Testing Serial Correlation in Fixed Effects Regression Models: the Ljung-Box Test for Panel Data

Okui, R.

Department of Economics, Hong Kong University of Science and Technology
Clear Water Bay, Kowloon, Hong Kong. Email: okui@ust.hk E-Mail: okui@ust.hk

Keywords: *panel data, testing serial-correlation, double asymptotics*

EXTENDED ABSTRACT

Testing the presence of serial correlation in the error terms of a fixed effects regression model is important for many reasons. While there have been a number of testing procedures developed so far (see, e.g., Bhargava, Franzini and Narendranathan (1982), Baltagi and Li (1995), Baltagi and Wu (1999), Bera et al. (2001), Wooldridge (2002), Drukker (2003), Hong and Kao (2004) and Inoue and Solon (2006)), testing for serial correlation has not been a standard practice in applied research that uses panel data, as recognized by Kédzi (2004) and Bertrand et al. (2004). As conjectured by Inoue and Solon (2006), the reason for this might be that there had been no simple testing procedure until very recently. Moreover, many simple testing procedures, such as those suggested by Wooldridge (2002), look only at the first order autocorrelation and are not portmanteau tests. Although portmanteau tests do exist for panel data, such as those proposed by Hong and Kao (2004) and Inoue and Solon (2006), there has been no test for serial correlation in micro economic panel data that is both portmanteau test and is as straightforward as the Ljung–Box or Box–Pierce test. (Fu et al. (2002) extend the Box–Pierce test to panel data settings. However, they consider only time effects and do not consider the presence of individual effects. In most applied economics research using panel data, we need to consider individual effects and their test is not readily applicable.) This paper seeks to provide a simple and straightforward portmanteau test to fill this gap.

The goal of this paper is to develop a test for serial correlation in fixed effects models. Our test is a natural extension of the well-known test by Ljung and Box (1978) to panel data settings. The Ljung–Box test is a modification of Box and Pierce's (1970) test and the basic idea is that we use a weighted sum of the squares of the estimated autocorrelations as the test statistics. This approach yields a test that is intuitive, easy to interpret and simple to compute, because the asymptotic variance matrix of the vector of the estimated auto-correlations is an identity matrix under the null hypothesis of no serial correlation even in panel data settings. The main issue that arises when we extend the idea of the Ljung–Box test to panel data settings is that sample auto-correlations computed with panel data might be severely biased when the length of the time series is not very large compared with the cross-sectional sample size (see, e.g., Solon (1984) and Okui (2007)). This bias distorts the size and is the main issue in applying the idea of the Ljung–Box test to fixed effects regression models. We modify the Ljung–Box test to take into account the bias of autocorrelation estimators, and our modification is based on Okui's results (2007) that proposes asymptotically unbiased autocorrelation estimators for long panel data. Okui (2007) observes that the leading term of the bias of within-group auto-covariances are proportional to the long-run variance. Given this observation, Okui (2007) proposes to eliminate the bias by estimating the long-run variance with the kernel estimator of Parzen (1957) and Andrews (1991). We then construct autocorrelation estimators using the bias-corrected autocovariance estimators. The modified Ljung–Box test for panel data analysis is based on these bias-corrected autocorrelation estimators.

We run Monte Carlo simulations to evaluate the performance of our new testing procedure and to compare our test with other existing tests. We find that our test yields a reasonable size even if the length of the time series is not very long. Our tests are powerful against a wide range of alternatives because our test is a portmanteau test as is the original Ljung–Box test. Note that many existing tests consider only first–order autocorrelations and they are not portmanteau tests. We also find that our test is more powerful than the test of Inoue and Solon (2006), which is also a portmanteau test.

1 THE NEW TESTING PROCEDURE

Suppose that we have panel data of (y_{it}, x_{it}) for $i = 1, \ldots, N$ and $t = 1, \ldots, T$. Let y_{it} be the scalar dependent variable and x_{it} be the vector of regressors. The relationship between y_{it} and x_{it} is modeled as the following fixed-effects model:

$$y_{it} = \beta' x_{it} + \eta_i + w_{it},$$

for $i = 1, \ldots, N$ and $t = 1, \ldots, T$, where β is the vector of parameters, η_i is the individual fixed effect for individual i and w_{it} is the time-varying unobservable error term. We assume that x_{it} is strictly exogenous: i.e., x_{it} is uncorrelated with w_{is} for any s.

We are interested in the dynamic structure of w_{it}. In particular, this paper focuses on testing the presence of serial correlation in w_{it}. Let ρ_k denote the k-th order autocorrelation of w_{it}: $\rho_k = Cov(w_{it}, w_{i,t-k})/var(w_{it})$. Our null hypothesis is:

$$H_0 : \rho_k = 0, \quad \text{for } k = 1, \ldots.$$

We consider Box and Pierce's (1970) approach for testing serial correlation. This approach is based on the sum of the squares of sample autocorrelations. We then apply Ljung and Box's (1978) modification in order to improve the finite sample properties of the test. To describe the test statistics, we first discuss how to estimate the autocovariances of w_{it}. Let $\hat{\beta}$ be the fixed effects estimator:

$$\begin{aligned} \hat{\beta} &= \left\{ \sum_{i=1}^{N} \sum_{t=1}^{T} (x_{it} - \bar{x}_i)(x_{it} - \bar{x}_i)' \right\}^{-1} \\ &\quad \times \sum_{i=1}^{N} \sum_{t=1}^{T} (x_{it} - \bar{x}_i)(y_{it} - \bar{y}_i), \end{aligned}$$

where $\bar{x}_i = \sum_{t=1}^{T} x_{it}/T$ and $\bar{y}_i = \sum_{t=1}^{T} y_{it}/T$.[1] Let u_{it} be the (un-centered) residuals from the regression:

$$u_{it} = y_{it} - \hat{\beta}' x_{it}.$$

Let

$$\hat{\gamma}_k = \frac{1}{N(T-k)} \sum_{i=1}^{N} \sum_{t=k+1}^{T} (u_{it} - \bar{u}_i)(u_{i,t-k} - \bar{u}_i),$$

which may be a natural (but naive) estimator of the k-th order autocovariance of w_{it}, where $\bar{u}_i = \sum_{t=1}^{T} u_{it}/T$. However, the estimator, $\hat{\gamma}_k$, may be severely biased when T is not very large relative to N. The main source of the bias is the estimation error of η_i. Note that when T is fixed, we cannot consistently estimate η_i. Even when T tends to infinity, the estimation of η_i is problematic if N is large compared with T. To see this, we observe that $\hat{\gamma}_k$ may be decomposed in the following form (see Okui (2007)):

$$\begin{aligned} \hat{\gamma}_k &= \frac{1}{N(T-k)} \sum_{i=1}^{N} \sum_{t=k+1}^{T} w_{it} w_{i,t-k} \\ &\quad - \frac{1}{N} \sum_{i=1}^{N} (\bar{w}_i)^2 + small. \end{aligned}$$

The term $\bar{w}_i (= (\bar{y}_i - \beta' \bar{x}_i) - \eta_i)$ can be understood as the estimation error for η_i. This estimation error is the main source of the bias even when T tends to infinity, because $\sum_{i=1}^{N} (\bar{w}_i)^2/N$ is of order $O_p(1/T)$.

An important observation is that $(T/N) \sum_{i=1}^{N} (\bar{w}_i)^2$ converges to the long-run variance of w_{it}. This observation motivates us to consider a bias correction method based on a long-run variance estimator. Okui (2007) proposes to estimate the long-run variance by the kernel estimator of Parzen (1957) and Andrews (1991) and, then, to use the proposed long-run variance estimator to correct the bias. The kernel estimator for the long-run variance is:

$$\tilde{V}_T = \sum_{j=-T+1}^{T-1} k\left(\frac{j}{S}\right) \frac{T - |j|}{T} \hat{\gamma}_j,$$

where $k(\cdot)$ is a kernel function and the scalar, S, is the bandwidth to be chosen by the researcher. In the simulations, we use the quadratic spectrum (QS) kernel:

$$k(x) = \frac{3}{(6\pi x/5)^2} \left\{ \frac{\sin(6\pi x/5)}{6\pi x/5} - \cos(6\pi x/5) \right\},$$

for $|x| \leq 1$ and $k(x) = 0$ otherwise. This choice of the kernel function follows the recommendation of Andrews (1991). A bias-corrected estimator of γ_k may be constructed by:

$$\tilde{\gamma}_k = \hat{\gamma}_k + \frac{1}{T} \tilde{V}_T.$$

Okui (2007) shows that $\tilde{\gamma}_k$ is asymptotically unbiased: its asymptotic distribution is centered around zero. We can iterate this bias correction: we update the estimate of V_T by using the bias-corrected estimators for γ_k in $k = 0, \ldots, T-1$; then, we re-estimate γ_k based on the updated estimate of V_T. As $\tilde{\gamma}_k$s are better estimates of γ_k, the bias may be better estimated using $\tilde{\gamma}_k$. This iteration may be repeated many times and it converges under very mild conditions. Let $\hat{\gamma}_{0,T-1} = (\hat{\gamma}_0, \ldots, \hat{\gamma}_{T-1})'$, I_T be the $T \times T$ identity matrix and ι_T be the $T \times 1$ vector of ones. Let

$$K_T = \left(k(0), \frac{2T-2}{T} k\left(\frac{1}{S}\right), .., \frac{2}{T} k\left(\frac{T-1}{S}\right) \right)'.$$

The condition for the convergence of this iteration is $\iota_T' K_T < T$, which is satisfied with the QS kernel. Then, the vector of the auto-covariance estimators

[1] Any estimator that satisfies the condition in Assumption 2 can be used.

obtained after the convergence, denoted $\tilde{\gamma}_{0,T-1}(\infty)$, is written as:

$$\begin{aligned}&\tilde{\gamma}_{0,T-1}(\infty)\\ =&\left(I_T-\frac{1}{T}\iota_T K_T'\right)^{-1}\hat{\gamma}_{0,T-1}\\ =&\left(I_T+\frac{1}{T-\iota_T' K_T}\iota_T K_T'\right)\hat{\gamma}_{0,T-1}.\end{aligned}$$

We also follow Okui (2007) in the choice of the bandwidth. For the QS kernel, we use the following bandwidth:

$$\hat{S}^* = 1.3221(\hat{\tau}^2 TN)^{1/5},$$

where $\hat{\tau} = 2\hat{\delta}/(1-\hat{\delta})^2$ and $\hat{\delta}$ is Hahn and Kuersteiner's (2002) estimator for the panel AR(1) model of u_{it}.[2] The bandwidth is chosen so that it achieves the minimum of the asymptotic mean square error of the long-run variance estimator when the true dynamics of w_{it} follows the AR(1) process. Our test statistic is based on $\hat{\gamma}_{0,T-1}(\infty)$.

We now construct the test statistics for the Ljung–Box test test modified for fixed-effects regression models. We estimate the autocorrelations based on $\hat{\gamma}_{0,T-1}(\infty)$:

$$\tilde{\rho}_k = \frac{\tilde{\gamma}_k(\infty)}{\tilde{\gamma}_0(\infty)}.$$

Note that $\tilde{\rho}_k$ is asymptotically unbiased as are the results of the asymptotically unbiasedness of $\tilde{\gamma}_k$s. Let p be an integer that the researcher specifies in advance. The test statistic takes the following form:

$$Q_{LB} = NT\sum_{k=1}^{p}\frac{T+2}{T-k}\left(\frac{T-k}{T}\tilde{\rho}_k\right)^2.$$

It is easy to see that the test-statistic Q_{LB} is a straightforward extension of the Ljung–Box test statistic. The ratio $(T-k)/T$ makes the denominator of the k-th autocorrelation estimator NT rather than $N(T-k)$. The ratio $(T+2)/(T-k)$ is important to have a good size property (see Ljung and Box (1978)). We see the effect of including these ratios in the test statistic in the Monte Carlo simulations reported in Section 4.3. Our testing procedure is:

$$\text{Reject } H_0 \text{ if } Q_{LB} > \chi^2_{p,1-\alpha},$$

where α is the nominal size of the test chosen by the researcher and $\chi^2_{p,1-\alpha}$ denotes the $1-\alpha$ quantile of the χ^2 distribution with p degrees of freedom.

[2] Hahn and Kuersteiner's (2002) estimator is

$$\begin{aligned}\hat{\delta} =& \frac{T}{T-1}\frac{\sum_{i=1}^{N}\sum_{t=2}^{T}(u_{i,t-1}-\bar{u}_{i-})(u_{i,t-1}-\bar{u}_{i+})}{\sum_{i=1}^{N}\sum_{t=2}^{T}(u_{i,t-1}-\bar{u}_{i-})^2}\\ &+\frac{1}{T-1},\end{aligned}$$

where $\bar{u}_{i-}=\sum_{t=1}^{T-1}u_{it}/(T-1)$ and $\bar{u}_{i+}=\sum_{t=2}^{T}u_{it}/(T-1)$.

2 THE ASYMPTOTIC DISTRIBUTION OF THE TEST STATISTICS UNDER THE NULL HYPOTHESIS

This section derives the asymptotic distributions of our test statistic, Q_{LB}, under the null hypothesis. We make two sets of assumptions. The first set of assumptions mainly concerns the probabilistic nature of w_{it}. This set of assumptions is important to deriving the asymptotic distribution of Q_{LB} even if β is known or there is no regressor. The second set of assumptions is mainly about the properties of the regressor, x_{it}, and is used to guarantee that the effect of estimating β may be ignored.

The first set of assumptions is:

Assumption 1. *1. $\{x_{it}, w_{it}\}_{t=1}^{T}$ are i.i.d. across individuals.*

2. w_{it} is strictly stationary within individuals and $\sum_{j=-\infty}^{\infty}|\gamma_j|<\infty$.

3. There exists $M<\infty$ such that $E(|w_{it}w_{ik}w_{im}w_{il}|)<M$ for any t, k, m and l.

4. $\sum_{j_1=-\infty}^{\infty}\cdots\sum_{j_7=-\infty}^{\infty}|cum(0,j_1,\ldots,j_7)|<\infty$, where $cum(0,j_1,\ldots,j_7)$ is the eighth-order cumulant of $(w_{i0},w_{i,j_1},\ldots,w_{i,j_7})$.

5. $\sum_{j_1=-\infty}^{\infty}\cdots\sum_{j_3=-\infty}^{\infty}|cum(0,j_1,j_2,j_3)|<\infty$, where $cum(0,j_1,j_2,j_3)$ is the fourth-order cumulant of $(w_{i0},w_{i,j_1},w_{i,j_2},w_{i,j_3})$.

Note that Assumption 1 does not impose any restriction on the probabilistic nature of η_i, as η_i is eliminated by the fixed-effects transformation. Assumptions 1.1–1.3 are standard. Assumption 1.4 guarantees the uniform integrability of $\sum_{t=k+1}^{T}(w_{it}w_{i,t-k}-\gamma_k)/\sqrt{T}$. It is one of the key conditions in applying the central limit theorem under double asymptotics by Phillips and Moon (1999) on which our asymptotic results are based. This assumption may be relaxed as long as the uniform integrable condition is met. Assumption 1.5 guarantees that the the asymptotic variance of $(\tilde{\rho}_1,\ldots,\tilde{\rho}_p)'$ exists and that it is the identity matrix under the null hypothesis.

Next, we consider the set of assumptions that allows us to ignore the estimation error of $\hat{\beta}$ when we derive the asymptotic distribution of the test statistics.

Assumption 2. *1. $E(w_{it_1}x_{it_2})=0$ for any t_1 and t_2.*

2. $\hat{\beta}-\beta=O_p((NT)^{-1/2})$.

3. *$E(||w_{it_1} w_{it_2} x_{it_3} x'_{it_4}||) < M$ for some $M < \infty$ for any combination of t_1, t_2, t_3 and t_4, where $||\cdot||$ is the Euclidean norm.*
4. *Let $x_{it,a}$ be the a-th element of x_{it}. Then $E(|x_{it_1,a} x_{it_2,b} x_{it_3,c} x_{it_4,d}|) < M$ for some $M < \infty$ for any combination of t_1, t_2, t_3, t_4, a, b, c and d.*

We assume that the regressor, x_{it}, is strictly exogenous in Assumption 2.1. Allowing regressors to be predetermined is not considered here and this would change the results. Assumption 2.2 says that $\hat{\beta}$ is $\sqrt{NT}$-consistent, which is satisfied by the fixed-effects estimator. Assumptions 2.3 and 2.4 impose the existence of fourth moments, which are important in asymptotic results. These assumptions are standard in the literature on fixed-effects estimation.

Let $\tilde{\rho}_{1,p} = (\tilde{\rho}_1, \ldots, \tilde{\rho}_p)'$.

Theorem 1. *Suppose that Assumptions 1 and 2 are satisfied. Under H_0, as $T \to \infty$ and $N \to \infty$ with $N/T^3 \to 0$, we have*

$$\sqrt{NT}\tilde{\rho}_{1,p} \to_d N(0, I_p).$$

Proof. This theorem is obtained as a corollary of the results in Okui (2007). Here, we present a brief sketch of the proof. First, Theorem 11 in Okui (2007) shows that the estimation error in the fixed-effects estimator $\hat{\beta}$ can be ignored. Second, Theorems 3 and 5 in Okui (2007) demonstrate that $\tilde{\gamma}_k(\infty)$s are asymptotically normal. The asymptotic variance of $\tilde{\gamma}_k(\infty)$ is $\sum_{j=-\infty}^{\infty}(\gamma_j^2 + \gamma_{k+j}\gamma_{k-j})$, which is γ_0^2 for $k \neq 0$ and $2\gamma_0^2$ for $k = 0$ under H_0. The asymptotic covariance between $\tilde{\gamma}_k(\infty)$ and $\tilde{\gamma}_j(\infty)$ for $k \neq j$ is $\sum_{l=-\infty}^{\infty}(\gamma_l\gamma_{l-k+j} + \gamma_{l+j}\gamma_{l-k})$, which is 0 for any $k \neq j$ under H_0 (see also Remark 3 in Okui (2007)). By applying the Delta method, we obtain the desired result. □

This result is analogous to the well-known result by Box and Pierce (1970). An important point is that the bias must be corrected to obtain this theorem. If we do not correct the bias, the bias remains even in the first-order asymptotics (see Remark 2 below). As a direct corollary of this theorem, we obtain our main result.

Corollary 1. *Suppose that Assumptions 1 and 2 are satisfied. Under H_0, as $T \to \infty$ and $N \to \infty$ with $N/T^3 \to 0$, we have*

$$Q_{LB} \to_d \chi_p^2.$$

This result justifies the use of our test.

Remark 1. It would be clear that our test is consistent against any alternative in which $\rho_k \neq 0$ for $k \leq q$. However, like the Box–Pierce and Ljung–Box tests in the time–series context, our test does not have power if serial correlation arises only at an order higher than p. Detecting higher-order autocorrelations requires a large value of p, but the size of the test might deviate from the nominal one when p is too large. We should pick a value of p by taking this trade-off into account. Unfortunately, there is no universally accepted method to choose p even for the original Box-Pierce test for single time-series (e.g., see Hayashi (2000, p144)). This problem is beyond the scope of the current paper.

Remark 2. Let $\hat{\rho}_k$ be the estimator of the k-th order autocorrelation based on the within-group autocovariances: $\hat{\rho}_k = \hat{\gamma}_k/\hat{\gamma}_0$. Let $\hat{\rho}_{1,p} = (\hat{\rho}_1, \ldots, \hat{\rho}_p)'$. As a special case of Theorem 2 in Okui (2007), the asymptotic distribution of $\hat{\rho}_k$ under the null hypothesis is

$$\sqrt{NT}\left(\hat{\rho}_{1,p} + \frac{1}{T}\iota_p\right) \to_d N(0, I_p),$$

where ι_p is the $p \times 1$ vector of ones. The asymptotic bias in $\hat{\rho}$ has a very simple form of ι_p/T under the null hypothesis. This result suggests another test for serial correlation that is based on correcting the bias by adding $1/T$ to the within-group auto-covariances. We denote such a test as "BP3" whose properties are examined below in the Monte Carlo simulations. This test is very easy to compute and might be attractive at first glance. However, the simulations show that this test has a severe size distortion problem. Thus, we do not recommend using this test.

3 MONTE CARLO EXPERIMENTS

This section reports the results of the Monte Carlo experiments. The simulations are conducted with the Ox 4.04 for Windows software (Doornik (2006)). The primary purpose of the simulations is to check the finite sample properties of our new testing procedure. Our testing procedure is based on asymptotic results. It is important to see if the asymptotic results provide a good approximation of the finite sample properties of the test. Another purpose is to compare our tests with existing testing procedures.

3.1 Design

Our data-generating process is:

$$y_{it} = w_{it} + \eta_i,$$

where $\eta_i \sim i.i.d.N(0,1)$. Note that the specification of η_i does not affect the results because η_i is eliminated in all the procedures examined in the experiments. For simplicity, we consider models with no regressor. We consider the following four specifications of w_{it}.

DGP 0 (null): $w_{it} \sim i.i.d.N(0,1)$.

DGP 1 (AR(1)): $w_{it} = 0.1w_{i,t-1} + \epsilon_{it}$, where $\epsilon_{it} \sim i.i.d.N(0,1)$. The initial observations are generated from the stationary distribution; i.e., $w_{i0} \sim i.i.d.N(0,(0.81)^{-1})$.

DGP 2 (MA(2)): $w_{it} = \epsilon_{it} + 0.1\epsilon_{i,t-2}$, where $\epsilon_{it} \sim i.i.d.N(0,1)$.

DGP 3 (incidental trends): $w_{it} = \epsilon_{it} + \alpha_i t$, where $\epsilon_{it} \sim i.i.d.N(0,1)$ and $\alpha_i \sim i.i.d.N(0,0.01)$.

The data-generating processes are similar to those considered by Inoue and Solon (2006), although the values of the parameters are different. Each experiment is characterized by the cross-sectional sample size, N, the length of the time series, T, and the data-generating process. We set $N = 20,100,200$; $T = 5,10,25$. The number of replications is 5000 for each specification.

We examine the properties of six tests in the experiments. The nominal size is set to 5% throughout the experiments. First, we consider our new testing procedure with $p = 4$ and $p = 9$. Our test is denoted "LB". We consider four testing procedures based on existing articles to check the relative performance of our tests. The test "ISS" is a modified version of the portmanteau test by Inoue and Solon (2006).[3] We modify the Inoue and Solon test in two ways: "ISS" is specialized for first- to p-order autocorrelations and it is based on the assumption that the process is covariance stationary.[4] We set $p = 4$ and $p = 9$ for ISS. We consider two tests suggested by Wooldridge (2002). Wooldridge's (2002, p. 275) test, denoted "W", is based on the first-order autoregression of $\hat{w}_{it}$. We conduct the t-test of the null hypothesis that the coefficient on $\hat{w}_{i,t-1}$ is $-1/(T-1)$. Note that we have to use an autocorrelation robust standard error to construct the test statistics. The test "WD" by Wooldridge (2002, pp 282-283) and Drukker (2003) is based on the first difference of the residuals. We test the hypothesis that the first-order autoregressive coefficient of the differenced residuals is $-1/2$.

[3]Results not shown here indicate that the original version of the Inoue and Solon test performs poorly when T is not small. This is unsurprising because the asymptotic distribution of their test statistics under an asymptotic sequence in which $N \to \infty$ and T is fixed is $\chi^2_{(T-1)(T-2)/2}$, which depends on T and diverges as $T \to \infty$.

[4]The test statistic is computed by the formula given in Inoue and Solon (2006, p839) with replacing $D_{k,T}$ in their formula by the $T^2 \times p$ matrix whose j-th column is equal to $vec(A_j)$ where A_j is the $T \times T$ matrix whose (l,m)-th element is 1 if $|l-m| = j$ and 0 otherwise. Note that the asymptotic null distribution of this modified test statistic is χ^2_p.

Table 1. Empirical rejection probability under the null hypothesis (DGP 0)

(N,T) p	LB (4)	LB (9)	ISS (4)	ISS (9)	W	WD
(20, 5)	0.054		0.011		0.090	0.096
(20, 10)	0.064	0.059	0.032	0.002	0.089	0.094
(20, 25)	0.054	0.057	0.035	0.003	0.079	0.082
(100, 5)	0.066		0.024		0.061	0.062
(100, 10)	0.066	0.066	0.053	0.030	0.061	0.055
(100, 25)	0.055	0.060	0.050	0.049	0.052	0.054
(200, 5)	0.062		0.025		0.051	0.049
(200, 10)	0.062	0.063	0.055	0.029	0.060	0.060
(200, 25)	0.052	0.060	0.053	0.050	0.054	0.049

3.2 Results

Table 1 reports the empirical sizes of the tests. All the tests have reasonable empirical sizes in all cases. We observe that the rejection probability of "LB" is close to the nominal size when T is small even though its theoretical justification is based on double asymptotics. However, we observe that the empirical size of "LB" becomes closer to the nominal size as T increases while N is not critical in determining the size property of "LB". The sizes of "ISS" are affected by both T and N. For "LB", the value of p does not affect the size much, but "ISS" tends to be very conservative when p is large ($p = 9$). The sizes of "W" and "WD" are better when N is large while they are not very sensitive to T.

Tables 2-4 report the powers of the tests. First of all, we recognize that both "LB" with $p = 4$ and with $p = 9$ show encouraging results. With DGP2 and DGP3, setting $p = 4$ yields better power than setting $p = 9$, but with DGP4, the opposite is observed. Compared with other tests, setting $p = 4$ tends to yield competent power in all cases while our test with $p = 9$ sometimes has a lower power than other existing tests.

The most powerful test with DGP1 (the AR (1) alternative) turns out to be "W", although we observe that other tests are also effective in detecting the AR(1) alternative. However, the result also indicates that "W" does not have strong power against the MA(2) alternative (DGP2). It is interesting to note that "W" has some power with DGP2 when T is small, although it may appear that "W" tests only the first-order autocorrelation. It turns out that "W" is consistent against some (but not all) alternatives with zero first–order autocorrelation under $N \to \infty$ and T fixed asymptotics. When T is fixed, the coefficient estimator on which "W" is based is not a consistent estimator for the first-order autocorrelation and converges to a function of all the autocorrelations (Solon (1984)). Therefore, "W" has some power even in DGP2 in which the first-order autocorrelation is zero. On the other hand, when T tends to infinity, "W" is consistent only against alternatives with non-zero first-order autocorrelation. This explains the fact

Table 2. Empirical rejection probability under AR(1) alternatives (DGP1)

(N,T) p	LB (4)	LB (9)	ISS (4)	ISS (9)	W	WD
(20, 5)	0.067		0.015		0.139	0.135
(20, 10)	0.166	0.120	0.062	0.003	0.249	0.180
(20, 25)	0.389	0.271	0.170	0.011	0.560	0.344
(100, 5)	0.224		0.098		0.242	0.221
(100, 10)	0.637	0.503	0.466	0.259	0.693	0.478
(100, 25)	0.985	0.950	0.966	0.871	0.996	0.900
(200, 5)	0.418		0.228		0.401	0.349
(200, 10)	0.913	0.832	0.829	0.629	0.929	0.744
(200, 25)	1.000	1.000	1.000	0.999	1.000	0.997

Table 3. Empirical rejection probability under MA(2) alternatives (DGP2)

(N,T) p	LB (4)	LB (9)	ISS (4)	ISS (9)	W	WD
(20, 5)	0.106		0.010		0.124	0.126
(20, 10)	0.149	0.127	0.049	0.001	0.091	0.185
(20, 25)	0.360	0.255	0.153	0.009	0.084	0.374
(100, 5)	0.249		0.077		0.171	0.230
(100, 10)	0.533	0.428	0.430	0.230	0.105	0.498
(100, 25)	0.970	0.925	0.958	0.863	0.080	0.909
(200, 5)	0.411		0.198		0.265	0.363
(200, 10)	0.854	0.770	0.810	0.594	0.164	0.779
(200, 25)	1.000	1.000	1.000	0.999	0.089	0.998

that "W" loses its power with DGP2 as T increases.

The "WD" test performs reasonably well with DGP1 and DGP2, while it fails to have power against the alternative with incidental trends (DGP3). With DGP3, the presence of serial correction is visible only through large values of the trend term, t, as the variance of α_i is very small. However, the trend is eliminated when we take the first difference. The "WD" test, which is based on the first difference, fails to detect this alternative.

Like our new test, "ISS" is also a portmanteau test and it has good power against all kinds of alternatives considered here. However, our tests are typically more powerful than "ISS". The power of "ISS" can be vastly low when we set $p = 9$.

Summing up, we observe that our new "LB" test performs very well throughout the experiments. Its size is good even when T is small and is not sensitive to the choice of p. Moreover, it is powerful against a wide range of alternatives because it is a portmanteau test.

3.3 Other modified Box–Pierce tests

This paper extends the Ljung–Box test to panel data settings. The Ljung–Box test has been considered to improve the finite sample properties of the Box–Pierce test. This subsection asks how the original version of the Box–Pierce test and other modifications of the Box–Pierce test perform with finite samples. In particular, we consider the following three versions of

Table 4. Empirical rejection probability under alternatives with incidental trends (DGP3)

(N,T) p	LB (4)	LB (9)	ISS (4)	ISS (9)	W	WD
(20, 5)	0.060		0.012		0.104	0.101
(20, 10)	0.214	0.262	0.053	0.003	0.155	0.091
(20, 25)	1.000	1.000	0.410	0.017	0.987	0.082
(100, 5)	0.080		0.030		0.063	0.065
(100, 10)	0.743	0.866	0.638	0.406	0.390	0.069
(100, 25)	1.000	1.000	1.000	1.000	1.000	0.082
(200, 5)	0.093		0.040		0.069	0.059
(200, 10)	0.964	0.993	0.963	0.915	0.657	0.077
(200, 25)	1.000	1.000	1.000	1.000	1.000	0.110

Table 5. Empirical rejection probability under the null hypothesis (DGP 0)

(N,T) p	LB (4)	BP1 (4)	BP2 (4)	BP3 (4)
(20, 5)	0.054	0.225	0.002	0.463
(20, 10)	0.064	0.088	0.012	0.143
(20, 25)	0.054	0.066	0.028	0.071
(100, 5)	0.066	0.232	0.003	0.690
(100, 10)	0.066	0.092	0.014	0.173
(100, 25)	0.055	0.068	0.027	0.075
(200, 5)	0.062	0.229	0.002	0.957
(200, 10)	0.062	0.094	0.008	0.207
(200, 25)	0.052	0.059	0.027	0.074

the Box–Pierce test modified for panel data analysis:

$$
\begin{aligned}
Q_1 &= NT\sum_{k=1}^{p}\tilde{\rho}_k^2, \\
Q_2 &= NT\sum_{k=1}^{p}\left(\frac{T-k}{T}\tilde{\rho}_k\right)^2, \\
Q_3 &= NT\sum_{k=1}^{p}(\hat{\rho}_k + 1/T)^2.
\end{aligned}
$$

Let "BP1" be the test based on the test statistic Q_1. The "BP2" and "BP3" tests are defined similarly. The "BP1" and "BP2" tests may be considered as natural extensions of the original Box–Pierce test. The difference between "BP1" and "BP2" is that Q_1 uses the autocovariance estimators in which the denominator is $T - k$, while that for Q_2 is T. The Q_3 test statistic is based on the observation given in Remark 1. Under the null hypothesis, the bias of each $\hat{\rho}_k$ is $-1/T$ and Q_3 corrects the bias by adding $1/T$ to each $\hat{\rho}_k$. Note that Q_1, Q_2 and Q_3 all possess the same asymptotic distribution (i.e., χ_p^2) under the null hypothesis.

Table 5 presents the empirical rejection probabilities of the tests under the null hypothesis. For reference, we also present the empirical size of "LB" in Table 5. We set $p = 4$ for all tests. The "BP1" and "BP3" tests suffer from substantial size distortion when T is small, although their sizes are not bad when $T = 25$. It seems that we cannot use "BP1" and "BP3" unless we correct the critical values or our panel data have a long time series. The "BP2" test appears to be conservative. This indicates that we need to consider correcting the critical value for "BP2" in order to obtain good power.

These results show the advantage of the Ljung–Box test, which provides the best size properties among four modifications of the Box–Pierce test.

4 CONCLUSION

This paper proposes a new portmanteau test for serial correlations in fixed effects regression models. Our test is a natural extension of the Ljung–Box test to panel data settings. The main point is that our test is based on asymptotically unbiased autocorrelation estimators. The new test behaves nicely in our Monte Carlo simulations. Our testing procedures should be helpful to applied researchers.

There are several points that should be investigated in the future. The first is how to choose p as discussed in a previous section. Another possible future study is to consider panel data models with predetermined regressors. Note that our current discussion considers only strictly exogenous regressors. When the model includes predetermined regressors, the estimation error of β might affect the asymptotic distribution of the autocorrelation estimators as demonstrated by Box and Pierce (1970). While this might complicate the analysis, considering predetermined regressors would also be an interesting extension.

5 ACKNOWLEDGEMENT

The author would like to thank In Choi and Songnian Chen for valuable comments. The author also acknowledges financial support from the Hong Kong University of Science and Technology under Project No. DAG05/06.BM16. The author is solely responsible for all errors.

6 REFERENCES

Andrews, D. W. K. (1991). Heteroskedasticity and autocorrelation consistent covariance matrix estimation, *Econometrica* **59**(3): 817–858.

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

Baltagi, B. H. and Wu, P. X. (1999). Unequally spaced panel data regressions with AR(1) disturbances, *Econometric Theory* **15**: 814–823.

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

Bertrand, M., Duflo, E. and Mullainathan, S. (2004). How much should we trust differences-in-differences estimates, *The Quarterly Journal of Economics* **119**: 249–275.

Bhargava, A., Franzini, L. and Narendranathan, W. (1982). Serial correlation and the fixed effects model, *The Review of Economic Studies* **49**(4): 533–549.

Box, G. E. P. and Pierce, D. A. (1970). Distribution of residual autocorrelations in autoregressive-integrated moving average time series models, *Journal of the American Statistical Association* **65**(332): 1509–1526.

Doornik, J. A. (2006). *An Object-oriented Matrix Programming Language - Ox 4*, Timberlake Consultants Ltd.

Drukker, D. M. (2003). Testing for serial correlation in linear panel-data models, *Stata Journal* **3**: 168–177.

Fu, B., Li, W. K. and Fung, W. K. (2002). Testing model adequacy for dynamic panel data with intercorrelation, *Biometrika* **89**(3): 591–601.

Hahn, J. and Kuersteiner, G. (2002). Asymptotically unbiased inference for a dynamic panel model with fixed effects when both n and T are large, *Econometrica* **70**(4): 1639–1657.

Hayashi, F. (2000). *Econometrics*, Princeton University Press.

Hong, Y. and Kao, C. (2004). Wavelet-based testing for serial correlation of unknown form in panel models, *Econometrica* **72**(5): 1519–1563.

Inoue, A. and Solon, G. (2006). A portmanteau test for serially correlated errors in fixed effects models, *Econometric Theory* **22**: 835–851.

Kézdi, G. (2004). Robust standard error estimation in fixed-effects panel models, *Hungarian Statistical Review* **9**: 95–116.

Ljung, G. M. and Box, G. E. P. (1978). On a measure of lack of fit in time series models, *Biometrika* **65**(2): 297–303.

Okui, R. (2007). Asymptotically unbiased estimation of autocovariances and autocorrelations with long panel data, unpublished manuscript.

Parzen, E. (1957). Consistent estimates of the spectrum of a stationary time series, *Annlas of Mathematical Statistics* **28**(2): 329–348.

Wooldridge, J. M. (2002). *Econometric analysis of cross section and panel data*, The MIT press.