Higher order expansions for error variance matrix estimates in the Gaussian AR(1) linear regression model

Karavias, Yiannis; Symeonides, Spyridon; Tzavalis, Elias

DOI:
10.1016/j.spl.2017.11.016

Document Version
Peer reviewed version

Citation for published version (Harvard):
Karavias, Y, Symeonides, S & Tzavalis, E 2018, 'Higher order expansions for error variance matrix estimates in the Gaussian AR(1) linear regression model', *Statistics and Probability Letters*, vol. 135, pp. 54-59. https://doi.org/10.1016/j.spl.2017.11.016

Higher Order Expansions for Error Variance Matrix Estimates in the Gaussian AR(1) Linear Regression Model

Yiannis Karavias* Spyridon D. Symeonides† Elias Tzavalis‡

November 27, 2017

Abstract

We derive a stochastic expansion of the error variance-covariance matrix estimator for the linear regression model under Gaussian AR(1) errors. The higher order accuracy terms of the refined formula are not directly derived from formal Edgeworth-type expansions but instead, the paper adopts Magadalinos' (1992) stochastic order of ω which is a convenient device to obtain the equivalent relation between the stochastic expansion and the asymptotic approximation of corresponding distribution functions. A Monte Carlo experiment compares tests based on the new estimator with others in the literature and shows that the new tests perform well.

Key words: Linear regression; AR(1) disturbances; stochastic expansions; asymptotic approximations; autocorrelation robust inference

The authors would like to thank Stelios Arvanitis, Anindya Banerjee, Liudas Giraitis, Menelaos Karanasos, Michalis Stamatogiannis and seminar participants at the University of Bath, the University of Birmingham, Brunel University London and Queen Mary University of London.

1 Introduction

When linear hypotheses are to be tested in the linear regression model with autocorrelated disturbances, the OLS estimators of the regression parameters need to be standardized by appropriate standard errors, which perform well only in very large samples. In this note

*Corresponding Author, Department of Economics, University of Birmingham. Address: JG Smith Building, University of Birmingham, Edgbaston, Birmingham B15 2TT, UK. *Email*: i.karavias@birmingham.ac.uk

†Department of Economics, University of Ioannina. *Email*: ssymeoni@cc.uoi.gr

‡Department of Economics, Athens University of Economics and Business. *Email*: etzavalis@aueb.gr

we propose estimators of the error variance which perform well in small samples, in OLS-estimated, single-equation econometric specifications with Gaussian AR(1) disturbances and no lagged dependent regressors.[1] These corrections aim at improving the finite sample properties of the t and F tests.

If the researcher can soundly hypothesize on the true autocorrelation scheme, the use of GLS estimators and of the conventional t and F testing procedures, provides a solution, under the implicit assumption that the number of observations is large enough to permit the normal and chi-square approximations, respectively, to the true distributions of the corresponding test statistics. However, since, in finite samples, the standard t and F tests remain oversized, refined asymptotic techniques have been proposed to correct the finite-sample size of these tests. Thus, Rothenberg (1984) suggested the use of Edgeworth expansions in terms of the chi-square and normal distributions to derive general formulae of corrected critical values for the Wald (or F) and t statistics, respectively. Edgeworth expansions have been previously used in dealing with autocorrelation but only in models with no explanatory variables, see Giraitis and Phillips (2012) and Velasco and Robinson (2001). Alternatively, for the linear regression model with first-order autocorrelated disturbances, Magdalinos and Symeonides (1995) suggested the use of degrees-of-freedom-adjusted Cornish-Fisher corrected t and F statistics, rather than the Edgeworth corrections of the critical values.

This paper considers an alternative approach; given that the autocorrelation process is of the AR(1) type, we develop a refined-asymptotics, second order, autocorrelation-consistent estimator of the true variance matrix of the OLS parameter estimator. We suggest the use of a second order asymptotic approximation to the true disturbance variance matrix in order to derive a better, more accurate, finite-sample estimate of the OLS-estimator variance matrix. The higher order accuracy terms of the refined formula are not directly derived from formal Edgeworth-type expansions, which are challenging but instead, the paper adopts Magdalinos'

[1] The recent paper by Muller (2014) renewed interest in manipulating specific autocorrelation schemes in regression hypothesis testing. In that paper, the errors were assumed to be Gaussian AR(1); an assumption frequently encountered in the literature.

(1992) stochastic order of ω which can be regarded as a convenient device to obtain the equivalent relation between the stochastic expansion and the asymptotic approximation of corresponding distribution functions.[2] By using this estimate, we can calculate t and F test statistics with better finite-sample distributional properties. We prefer to use OLS over feasible GLS in small samples because the latter is biased. The new estimator can be seen as a refined parametric estimator of Den Haan and Levin (1994) when the errors are Gaussian AR(1) errors.

In an extensive Monte Carlo experiment, we show that the new estimator leads to tests with correct size and good power when serial correlation is strong. The experiment further compares the performance of some standard and some newly proposed estimators in terms of size and power of t-tests, including those proposed by Rothenberg (1984), Andrews (1991), Newey and West (1994), Kiefer, Vogelsang and Bunzel (2000), Goncalves and Vogelsang (2011) and Muller (2014). Because there is no uniformly most powerful test in our problem, it is unavoidable that different tests will perform better at various subsets of the parameter space. We find that the t-tests based on the new error variance estimator have good properties in many cases; for relatively strong autocorrelation they offer better size control and more power.

The paper is organized as follows. Section 2 provides some preliminary notations and the assumptions needed in our expansions. Section 3 develops the first-order asymptotic approximation of the true error variance matrix and gives the analytic first-order estimation of the OLS-estimator variance matrix. A Monte Carlo evaluation of the suggested size corrections is reported in Section 3. Section 4 concludes. The Appendix collects all proofs. Lemmas needed in the proofs are included in the Supplementary Material.

A word on notation. For any matrix X with T rows, $P_X = X(X'X)^{-1}X'$ and $M_X = I_T - P_X = I_T - X(X'X)^{-1}X'$. Further, for any stochastic quantity (scalar, vector, or matrix) the symbols $E(\cdot)$ and $var(\cdot)$ denote the expectation and variance operators, respectively, and

[2] This approach was also used by Symeonides et al. (2016) in deriving tests for SUR models.

the stochastic order $\omega(\cdot)$, defined in the Appendix, has the same operational properties as order $O(\cdot)$. Finally, for any sample size T, the "asymptotic scale" of our expansions is denoted as $\tau = 1/\sqrt{T}$.

2 Model and Assumptions

Consider the single-equation econometric specification

$$y = X\beta + \sigma u, \tag{1}$$

where y is the $T \times 1$ vector of observations on the dependent variable, X is the $T \times k$ matrix of observations on a set of k non-stochastic regressors, β is a $k \times 1$ vector of unknown structural parameters, and σu is the $T \times 1$ vector of non-observable stochastic disturbances with variance-covariance matrix $\sigma^2\Omega$. The elements of the stochastic vector u are assumed to be generated by a stationary first-order autoregressive (AR(1)) stochastic process, $u_t = \rho u_{t-1} + \epsilon_t, \quad (t = 1, \ldots, T)$,where $0 \leq |\rho| < 1$, $u_0 \sim N(0, 1/(1-\rho^2))$, and ϵ_t are independent normal variables with variance σ^2 $(\sigma > 0)$. The random vector u is distributed as $N(0, \Omega)$ with, $\Omega = R/(1-\rho^2)$, where $R = [(\rho^{|t-t'|})_{t,t'=1,\ldots,T}]$.[3]

Let $\hat{\rho}$ be a consistent estimator of ρ. For any function $f = f(\rho)$, we write $\hat{f} = f(\hat{\rho})$. The ordinary least squares (LS) estimators of β and σ^2 are $\hat{\beta}_{LS} = (X'X)^{-1}X'y$, and $\hat{\sigma}^2_{LS} = (y - X\hat{\beta}_{LS})'(y - X\hat{\beta}_{LS})/(T-k)$.

In this paper we provide a novel, refined estimator of Ω. To derive the first-order approximation of matrix Ω, we start by finding the first-order approximation of Ω^{-1}, based on a number of assumptions concerning its elements. To this end, we denote as Ω^{-1}_{ρ} the $T \times T$ matrix of first-order derivatives of the elements of matrix Ω^{-1} with respect to the autocorrelation coefficient ρ. For any estimator $\hat{\rho}$, we define the scalar $\delta_\rho = \frac{1}{\tau}(\hat{\rho} - \rho)$,which

[3]The restrictive assumptions of AR(1) and Gaussian errors can be relaxed. Deriving the necessary estimators for this case however, is computationally burdensome. The extension to AR(p) errors is straightforward and the extension to non-Gaussian distributions can be done along the lines of Phillips (1980).

can be interpreted as a "measure" of the sampling error of $\hat{\rho}$.

Following Magdalinos (1992, page 3444), let the open interval $J = (0, 1)$ be a set of indexes. For any collection of stochastic quantities (scalars, vectors, or matrices) Y_τ $(\tau \in J)$, we write $Y_\tau = \omega(\tau^i)$, if for any given $n > 0$, there exists a $0 < \epsilon < \infty$ such that

$$\Pr\left[\|Y_\tau/\tau^i\| > (-\ln \tau)^\epsilon\right] = o(\tau^n), \quad \text{as} \quad \tau \to 0, \tag{2}$$

where the $\|\cdot\|$ is the Euclidean norm. Further, if (2) is valid for any $n > 0$, we write $Y_\tau = \tau(\infty)$. To justify the use of order $\omega(\cdot)$, notice that if two stochastic quantities differ by a quantity of order $\omega(\tau^i)$, then, under general conditions, the distribution function of the first provides an asymptotic approximation to the distribution function of the second, with an error of order $O(\cdot)$. Also, notice that orders $\omega(\cdot)$ and $O(\cdot)$ have similar operational properties.

The first-order approximation of matrix Ω can be estimated when the following assumptions hold:

(i) The elements of Ω and Ω^{-1} are bounded for all T and all $-1 < \rho < 1$, and matrices

$$A_\Omega = \frac{1}{T} X'\Omega X, \quad A_{\Omega^{-1}} = \frac{1}{T} X'\Omega^{-1} X, \quad F = \frac{1}{T} X'X, \tag{3}$$

converge to non-singular limits, as $T \to \infty$.

(ii) Up to the fourth order, the partial derivatives of the elements of Ω^{-1} with respect to ρ are bounded for all T and all $-1 < \rho < 1$.

(iii) The estimator $\hat{\rho}$ is an even function of u, and it is functionally unrelated to the parameter vector β, i.e., it can be written as a function of X and σu only.

(iv) The nuisance parameter δ_ρ admits a stochastic expansion of the form $\delta_\rho = d_{1\rho} + \tau d_{2\rho} + \omega(\tau^2)$,and the expectation $E(\sqrt{T} d_{1\rho} + d_{2\rho})$ exists and has finite limit, as $T \to \infty$.

The first two assumptions imply that matrix

$$\frac{1}{T}X'\Omega\Omega_{\rho}^{-1}\Omega X \tag{4}$$

is bounded. Moreover, since the autocorrelation coefficient is functionally unrelated to the regression parameters, assumption (iii) is satisfied for a wide class of estimators of ρ, which includes the maximum likelihood estimator and the simple or iterative estimators based on the regression residuals (see Breusch (1980)). Note that we do not need to assume that the estimator $\hat{\rho}$ is asymptotically efficient. Moreover, assumptions (i)–(iv) are satisfied by all the estimators of ρ, considered in the paper, which are the least squares (LS), Durbin-Watson (DW), generalized least squares (GL), Prais-Winsten (PW) and maximum likelihood (ML) estimators.[4]

Finally, for any estimator $\hat{\rho}_I$, indexed by I=$\{LS,\ DW,\ GLS,\ PW,\ ML\}$, we define the scalar $\kappa_{\rho}^{I} = \lim_{T\to\infty} E(\sqrt{T}d_{1\rho}+d_{2\rho})$, which can be interpreted as a "measure" of the accuracy of the expansion of estimator $\hat{\rho}_I$ around its true value.

3 Main results

In this section, we present the main theorems which provide the asymptotic expansions of $\hat{\Omega}$ and $\hat{\Omega}^{-1}$, as well as the first-order approximations the variance of the LS estimator $\hat{\beta}_{LS}$.

[4]The closed forms of these estimators of ρ are given as follows:

(i) LS: $\hat{\rho}_{LS} = \left(\sum_{t=1}^{T}\tilde{u}_t^2\right)^{-1}\left(\sum_{t=2}^{T}\tilde{u}_t\tilde{u}_{t-1}\right)$,where $\tilde{u}_t$ are the LS residuals of regression model (1). (ii) DW: $\hat{\rho}_{DW} = 1-(DW/2)$,where the DW is the Durbin-Watson statistic. (iii) GLS: $\hat{\rho}_{GL} = \left(\sum_{t=1}^{T}\hat{u}_t^2\right)\left(\sum_{t=2}^{T}\hat{u}_t\hat{u}_{t-1}\right)$, where $\hat{u}_t$ denote the GLS estimates of u_t, based on the autocorrelation-correction of regression model (1), using any asymptotically efficient estimator of ρ. (iv) PW: This estimator of ρ, denoted as $\hat{\rho}_{PW}$, together with the PW estimator of β, denoted as $\hat{\beta}_{PW}$, minimize the sum of squared GLS residuals (see Prais and Winsten (1954)). (v) ML: This estimator, denoted as $\hat{\rho}_{ML}$, satisfies a cubic equation with coefficients defined in terms of the ML residuals (see Beach and MacKinnon (1978)).

3.1 Asymptotic expansions of $\hat{\Omega}^{-1}$ and $\hat{\Omega}$

In this section, we derive the first-order approximation of matrix Ω based on the following asymptotic expansions of matrices $\hat{\Omega}^{-1}$ and $\hat{\Omega}$:

Theorem 1 *Any consistent estimator $\hat{\Omega}^{-1}$ of matrix Ω^{-1} admits a stochastic expansion of the form $\hat{\Omega}^{-1} = \Omega^{-1} + \tau\Omega_\rho^{-1}\delta_\rho + \omega(\tau^2)$.*

Theorem 2 *Any consistent estimator $\hat{\Omega}$ of matrix Ω admits a stochastic expansion of the form $\hat{\Omega} = \Omega - \tau\Omega\Omega_\rho^{-1}\Omega\delta_\rho + \omega(\tau^2)$.*

Define the scalars $c_1 = tr(A_{\Omega^{-1}}F^{-1}\Gamma F^{-1}) + (1-\rho^2)tr(F^{-1}\Gamma)$, and $c_2 = (1-\rho^2)tr(FG_{\Omega^{-1}})$, where $G_{\Omega^{-1}} = A_{\Omega^{-1}}^{-1}$ and $\Gamma = (X'RX)/T$ are $k \times k$ matrices. The first-order approximation of matrix Ω is derived in the next theorem.

Theorem 3 *According to the estimator $\hat{\rho}_I$ used, the first-order approximation of matrix Ω can be written as*

$$\Omega = \hat{\Omega} + \tau\hat{\Omega}\hat{\Omega}_\rho^{-1}\hat{\Omega}\kappa_\rho^I + \omega(\tau^2), \tag{5}$$

where $\kappa_\rho^{LS} = -\left[(k+3)\rho + \frac{(c_1-2k)}{2\rho}\right]$, $\kappa_\rho^{DW} = \kappa_\rho^{LS} + 1$, $\kappa_\rho^{GL} = \kappa_\rho^{PW} = \kappa_\rho^{LS} - \frac{(1-\rho^2)c_2}{2\rho} + \frac{(c_1-(1-\rho^2)k)}{2\rho}$ *and* $\kappa_\rho^{ML} = \kappa_\rho^{GL} + \rho$.

In the above expression, $\hat{\Omega} = \hat{R}/(1-\hat{\rho}_I^2)$, $\hat{R} = [(\hat{\rho}_I^{|t-t'|})_{t,t'=1,\ldots,T}]$, and $\hat{\Omega}_\rho^{-1} = 2\hat{\rho}_I I_T - D - 2\hat{\rho}_I\Delta$, with Δ a $T \times T$ matrix with 1 in $(1,1)$-st nd (T,T)-th positions and 0's elsewhere. Also, D is the $T \times T$ band matrix whose (t,t')-th element is equal to 1 if $|t-t'| = 1$ and 0 otherwise and I_T is the $T \times T$ identity matrix. In the above expressions, $\hat{\Omega}$ and $\hat{R}$ are functions of $\hat{\rho}_I$ and thus depend on I but we suppress this dependence to avoid burdensome notation.

3.2 Example: First order approximation of $var(\hat{\beta}_{LS})$

Let $S = var(\hat{\beta}_{LS})$ be the variance matrix of the least squares estimator of β. Since the disturbances in model (1) are AR(1)-autocorrelated, the true variance matrix of the estimator $\hat{\beta}_{LS}$ is $S = \sigma^2(X'X)^{-1}X'\Omega X(X'X)^{-1}$.We suggest employing the first-order approximation of Ω given by (5). This will give the following first-order approximation of matrix S, derived in the next theorem.

Proposition 4 *For all $\hat{\rho}_I$, the first-order approximation of matrix S can be written as $S = \hat{S} + \tau\hat{S}_\rho\kappa_\rho^I + \omega(\tau^2)$,where*

$$\hat{S}_\rho = \hat{\sigma}_{LS}^2(X'X)^{-1}X'\hat{\Omega}\hat{\Omega}_\rho^{-1}\hat{\Omega}X(X'X)^{-1}. \tag{6}$$

Note that matrix S_ρ is bounded. This is straightforward, since (3), (4), and (6) imply that

$$\hat{S}_\rho = \frac{\hat{\sigma}_{LS}^2}{T}\left(\frac{X'X}{T}\right)^{-1}\left(\frac{X'\hat{\Omega}\hat{\Omega}_\rho^{-1}\hat{\Omega}X}{T}\right)\left(\frac{X'X}{T}\right)^{-1} = \frac{\hat{\sigma}_{LS}^2}{T}F^{-1}\left(\frac{X'\hat{\Omega}\hat{\Omega}_\rho^{-1}\hat{\Omega}X}{T}\right)F^{-1}, \tag{7}$$

i.e., $\hat{S}_\rho$ is a function of bounded matrices.

4 Simulation

In this section, we examine the small sample performance of the proposed refined estimator in a t-test framework like the one presented in Section 3.2 (henceforth denoted by BC from "bias corrected"), by conducting a Monte Carlo study. We choose to compare methods using t-testing because of the attention it has received from the literature (see e.g. Muller (2014)). We compare our new procedure with extant ones from the literature, including those proposed by Rothenberg (1984) (denoted by R), Andrews (1991) (A), Newey and West (1994) (NW), Den Haan and Levin (1994) (VARHAC), Kiefer, Vogelsang and Bunzel

(2000) (KVB), the naive bootstrap of Goncalves and Vogelsang (2011) (GV) and Muller (2014) (MUL). For completeness we have also included the t-statistic based on the OLS estimator. We denote t-tests based on the above estimators as t_{BC}, t_{MUL}, t_R, etc.

All experiments are conducted based on 5000 iterations and consider sample sizes of $T = \{15, 30, 60\}$ observations, respectively. The data generating process is given as $y_t = \beta_0 x_{0t} + \beta_1 x_{1t} + \beta_2 x_{2t} + u_t$, where $u_t = \rho u_{t-1} + e_t$ and where $\rho = \{0.6, 0.7, 0.8, 0.9\}$, $e_t \sim N(0,1)$, $y_1 \sim N(0, (1-\rho^2)^{-1})$ and $\beta_0 = \beta_1 = \beta_2 = 1$. For the regressors, we assume that $x_{0t} = 1$, and $x_{jt} = a^{1/2}\zeta_{1t} + (1-a)^{1/2}\zeta_{jt}$, for $j = 1, 2$, where ζ_{jt} follow an $AR(1)$ model with parameter ρ and $N(0,1)$ errors. Finally, $a = \{0.5\}$ determines the correlation between regressors x_{jt}.[5]

The hypothesis of interest tested in our simulation exercise is H_0: $\beta_1 = 1$ against its one sided alternative H_1: $\beta_1 > 1$. In the simulations the alternative is set to $\beta = 1.2$. The method of Andrews (1991) is implemented by assuming an AR(1) model and using both the quadratic spectral kernel (A-QS) and the Bartlett kernel (A-B). For the t_{NW} test, we choose the value $0.75T^{1/3}$ and the Bartlett kernel. We use the S_{12} version of Muller's (2014) test, as we found that this version is almost equally powerful to versions S_{24} and S_{48} and is not size distorted for small samples, as the latter. For the bootstrap procedures we draw 499 bootstrap samples and we employ blocks of length 3 for samples with 15 observations, and blocks of length 5 for the other cases.

Table 1 presents the size of t-tests. Overall performance deteriorates as persistence increases. The t_{NW}, t_R, t_{A-B}, t_{A-QS} and t_{GV} tests are oversized everywhere in the table. The t_{MUL} and t_{KVB} perform better. For $T = \{15, 30\}$ and $\rho = \{0.6, 0.7, 0.8\}$ the new t_{BC} tests have overall the best size. When $T = 60$ tests based on t_{VARHAC} have a smaller size. From the rest of the tests t_{MUL} has size very close to the nominal. For $\rho = 0.9$ it surpasses t_{BC} and t_{VARHAC}. Table 2 presents the size-adjusted power of the t-tests. The most powerful test is t_R which is based on the asymptotically efficient feasible GLS. The tests that had good size properties are also the least powerful. For $T = \{15, 30\}$ and $\rho = \{0.6, 0.7, 0.8\}$ the new

[5] Note that in our simulation exercise we have also tried different values of a like $a = \{0.1, 0.9\}$, but these do not change the results. These results are available upon request.

t_{BC} tests have more power than t_{MUL}, sometimes almost twice as much. The t_{VARHAC} tests have more power than the t_{BC} ones.

To evaluate the robustness of our method we provide, in the supplementary material, the results of two other experiments, in which the errors are AR(2) and ARMA(1,1). We find that the t_{BC} is robust to such departures from the AR(1) case. These results show that our methodology is useful for applied work and warrant the extension to AR(p) and non-Gaussian errors. This is left for future work.

5 Appendix

In this appendix, we provide the proofs of the theorems presented in the main text. These proofs rely on a number of lemmas, which are proved in the Supplementary material.

The following proofs are based on the expansion of $\hat{\Omega}^{-1}$ around Ω^{-1} and the asymptotic expansion of the scalar δ_ρ, defined by $\delta_\rho = \frac{1}{\tau}(\hat{\rho} - \rho)$.

Proof of Theorem 1: Using Corollary 2 of Magdalinos (1992) and the definition of δ_ρ, we can write $\hat{\Omega}^{-1} = \Omega^{-1} + \frac{\partial \Omega^{-1}}{\partial \rho}(\hat{\rho} - \rho) + \omega(\tau^2)$

$$= \Omega^{-1} + \tau \Omega_\rho^{-1} \frac{(\hat{\rho} - \rho)}{\tau} + \omega(\tau^2) = \Omega^{-1} + \tau \Omega_\rho^{-1} \delta_\rho + \omega(\tau^2). \tag{8}$$

Proof of Theorem 2: Re-write the expansion of matrix $\hat{\Omega}^{-1}$, given by Theorem 1, as follows:

$$\hat{\Omega}^{-1} = \Omega^{-1} + \tau \Omega_*^{-1}, \text{ where } \Omega_*^{-1} = \Omega_\rho^{-1} \delta_\rho + \omega(\tau). \tag{9}$$

Using Magdalinos' (1992) Corollary 1, we can derive $\hat{\Omega} = \left[\Omega^{-1} + \tau \Omega_*^{-1}\right]^{-1} = \Omega - \tau \Omega \Omega_*^{-1} \Omega + \omega(\tau^2)$ Using (9), the last relationship implies

$$\hat{\Omega} = \Omega - \tau \Omega [\Omega_\rho^{-1} \delta_\rho + \omega(\tau)] \Omega + \omega(\tau^2) = \Omega - \tau \Omega \Omega_\rho^{-1} \Omega \delta_\rho + \omega(\tau^2) \tag{10}$$

Proof of Theorem 3: Based on the results of Theorem 2 and the results of lemmas which

appear in the supplementary material, we can see that, for all estimators considered (i.e., $\hat{\rho}_I$, indexed by $I = \{LS, DW, GL, PW, ML\}$), we can write $\Omega = \hat{\Omega} + \tau\hat{\Omega}\hat{\Omega}_\rho^{-1}\hat{\Omega}\delta_\rho^I + \omega(\tau^2)$,where the sampling errors $\delta_\rho^I = \sqrt{T}(\hat{\rho}_I - \rho)$ admit the stochastic expansion, for all I:

$$\begin{aligned}
\delta_\rho^{LS} &= -\frac{(1-\rho^2)u'\Omega_2^{-1}u}{2\sqrt{T}} - \frac{\tau(1-\rho^2)}{2}\left[u'M_X\Omega_2^{-1}M_Xu - \frac{(1-\rho^2)u'uu'\Omega_2^{-1}u}{T}\right] + \omega(\tau^2) \\
\delta_\rho^{DW} &= \delta_\rho^{LS} + \frac{\tau^2(1-\rho^2)}{2}(u_1^2 + u_T^2) + \omega(\tau^2) \\
\delta_\rho^{GL} &= \delta_\rho^{PW} = \delta_\rho^{LS} - \tau^2(1-\rho^2)\left[u'M_X\Omega_2^{-1}P_X\Sigma\Omega^{-1}u + \frac{1}{2}u'\Omega^{-1}\Sigma P_X\Omega_2^{-1}P_X\Sigma\Omega^{-1}u\right] + \omega(\tau^2) \\
\delta_\rho^{ML} &= \delta_\rho^{GL} + \tau\rho\left[(1-\rho^2)(u_1^2 + u_T^2) - 1\right] + +\omega(\tau^2).
\end{aligned} \tag{11}$$

Taking expectations and collecting terms of order $O(1)$ of the above relationships, we can complete the proof of the theorem.

References

[1] Andrews, D.W.K., 1991, Heteroskedasticity and autocorrelation consistent covariance matrix estimation, Econometrica 59, pp.817-858.

[2] Beach, C.M. and MacKinnon, J.G. (1978). A Maximum Likelihood Procedure for Regression with Autocorrelated Errors, Econometrica, 46, 51-58.

[3] Breusch, T.S. (1980), Useful Invariance Results for Generalized Regression Models, Journal of Econometrics, 13, 327–340.

[4] Den Haan, W. J. and A. Levin (1994). Inferences from parametric and non-parametric covariance matrix estimation procedures. Manuscript, UCSD and Federal Reserve Board.

[5] Giraitis, L. and Phillips, P.C.B. (2012). Mean and autocovariance function estimation near the boundary of stationarity. Journal of Econometrics, 169, 166—178.

[6] Goncalves, S., and Vogelsang, T. (2011), "Block Bootstrap and HAC Robust Tests: The Sophistication of the Naive Bootstrap," Econometric Theory, 27, 745–791.

[7] Kiefer, N.M., Vogelsang, T.J., and Bunzel, H. (2000), Simple Robust Testing of Regression Hypotheses, Econometrica, 68, 695–714.

[8] Magdalinos, M.A. (1992), Stochastic Expansions and Asymptotic Approximations, Econometric Theory, 8, 343–367.

[9] Magdalinos, M. A., and S. D. Symeonides. 1995. "Alternative Size Corrections for Some GLS Test Statistics: The Case of the AR(1) Model." Journal of Econometrics 66:35–59.

[10] Muller, U.K. (2014), HAC Corrections for Strongly Autocorrelated Time Series, Journal of Business & Economic Statistics, 32, 311–322.

[11] Newey, W.K. and West, K. (1994), Automatic Lag Selection in Covariance Matrix Estimation. Review of Economic Studies, 61, 631-653.

[12] Newey, W.K. and West, K. (1987), A Simple Positive Semi-Definite, Heteroskedasticity and Autocorrelation Consistent Covariance Matrix, Econometrica, 55, 703–708.

[13] Phillips, P.C.B. (1980), Finite sample theory and the distributions of alternative estimators of the marginal propensity to consume. Review of Economic Studies 47, 183-224.

[14] Prais, S.J. and Winsten, C.I. (1954), Trend Estimators and Serial Correlation, Cowles Commission Discussion Paper, Statistics No. 383.

[15] Rothenberg, T.J. (1984), Approximate Normality of Generalized Least Squares Estimates, Econometrica, 52, 811–825.

[16] Symeonides, S., Karavias, Y. & Tzavalis, E. (2016). Size corrected Significance Tests in Seemingly Unrelated Regressions with Autocorrelated Errors. Journal of Time Series Econometrics, 9(1), doi:10.1515/jtse-2015-0014

Table 1: Size of t-tests

	T	t_{OLS}	t_R	t_{A-B}	t_{A-QS}	t_{NW}	t_{GV}	t_{KVB}	t_{MUL}	t_{VARHAC}	t_{BC}
ρ=0.6	15	0.114	0.078	0.382	0.192	0.183	0.109	0.085	0.065	0.117	0.052
	30	0.128	0.068	0.341	0.163	0.148	0.118	0.068	0.055	0.056	0.042
	60	0.130	0.060	0.249	0.117	0.187	0.070	0.057	0.057	0.035	0.044
ρ=0.7	15	0.133	0.092	0.391	0.205	0.200	0.111	0.096	0.071	0.122	0.062
	30	0.148	0.075	0.341	0.179	0.157	0.117	0.071	0.066	0.067	0.042
	60	0.160	0.062	0.263	0.145	0.195	0.069	0.061	0.061	0.037	0.043
ρ=0.8	15	0.151	0.094	0.395	0.231	0.216	0.101	0.100	0.076	0.133	0.070
	30	0.196	0.086	0.360	0.215	0.199	0.107	0.091	0.068	0.082	0.062
	60	0.210	0.071	0.280	0.162	0.221	0.095	0.079	0.064	0.049	0.052
ρ=0.9	15	0.169	0.103	0.410	0.256	0.231	0.118	0.109	0.073	0.152	0.089
	30	0.229	0.099	0.376	0.251	0.224	0.125	0.106	0.065	0.097	0.086
	60	0.268	0.103	0.310	0.224	0.254	0.131	0.117	0.061	0.075	0.076

Table 2: Size-adjusted power of t-tests

	T	t_{OLS}	t_R	t_{A-B}	t_{A-QS}	t_{NW}	t_{GV}	t_{KVB}	t_{MUL}	t_{VARHAC}	t_{BC}
ρ=0.6	15	0.173	0.170	0.246	0.203	0.198	0.145	0.170	0.070	0.169	0.110
	30	0.227	0.234	0.284	0.242	0.233	0.220	0.212	0.098	0.193	0.123
	60	0.338	0.368	0.374	0.315	0.351	0.349	0.303	0.124	0.219	0.169
ρ=0.7	15	0.179	0.154	0.233	0.207	0.201	0.132	0.178	0.058	0.163	0.106
	30	0.238	0.229	0.279	0.240	0.238	0.231	0.222	0.078	0.169	0.122
	60	0.317	0.367	0.345	0.283	0.325	0.333	0.299	0.109	0.186	0.150
ρ=0.8	15	0.168	0.175	0.219	0.197	0.183	0.142	0.163	0.064	0.175	0.115
	30	0.204	0.242	0.238	0.217	0.206	0.212	0.208	0.078	0.182	0.120
	60	0.284	0.366	0.294	0.279	0.287	0.317	0.281	0.097	0.176	0.140
ρ=0.9	15	0.177	0.172	0.209	0.188	0.183	0.151	0.164	0.062	0.154	0.119
	30	0.205	0.242	0.203	0.199	0.203	0.219	0.218	0.070	0.165	0.132
	60	0.230	0.342	0.248	0.222	0.220	0.251	0.230	0.077	0.155	0.135

[17] Velasco, Carlos and Robinson, Peter M. (2001). Edgeworth expansions for spectral density estimates and studentized sample mean Econometric Theory, 17 (3). 497-539. ISSN 1469-4360