Testing for Serial Correlation in Fixed-Effects Panel Data Models

Benjamin Born
University of Mannheim

Jörg Breitung*
University of Cologne

Abstract

In this paper we propose various tests for serial correlation in fixed-effects panel data regression models with a small number of time periods. First, a simplified version of the test suggested by Wooldridge (2002) and Drukker (2003) is considered. The second test is based on the Lagrange Multiplier (LM) statistic suggested by Baltagi and Li (1995), and the third test is a modification of the classical Durbin-Watson statistic. Under the null hypothesis of no serial correlation, all tests possess a standard normal limiting distribution as N tends to infinity and T is fixed. Analyzing the local power of the tests, we find that the LM statistic has superior power properties. Furthermore, a generalization to test for autocorrelation up to some given lag order and a test statistic that is robust against time dependent heteroskedasticity are proposed.

JEL Classification: C12, C23
Keywords: fixed-effects panel data; serial correlation; hypothesis testing;

*Corresponding author: Center of Econometrics and Statistics, University of Cologne, Meister-Ekkehart-Str. 9, 50923 Cologne, Germany. *Email:* breitung@statistik.uni-koeln.de

1 Introduction

Panel data models are increasingly popular in applied work as they have many advantages over cross-sectional approaches (see e.g. Hsiao, 2003, p. 3ff). The classical linear panel data model assumes serially uncorrelated disturbances. As argued by Baltagi (2008, p. 92), this assumption is likely to be violated as the dynamic effect of shocks to the dependent variable is often distributed over several time periods. In such cases, serial correlation leads to inefficient estimates and biased standard errors. Nowadays, (cluster) robust standard errors are readily available in econometric software like Stata, SAS or EViews that allow for valid inference under heteroskedasticity and autocorrelation. Therefore, many practitioners consider a possible autocorrelation as being irrelevant, in particular, in large data sets where efficiency gains are less important. On the other hand, a significant test statistic provides evidence that a static model may not be appropriate and should be replaced by a dynamic panel data model.

A number of tests for serial error correlation in panel data models have been proposed in the literature. Bhargava et al. (1982) generalized the Durbin-Watson statistic to the fixed-effects panel model. Baltagi and Li (1991, 1995) and Baltagi and Wu (1999) derived Lagrange Multiplier (LM) statistics for first order serial correlation. Drukker (2003), elaborating on an idea originally proposed by Wooldridge (2002), developed an easily implementable test for serial correlation based on the ordinary least-squares (OLS) residuals of the first-differenced model. This test is referred to as the Wooldridge-Drucker (henceforth: WD) test. Inoue and Solon (2006) suggested a portmanteau statistic to test for autocorrelations at *any* lag order.

However, all these tests have their limitations. A serious problem of the Bhargava et al. (1982) statistic is that the distribution depends on N (the number of cross-section units) and T (the number of time periods) and, therefore, the critical values have to be provided in large tables depending on both dimensions. Baltagi and Li (1995) noted that, for fixed T, their test statistic does not possess the usual χ^2 limiting distribution due to the (Nickell) bias in the estimation of the autocorrelation coefficient. The WD test is based on first differences, which may imply a loss of power against some particular alternatives. Furthermore, these tests are not robust against temporal heteroskedasticity and are not applicable to unbalanced panels (with the exception of the WD test and the Baltagi and Wu (1999)-statistic). Inoue and Solon (2006) proposed an LM statistic for the more general null hypothesis: $E(u_{it}u_{is}) = 0$ for all $t < s$, where u_{it} denotes the error of the panel regression. This portmanteau test is attractive for panels with small T as it is not designed to test against a particular alternative. If T is moderate or large, however, the size and power properties suffer from the fact that the dimension of the null hypothesis increases with T^2.

In this paper, we propose new test statistics and modifications of existing test statistics that sidestep some of these limitations. In Section 2, we first present the model framework and briefly review the existing tests. Our new test procedures are considered in Sections 3–5 and the small sample properties of the tests are studied in Section 6. Section 7 concludes.

2 Preliminaries

Consider the usual fixed effects panel data model with serially correlated disturbances

$$y_{it} = x_{it}'\beta + \mu_i + u_{it} \tag{1}$$

$$u_{it} = \rho u_{i,t-1} + \varepsilon_{it} \ , \tag{2}$$

where $i = 1, \ldots, N$ denotes the cross-section dimension and $t = 1, \ldots, T$ is the time dimension, y_{it} is the dependent variable, x_{it} is a $k \times 1$ vector of regressors, β is the unknown vector of associated coefficients, μ_i is the (fixed) individual effect, and u_{it} denotes the regression error. In our benchmark situation, the following assumption is imposed:

Assumption 1. *(i) The error ε_{it} is independently distributed across i and t with $E(\varepsilon_{it}) = 0$, $E(\varepsilon_{it}^2) = \sigma_i^2$ with $c^{-1} < \sigma_i^2 < c$ for all $i = 1, \ldots, N$ and some finite positive constant c. Furthermore, $E\left|\varepsilon_{it}\right|^{4+\delta} < \infty$ for some $\delta > 0$. (ii) The vector of coefficients β is estimated by a consistent estimator $\widehat{\beta}$ with $\widehat{\beta} - \beta = O_p(N^{-1/2})$.*

(iii) For the regressors it holds that for fixed T and $N \to \infty$

$$\frac{1}{N}\sum_{t=1}^{T}\sum_{i=1}^{N} E(x_{it}x_{it}') \ \to \ \overline{C} \ ,$$

where $||\overline{C}|| = \sqrt{tr(\overline{C}'\overline{C})} < \infty$.

All tests are based on the residual $\widehat{e}_{it} = y_{it} - x_{it}'\widehat{\beta} = \mu_i + u_{it} - x_{it}'(\widehat{\beta} - \beta)$. Assumption 1 (ii) allows us to replace the residuals $\widehat{e}_{it}$ by $e_{it} = \mu_i + u_{it}$ in our asymptotic considerations. If x_{it} is assumed to be strictly exogenous, the usual within-group (or least-squares dummy-variable, cf. Hsiao, 2003) estimator or the first difference estimator (cf. Wooldridge, 2002) may be used. If x_{it} is predetermined or endogenous, suitable instrumental variable (or GMM) estimators of β satisfy Assumption 1 (ii) (e.g. Wooldridge, 2002). It is important to notice that the efficiency of the estimator $\widehat{\beta}$ does not affect the asymptotic properties (i.e. size or local power) of tests based on $\widehat{e}_{it}$.

To test the null hypothesis $\rho = 0$, Bhargava et al. (1982) proposed the pooled Durbin-Watson statistic given by

$$p\text{DW} = \frac{\sum\limits_{i=1}^{N}\sum\limits_{t=2}^{T}(\widehat{u}_{it} - \widehat{u}_{i,t-1})^2}{\sum\limits_{i=1}^{N}\sum\limits_{t=1}^{T}\widehat{u}_{it}^2} \ ,$$

where $\widehat{u}_{it} = y_{it} - x_{it}'\widehat{\beta} - \widehat{\mu}_i$, and $\widehat{\beta}$ and $\widehat{\mu}_i$ denote the least-squares dummy-variable (or within-group) estimator of β and μ_i, respectively. A serious problem of this test is that its null distribution depends on N and T and, therefore, the critical values are provided in large tables depending on both dimensions in Bhargava et al. (1982). Furthermore, no critical values are available for *unbalanced* panels.

Baltagi and Li (1995) derive the LM test statistic for the hypothesis $\rho = 0$ assuming normally distributed errors. The resulting test statistic is equivalent to (the LM version of) the t-statistic of ϱ in the regression

$$\widehat{u}_{it} = \varrho \widehat{u}_{i,t-1} + \nu_{it} \ . \tag{3}$$

The LM test statistic results as

$$\text{LM} = \sqrt{\frac{NT^2}{T-1}} \left(\frac{\sum\limits_{i=1}^{N} \sum\limits_{t=2}^{T} \widehat{u}_{it} \widehat{u}_{i,t-1}}{\sum\limits_{i=1}^{N} \sum\limits_{t=1}^{T} \widehat{u}_{it}^2} \right).$$

Baltagi and Li (1995) show that if $N \to \infty$ and $T \to \infty$, the LM statistic has a standard normal limiting distribution. However, if T is fixed and $N \to \infty$, the application of the respective critical values leads to severe size distortions (see Section 3.2).

3 Test statistics for fixed T

In this section, we consider test procedures that are valid for fixed T and $N \to \infty$. All statistics have the general form

$$\lambda_{NT} \;=\; \frac{\sum\limits_{i=1}^{N} \widehat{z}_{Ti}}{\sqrt{\sum\limits_{i=1}^{N} \widehat{z}_{Ti}^2 - \frac{1}{N} \left(\sum\limits_{i=1}^{N} \widehat{z}_{Ti} \right)^2}}\,, \tag{4}$$

where $\widehat{z}_{Ti} = \widehat{e}_i' A_T \widehat{e}_i$, $\widehat{e}_i = [\widehat{e}_{i1}, \ldots, \widehat{e}_{iT}]'$, $\widehat{e}_{it} = y_{it} - x_{it}'\widehat{\beta}$ and A_T is a deterministic $T \times T$ matrix. The matrix A_T is constructed such that it eliminates the individual effects from $\widehat{e}_i$. Specifically we make the following assumption with respect to the matrix A_T:

Assumption 2. *(i) Let ι_T denote the $T \times 1$ vector of ones. The $T \times T$ matrix A_T obeys $A_T \iota_T = 0$, $A_T' \iota_T = 0$ and $tr(A_T) = 0$. (ii) Let $X_i = [x_{i1}, \ldots, x_{iT}]'$ and $u_i = [u_{i1}, \ldots, u_{iT}]'$. For all $i = 1, \ldots, N$ it holds that $E[X_i'(A_T + A_T')u_i] = 0$.*

In Lemma A.1 of the appendix it is shown that under Assumptions 1 – 2 the estimation error $\widehat{\beta} - \beta$ is asymptotically negligible. Note that, in general, Assumption 2 (ii) is violated if x_{it} is predetermined or endogenous.

The following lemma presents the limiting distribution of the test statistic under a sequence of local alternatives.

Lemma 1. *(i) Under Assumptions 1–2, $\rho = 0$, fixed T, and $N \to \infty$, the test statistic λ_{NT} has a standard normal limiting distribution. (ii) Under the local alternative $\rho_N = c/\sqrt{N}$, $\varepsilon_{it} \sim \mathcal{N}(0, \sigma_i^2)$ and Assumptions 1–2 the limiting distribution is given by*

$$\lambda_{NT} \xrightarrow{d} \mathcal{N}\left(\frac{c\kappa \left(\sum_{i=1}^{N} tr(A_T H_T) \right)}{\sqrt{tr(A_T^2 + A_T' A_T)}}, 1 \right),$$

where $\kappa = m_2/\sqrt{m_4}$ with $m_k = \lim_{N\to\infty} N^{-1} \sum_{i=1}^{N} \sigma_i^k$ and H_T is a $T \times T$ matrix with (t, s)-element $H_{T,ts} = 1$ for $|t - s| = 1$ and zeros elsewhere.

A straightforward extension to unbalanced panel data with T_i consecutive observations in group i and $\widehat{e}_i' A_{T_i} \widehat{e}_i$ yields[1]

$$\lambda_{NTi} \xrightarrow{d} \mathcal{N}\left(c \frac{\lim\limits_{N\to\infty} N^{-1} \sum_{i=1}^{N} \sigma_i^2 tr(A_{T_i} H_{T_i})}{\sqrt{\lim\limits_{N\to\infty} N^{-1} \sum_{i=1}^{N} \sigma_i^4 tr(A_{T_i}^2 + A_{T_i}' A_{T_i})}}, 1 \right).$$

Hence, to accommodate unbalanced panel data, the tests are computed with individual specific transformation matrices A_{T_i} instead of a joint matrix A_T.

3.1 The WD test statistic

To obtain a valid test statistic for fixed T, Wooldridge (2002, p. 282f) suggests to run a least squares regression of the differenced residuals $\Delta\widehat{e}_{it} = \widehat{e}_{it} - \widehat{e}_{i,t-1}$ on the lagged differences $\Delta\widehat{e}_{i,t-1}$.[2] Under the null hypothesis $\rho = 0$, the first order autocorrelation of $\Delta\widehat{e}_{it}$ converges in probability to -0.5. Since $\Delta\widehat{e}_{it}$ is serially correlated, Drukker (2003) suggests to employ heteroskedasticity and autocorrelation consistent (HAC)

[1] As pointed out by Baltagi and Wu (1999), the case of gaps within the sequence of observations is more complicated.

[2] Since this test employs the least-squares estimator in first differences, this test assumes $E(\Delta x_{it} \Delta u_{it}) = 0$ or $E(x_{it} u_{is}) = 0$ for all t and $s \in \{t-1, t, t+1\}$.

standard errors,[3] yielding the test statistic

$$\text{WD} \;=\; \frac{\widehat{\theta}+0.5}{\widehat{s}_\theta}\;.$$

Here, $\widehat{\theta}$ denotes the least-squares estimator of θ in the regression

$$\Delta\widehat{e}_{it} = \theta\Delta\widehat{e}_{i,t-1} + \eta_{it}\;, \tag{5}$$

and $\widehat{s}^2_\theta$ is the HAC variance estimator computed as

$$\widehat{s}^2_\theta = \frac{\sum\limits_{i=1}^{N}\left(\sum\limits_{t=3}^{T}\Delta\widehat{e}_{i,t-1}\widehat{\eta}_{it}\right)^2}{\left(\sum\limits_{i=1}^{N}\sum\limits_{t=3}^{T}\Delta\widehat{e}^2_{i,t-1}\right)^2}\;,$$

with $\widehat{\eta}_{it}$ as the pooled OLS residual from the autoregression (5).

In the following theorem, we propose a simplified and asymptotically equivalent version of the WD test.

Theorem 1. *Let* $\widehat{z}_{Ti} = \sum\limits_{t=3}^{T}(\widehat{e}_{it} - \frac{1}{2}\widehat{e}_{i,t-1} - \frac{1}{2}\widehat{e}_{i,t-2})(\widehat{e}_{i,t-1} - \widehat{e}_{i,t-2})$ *and denote by* $\widetilde{WD}$ *the corresponding test statistic constructed as in* (4)*.*

(i) Under Assumptions 1–2, $u_{it} \sim \mathcal{N}(0,\sigma_i^2)$*,* $T \geq 3$*, and* $\rho_N = c/\sqrt{N}$*, it follows that*

$$\widetilde{WD} \;\xrightarrow{d}\; \mathcal{N}\left(c\,\frac{\kappa(T-2)}{\sqrt{2(T-3)+3}},1\right),$$

where κ *is defined in Lemma 1.*

(ii) The test statistic $\widetilde{WD}$ *is asymptotically equivalent to WD in the sense that* $\widetilde{WD} - WD \xrightarrow{p} 0$ *as* $N \to \infty$*.*

Remark 1: The form of the test statistic results from

$$\widehat{\theta}+0.5 \;=\; \frac{\sum\limits_{i=1}^{N}\sum\limits_{t=3}^{T}(\Delta\widehat{e}_{it}\Delta\widehat{e}_{i,t-1} + \frac{1}{2}\Delta\widehat{e}^2_{i,t-1})}{\sum\limits_{i=1}^{N}\sum\limits_{t=3}^{T}(\Delta\widehat{e}_{i,t-1})^2} \;=\; \frac{\sum\limits_{i=1}^{N}\widehat{z}_{Ti}}{\sum\limits_{i=1}^{N}\sum\limits_{t=3}^{T}(\Delta\widehat{e}_{i,t-1})^2}\;.$$

[3]This approach is also known as “robust cluster” or “panel corrected” standard errors.

The $\widetilde{\text{WD}}$ statistic deviates from WD by computing the HAC variance based on the residuals obtained under the null hypothesis, i.e., $\eta^0_{it} = \Delta\widehat{e}_{it} + 0.5\Delta\widehat{e}_{i,t-1}$ instead of the residuals $\widehat{\eta}_{it}$ used to compute the original WD statistic.

Remark 2: It is interesting to note that under the alternative we have as $N \to \infty$

$$\widehat{\theta} + 0.5 \xrightarrow{p} \frac{r_1 - r_2}{2(1-r_1)} + O(T^{-1}) \ ,$$

where r_j is the j'th autocorrelation of u_{it}. This suggests that WD (and $\widetilde{\text{WD}}$) is a test against the *difference* between the first and second order autocorrelation of the errors. Therefore, the test is expected to have poor power against alternatives with $r_1 \approx r_2$.

Remark 3: The test statistic is robust against cross-sectional heteroskedasticity. However, using $\theta = -0.5$ requires that the variances do not change in time, i.e., $E(u^2_{it}) = \sigma^2_i$ for all t. Thus, this test rules out time dependent heteroskedasticity.

3.2 The LM test

An important feature of the LM test suggested by Baltagi and Li (1995) is that the limit distribution depends on T. This is due to the fact that the least-squares estimator of ϱ in regression (3) is biased and the errors ν_{it} in (3) are autocorrelated. Under fairly restrictive assumptions the following asymptotic null distribution is obtained:

Lemma 2. *Under Assumption 1 and $u_{it} \overset{iid}{\sim} \mathcal{N}(0,\sigma^2)$ for all i and t, it holds for fixed T and $N \to \infty$ that*

$$\left(LM + \sqrt{\frac{N}{T-1}} \right) \xrightarrow{d} \mathcal{N}\left(0, \frac{(T+1)(T-2)^2}{(T-1)^3}\right) .$$

It follows that applying the usual critical values derived from the χ^2_1 distribution to the (two-sided) test statistic LM^2 yields a test with actual size tending to unity as $N \to \infty$. Note also that the limit distribution of the LM statistic tends to a standard normal distribution if $T \to \infty$ and $N/T \to 0$. To obtain an asymptotically valid test for fixed T, the transformed statistic

$$\widetilde{LM} \;=\; \sqrt{\frac{(T-1)^3}{(T+1)(T-2)^2}}\left(LM + \sqrt{\frac{N}{T-1}}\right)$$

may be used, which has a standard normal limiting null distribution. An important limitation of this test statistic is, however, that the result requires the errors to be normally distributed with identical variances for all $i \in \{1,\ldots,N\}$. To obtain a test statistic that is valid in more general (and realistic) situations, we follow (Wooldridge, 2002, p. 275) and construct a test of the least squares estimator $\widehat{\varrho}$ of ϱ in the regression

$$\widehat{e}_{it} - \frac{1}{T}\sum_{t=1}^{T}\widehat{e}_{it} \;=\; \varrho\left(\widehat{e}_{i,t-1} - \frac{1}{T}\sum_{t=1}^{T}\widehat{e}_{it}\right) + \nu_{it} \;. \tag{6}$$

Let M_1 and M_T denote matrices that result from $M = I_T - T^{-1}\iota_T\iota_T'$ (with I_T indicating the $T \times T$ identity matrix) by dropping the first or the last row, respectively. Using this matrix notation, we obtain

$$\widehat{\varrho} \;\xrightarrow{p}\; \varrho_0 = \frac{tr(M_1'M_T)}{tr(M_T'M_T)} = -\frac{(T-1)/T}{(T-1)^2/T} = -\frac{1}{T-1} \;. \tag{7}$$

Thus the regression t-statistic is employed to test the modified null hypothesis $H_0' : \varrho = \varrho_0 = -1/(T-1)$. To account for autocorrelation in ν_{it}, (HAC) robust standard errors are employed, yielding the test statistic

$$\mathrm{LM}^* = \frac{\widehat{\varrho} - \varrho_0}{\widetilde{v}_\rho} \;,$$

where

$$\widetilde{v}_\rho^2 = \frac{\sum_{i=1}^{N}\widehat{e}_i'M_T'\widetilde{\nu}_i\widetilde{\nu}_i'M_T\widehat{e}_i}{\left(\sum_{i=1}^{N}\widehat{e}_i'M_T'M_T\widehat{e}_i\right)^2} \;=\; \frac{\sum_{i=1}^{N}\widehat{e}_i'M_T'(M_1 - \widehat{\varrho}M_T)\widehat{e}_i\widehat{e}_i'(M_1 - \widehat{\varrho}M_T)'M_T\widehat{e}_i}{\left(\sum_{i=1}^{N}\widehat{e}_i'M_T'M_T\widehat{e}_i\right)^2} \;.$$

As for the WD test, we propose a simplified version of this test that is asymptotically equivalent to the statistic LM*.

Theorem 2. *Let*

$$\widehat{z}_{Ti} = \sum_{t=2}^{T} \left[\left(\widehat{e}_{it} - \frac{1}{T}\sum_{t=1}^{T}\widehat{e}_{it} \right) \left(\widehat{e}_{i,t-1} - \frac{1}{T}\sum_{t=1}^{T}\widehat{e}_{it} \right) + \frac{1}{T-1} \left(\widehat{e}_{i,t-1} - \frac{1}{T}\sum_{t=1}^{T}\widehat{e}_{it} \right)^2 \right]$$

and denote by $\widetilde{LM}^*$ *the corresponding test statistic constructed as in* (4)*.*

(i) Under Assumptions 1–2, $u_{it} \sim \mathcal{N}(0, \sigma_i^2)$*,* $T \geq 3$*, and the local alternative* $\rho_N = c/\sqrt{N}$*, the test statistic* $\widetilde{LM}^*$ *is asymptotically distributed as*

$$\mathcal{N}\left(c\,\kappa\sqrt{T - 3 + \frac{2}{T^2 - T}}\,,\, 1 \right).$$

(ii) Under the the local alternative, the statistic $\widetilde{LM}^*$ *is asymptotically equivalent to* LM^**.*

3.3 A modified Durbin-Watson statistic

The pDW statistic suggested by Bhargava et al. (1982) is the ratio of the sum of squared differences and the sum of squared residuals. Instead of the ratio (which complicates the theoretical analysis), our variant of the Durbin-Watson test is based on the *linear combination* of the numerator and denominator of the Durbin-Watson statistic. Let

$$\widehat{z}_{Ti} = \widehat{e}_i' M D' D M \widehat{e}_i - 2\,\widehat{e}_i' M \widehat{e}_i \,, \tag{8}$$

where D is a $(T-1) \times T$ matrix producing first differences, i.e.

$$D \;=\; \begin{pmatrix} -1 & 1 & 0 & \cdots & & 0 \\ 0 & -1 & 1 & \cdots & & 0 \\ \vdots & & \ddots & \ddots & & \vdots \\ \vdots & & & \ddots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & -1 & 1 \end{pmatrix}, \tag{9}$$

and M as defined above. Using $tr(MD'DM) = 2(T-1)$ and $tr(M) = T-1$, it follows that $E(z_{Ti}) = 0$ for all i, where $z_{Ti} = e_i' M D' D M e_i - 2\,e_i' M e_i$.

The panel test statistic is constructed as in (4), where $\widehat{z}_{Ti}$ is given by (8). The resulting test is denoted by mDW. The following theorem presents the asymptotic distribution of the test statistic.

Theorem 3. *Under Assumptions 1–2, $u_{it} \sim \mathcal{N}(0, \sigma_i^2)$, $T \geq 3$, and the local alternative $\rho_N = c/\sqrt{N}$, the limiting distribution is*

$$mDW \xrightarrow{d} \mathcal{N}\left(-c\kappa \frac{T-1}{T}\sqrt{T-2}, 1\right).$$

3.4 Relative local power

Theorems 1–3 allow us to compare the asymptotic power of the three different tests. Let $\psi_{WD}(T) = \kappa(T-2)/\sqrt{2(T-3)+3}$ denote the Pitman drift of the local power resulting from Theorem 1(i) and denote by $\psi_{LM}(T)$ and $\psi_{mDW}(T)$ the respective terms resulting from Theorem 2(i) and Theorem 3. The test A is asymptotically more powerful than test B for some given value of T if $\psi_A(T)/\psi_B(T) > 1$, where $A, B \in \{WD, LM, mDW\}$. In what follows we use the Pitman drift of the WD test as a benchmark and define the relative Pitman drift as $\tau_{LM}(T) = \psi_{LM}(T)/\psi_{WD}(T)$ and $\tau_{mDW}(T) = \psi_{mDw}(T)/\psi_{WD}(T)$. Figure 1 presents the relative Pitman drifts of both tests as a function of T. It turns out that for the minimum value of $T = 3$ the mDW test is asymptotically more powerful among the set of three alternative tests. For $T \geq 4$, however, the LM test is asymptotically most powerful, where the difference between the LM and the mDW tests diminish as T gets large. The power gain of both tests relative to the WD test approaches 40 percent in terms of the Pitman drift of the WD test.

Note that the asymptotic power of some test $A \in \{WD, LM, mDW\}$ is a nonlinear function of $\psi_A(T)$. For small values of c, however, the asymptotic power is approximately linear. In Table 1 we report the asymptotic power of the tests (in parentheses) for various values of c and T. For example, consider $c = 0.5$ and $T = 50$. The asymptotic power of the LM test relative to the WD test is $0.929/0.683 = 1.36$ and the relative asymptotic power of the mDW test results as $0.923/0.683 = 1.35$, which roughly corresponds to the relative Pitman drifts $\tau_{LM}(50)$ and $\tau_{LM}(50)$.

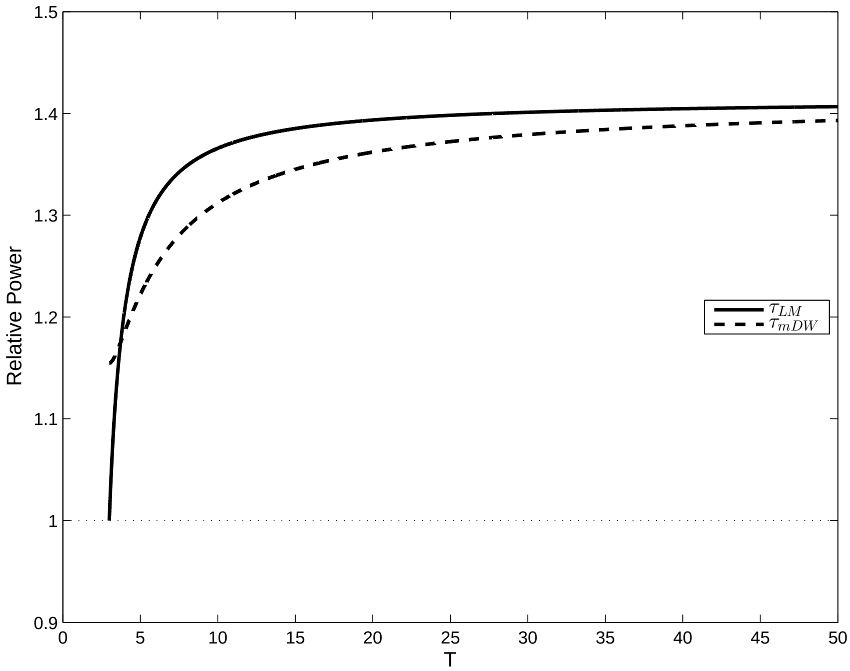

Figure 1: Comparison of local power functions

4 Higher order autocorrelation

Since the estimation error $x_{it}(\widehat{\beta} - \beta)$ in the residuals $\widehat{e}_{it} = e_{it} - x_{it}(\widehat{\beta} - \beta)$ is asymptotically negligible we simplify our notation and write e_{it} instead of $\widehat{e}_{it}$ in this section.

4.1 The Inoue-Solon statistic

A test for higher order serial correlation in panel data models was proposed by Inoue and Solon (2006). Due to the fact that the covariance matrix of the demeaned error vector $\Sigma_i^* = E(Me_ie_i'M)$ is singular with rank $T-1$, we drop the k'th row and column of Σ_i^* yielding the invertible matrix $\Sigma_{i,k}^*$. The null hypothesis of no autocorrelation implies that $\Sigma_{i,k}^* = \sigma_i^2 M_k$, where M_k results from M by dropping the k'th row and column of M. Denote by $S_{i,k}$ the matrix that results from deleting the k'th row and columns of $S_i = Me_ie_i'M$. The LM statistic is given by

$$\text{IS}_k \;=\; \left(\sum_{i=1}^{N} \widetilde{s}_{i,k}'\right)\left(\sum_{i=1}^{N} \widetilde{s}_{i,k}\widetilde{s}_{i,k}'\right)^{-1}\left(\sum_{i=1}^{N} \widetilde{s}_{i,k}\right)\;, \tag{10}$$

where

$$\widetilde{s}_{i,k} \;=\; \text{vech}(S_{i,k} - \sigma_i^2 M_k)$$

and vech(A) denotes the linear operator that stacks all elements of the $m \times m$ matrix A below the main diagonal into a $m(m-1)/2$ dimensional vector. In empirical practice the unknown variances $\sigma_1^2, \ldots, \sigma_N^2$ are replaced by the unbiased estimate $\widehat{\sigma}_i^2 = (T-1)^{-1}e_i'Me_i$.

An important problem with this test statistic is that it essentially tests a $(T-2)(T-1)/2$ dimensional null hypothesis. Thus, to yield a reliable asymptotic approximation of the null distribution, N should be large relative to $T^2/2$. Another problem is that the power of the test is poor if only a few autocorrelation coefficients are substantially different from zero. For example, if the errors are generated by a first order autoregression with $u_{it} = 0.1u_{i,t-1} + \varepsilon_{it}$, then the second and third order autocorrelations are as small as 0.01 and 0.001. In such cases, a test that includes all autocorrelations up to the order $T-1$ lacks power. In order to improve the power of the test statistic in such cases, the test statistic should focus on the first order autocorrelations. Inoue and Solon (2006) suggest to construct a test statistic based on the first p autocorrelations by forming the lower-dimensional vector $\widetilde{s}_i(p) = \{\widetilde{s}_i^{(ts)}\}_{1\le t-s\le p}$ with elements

$$\widetilde{s}_i^{(ts)} = (e_{it} - \overline{e}_i)(e_{is} - \overline{e}_i) + \sigma_i^2/T$$

for all $t, s \in \{1, \ldots, T\}$ with $1 \le t - s \le p$. The LM statistic is obtained as in (10) where the vector $\widetilde{s}_{i,k}$ is replaced by $\widetilde{s}_i(p)$. The resulting test statistic is denoted by IS(p) and has an asymptotic χ^2 distribution with $pT - p(p+1)/2$ degrees of freedom.

4.2 Bias-corrected test statistics

An alternative approach to test against autocorrelation of order k is to consider the k-th order autoregression

$$e_{it} - \overline{e}_i = \varrho_k(e_{i,t-k} - \overline{e}_i) + \epsilon_{it} \ , \quad t = k+1, \ldots, T; \quad i = 1, \ldots, N \ . \tag{11}$$

In the following lemma, we present the probability limit of the OLS estimator of ϱ_k for fixed T and $N \to \infty$.

Lemma 3. *Let $\widehat{\varrho}_k$ be the pooled OLS estimator of ϱ_k in* (11). *Under Assumption 1 and $\rho = 0$, we have*

$$\widehat{\varrho}_k \ \stackrel{p}{\longrightarrow} \ -\frac{1}{T-1} \ ,$$

for all $k \in \{1, \ldots, T-2\}$.

Using this lemma, a test for zero autocorrelation at lag k can be constructed based on the (heteroskedasticity and autocorrelation robust) t-statistic of the hypothesis $\varrho_k = -1/(T-1)$. An asymptotically equivalent test statistic is obtained by letting

$$z_{Ti}^{(k)} = \sum_{t=k+1}^{T} \left[(e_{it} - \overline{e}_i)(e_{i,t-k} - \overline{e}_i) + \frac{1}{T-1}(e_{i,t-k} - \overline{e}_i)^2 \right]$$

and

$$\widetilde{\mathrm{LM}}_k^* = \frac{\sum_{i=1}^{N} z_{Ti}^{(k)}}{\sqrt{\sum_{i=1}^{N} z_{Ti}^{(k)2} - \frac{1}{N}\left(\sum_{i=1}^{N} z_{Ti}^{(k)}\right)^2}} \ . \tag{12}$$

Along the lines of the proof of Theorem 2, it is not difficult to show that, under the null hypothesis, $\widetilde{\mathrm{LM}}_k^*$ has a standard normal limiting distribution.

In empirical practice it is usually more interesting to test the hypothesis that the errors are not autocorrelated *up to order p*. Unfortunately, the probability limit of the autoregressive coefficients $\phi_1, \ldots, \phi_p$ in the autoregression

$$e_{it} - \overline{e}_i = \phi_1(e_{i,t-1} - \overline{e}_i) + \cdots + \phi_p(e_{i,t-p} - \overline{e}_i) + \upsilon_{it}$$

is a much more complicated function of T, since the regressors are mutually correlated. Instead of presenting the respective probability limits, we therefore suggest a simple transformation of the regressors such that the autoregressive coefficients can be estimated unbiasedly under the null hypothesis. Specifically, we propose an implicit correction that eliminates the bias. In matrix notation, the (transformed) autoregression is given by

$$Me_i = \phi_1 L_1 e_i + \phi_2 L_2 e_i + \cdots + \phi_p L_p e_i + v_i \ , \tag{13}$$

where

$$L_k e_i \ = \ \begin{pmatrix} 0_{k\times 1} \\ e_{i1} - \overline{e}_i \\ \vdots \\ e_{i,T-k} - \overline{e}_i \end{pmatrix} + \left(\frac{T-k}{T^2 - T} \right) e_i \ ,$$

where the first vector on the right hand side represents the lagged values of the demeaned residuals, whereas the second vector is introduced to remove the bias. It is not difficult to show that

$$E(e_i' M L_k e_i) \ = \ -\left(\frac{T-k}{T}\right)\sigma_i^2 + \left(\frac{T-1}{T}\right)\left(\frac{T-k}{T-1}\right)\sigma_i^2 \ = \ 0$$

for all $k = 1, \ldots, p$ and, therefore, under the null hypothesis, the least-squares estimators of $\phi_1, \ldots, \phi_p$ converge to zero in probability for all T and $N \to \infty$. The null hypothesis $\phi_1 = \cdots = \phi_p = 0$ can therefore be tested using the associated Wald statistic

$$Q(p) \ = \ \widehat{\phi}' \widehat{S}^{-1} \widehat{\phi} \ , \tag{14}$$

where

$$\widehat{S} = \left(\sum_{i=1}^{N} Z_i' Z_i \right)^{-1} \left(\sum_{i=1}^{N} Z_i' \widehat{v}_i \widehat{v}_i' Z_i \right) \left(\sum_{i=1}^{N} Z_i' Z_i \right)^{-1}$$

is the HAC estimator for the covariance matrix of the vector of coefficients with $\widehat{\phi} = \left[\widehat{\phi}_1, \ldots, \widehat{\phi}_p\right]'$, $Z_i = [L_1 e_i, \ldots, L_p e_i]$ and $\widehat{v}_i = Me_i - Z_i \widehat{\phi}$. The $Q(p)$ statistic is asymptotically equivalent to the statistic

$$\widetilde{Q}(p) \ = \ \sum_{i=1}^{N} e_i' M Z_i \left[\sum_{i=1}^{N} Z_i' M e_i e_i' M Z_i - \frac{1}{N} \left(\sum_{i=1}^{N} Z_i' M e_i \right) \left(\sum_{i=1}^{N} e_i' M Z_i \right) \right]^{-1} \sum_{i=1}^{N} Z_i' M e_i \ .$$

Along the lines of Theorem 2, it is straightforward to show that the statistics $Q(p)$ and $\widetilde{Q}(p)$ are asymptotically equivalent and have an asymptotic χ^2 distribution with p degrees of freedom.

5 A heteroskedasticity robust test statistic

An important drawback of all test statistics considered so far is that they are not robust against time dependent heteroskedasticity. This is due to the fact that the implicit or explicit bias correction of the autocovariances depends on the error variances. To overcome this drawback of the previous test statistics, we construct an unbiased estimator of the autocorrelation coefficient. The idea is to apply backward and forward transformations such that the products of the transformed series are uncorrelated under the null hypothesis. Specifically, we employ the following transformations for eliminating the individual effects:

$$\begin{aligned}\widetilde{e}_{it}^{f} &= \widehat{e}_{it} - \frac{1}{T-t+1}\left(\widehat{e}_{it} + \cdots + \widehat{e}_{iT}\right)\\ \widetilde{e}_{it}^{b} &= \widehat{e}_{it} - \frac{1}{t}\left(\widehat{e}_{i1} + \cdots + \widehat{e}_{it}\right) \ .\end{aligned}$$

The hypothesis can be tested based on the regression

$$\widetilde{e}_{it}^{f} = \psi \widetilde{e}_{i,t-1}^{b} + \omega_{it} \ , \quad t = 3, \ldots, T-1 \ , \tag{15}$$

or, in matrix notation,

$$V_0 \widehat{e}_i = \psi V_1 \widehat{e}_i + w_i \ ,$$

where the $(T-3) \times T$ matrices V_0 and V_1 are defined as

$$V_0 = \begin{bmatrix} -\frac{1}{2} & \frac{1}{2} & 0 & 0 & \ldots & 0 & 0 & 0\\ -\frac{1}{3} & -\frac{1}{3} & \frac{2}{3} & 0 & \ldots & 0 & 0 & 0\\ \vdots & & & & & & & \vdots\\ -\frac{1}{T-2} & -\frac{1}{T-2} & -\frac{1}{T-2} & -\frac{1}{T-2} & \ldots & \frac{T-3}{T-2} & 0 & 0 \end{bmatrix}$$

and

$$V_1 = \begin{bmatrix} 0 & 0 & \frac{T-3}{T-2} & -\frac{1}{T-2} & -\frac{1}{T-2} & \cdots & -\frac{1}{T-2} & -\frac{1}{T-2} \\ 0 & 0 & 0 & \frac{T-4}{T-3} & -\frac{1}{T-3} & \cdots & -\frac{1}{T-3} & -\frac{1}{T-3} \\ \vdots & & & & & & & \vdots \\ 0 & 0 & 0 & 0 & 0 & \cdots & \frac{1}{2} & -\frac{1}{2} \end{bmatrix}.$$

Under the null hypothesis of no (first order) autocorrelation, we have $\psi = 0$. Since the error vector w_i is heteroscedastic and autocorrelated, we again employ robust (HAC) standard errors. The resulting test statistic is equivalent to the test statistic

$$\widetilde{\mathrm{HR}} = \frac{\sum\limits_{i=1}^{N} \widehat{z}_{Ti}}{\sqrt{\sum\limits_{i=1}^{N} \widehat{z}_{Ti}^2 - \frac{1}{N}\left(\sum\limits_{i=1}^{N} \widehat{z}_{Ti}\right)^2}},$$

where

$$\widehat{z}_{Ti} = \widehat{e}_i' A_T \widehat{e}_i = \sum_{t=3}^{T-1} \widehat{e}_{it}^f \widehat{e}_{i,t-1}^b \quad \text{with } A_T = V_0' V_1.$$

Since $V_0 \iota_T$, $V_1 \iota_T = 0$ and $tr(V_0' V_1) = 0$ it follows from Lemma 1 (i) that the test statistic $\widetilde{\mathrm{HR}}$ has a standard normal limiting null distribution.

6 Monte Carlo Study

This section presents the finite sample performance of the test statistics for serial correlation and assesses the reliability of the asymptotic results derived in Section 3. We also conduct experiments with different forms of serial correlation and heteroskedasticity.

The benchmark data generating process for all simulations is a linear panel data model of the form

$$y_{it} = x_{it}\beta + \mu_i + u_{it},$$

where $i = 1, \ldots, 500$[4] and various values of T. We set β to 1 in all simulations and draw the individual effects μ_i from a $\mathcal{N}(0, 2.5^2)$ distribution. To create correlation

[4]The cross-sectional dimension is held constant at $N = 500$, results for smaller N are qualitatively similar. Values of $c = 0$, $c = 0.5$, and $c = 1$ therefore imply $\rho = 0$, $\rho = 0.0244$, and $\rho = 0.0447$, respectively.

between the regressor and the individual effect, we follow Drukker (2003) by drawing x_{it}^0 from a $\mathcal{N}(0, 1.8^2)$ distribution and computing $x_{it} = x_{it}^0 + 0.5\mu_i$. The regressor is drawn once and then held constant for all experiments. In our benchmark model, the disturbance term follows an autoregressive process of order 1:

$$u_{it} = \rho u_{i,t-1} + \varepsilon_{it} \ ,$$

where $\varepsilon_{it} \sim \mathcal{N}(0, 1)$ and we discard the first 100 observations to eliminate the influence of the initial value.[5]

Table 1 compares the size and power of alternative test statistics where, according to our theoretical analysis, we let $\rho_N = c/\sqrt{N}$. The empirical values are computed using 10,000 Monte Carlo simulations, while the asymptotic values (given in parentheses) are computed according to the results given in Theorems 1 – 3. All tests exhibit good size control (i.e. for $c = 0$ the actual size is close to the nominal size). They reject the null hypothesis of no serial correlation slightly more often than the nominal size of 5% but are never above 6%. In all cases, the LM statistic has superior power compared to the competing statistics. While the modified Durbin-Watson statistic performs similar to the LM statistic, the WD test exhibit considerably less power.

To evaluate the small-sample properties of the test for higher-order autocorrelation $\widetilde{Q}(p)$, we simulate a model with disturbances following a second-order autoregressive process:

$$\text{AR(2):} \qquad u_{it} = \alpha_1 u_{i,t-1} + \alpha_2 u_{i,t-2} + \varepsilon_{it} \ ,$$

where again $\varepsilon_{it} \sim \mathcal{N}(0, 1)$ and we discard the first 100 observations to eliminate the influence of the initial value. The results for three different AR(2) processes are presented in Table 2. The $\widetilde{Q}(2)$ statistic has good power properties in all configurations considered. Not surprisingly, the WD test exhibit high power when $\alpha_1 = 0.03$ and $\alpha_2 = -0.03$ or $\alpha_1 = 0$ and $\alpha_2 = 0.08$. As mentioned in Remark 2, however, the WD (and $\widetilde{\text{WD}}$) statistic is a test of the *difference* between the first and second order autocorrelation of the errors. As a consequence, it also has

[5]As suggested by a referee, we also consider non-Gaussian error distributions. The results were very similar and can be obtained on request.

power against most AR(2) alternatives. However, choosing $\alpha_1 = \alpha_2 = 0.03$, which implies equal first- and second-order autocorrelations of u_{it}, leads to a complete loss of power of the WD test, confirming our considerations in Remark 2. The LM statistic has less power than the $\widetilde{Q}(2)$ statistic, in particular for the setup with $\alpha_1 = 0$ and $\alpha_2 = 0.08$.

In the last two columns of Table 2 we present the original portmanteau statistic IS_k with $k = 1$ and the IS(2) statistic that is based on all autocorrelation up to to the lag order $p = 2$ (see section 4). Note that for a sample with $N = 500$ the inverse of IS_k does not exist for $T > 20$. It turns out that for the AR(2) alternatives considered in this Monte Carlo experiment[6] the $\widetilde{Q}(2)$ statistics has substantially more power than the variants of the IS statistic.

Next, we consider four different types of time-dependent heteroskedasticity by using the following error process:

$$u_{it} = \sqrt{h_t}\varepsilon_{it} \ ,$$

where $\varepsilon_{it} \sim \mathcal{N}(0, 1)$. The first model implies a break in the variance function, i.e. $h_t = 10$ for $t = 1, \ldots, T/5$ and $h_t = 1$ otherwise. The second specification is a U-shaped variance function of the form $h_t = (t - T/2)^2 + 1$. The third and fourth types of heteroskedasticity are positive and negative exponential variance functions, i.e. $h_t = \exp(-0.2t)$ and $h_t = \exp(0.2t)$, respectively. Table 3 shows that the modified Durbin-Watson test has massive size distortions under all four types of heteroskedasticity. While the WD test has the proper size in case of a U-shaped variance function, the break in the variance function and the exponential variance functions lead to large size distortions. The modified LM statistic does surprisingly well under heteroskedasticity, but for small T we observe also some size distortions, especially for the negative exponential variance function. On the other hand our proposed heteroskedasticity-robust test statistic retains its correct size under all types of heteroskedasticity considered here.

[6] We also considered a wide range of other combinations of α_1 and α_2. In all cases the $\widetilde{Q}(2)$ turns out to be more powerful than the IS statistics.

7 Conclusion

This paper proposes various tests for serial correlation in fixed-effects panel data regression models with a small number of time periods. First, a simplified version of the test suggested by Wooldridge (2002) and Drukker (2003) is considered. The second test is based on the LM statistic suggested by Baltagi and Li (1995), and the third test is a modification of the classical Durbin-Watson statistic. Under the null hypothesis of no serial correlation, all tests possess a standard normal limiting distribution as $N \to \infty$ and T is fixed. Analyzing the local power of the tests, we find that the LM statistic has superior power properties. We also propose a generalization to test for autocorrelation up to some given lag order and a test statistic that is robust against time dependent heteroskedasticity.

References

Baltagi, B. H. (2008). *Econometric Analysis of Panel Data* (4th ed.). Wiley.

Baltagi, B. H. and Q. Li (1991). A Joint Test for Serial Correlation and Random Individual Effects. *Statistics & Probability Letters 11*(3), 277–280.

Baltagi, B. H. and Q. Li (1995). Testing AR(1) against MA(1) Disturbances in an Error Component Model. *Journal of Econometrics 68*(1), 133–151.

Baltagi, B. H. and P. X. Wu (1999). Unequally Spaced Panel Data Regressions With AR(1) Disturbances. *Econometric Theory 15*(06), 814–823.

Bhargava, A., L. Franzini, and W. Narendranathan (1982). Serial Correlation and the Fixed Effects Model. *Review of Economic Studies 49*(4), 533–49.

Cox, D. R. and P. J. Solomon (1988). On Testing for Serial Correlation in Large Numbers of Small Samples. *Biometrika 75*(1), pp. 145–148.

Drukker, D. M. (2003). Testing for Serial Correlation in Linear Panel-Data Models. *Stata Journal 3*(2), 168–177.

Hsiao, C. (2003). *Analysis of Panel Data* (2nd ed.). Cambridge University Press.

Inoue, A. and G. Solon (2006). A Portmanteau Test For Serially Correlated Errors In Fixed Effects Models. *Econometric Theory 22*(05), 835–851.

Searle, S. (1971). *Linear Models*. New York: Wiley.

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

Appendix: Proofs

Proof of Lemma 1:

(i) Let $y_i = [y_{i1}, \ldots, y_{iT}]'$ and $X_i = [x_{i1}, \ldots, x_{iT}]'$. The residual vector can be written as

$$\widehat{e}_i = e_i - X_i(\widehat{\beta} - \beta) \ .$$

The following lemma implies that the estimation error in the residuals $\widehat{e}_{it}$ is asymptotically negligible.

Lemma A.1. Under Assumptions 1 and 2 it holds that

$$(i) \quad \frac{1}{\sqrt{N}} \sum_{i=1}^{N} \widehat{e}_i' A_T \widehat{e}_i = \frac{1}{\sqrt{N}} \sum_{i=1}^{N} e_i' A_T e_i + O_p(N^{-1})$$

$$(ii) \quad \frac{1}{N} \sum_{i=1}^{N} (\widehat{e}_i' A_T \widehat{e}_i)^2 = \frac{1}{N} \sum_{i=1}^{N} (e_i' A_T e_i)^2 + O_p(N^{-1/2}).$$

Proof: Let

$$\sum_{i=1}^{N} \widehat{e}_i' A_T \widehat{e}_i = \sum_{i=1}^{N} e_i' A_T e_i + \sum_{i=1}^{N} \delta_i$$

where

$$\begin{aligned} \delta_i &= -e_i'(A_T + A_T')X_i(\widehat{\beta} - \beta) + (\widehat{\beta} - \beta)' X_i' A_T X_i (\widehat{\beta} - \beta) \\ &= -u_i'(A_T + A_T')X_i(\widehat{\beta} - \beta) + (\widehat{\beta} - \beta)' X_i' A_T X_i (\widehat{\beta} - \beta) \end{aligned}$$

(i) We have

$$||\tfrac{1}{N} \sum_{i=1}^{N} e_i'(A_T + A_T')X_i(\widehat{\beta} - \beta)|| \quad \leq \quad ||\tfrac{1}{N} \sum_{i=1}^{N} u_i'(A_T + A_T')X_i|| \cdot ||\widehat{\beta} - \beta||$$

Using Assumption 2 and Chebyshev's inequality yields

$$\frac{1}{N} \sum_{i=1}^{N} u_i'(A_T + A_T')X_i = O_p(N^{-1/2})$$

and, therefore, from Assumption 1 (c) we conclude

$$\left\| \frac{1}{N} \sum_{i=1}^{N} u_i' A_T X_i \right\| \cdot ||\widehat{\beta} - \beta|| = O_p(N^{-1}) \ .$$

(ii) Consider

$$\begin{aligned} |\widehat{e}_i' A_T \widehat{e}_i| - |e_i' A_T e_i| \ &\leq \ |\delta_i|, \\ (\widehat{e}_i' A_T \widehat{e}_i)^2 - (e_i' A_T e_i)^2 &\leq 2\,|\delta_i| \cdot |u_i' A_T u_i| + \delta_i^2 \ , \end{aligned}$$

with

$$\delta_i^2 \leq 2 \left\{ \left[u_i'(A_T + A_T') X_i (\widehat{\beta} - \beta) \right]^2 + \left[(\widehat{\beta} - \beta)' X_i' A_T X_i (\widehat{\beta} - \beta) \right]^2 \right\} \ .$$

We have

$$\begin{aligned} \frac{1}{N} \sum_{i=1}^{N} |u_i'(A_T + A_T') X_i (\widehat{\beta} - \beta)|^2 &\leq ||A_T + A_T'||^2 \left(\frac{1}{N} \sum_{i=1}^{N} ||u_i' X_i||^2 \right) (\widehat{\beta} - \beta)'(\widehat{\beta} - \beta) \\ &= O_p(N^{-1}) \end{aligned}$$

and

$$\begin{aligned} \frac{1}{N} \sum_{i=1}^{N} \left[(\widehat{\beta} - \beta)' X_i' A_T X_i (\widehat{\beta} - \beta) \right]^2 &\leq ||A_T||^2 \left(\frac{1}{N} \sum_{i=1}^{N} ||X_i' X_i||^2 \right) ||\widehat{\beta} - \beta)||^4 \\ &= O_p(N^{-2}) \ . \end{aligned}$$

It follows that

$$\frac{1}{N} \sum_{i=1}^{N} \delta_i^2 = O_p(N^{-1}).$$

Since

$$\frac{1}{N} \sum_{i=1}^{N} |\delta_i| \ |u_i' A_T u_i| \leq \left(\frac{1}{N} \sum_{i=1}^{N} \delta_i^2 \right)^{1/2} \left(\frac{1}{N} \sum_{i=1}^{N} (u_i' A_T u_i)^2 \right)^{1/2} = O_p(N^{-1/2})$$

we finally obtain

$$\frac{1}{N}\sum_{i=1}^{N}\left(\widehat{e}_i' A_T \widehat{e}_i\right)^2 - \frac{1}{N}\sum_{i=1}^{N}\left(e_i' A_T e_i\right)^2 \; = \; O_p(N^{-1/2}).$$

■

Since $E(u_i' A_T u_i) = \sigma_i^2 tr(A_T) = 0$ and there exist some real number M such that $1/M < var(u_i' A_T u_i) < M$ it follows from the Lindeberg-Feller central limit theorem that

$$\lambda_{NT} = \frac{\sum_{i=1}^{N} \widehat{z}_{Ti}}{\sqrt{\sum_{i=1}^{N} \widehat{z}_{Ti}^2}} + o_p(1) \;\xrightarrow{d}\; \mathcal{N}(0,1).$$

(ii) Under the sequence of local alternatives, we have

$$\begin{aligned}
\Omega = E(u_i u_i') = \frac{\sigma_i^2}{1-\rho_N^2} & \begin{pmatrix} 1 & \rho_N & \rho_N^2 & \cdots & \rho_N^{T-1} \\ \rho_N & 1 & \rho_N & \cdots & \rho_N^{T-2} \\ \vdots & & & & \vdots \\ \rho_N^{T-1} & \rho_N^{T-2} & \rho_N^{T-3} & \cdots & 1 \end{pmatrix} \\
= \frac{\sigma_i^2}{1-{}^{c}\!/\!_{N}} I_T + \frac{c\,\sigma_i^2}{(1-{}^{c}\!/\!_{N})\sqrt{N}} H_T + O(N^{-1})C_T & \\
= \sigma_i^2 I_T + \frac{c\,\sigma_i^2}{\sqrt{N}} H_T + O(N^{-1})C_T^* &
\end{aligned}$$

where C_T and C_T^* are $T \times T$ matrices with bounded elements and H_T is a matrix with elements $h_{ts} = 1$ if $|t-s| = 1$ and zeros elsewhere. Since $tr(A_T) = 0$ it follows that

$$\begin{aligned}
E(u_i' A_T u_i) = tr(A_T \Omega) \;\; &= \;\; \sigma_i^2 tr(A_T) + \frac{\sigma_i^2 c}{\sqrt{N}} tr(A_T H_T) + O(N^{-1}) \\
&= \;\; \frac{\sigma_i^2 c}{\sqrt{N}} tr(A_T H_T) + O(N^{-1}) \; .
\end{aligned}$$

Using standard results for the variance of quadratic forms of normally distributed random variables[7] (e.g. Searle, 1971), we have

$$var(u_i' A_T u_i) = \sigma_i^4 \, tr(A_T^2 + A_T' A_T) + O(N^{-1/2}) \ .$$

It follows from the Lindeberg-Feller central limit theorem

$$\frac{1}{\sqrt{N}} \sum_{i=1}^{N} u_i' A_T u_i \ \xrightarrow{p} \ \mathcal{N}\Big(c\, m_2 tr(A_T H_T), m_4 \, tr(A_T^2 + A_T' A_T)\Big),$$

where m_2 and m_4 are defined in Lemma 1.

For the second moment we obtain from the law of large numbers

$$\begin{aligned} \frac{1}{N} \sum_{i=1}^{N} (\widehat{e}_i' A_T \widehat{e}_i)^2 &= \frac{1}{N} \sum_{i=1}^{N} \Big(u_i' A_T u_i + O_p(N^{-1/2}) \Big)^2 \\ &= \frac{1}{N} \sum_{i=1}^{N} (u_i' A_T u_i)^2 + O_p(N^{-1/2}) \\ &\xrightarrow{p} m_4 \, tr(A_T^2 + A_T' A_T) \ . \end{aligned}$$

Furthermore, under the sequence of local alternatives

$$\frac{1}{N} \left(\frac{1}{\sqrt{N}} \sum_{i=1}^{N} \widehat{e}_i' A_T \widehat{e}_i \right)^2 \ = \ O_p(N^{-1})$$

and, therefore, the second term of the denominator in (4) does not affect the limiting distribution.

It follows that the limiting distribution of the test statistic is given by

$$\lambda_{NT} \ = \ \frac{\frac{1}{\sqrt{N}} \sum\limits_{i=1}^{N} u_i' A_T u_i}{\sqrt{\frac{1}{N} \sum_{i=1}^{N} (u_i' A_T u_i)^2}} + o_p(1) \ \xrightarrow{d} \ \mathcal{N}\left(c \, \frac{\kappa \, tr(A_T H_T)}{\sqrt{tr(A_T^2 + A_T' A_T)}}, 1 \right) \ .$$

Proof of Theorem 1:

(i) Let D_k denote a $(T-2) \times T$ matrix that results from dropping the k'th row of the first difference matrix D defined in (9). Using Lemma A.1 we have

[7] Note that for non-symmetric A_T the expression can be rewritten in a symmetric format as $u_i' A_T u_i = 0.5[u_i (A_T + A_T') u_i]$.

$N^{-1/2}\sum_{i=1}^{N}\widehat{z}_{Ti} = N^{-1/2}\sum_{i=1}^{N} z_{Ti} + O_p(N^{-1})$, where

$$z_{Ti} \;=\; e_i'(D_{T-1}'D_1 + 0.5D_{T-1}'D_{T-1})e_i \;\equiv\; e_i'A_Te_i \;.$$

For $T \geq 3$

$$A_T = \begin{bmatrix} 0.5 & -0.5 & 0 & 0 & 0 & 0 & \cdots & 0 & 0 & 0 & 0 & 0 & 0 \\ 0.5 & 0 & -0.5 & 0 & 0 & 0 & \cdots & 0 & 0 & 0 & 0 & 0 & 0 \\ -1 & 1.5 & 0 & -0.5 & 0 & 0 & \cdots & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & -1 & 1.5 & 0 & -0.5 & 0 & \cdots & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & -1 & 1.5 & 0 & -0.5 & \cdots & 0 & 0 & 0 & 0 & 0 & 0 \\ \vdots & & & & & \vdots & \cdots & \vdots & & & & & \vdots \\ 0 & 0 & 0 & 0 & 0 & 0 & \cdots & -1 & 1.5 & 0 & -0.5 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & \cdots & 0 & -1 & 1.5 & 0 & -0.5 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & \cdots & 0 & 0 & -1 & 1.5 & -0.5 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & \cdots & 0 & 0 & 0 & -1 & 1 & 0 \end{bmatrix}.$$

For example,

$$\begin{aligned} A_3 &= \begin{bmatrix} 0 \\ -1 \\ 1 \end{bmatrix} \begin{bmatrix} -1 & 1 & 0 \end{bmatrix} + 0.5 \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix} \begin{bmatrix} -1 & 1 & 0 \end{bmatrix} \\ &= \begin{bmatrix} -0.5 \\ -0.5 \\ 1 \end{bmatrix} \begin{bmatrix} -1 & 1 & 0 \end{bmatrix} = \begin{bmatrix} 0.5 & -0.5 & 0 \\ 0.5 & -0.5 & 0 \\ -1 & 1 & 0 \end{bmatrix}. \end{aligned}$$

Obviously, $tr(A_T) = 0$, $A_T\iota_T = 0$ and $\iota_T'A_T = 0$. If $u_{it} \overset{iid}{\sim} N(0,\sigma_i^2)$, the variance of the quadratic form $e_i'A_Te_i = u_i'A_Tu_i$ is $\sigma_i^4 tr(A_T^2 + A_T'A_T)$, where

$$\begin{aligned} tr(A_T^2) &= -\frac{3}{2}(T-3) \\ tr(A_T'A_T) &= 3 + \frac{7(T-3)}{2} \;. \end{aligned}$$

It follows that $var(e_i' A_T e_i) = \sigma_i^4[2(T-3)+3]$. Furthermore,

$$tr(A_T H_T) = T - 2$$

and, by using Lemma 1, we obtain

$$\widetilde{\text{WD}} \xrightarrow{d} \mathcal{N}\left(c\,\frac{\kappa(T-2)}{\sqrt{2(T-3)+3}}, 1\right).$$

(ii) The original version of the WD test can be written as

$$\begin{aligned}
\text{WD} &= \frac{\sum\limits_{i=1}^{N} \widehat{z}_{Ti}}{\sqrt{\sum\limits_{i=1}^{N}\left(\sum\limits_{t=3}^{T} \Delta\widehat{e}_{i,t-1}(\Delta\widehat{e}_{it} - \widehat{\theta}\Delta\widehat{e}_{i,t-1})\right)^2}} \\
&= \frac{\frac{1}{\sqrt{N}}\sum\limits_{i=1}^{N} \widehat{z}_{Ti}}{\sqrt{\frac{1}{N}\sum\limits_{i=1}^{N}\left(\sum\limits_{t=3}^{T} \Delta\widehat{e}_{i,t-1}(\Delta\widehat{e}_{it} + 0.5\Delta\widehat{e}_{i,t-1}) - (\widehat{\theta}+0.5)\Delta\widehat{e}_{t-1}^2\right)^2}} \\
&= \frac{\frac{1}{\sqrt{N}}\sum\limits_{i=1}^{N} \widehat{z}_{Ti}}{\sqrt{\frac{1}{N}\sum\limits_{i=1}^{N}\left(\sum\limits_{t=3}^{T} \Delta\widehat{e}_{i,t-1}(\Delta\widehat{e}_{it} + 0.5\Delta\widehat{e}_{i,t-1})\right)^2}} + O_p(N^{-1/2}) \\
&= \frac{\frac{1}{\sqrt{N}}\sum\limits_{i=1}^{N} \widehat{z}_{Ti}}{\sqrt{\frac{1}{N}\sum\limits_{i=1}^{N} \widehat{z}_{Ti}^2}} + O_p(N^{-1/2}) \ ,
\end{aligned}$$

where $\widehat{\theta} + 0.5 = O_p(N^{-1/2})$ under the null hypothesis and local alternative. Furthermore, by using the results in the proof of Lemma 1:

$$\widetilde{\text{WD}} = \frac{\sum\limits_{i=1}^{N} \widehat{z}_{Ti}}{\sqrt{\sum\limits_{i=1}^{N} \widehat{z}_{Ti}^2 - \frac{1}{N}\left(\sum\limits_{i=1}^{N} \widehat{z}_{Ti}\right)^2}}$$

$$= \frac{\frac{1}{\sqrt{N}} \sum\limits_{i=1}^{N} \widehat{z}_{Ti}}{\sqrt{\frac{1}{N} \sum\limits_{i=1}^{N} \widehat{z}_{Ti}^2}} + O_p(N^{-1/2})$$

It follows that $\text{WD} - \widetilde{\text{WD}} = O_p(N^{-1/2})$.

Proof of Lemma 2:

For the first order autocorrelation, Cox and Solomon (1988) derive the following asymptotic approximation:

$$\sqrt{N}\left(\frac{\sum\limits_{i=1}^{N}\sum\limits_{t=2}^{T}(u_{it} - \overline{u}_i)(u_{i,t-1} - \overline{u}_i)}{\sum\limits_{i=1}^{N}\sum\limits_{t=1}^{T}(u_{it} - \overline{u}_i)^2}\right) \xrightarrow{d} \mathcal{N}\left(-\frac{1}{T}, \frac{(T+1)(T-2)^2}{T^2(T-1)^2}\right) .$$

Using this result, the limiting distribution of the LM statistic is easily derived.

Proof of Theorem 2:

(i) As shown in the proof of Lemma 1 we have $\widehat{z}_{Ti} = e_i' A_T e_i + o_p(1)$. Using the matrix notation introduced in section 3.2, $A_T = M_1' M_T + \frac{1}{T-1} M_T' M_T$. Let S_k denote a selection matrix that is obtained from dropping the k'th row of I_T. Since $M_T \iota_T = S_T M \iota_T = 0$ and $M_1 \iota_T = S_1 M \iota_T = 0$, it follows that $A_T \iota_T = 0$ and $A_T' \iota_T = 0$. Since $tr(M_1' M_T) = tr(M S_1' S_T M) = tr(S_T M S_1') = -(T-1)/T$ and $tr(M_T' M_T) = tr(S_T M S_T') = (T-1) - (T-1)/T$ it follows that $tr(A_T) = 0$. Thus,

$$z_{Ti} = e_i' A_T e_i = u_i' A_T u_i .$$

The variance of z_{Ti} is obtained as $var(z_{Ti}) = \sigma_i^4 tr(A^2 + AA')$ by using the following results:

$$\begin{aligned} tr(M_1' M_T M_1' M_T) &= -(T^2 - 2T - 1)/T^2 \\ tr(M_1' M_T M_T' M_1) &= tr(M_T' M_T M_T' M_T) = (T^3 - 2T^2 + 1)/T^2 \\ tr(M_1' M_T M_T' M_T) &= tr(M_T' M_T M_1' M_T) = tr(M_T' M_T M_T' M_1) = -(T^2 - T - 1)/T^2 \ . \end{aligned}$$

It follows that

$$var(z_{Ti}) = \sigma_i^4 \left(T - 3 + \frac{2}{T(T-1)} \right).$$

Furthermore,

$$tr(M_1' M_T H_T) = tr(S_T' M H_T M S_1) = tr(S_T' H_T S_1) - 2T^{-1} tr(S_T' \iota_T \iota_T' H_T S_1) + T^{-2} tr(S_T' \iota_T \iota_T' H_T \iota_T \iota_T' S_1) \ .$$

Using

$$\begin{aligned} & tr(S_T' H_T S_1) = T - 1 \\ & tr(S_T' H_T S_T) = 0 \\ & tr(S_T' \iota_T \iota_T' H_T S_1) = tr(S_T' \iota_T \iota_T' H_T S_T) = 2(T-1) - 1 \\ & tr(S_T' \iota_T \iota_T' H_T \iota_T \iota_T' S_1) = tr(S_T' \iota_T \iota_T' H_T \iota_T \iota_T' S_T) = 2(T-1)^2 \end{aligned}$$

yields

$$tr(A_T H_T) \ = \ tr \left(M_1' M_T H_T - \frac{1}{T-1} M_1' M_T H_T \right) = T - 3 + \frac{2}{T(T-1)} \ ,$$

which is identical to $tr(A_T^2 + A_T' A_T)$. With these results the limiting distribution follows.

(ii) Using $\widehat{\varrho} + \frac{1}{T-1} = O_p(N^{-1/2})$ and $\overline{\widehat{e}_i} = T^{-1} \sum_{t=1}^{T} \widehat{e}_{it}$ the LM* statistic can be written as

$$
\begin{aligned}
\mathrm{LM}^* &= \frac{\sum\limits_{i=1}^{N} \widehat{z}_{Ti}}{\sqrt{\sum\limits_{i=1}^{N} \left[\sum\limits_{t=2}^{T} (\widehat{e}_{i,t-1} - \overline{\widehat{e}_i}) \Big(\widehat{e}_{it} - \widehat{\varrho} \widehat{e}_{i,t-1} - (1-\widehat{\varrho})\overline{\widehat{e}_i} \Big) \right]^2}} \\
&= \frac{\frac{1}{\sqrt{N}} \sum\limits_{i=1}^{N} \widehat{z}_{Ti}}{\sqrt{\frac{1}{N} \sum\limits_{i=1}^{N} \left[\sum\limits_{t=2}^{T} (\widehat{e}_{it} - \overline{\widehat{e}_i})(\widehat{e}_{i,t-1} - \overline{\widehat{e}_i}) + (T-1)^{-1} \sum\limits_{t=2}^{T} (\widehat{e}_{it} - \overline{\widehat{e}_i})^2 + O_p(N^{-1/2}) \right]}} \\
&= \frac{\frac{1}{\sqrt{N}} \sum\limits_{i=1}^{N} \widehat{z}_{Ti}}{\sqrt{\frac{1}{N} \sum\limits_{i=1}^{N} \widehat{z}_{Ti}^2}} + O_p(N^{-1/2}).
\end{aligned}
$$

Furthermore, under the null hypothesis and local alternative

$$
\widetilde{\mathrm{LM}}^* = \frac{\frac{1}{\sqrt{N}} \sum\limits_{i=1}^{N} z_{Ti}}{\sqrt{\frac{1}{N} \sum\limits_{i=1}^{N} z_{Ti}^2}} + O_p(N^{-1/2}) \; .
$$

Proof of Theorem 3:

By Lemma A.1 we have

$$
N^{-1/2} \sum_{i=1}^{N} \widehat{z}_{Ti} \;=\; N^{-1/2} \sum_{i=1}^{N} e_i' M (D'D - 2I_T) M e_i + O_p(N^{-1}) \;=\; u_i' A_T u_i + O_p(N^{-1}) \; ,
$$

where

$$
D'D - 2I_T = \begin{bmatrix}
-1 & -1 & 0 & 0 & \cdots & 0 & 0 & 0 \\
-1 & 0 & -1 & 0 & \cdots & 0 & 0 & 0 \\
0 & -1 & 0 & -1 & \cdots & 0 & 0 & 0 \\
\vdots & & & & & & & \vdots \\
0 & 0 & 0 & 0 & \cdots & -1 & 0 & -1 \\
0 & 0 & 0 & 0 & \cdots & 0 & -1 & -1
\end{bmatrix}
$$

is a symmetric $T \times T$ matrix. Since $M\iota_T = M'\iota_T = 0$ it follows that $A_T\iota_T = A'_T\iota_T = 0$. Furthermore, $tr[M(D'D-2I_T)M] = tr(D'D-2I_T) - \iota'_T(D'D-2I_T)\iota_T = -2+2 = 0$.

The variance is obtained as

$$\begin{aligned} var(z_{Ti}) &= 2\sigma_i^4 tr(A_T^2) \\ &= 4\sigma_i^4 (T-2) \ . \end{aligned}$$

Furthermore,

$$tr(A_T H_T) = -\frac{T-1}{T} 2(T-2) \ .$$

It follows from Lemma 1 that under the local alternative

$$\text{mDW} \xrightarrow{d} \mathcal{N}\left(c\,\kappa \frac{T-1}{T}\ \sqrt{T-2}, 1\right).$$

Proof of Lemma 3:

Rewrite equation (11) in matrix terms as

$$\widetilde{L}_0 M \widehat{e}_i = \varrho_k \widetilde{L}_k M \widehat{e}_i + \epsilon_i \ ,$$

where

$$\widetilde{L}_0 = \begin{pmatrix} 0 & \cdots & 0 & 1 & 0 & \cdots & 0 \\ 0 & \cdots & 0 & 0 & 1 & & \vdots \\ \vdots & & \vdots & \vdots & & \ddots & 0 \\ 0 & \cdots & 0 & 0 & \cdots & 0 & 1 \end{pmatrix} \quad \widetilde{L}_k = \begin{pmatrix} 1 & 0 & \cdots & 0 & 0 & \cdots & 0 \\ 0 & 1 & & \vdots & 0 & \cdots & 0 \\ \vdots & & \ddots & 0 & \vdots & & \vdots \\ 0 & \cdots & 0 & 1 & 0 & \cdots & 0 \end{pmatrix}$$

are $(T-k)\times T$ matrices. Under Assumption (1) and using the following results:

$$\begin{aligned}
tr(M\widetilde{L}_k'\widetilde{L}_0M) &= tr\left[\left(\widetilde{L}_k' - \frac{1}{T}\iota_T\iota_T'\widetilde{L}_k'\right)\widetilde{L}_0\right] \\
&= tr\left(\widetilde{L}_k'\widetilde{L}_0\right) - \frac{1}{T}tr\left(\iota_T'\widetilde{L}_k'\widetilde{L}_0\iota_T\right) \\
&= -\frac{T-k}{T}\ , \\
tr(M\widetilde{L}_k'\widetilde{L}_kM) &= tr\left[\left(\widetilde{L}_k' - \frac{1}{T}\iota_T\iota_T'\widetilde{L}_k'\right)\widetilde{L}_k\right] \\
&= tr\left(\widetilde{L}_k'\widetilde{L}_k\right) - \frac{1}{T}tr\left(\iota_T'\widetilde{L}_k'\widetilde{L}_k\iota_T\right) \\
&= (T-1)\frac{T-k}{T}\ ,
\end{aligned}$$

it follows as $N\to\infty$:

$$\widehat{\varrho}_k \ \xrightarrow{p}\ \frac{tr(M\widetilde{L}_k'\widetilde{L}_0M)}{tr(M\widetilde{L}_k'\widetilde{L}_kM)} \ = \ = \ -\frac{1}{T-1}\ ,$$

for all $k\in\{1,\dots,T-2\}$.

Table 1: Size and power of alternative tests

T	WD	$\widetilde{\text{WD}}$	LM*	$\widetilde{\text{LM}}^*$	mDW
			$c = 0$		
5	0.050	0.049	0.054	0.052	0.055
10	0.050	0.050	0.054	0.054	0.051
20	0.049	0.049	0.047	0.047	0.045
30	0.052	0.051	0.052	0.052	0.052
50	0.049	0.049	0.047	0.047	0.048
			$c = 0.5$		
5	0.092	0.097	0.111	0.109	0.107
	(0.088)	(0.088)	(0.112)	(0.112)	(0.107)
10	0.169	0.177	0.264	0.263	0.251
	(0.163)	(0.163)	(0.263)	(0.263)	(0.247)
20	0.311	0.320	0.532	0.531	0.509
	(0.316)	(0.316)	(0.541)	(0.541)	(0.522)
30	0.449	0.457	0.735	0.735	0.720
	(0.458)	(0.458)	(0.738)	(0.738)	(0.725)
50	0.672	0.679	0.929	0.929	0.923
	(0.683)	(0.683)	(0.929)	(0.929)	(0.924)
			$c = 1$		
5	0.210	0.219	0.292	0.288	0.282
	(0.205)	(0.205)	(0.305)	(0.305)	(0.283)
10	0.493	0.502	0.751	0.750	0.718
	(0.492)	(0.492)	(0.755)	(0.755)	(0.721)
20	0.833	0.839	0.987	0.987	0.983
	(0.841)	(0.841)	(0.985)	(0.985)	(0.981)
30	0.954	0.955	1.000	1.000	0.999
	(0.960)	(0.960)	(0.999)	(0.999)	(0.999)
50	0.998	0.998	1.000	1.000	1.000
	(0.998)	(0.998)	(1.000)	(1.000)	(1.000)

Note: $N = 500$. Empirical values are computed using 10,000 Monte Carlo simulations. Asymptotic values (given in parentheses) are computed according to the results in Theorems 1 – 3.

Table 2: AR(2) processes

T	WD	$\widetilde{\text{WD}}$	LM*	$\widetilde{\text{LM}}^*$	mDW	$\widetilde{Q}(2)$	IS_k	IS(2)
				$\alpha_1 = 0,\ \alpha_2 = 0$				
5	0.048	0.048	0.049	0.047	0.051	0.052	0.049	0.050
10	0.050	0.049	0.051	0.051	0.049	0.050	0.063	0.053
20	0.050	0.049	0.049	0.049	0.048	0.049	0.067	0.045
30	0.053	0.052	0.055	0.055	0.055	0.053	*na*	0.041
50	0.054	0.054	0.056	0.056	0.055	0.057	*na*	0.036
				$\alpha_1 = 0.03,\ \alpha_2 = -0.03$				
5	0.342	0.353	0.351	0.346	0.286	0.322	0.165	0.178
10	0.745	0.754	0.600	0.598	0.554	0.696	0.212	0.310
20	0.976	0.977	0.862	0.861	0.847	0.962	0.189	0.461
30	0.999	0.999	0.957	0.957	0.951	0.996	*na*	0.579
50	1.000	1.000	0.997	0.997	0.997	1.000	*na*	0.703
				$\alpha_1 = \alpha_2 = 0.03$				
5	0.049	0.048	0.068	0.066	0.073	0.096	0.068	0.081
10	0.048	0.049	0.266	0.264	0.267	0.436	0.134	0.179
20	0.049	0.048	0.687	0.686	0.675	0.913	0.152	0.327
30	0.052	0.052	0.892	0.892	0.888	0.988	*na*	0.450
50	0.054	0.054	0.990	0.990	0.989	1.000	*na*	0.611
				$\alpha_1 = 0,\ \alpha_2 = 0.08$				
5	0.524	0.500	0.376	0.371	0.252	0.591	0.289	0.369
10	0.932	0.926	0.226	0.225	0.175	0.984	0.542	0.775
20	0.999	0.999	0.137	0.137	0.124	1.000	0.586	0.986
30	1.000	1.000	0.106	0.106	0.101	1.000	*na*	0.999
50	1.000	1.000	0.085	0.085	0.081	1.000	*na*	1.000

Note: $N = 500$. Rejection frequencies are computed using 10,000 Monte Carlo replications. Note that the IS test statistics cannot be computed if the degrees of freedom are larger than N (*na*).

Table 3: Size under temporal heteroskedasticity

T	$\widetilde{\text{WD}}$	$\widetilde{\text{LM}}^*$	mDW	HR	$\widetilde{\text{WD}}$	$\widetilde{\text{LM}}^*$	mDW	HR
	break in variance				u-shaped variance			
5	1.000	1.000	1.000	0.049	0.052	0.169	1.000	0.048
10	1.000	0.374	1.000	0.052	0.053	0.119	1.000	0.049
20	0.993	0.081	0.927	0.051	0.053	0.063	1.000	0.051
30	0.905	0.062	0.751	0.050	0.051	0.053	1.000	0.050
50	0.670	0.051	0.504	0.050	0.049	0.050	0.996	0.049
	negative exponential trend (-0.2)				positive exponential trend (0.2)			
5	0.798	0.185	0.080	0.054	0.591	0.122	0.080	0.053
10	0.992	0.125	0.353	0.051	0.931	0.069	0.361	0.049
20	1.000	0.088	0.924	0.049	0.988	0.054	0.922	0.053
30	1.000	0.075	0.993	0.053	0.990	0.051	0.993	0.049
50	1.000	0.057	1.000	0.051	0.990	0.047	1.000	0.049

Note: $N = 500$. Rejection frequencies are computed using 10,000 Monte Carlo replications.