Economics Letters 155 (2017) 104–110

Contents lists available at ScienceDirect

Economics Letters

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

# Tests for serial correlation of unknown form in dynamic least squares regression with wavelets[☆]

Meiyu Li, Ramazan Gençay [*]

*Department of Economics, Simon Fraser University, 8888 University Drive, Burnaby, BC, V5A 1S6, Canada*

ARTICLE INFO

*Article history:*
Received 12 March 2017
Accepted 18 March 2017
Available online 29 March 2017

*JEL classification:*
C1
C2
C12
C22
C26

*Keywords:*
Dynamic least squares regression
Serial correlation
Conditional heteroscedasticity
Maximum overlap discrete wavelet transformation

ABSTRACT

This paper extends the multi-scale serial correlation tests of Gençay and Signori (2015) for observable time series to unobservable errors of unknown forms in a linear dynamic regression model. Our tests directly build on the variance ratio of the sum of squared wavelet coefficients of residuals over the sum of squared residuals, utilizing the equal contribution of each frequency of a white noise process to its variance and delivering higher empirical power than parametric tests. Our test statistics converge to the standard normal distribution at the parametric rate under the null hypothesis, faster than the nonparametric test using kernel estimators of the spectrum.



## 1. Introduction

Existing residual based serial correlation tests for dynamic models fall into two categories, the parametric tests in the time domain and the nonparametric tests in the frequency domain. The residual based parametric tests estimate autocorrelation coefficients directly, delivering easy implementation and desirable finite sample performance (Hayashi, 2000; Box and Pierce, 1970; Godfrey, 1978; Wooldridge, 1990, 1991), whereas their power relies on the choice of lag length heavily. In contrast, the residual based nonparametric tests compare the kernel-based spectrum estimator with the white noise spectrum and have more power (Hong, 1996; Hong and Lee, 2003). Nevertheless, the nonparametric estimation of the underlying spectrum sacrifices the convergence rate, which is slower than the parametric rate and often associated with poor finite sample performance (Chen and Deo, 2004a, b).

Based on the fact of the equal contribution of each frequency of a white noise process to its variance, Gençay and Signori (2015) proposes a family of wavelet-based serial correlation tests for observable time series, *GS* and *GSM*, utilizing only wavelet coefficients of the observed time series data. This paper extends the *GS* and *GSM* tests into a linear regression framework, particularly when lagged dependent variables are included in regression equations. Our modified tests are robust to models with conditionally heteroscedastic errors of unknown form. Inheriting the test design of Gençay and Signori (2015) by using the additive variance decomposition of the wavelet and the scaling coefficients, instead of any nonparametric estimation of the underlying spectrum, our test statistics converge to the normal distribution at the parametric rate under the null hypothesis (faster than the nonparametric test) and display higher power than the parametric test. In addition, contrary to the sensitiveness of finite sample performance to the choice of lag length for the parametric test, and the choice of bandwidth for the nonparametric test, our tests are rather stable when different wavelet decomposition levels are utilized.

This paper proceeds as follows. Section 2 illustrates the modeling environment for our residual-based tests. Section 3 proposes The corresponding modified test statistics. Section 4 reports comprehensive Monte Carlo simulations for finite sample performance of our tests in comparison to commonly used tests. Section 5 contains concluding remarks of this paper.

☆ We gratefully acknowledge Bertille Antoine for providing helpful comments and suggestions. All errors are attributed to the authors.

* Corresponding author.
*E-mail addresses:* meiyul@sfu.ca (M. Li), rgencay@sfu.ca (R. Gençay).

http://dx.doi.org/10.1016/j.econlet.2017.03.021

## 2. Model setting

Consider the dynamic regression model

$$y_t = Y_t'\alpha + W_t'\gamma + u_t = X_t'\beta + u_t \quad (1)$$

where: $Y_t' = (y_{t-1}, \ldots, y_{t-P})$ with $P \geq 1$; $\alpha' = (\alpha_1, \ldots, \alpha_P)$ having the values such that the roots of $z^P - \alpha_1 z^{P-1} - \cdots - \alpha_P = 0$ are all strictly inside the unit circle; $W_t$ is an $L$-vector of regressors which may include lagged exogenous variables and constants; $X_t' \equiv (Y_t', W_t')$; and $\beta' \equiv (\alpha', \gamma')$. Let $K = P + L$ denotes the number of regression coefficients, and $T$ denotes the number of observations available for estimation. All of the data generation processes are subject to the following two assumptions.

**Assumption 2.1** (*Ergodic Stationary*). The $(K+1)$ dimensional vector stochastic process $\{y_t, X_t\}$ is jointly stationary and ergodic.

**Assumption 2.2** (*Rank Condition*). The $K \times K$ matrix $E(X_t X_t')$ is non-singular (and hence finite), which is denoted by $\Sigma_{XX}$.

Assumption 2.1 implies the stationarity and ergodicity of error process, $\{u_t\}$. Nevertheless, the unconditional homoscedasticity does not exclude the models with a conditional heteroscedastic error term. To be specific, for any stationary and ergodic process, Assumption 2.1 allows any form of unconditional higher order moment dependence which disappears as time lag increases, and any form of time varying conditional higher order moments. We consider tests for the following null hypothesis:

**Assumption 2.3** (*Stronger form of Predeterminedness*). All regressors are predetermined such that: $E(u_t | u_{t-1}, u_{t-2}, \ldots, X_t, X_{t-1}, \ldots) = 0$.[1]

Assumption 2.3 is stronger than the assumption of white noise and needed for the derivation of the null distribution of our test statistic. If Assumptions 2.1, 2.2 and 2.3 are satisfied, the OLS coefficient estimator in Eq. (1) is consistent, namely $\hat{\beta} \xrightarrow{p} \beta$ (Hayashi, 2000 Proposition 2.1(a), page 124). Throughout the paper, we assume Assumptions 2.1 and 2.2 hold and impose periodic boundary conditions on all time series, like $\{u_t\}$, where $u_t \equiv u_{t \bmod T}$.[2]

## 3. Asymptotic null distribution

Throughout this paper, all the statistics constructed with wavelets are based on the Maximum Overlap Discrete Wavelet Transform (MODWT) filter.[3] Define the $m$th level of wavelet coefficients as $w_{m,t} = \sum_{l=0}^{L_m-1} \tilde{h}_{m,l} u_{t-l \bmod T}$, where $L_m := (2^m - 1)(L-1) + 1$, $L$ is the length of the initial MODWT filter, and $\{\tilde{h}_{m,l}\}$ is the $m$th level MODWT filter.

Assumption 2.3 indicates that $\text{cov}(X_{t-j}, u_t) = 0$, while it allows $\text{cov}(X_{t+j}, u_t) \neq 0$ for some $j > 0$. The correlation between the future lagged dependent variable and current error term makes the asymptotic variance estimator in the *GS* and *GSM* tests inconsistent when the realizations of error terms $\{u_t\}$ are replaced with residuals $\{\hat{u}_t\}$ directly. As a result, the test statistics of Gençay and Signori (2015) distribute more tightly around zero than a standard normal distribution in dynamic models, more difficult to reject under the null hypothesis and less powerful. For dynamic linear models, we modify the statistics to restore the asymptotic distributions and improve efficiency in small samples.

**Assumption 3.1** (*Finite Fourth Moments for Regressors*). $E(u_t^2 u_{t-j} u_{t-k})$, $E(u_t u_{t-j} u_{t-k} x_{t-l,i})$, $E(u_t u_{t-j} x_{t-k,i} x_{t-l,n})$, $E(u_{t-j} x_{t,i} x_{t-k,n} x_{t-l,s})$, $E(x_{t,i} x_{t,n} x_{t-j,s} x_{t-k,v})$ exist and are finite for all $j, k, l = 0, 2, \ldots, L_m - 1$ and $i, n, s, v = 1, 2, \ldots, K$.

**Theorem 3.1.** *If Assumptions 2.1, 2.2 and 2.3 are satisfied, then the sample analogue $\hat{\mathcal{E}}_{m,T}$ constructed with residuals $\hat{u}_t$ from the model in Eq. (1) is a consistent estimator of the wavelet variance ratio $\mathcal{E}_m$*

$$\hat{\mathcal{E}}_{m,T} = \frac{\text{wvar}_m(\hat{u})}{\text{var}(\hat{u})} = \frac{\sum_{t=1}^T \hat{w}_{m,t}^2}{\sum_{t=1}^T \hat{u}_t^2} \xrightarrow{p} \frac{1}{2^m}.$$

*Further, if Assumption 3.1 is satisfied, then*

$$LG_m = \sqrt{\frac{Ts^4}{4H_m' \hat{C}_m' \hat{\Psi}_m \hat{C}_m H_m}} \left( \hat{\mathcal{E}}_{m,T} - \frac{1}{2^m} \right) \xrightarrow{d} N(0,1),$$

*where $\{\tilde{h}_{m,l}\}$ is the wavelet filter used in the construction of $\hat{\mathcal{E}}_{m,T}$ and*

$$\underset{((L_m-1)\times 1)}{H_m} \equiv \begin{bmatrix} \tilde{h}_{m,1} & \tilde{h}_{m,2} & \tilde{h}_{m,3} & \cdots & \tilde{h}_{m,L_m-1} \\ \tilde{h}_{m,2} & \tilde{h}_{m,3} & \cdots & \tilde{h}_{m,L_m-1} & 0 \\ \tilde{h}_{m,3} & \cdots & \tilde{h}_{m,L_m-1} & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \tilde{h}_{m,L_m-1} & 0 & 0 & \cdots & 0 \end{bmatrix} \times \begin{bmatrix} \tilde{h}_{m,0} \\ \tilde{h}_{m,1} \\ \tilde{h}_{m,2} \\ \vdots \\ \tilde{h}_{m,L_m-2} \end{bmatrix},$$

$$\underset{((L_m-1)(K+1)\times(L_m-1))}{\hat{C}_m} \equiv \begin{bmatrix} \hat{c}_1 & & \\ & \ddots & \\ & & \hat{c}_{L_m-1} \end{bmatrix}, \quad \underset{((K+1)\times 1)}{\hat{c}_j} \equiv \begin{bmatrix} 1 \\ -S_{XX}^{-1}\bar{\mu}_j \end{bmatrix},$$

$$\text{the } (j,k) \text{ block of } \underset{((K+1)\times(K+1))}{\hat{\Psi}_m} \equiv \begin{bmatrix} \frac{1}{T}\sum_{t=1}^T \hat{u}_t^2 \hat{u}_{t-j} \hat{u}_{t-k} & \frac{1}{T}\sum_{t=1}^T X_t' \hat{u}_t^2 \hat{u}_{t-j} \\ \frac{1}{T}\sum_{t=1}^T X_t \hat{u}_t^2 \hat{u}_{t-k} & \frac{1}{T}\sum_{t=1}^T \hat{u}_t^2 X_t X_t' \end{bmatrix},$$

$$s^2 \equiv \frac{1}{T-K}\sum_{t=1}^T \hat{u}_t^2, \; S_{XX} = \frac{1}{T}\sum_{t=1}^T X_t X_t', \; \bar{\mu}_j \equiv \frac{1}{T}\sum_{t=1}^T X_t \hat{u}_{t-j}.$$

Combining these multiple-scale tests to gain power against a wide range of alternatives, we derive the asymptotic joint distribution of these tests as follows.

**Theorem 3.2.** *If Assumptions 2.1, 2.2, 2.3 and 3.1 are satisfied, then the vector*

$$\sqrt{\frac{Ts^4}{4}} \left( \hat{\mathcal{E}}_{1,T} - \frac{1}{2^1}, \hat{\mathcal{E}}_{2,T} - \frac{1}{2^2}, \ldots, \hat{\mathcal{E}}_{N,T} - \frac{1}{2^N} \right)' \xrightarrow{d} N(0, \Upsilon' \Psi_N \Upsilon),$$

*where the mth column of $\underset{((L_N-1)(K+1)\times N)}{\Upsilon}$ is $\underset{((L_m-1)(K+1)\times 1)}{C_m H_m}$ followed by $\underset{((L_N-L_m)(K+1)\times 1)}{0}$.*

[1] Assumption 2.3 is a sufficient condition for $g_{jt} \equiv [u_{t-j}u_t, u_t X_t']'$ to be a martingale difference sequence, stronger than we need. We maintain this stronger assumption for the convenience of interpretation.

[2] The notation $a - b \bmod T$ stands for '$a - b$ modulo $T$'. If $j$ is an integer such that $1 \leq j \leq T$, then $j \bmod T \equiv j$. If $j$ is another integer, then $j \bmod T \equiv j + nT$ where $nT$ is the unique integer multiple of $T$ such that $1 \leq j + nT \leq T$. The periodic boundary conditions have trivial impact on the finite sample distribution due to the short length of most discrete wavelet filters.

[3] For details of MODWT filter, see Gençay et al. (2001) and Percival and Walden (2000).