

The rest of the paper is organized as follows: in Section 2 we introduce our test statistic and derive its limit distribution. In Section 3 we do some Monte Carlo simulations to study the finite sample performance of the proposed test. In Section 4 we extend our test to large N large T case. In Section 5 we conclude and suggest some directions for future research. The proofs are gathered in the Appendix.

2. Test statistic

In the sequel, we simplify the notations as follows: $Y_t = (y_{1t}, y_{2t}, \ldots, y_{Nt})'$, $X_t = (X_{1t}, X_{2t}, \ldots, X_{Nt})'$, $u_t = u_t(\theta_0) = (u_{1t}, u_{2t}, \ldots, u_{Nt})'$, and $\widehat{u}_t = u_t(\widehat{\theta}_T)$, where $\widehat{\theta}_T$ is a $\sqrt{T}$-consistent estimator for θ_0. Our asymptotic results in this section are obtained as T goes to infinity. In Section 4 we discuss the case where both N and T go to infinity. Furthermore, let $\varphi_j(x, y) = E[\exp(ix'u_t + iy'u_{t-j})]$ and $\varphi(x) = E[\exp(ix'u_t)]$ denote the joint and marginal characteristic functions of (u_t, u_{t-j}), respectively. Let $\|\cdot\|$ be the Euclidean norm; and let ^{c} denote the complex conjugate. Finally, let $\widehat{\varphi}_{j,T}(x) = 1/(T-j)\sum_{t=1+j}^{T}\exp(ix'\widehat{u}_t)$, $\widehat{\varphi}_{1,T-j}(y) = 1/(T-j)\sum_{t=1+j}^{T}\exp(iy'\widehat{u}_{t-j})$, $\varphi_{j,T}(x) = 1/(T-j)\sum_{t=1+j}^{T}\exp(ix'u_t)$ and $\varphi_{1,T-j}(y) = 1/(T-j)\sum_{t=1+j}^{T}\exp(iy'u_{t-j})$.

Please cite this article in press as: Du, Z., Testing for serial independence of panel errors. Computational Statistics and Data Analysis (2013), http://dx.doi.org/10.1016/j.csda.2013.07.031