Implications of Two Measures of Persistence for Correlation Between Permanent and Transitory Shocks in U.S. Real GDP*

Daisuke Nagakura [†] and Eric Zivot [‡]

revised January, 2007

Abstract

Conventionally, shocks to permanent and transitory components in the unobserved components (UC) model for the log of real GDP are assumed to be uncorrelated. This assumption is mainly for identification of model parameters. In this paper, we show important implications of two popular measures of persistence for the correlation between permanent and transitory shocks in the UC model, and demonstrate that the correlation is negative for the log of U.S. real GDP under a very general specification of the cycle process.

Key Words: Persistence; Impulse Response; Variance Ratio; Unobserved Components Model; Nonparametric; Prediction Error;

JEL Classification Numbers: C14, C22, E32.

---

*We would like to thank Richard Startz, Charles Nelson, KumHwa Oh, and Don Perceival for their helpful comments. Any remaining errors are ours.

†Department of Economics, University of Washington, E-mail: dn2@u.washington.edu

‡Department of Economics, University of Washington, E-mail: ezivot@u.washington.edu

1 Introduction

It is a widely accepted view that real GDP can be thought to consist of two components, a permanent and a transitory component. Shocks to the permanent component have long-lasting effects, whereas shocks to the transitory component are temporary and vanish in the long-run. The *unobserved components* (UC) model has proved to be useful to analyze these two components in real GDP. See, for example, Watson (1986) and Clark (1987). In this model, a time series is represented by the sum of two unobserved components, namely, stochastic trend (permanent component), which follows a random walk process, and cycle (transitory component), which is a stationary process.

In most of the UC literature, it is assumed that the correlation between shocks to trend and cycle is zero. The zero correlation assumption is often imposed not because it is reasonable, but mainly for identification of model parameters. However, imposing such restrictions may distort estimates of trend and cycle. This issue has been recently raised in Morley, Nelson, and Zivot (2003, hereafter MNZ); they pointed out that the correlation can be identified for UC models with an AR(2) cycle process and they estimated the correlation for U. S. quarterly real GDP data. Their estimate of the correlation is -0.9062 and significantly far away from zero (see also Oh at al., 2006; Proietti, 2006).

Our paper shows that the correlation is negative for U. S. quarterly real GDP under a much more general cycle process than an AR(2) process. For this purpose, we extend the theorem by Lippi and Reichlin (1992). Suppose that real GDP can also be represented by a difference-stationary process $\{y_t\}_{t=-\infty}^{\infty}$, whose first difference admits a MA(∞) representation $\Delta y_t = a_0 + A(L)u_t$ with white noise u_t and absolutely summable coefficients. Campbell and Mankiw (1987a,b) defined $A(1)$ as a measure of persistence for a shock to y_t; we call it the *impulse response* (IR) measure. Lippi and Reichlin (1992) showed that if $A(1)^2$ is greater than or equal to 1, then the correlation between trend and cycle is not zero.

We show that if $A(1)^2$ is greater than or equal to 1, then the correlation is negative. This suggests a useful way of examining the sign of correlation; if an estimate of $A(1)^2$ is significantly greater than or equal to 1 then we may conclude that the correlation is negative. Note that the method does not require parameters of the UC model to be identified and thus can be applied even for unidentified UC models.

There is another popular measure of persistence called the *variance ratio* (VR) measure, denoted by V, which was introduced by Cochrane (1988). In addition to the above results on the IR measure, we show that the VR measure also has an important implication for the correlation ρ. Specifically, we show that if V is greater than one, then there exists an upper bound for the correlation ρ which is a function of the VR measure. This upper bound is denoted by ρ_{ub}.

For estimating $A(1)$, the most straightforward way is to assume that the time series follows an ARIMA(p, 1, q) process and then estimate $A(1)$ using estimates of the AMRA(p, q) coefficients of the first difference, as Campbell and Mankiw (1987a,b) did. However, it is known that the estimate by this approach is very sensitive to the order of the fitted ARIMA process (Christiano and Eichenbaum, 1989; Hauser *et al.*, 1999).

We propose non-parametric estimators for $A(1)^2$ and ρ_{ub}, and discuss methods for constructing their confidence intervals. Our estimator for $A(1)^2$ is free from the order selection problem of ARIMA models although we need to choose a tuning or bandwidth parameter. Properties of the estimators and accuracy of the confidence intervals are examined by Monte Carlo experiments. It is found that the nonparametric estimator for $A(1)^2$ is comparable with the parametric estimator for a correctly parameterized ARIMA model.

We estimate $A(1)^2$ and ρ_{ub} for US quarterly real GDP applying our non-parametric estimators. The estimates of $A(1)^2$ with various values of bandwidth parameters are all greater than 1; however, one of the valid upper 95 % confidence intervals does not include 1. This implies that the correlation of the trend and cycle innovations for U. S. quarterly real GDP is negative. The estimates of the upper bounds range from -0.3409 to -0.7481 depending on the bandwidth parameter.

The rest of the paper is organized as follows. Section 2 gives some theoretical results, where we derive relationships between $A(1)^2$, V and the correlation ρ. In Section 3, we propose non-parametric estimators for $A(1)^2$ and the upper bound of ρ implied from V, and illustrate how we construct their confidence intervals. In Section 4, we conduct several Monte Carlo experiments to investigate the properties of our non-parametric estimators and the accuracy of the confidence intervals. We also examine the performance of conventional parametric estimators based on various ARIMA models; especially, we examine their behaviors under a misspecification of orders of a fitted ARIMA model. Section 5 reports some empirical results for US quarterly real GDP applying our non-parametric estimators, where we find that the estimate of $A(1)^2$ is significantly greater than 1. Some concluding remarks are given in Section 6. All proofs are relegated to the Appendix.

2 Measures of persistence and the UC model

2.1 Relationships of two measures of persistence to the UC model

Consider a difference-stationary process $\{y_t\}_{t=-\infty}^{\infty}$ (hereafter we abbreviate this to $\{y_t\}$ for notational simplicity) whose first difference admits a MA(∞) representation:

$$\Delta y_t = a_0 + A(L)u_t, \quad u_t \sim WN(0, \sigma_u^2), \quad t = 0, \pm 1, \pm 2, ... \tag{1}$$

where $\Delta y_t \equiv y_t - y_{t-1}$, $A(L) = 1 + a_1 L + a_2 L^2 + \cdots$, $\sum_{j=1}^{\infty} |a_j| < \infty$, $A(1) \neq 0$, and $WN(0, \sigma_u^2)$ denotes a white noise process with mean zero and variance $\sigma_u^2 > 0$.[1] Various measures of persistence for a shock to y_t have been proposed. Two of the most popular ones are the impulse response (hereafter IR) measure suggested by Campbell and Mankiw (1987a,b), and the variance ratio

[1]Thus, the MA(∞) process has absolutely summable autocovarinces; see Hamilton (1994, p 52).

(hereafter VR) measure introduced by Cochrane (1988). The former is defined as $A(1)$. This definition is motivated by the fact that the cumulated responses of y_t in the infinite future to a unit innovation is the sum of the MA coefficients, $A(1)$. Beveridge and Nelson (1981) pointed out that

$$\lim_{k\to\infty}[E_t(y_{t+k}) - E_{t-1}(y_{t+k})] = A(1)u_t, \tag{2}$$

where $E_t(y_s)$ denotes the conditional expectation of y_s conditioned on $(y_t, u_t, u_{t-1}, ...)$ assuming that we know the values of $(a_1, a_2, ...)$. This shows that $A(1)$ can also be interpreted as a revision in the long run prediction of y_t due to the occurrence of a unit shock at t. Because of this, the measure is sometimes called Beverage-Nelson persistence measure (Lippi and Reichlin, 1992, 1994).

The latter measure is defined as

$$V \equiv \frac{\sigma^2_{lrv}}{\sigma^2_{\Delta y}} \tag{3}$$

where σ^2_{lrv} and $\sigma^2_{\Delta y}$ are the long-run and the unconditional variances of Δy_t, which are defined as $\sigma^2_{lrv} \equiv A(1)^2\sigma^2_u$ and $\sigma^2_{\Delta y} \equiv (1 + \sum_{j=1}^{\infty} a_j^2)\sigma^2_u$, respectively.[2] It is easily shown that $A(1)^2 \geq V$ for the process in (1); that is, Cochrane's persistence measure is a lower bound of $A(1)^2$, and equality holds when $\{\Delta y_t\}$ is a white noise process.

Consider the following representation for y_t:

$$y_t = \tau_t + c_t, \quad t = 0, \pm 1, \pm 2, ..., \tag{4}$$

where (a) $c_t = B(L)\epsilon_t$, $\epsilon_t \sim WN(0, \sigma^2_\epsilon)$, $\sigma^2_\epsilon > 0$ is a zero mean stationary MA(∞) process with absolutely summable coefficients (b) τ_t is a random walk process with drift, i.e., $\tau_t = \mu + \tau_{t-1} + \eta_t$

[2]More precisely, Cochrane (1988) proposed the ratio of the variance of kth difference of the series to the variance of the first difference. When $k \to \infty$, it will converge to the quantity V defined above. This quantity is used most often in the literature. Cochrane (1988) also proposed the ratio of the sample variances of kth and the first differences as the estimator for V.

with $\eta_t \sim WN(0, \sigma_\eta^2)$ and $\sigma_\eta^2 > 0$; (c) $\mathrm{cov}(\epsilon_t, \eta_s) = \sigma_{\eta\epsilon}$ for $t = s$ and zero otherwise; (d) $|\rho| < 1$, where $\rho \equiv \mathrm{corr}(\epsilon_t, \eta_t) = \sigma_{\eta\epsilon}/(\sigma_\epsilon \sigma_\eta)$. This representation is known as the unobserved components (hereafter UC) model with correlated shocks. Note that if the representation exists then it implies that V cannot be one, which in turn implies that the process $\{\Delta y_t\}$ is not a white noise process [3] and $A(1)^2$ is always strictly greater than V; the proof of this result can be found in the footnote of the proof of Proposition 2 in Appendix. Note also that when $\{\Delta y_t\}$ is a white noise process, we have $A(1)^2 = 1$; however, $A(1)^2 = 1$ does not necessarily imply that $\{\Delta y_t\}$ is a white noise process.[4]

Lippi and Reichlin (1992) showed that $A(1)^2 < 1$ when the correlation ρ is zero. First, we generalize the Lippi and Reichlin's result in Proposition 1 below.

Proposition 1 *Let $\{y_t\}$ be the difference-stationary process defined in (1). Assume further that $\{y_t\}$ admits the UC model representation defined in (4). Then $A(1)^2 < (1 + \rho\delta)^{-2}$ when $\rho\delta \neq -1$, and $A(1)^2 < \delta^2/(1 - \delta^2)^2$ when $\rho\delta = -1$, where $\delta \equiv \sigma_\epsilon/\sigma_\eta$.*

Note that $\rho\delta = -1$ automatically implies that $\rho < 0$ and $\delta > 1$.

Remark 1 Proposition 1 includes Lippi and Reichlin (1992)'s theorem as a special case ($\rho = 0$).

[3]This implies that if the process is a white noise process, then this UC representation does not exist. However, it is not obvious whether this UC representation always exists for an arbitrary MA(∞) process that is not a white noise process, since we assume $|\rho| < 1$, which excludes the famous "Beveridge-Nelson decomposition" representation (see p 504, Hamilton, 1994, it actually requires a more strict condition, i.e., the one-summmability condition).

[4]Consider, for example, ARIMA(0,1,2) process such as $\Delta y_t = u_t + 0.2u_{t-1} - 0.2u_{t-2}$. The coefficients of the MA process satisfy the invertible condition and we have $A(1)^2 = (1+0.2-0.2)^2 = 1$.

From this proposition, we immediately obtain the following corollary:

Corollary 1 *Assume further that* $0 \leq \rho < 1$ *in Proposition 1. Then* $A(1)^2 < 1$.

Remark 2 The above corollary can be equivalently stated as: if $A(1)^2 \geq 1$, then $\rho < 0$.

Remark 2 can be used to examine the sign of the correlation; it implies that if an estimate of $A(1)^2$ is significantly greater than or equal to 1, then we may conclude that the correlation ρ is negative.

The result in Remark 2 is useful for examining the sign of the correlation. However, it is not informative about the magnitude of correlation since we cannot directly observe nor estimate δ (without further assumptions) and so the correlation may be arbitrary close to zero. The next proposition is useful for examining the magnitude of the negative correlation when $A(1)^2 \geq 1$.

Proposition 2 *Assume that* $\{y_t\}$ *satisfies the conditions in Proposition 1. If* $V > 1$, *then* $\rho < \rho_{ub}$, *where* $\rho_{ub} \equiv -\sqrt{1 - V^{-1}}$.

Note that to apply this upper bound, we need the condition that $V > 1$. This condition is stronger than $A(1)^2 \geq 1$; $V > 1$ implies $A(1)^2 \geq 1$ but the converse is not true. Then it makes sense that this stronger condition gives the result on ρ (i.e., the negative upper bound) which is stronger than the result implied by $A(1)^2 \geq 1$ (i.e., negative correlation). This upper bound is applicable regardless of the underlying cycle process. Unfortunately, because of this wide applicability, this upper bound may not be tight; the true value of ρ may be much less than the upper bound. We discuss this issue briefly in the next subsection using a particular UC model.

2.2 The upper bound for a particular UC model

As an illustrative example, consider the following simple UC model:

$$
\begin{aligned}
y_t &= \tau_t + c_t, \\
\tau_t &= \mu + \tau_t + \eta_t, \\
(1-\phi_1)c_t &= \epsilon_t,
\end{aligned}
\qquad
\begin{pmatrix} \eta_t \\ \epsilon_t \end{pmatrix} \sim i.i.d. \left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} \sigma_\eta^2 & \rho\sigma_\eta\sigma_\epsilon \\ \cdot & \sigma_\epsilon^2 \end{bmatrix} \right), \tag{5}
$$

where $|\phi| < 1$. The reduced form of the UC model in (5) is an ARIMA(1,1,1) model. When $\phi = 0$, the model reduces to the *random walk plus noise process*, which was considered in Oh at al. (2006) and Harvey and Koopman (2000). There are four parameters in this model; however only the AR(1) coefficient of the cycle process is identified and the other three parameters are not uniquely identified, as emphasized in Oh at al. (2006) for a special case of the model, namely the random walk plus noise process. However, we can show that[5]

$$
V = \frac{\sigma_\eta^2}{\sigma_\eta^2 + 2\rho\sigma_\eta\sigma_\epsilon + 2\sigma_\epsilon^2(1+\phi_1)^{-1}}.
$$

Note that the condition $V > 1$ is equivalent to $\rho < -\delta(1+\phi_1)^{-1}$, where $\delta \equiv \sigma_\epsilon/\sigma_\eta$. Note also that this immediately implies that $\rho < 0$ since $|\phi| < 1$. When this condition is satisfied, we can show that the upper bound derived in Proposition 2, i.e., ρ_{ub}, is

$$
\rho_{ub} = -\sqrt{-2\delta[\ \rho + \delta(1+\phi_1)^{-1}]} = -\sqrt{2\delta[\ |\rho| - \delta(1+\phi_1)^{-1}]}. \tag{6}
$$

Thus, the difference between ρ_{ub} and the true value ρ is

$$
d \equiv \rho_{ub} - \rho = -\sqrt{2\delta[\ |\rho| - \delta(1+\phi_1)^{-1}]} + |\rho| \geq |\rho| \left(1 - \sqrt{\frac{1+\phi_1}{2}} \right). \tag{7}
$$

The last inequality is obtained by minimizing d with respect to δ; the equality holds when $\delta = |\rho|(1+\phi_1)/2$.

[5] First, it can be shown that the long-run variance of Δy_t is always equal to the trend innovation variance σ_η^2 (see Appendix). The denominator is obtained by taking the variance of $\Delta y_t = \mu + \eta_t + c_t - c_{t-1}$.

These equations show that for a fixed ϕ_1, the minimum of d with respect to δ becomes larger in proportion to $|\rho|$, and d becomes as large as the absolute value of ρ as δ gets closer to 0 or $|\rho|(1+\phi_1)$. Consider the case in which $\phi = 0$. In this case, when $\rho = -0.8$ and $\delta = 0.4$, then $\rho_{ub} \approx -0.5657$ and thus $d = 0.2343$; we can see that ρ_{ub} is not very tight. Figure 1 plots (ρ_{ub}, δ) according to (6) with ρ fixed at -0.8 and various values of ϕ_1. The last remark is that from the derivation of ρ_{ub} in the Appendix, we can see that ρ_{ub} would become tighter as the underlying cycle process gets closer to a random walk process, which is also observed in Figure 1.

This example makes it very clear that ρ_{ub} is merely an upper bound; it does not imply that the true value of ρ is close to ρ_{ub}. However ρ_{ub} can be calculated from a reduced form ARIMA model and can be applied to any UC model, even to unidentified UC models like the one above. Furthermore, we do not need to know the true orders of the reduced form ARIMA model since ρ_{ub} can be estimated nonparametrically as shown in Section 3.

3 Non parametric estimators for $A(1)^2$ and ρ_{ub}

3.1 Estimation of $A(1)^2$

In this subsection, we consider various estimators of $A(1)^2$. First, we consider a parametric approach used in Campbell and Mankiw (1987a,b). They assumed that the stationary process $\{\Delta y_t\}$ can be represented by an ARMA(p, q) process,

$$\Phi(L)(\Delta y_t - a_0) = \Theta(L)u_t, \quad u_t \sim WN(0, \sigma_u^2), \tag{8}$$

where $\Phi(L)$ and $\Theta(L)$ are the lag polynomials of AR and MA coefficients, respectively. Then it can be shown that $A(1)^2 = (\Theta(1)/\Phi(1))^2$. Campbell and Mankiw (1987a,b) estimated $\Phi(1)$ and $\Theta(1)$ by substituting Gaussian MLEs of the coefficients into these polynomials.

This approach was, however criticized by Cochrane (1988); Cochrane argued that the estimates of persistence obtained by fitting ARMA (or any parametric) model with Gaussian MLE would have upward bias if the fitted model is misspecified and the true value of persistence is small (see also Christiano and Eichenbaum, 1989, who pointed that estimates by this approach are sensitive to the choice of ARMA orders). Hauser *et al.* (1999) questioned this criticism insisting that there is flaw in Cochrane's argument. They argued that the estimates may actually be downwardly biased if the orders of fitted ARMA models are greater than the true orders.

While the parametric ARMA approach is simple and can be easily implemented, it is also possible to estimate $A(1)^2$ in non-parametric ways, which is free from the choice of ARIMA orders but subject to bandwidth estimator issues. In this subsection we propose a non-parametric estimator for $A(1)^2$. Note that $A(1)^2$ can be rewritten as

$$A(1)^2 = \frac{A(1)^2\sigma_u^2}{\sigma_u^2} = \frac{\sigma_{lrv}^2}{\sigma_u^2}. \tag{9}$$

That is, it is the ratio of the long-run variance to the prediction error variance σ_u^2. This suggests the ratio of two consistent nonparametric estimators for σ_{lrv}^2 and σ_u^2 as a natural estimator for $A(1)^2$. Below first we explain some existing estimators for σ_{lrv}^2 and σ_u^2. Then we define the ratio of them as our nonparametric estimator. Estimating $A(1)^2$ in this way has not been considered in the literature.

Hereafter, we assume that the innovation process $\{u_t\}$ in (1) is i.i.d. with finite fourth moment and $s(1) < \infty$, where $s(q) \equiv \sum_{j=-\infty}^{\infty} |j|^q|\gamma_j|$, $\gamma_j = \mathrm{cov}(\Delta y_t, \Delta y_{t-j}), j = 0, 1, 2...$, and $\gamma_{-j} = \gamma_j$. This assumption is needed to show the asymptotic normality of the spectral density estimator with the Bartlett kernel. Also, we denote Δy_t by x_t and $A(1)^2$ by W for the sake of notational simplicity. Let $f_x(\omega)$ be the spectral density function of the stationary process $\{x_t\}$, i.e.,

$$f_x(\omega) \equiv \frac{1}{2\pi}\sum_{j=-\infty}^{\infty} \gamma_j e^{-ij\omega}. \tag{10}$$

Note that σ_{lrv}^2 is equal to 2π times the spectral density at zero frequency, i.e., $\sigma_{lrv}^2 = 2\pi f_x(0)$. This motivates the use of spectral density estimators to estimate σ_{lrv}^2. The most popular spectral density estimators are the non-parametric kernel spectral density estimators.

Following Andrews (1991), given T observations x_t $t = 1, ..., T$, we consider the class of kernel estimators of the form

$$\widehat{\sigma^2_{lrv,T}} = \frac{T}{T-1} \sum_{j=-T+1}^{T-1} k\left(\frac{j}{S_T}\right) \widehat{\gamma}_{j,T}, \quad \text{where}$$

$$\widehat{\gamma}_{j,T} = \frac{1}{T} \sum_{t=1}^{T-|j|} (x_t - \widehat{a}_0)(x_{t+|j|} - \widehat{a}_0) \quad \text{for} \quad j = 0, \pm 1, ..., \pm (T-1), \tag{11}$$

$\widehat{a}_0 = \frac{1}{T} \sum_{t=1}^{T} x_t$ is the estimate of the unconditional mean, $T/(T-1)$ is the factor for a small sample degrees of freedom adjustment to offset the effect of estimation of the unconditional mean, $k(\cdot)$ is a real-valued kernel in the set K_1 defined in equation (2.6) in Andrews (1991), and S_T is a band-width parameter.

The stationary process x_t satisfies the assumptions of Theorem 1(b) in Andrews (1991), and thus under a suitable choice of S_T, we have

$$\sqrt{T/S_T}(\widehat{\sigma^2_{lrv,T}} - \sigma^2_{lrv}) = O_p(1). \tag{12}$$

Considerable research has been conducted on the properties of kernel spectral density estimators. See, for example, Newey and West (1987), Andrews (1991), Newey and West (1994), Hannan (1970), Percival and Walden (1994) and Priestley (1981) among many others. In this paper, we consider only the *Bartlett kernel*:

$$k(x) = \begin{cases} 1 - |x| & \text{for} \quad |x| \leq 1, \\ 0 & \text{otherwise} \end{cases}.$$

However, the arguments below can be immediately extended to other kernels.

Andrews (1991) showed that, for the Bartlett (hereafter BT) kernel, the asymptotically optimal growth rate of the bandwidth parameter, which minimizes the asymptotic mean square error, is

$O(T^{1/3})$.[6] Andrews (1991) also suggested a plug-in type automatic bandwidth selection procedure. For the BT kernel, the method determines a bandwidth parameter according to:

$$S_T = 1.1447(\widehat{\nu}T)^g, \tag{13}$$

where $\widehat{\nu}$ is a value estimated with the formula in Andrews (1991, p835), T is the sample size, and g is the growth rate of bandwidth parameter, which is typically chosen to be 1/3 for the BT kernel since it is optimal in the sense that it minimizes the asymptotic MSE. Note that S_T determined by (13) is a positive real value but may not be a positive integer.[7] We apply this procedure when we estimate σ^2_{lrv} by kernel spectral density estimations.

Next we describe an estimator for the prediction error variance σ_u^2. At this point, we further assume that $\{u_t\}$ in (1) is normal, which assures the asymptotic normality and, in particular, $\sqrt{T}$ consistency of the estimator. It is well known that for a stochastic process admitting a Wold representation with the prediction error variance σ_u^2 and a spectral density function $f(\omega)$ that is positive almost everywhere, the following equation, which is known as Kolmogorov's formula, holds:

$$\sigma_u^2 = \exp\frac{1}{2\pi}\int_{-\pi}^{\pi}\log 2\pi f(\omega)\mathrm{d}\omega = \exp\frac{1}{\pi}\int_{0}^{\pi}\log 2\pi f(\omega)\mathrm{d}\omega. \tag{14}$$

See Hannan (1970, p137). Using (14), Davis and Jones (1968) proposed a nonparametric estimator for σ_u^2, which replaces the unknown $f(\omega)$ by the periodogram and the integral by a finite Riemann sum. In our case, it is given by

$$\widehat{\sigma^2_{u,T}} = \exp\left\{\frac{1}{M}\sum_{k=1}^{M}\log I_T(\omega_k) + \gamma\right\}, \tag{15}$$

[6] Strictly speaking, to get this result we need an additional assumption (Assumption D in Andrews (1991)), which is satisfied by the additional assumption of normality below.

[7] As Andrews (1991) argued, bandwidth parameter does not have to be a positive integer.

where $I_T(\omega) = (1/T)|\sum_{t=1}^T x_t e^{-i\omega t}|^2$ is the periodogram of the stationary process x_t, [8] $M = \lfloor (T-1)/2 \rfloor$, $\omega_k = 2\pi k/T$, and $\gamma \approx 0.57721$ is Euler's constant (not an autocovariance) used for bias correction.

Davis and Jones (1968) proved the asymptotic normality of their estimator assuming that the process is independent normal. Hannan and Nicholls (1977, 1979) established the strong consistency and asymptotic normality of the Davis - Jones estimator, namely, $\sqrt{T}(\widehat{\sigma^2_{u,T}} - \sigma^2_u) \to_d N(0, \psi'(1)\sigma^4_u)$, where $\psi(x)$ is the digamma function, without assuming independence (but assuming normality) and with fairly minimal assumptions on $f_x(w)$, which is shown to be slightly weaker than our assumptions here. In the appendix, we show that their assumptions are indeed satisfied under our assumptions.

Some other estimators were suggested by Chen and Hannan (1980) and An (1981). Chen and Hannan (1980) proved the strong consistency of their estimator under weaker assumptions than Hannan and Nicholls (1977); however they did not derive its asymptotic distribution nor its convergence rate. An (1981) proposed an estimator and established its strong consistency and asymptotic normality under conditions weaker than the conditions of Hannan and Nicholls (1979). The estimator, however, requires a numerical evaluation of an integral and is more difficult to calculate than the Davis - Jones estimator. See also Pukkila and Nyquist (1985), Mohanty and Pourahmadi (1996), Janacek (1975), Bhansali (1974), Taniguchi (1980), and Walden (1995) for related works.

We use the Davis - Jones estimator with the above assumptions, which are slightly stronger

[8] Note that the periodogram can be also written as $(1/T)\sum_{j=1}^T \sum_{t=1}^T x_t x_j \cos(\omega(t-j))$. Here we do not have to center the time series unlike (11) since even if we centered the time series by true a_0, the resulting periodogram is numerically the same as in (15) at Fourier frequencies $\omega_k = 2\pi k/T$, $k = 1, ..., M$. See Percival and Walden (1994, p196, p204) for more details.

than the ones assumed in Hannan and Nicholls (1977). Although Hannan and Nicholls (1977, p835) conjectured that their result would hold for non-normal cases, it seems that no formal proof for that claim has been available.

Finally, we define the ratio of $\widehat{\sigma^2_{lrv,T}}$ in (11) to $\widehat{\sigma^2_{u,T}}$ in (15), i.e.,

$$\widehat{W}_T = \frac{\widehat{\sigma^2_{lrv,T}}}{\widehat{\sigma^2_{u,T}}}, \tag{16}$$

as our non-parametric estimator for W. We examine the finite sample properties of this estimator by Monte Carlo experiment in Section 4.

3.2 Confidence intervals for W

Suppose that $\sqrt{T/S_T}(\widehat{\sigma^2_{lrv,T}} - \sigma^2_{lrv}) \to_d Z$, where Z is a random variable with non-degenerated distribution. Then we can show that $\sqrt{T/S_T}(\widehat{W}_T - W) \to_d Z/\sigma^2_u$.[9] Thus, the asymptotic distribution of $\widehat{W}_T$ is essentially the same as that of $\widehat{\sigma^2_{lrv}}$. It can be shown that the distribution of Z is normal whose variance depends on the true spectral density. However, its mean may or may not be zero depending on the growth rate of the bandwidth parameter S_T; when S_T increases at a sufficiently fast rate, the mean is zero; however, when S_T is increased at a slower rate, then the mean is an unknown nonzero constant, which differs depending on the true process.

More precisely, under our assumption that $s(1) < \infty$, for the BT kernel with $S_T = O(T^g)$, $1/3 \leq g < 1$, we can show that $Z \sim N(\mu_{BT}, (4/3)\sigma^4_{lrv})$, where $\mu_{BT} = -(T^{1/2}S_T^{-3/2})s(1)$. See Priestley (1981, p469) on the asymptotic normality of kernel spectral density estimators. Notice that if $1/3 < g$, then μ_{BT} goes to zero as $T \to \infty$; however, it is a non-zero constant when $g = 1/3$. Although μ_{BT} may be estimated, we construct confidence intervals for W assuming that $\mu_{BT} = 0$. Then, we have $\sqrt{T/S_T}(\widehat{W}_T - W) \to_d N(0, (4/3)W^2)$. The point in this argument is that although for the BT kernel setting $g = 1/3$ is minimizing the asymptotic MSE, it would not be desirable

[9] A simple application of Slutsky Theorem leads to the result.

for the purpose of constructing confidence intervals. This issue will be examined in Monte Carlo experiments in the next section.

Next we consider the construction of confidence intervals for W. Since our objective is to provide evidence that W is greater than 1 for U. S. real GDP, it is natural to consider one-sided upper confidence intervals; if they do not contain 1, then we may reject that $W < 1$. Let c_α be the 100α % point of the standard normal distribution; i.e., $\Pr(N(0,1) \leq c_\alpha) = \alpha$. For example, if $\alpha = 0.95$, then $c_\alpha = 1.64$. Then, for the BT kernel, the one-sided upper 100α % confidence interval is given by

$$\left[\frac{\widehat{W}_T}{1 + c_\alpha/\kappa_T}, \quad \infty\right), \tag{17}$$

where $\kappa_T = \sqrt{3T/4S_T}$, since

$$\begin{aligned}\alpha &= \Pr\left(\kappa_T\ W^{-1}(\widehat{W}_T - W) \leq c_\alpha\right) = \Pr\left(\kappa_T\ W^{-1}\widehat{W}_T - \kappa_T \leq c_\alpha\right) \\ &= \Pr\left(\frac{\kappa_T\ \widehat{W}_T}{\kappa_T + c_\alpha} \leq W\right) = \Pr\left(\frac{\widehat{W}_T}{1 + c_\alpha/\kappa_T} \leq W\right).\end{aligned}$$

The first equality comes from the fact that $\kappa_T W^{-1}(\widehat{W}_T - W)$ (asymptotically) follows the standard normal distribution. Note that for a fixed T, the confidence interval becomes wider as S_T increases (as κ_T decreases). This implies that there is a trade off between obtaining a more accurate confidence interval and obtaining a tighter confidence interval. Later, we examine the accuracy of the confidence intervals with various values of g by simulation.

3.3 An Estimator and confidence interval for ρ_{ub}

In Section 2.1, we derived an upper bound of the correlation ρ, which is given by $\rho_{ub} \equiv -\sqrt{1 - V^{-1}}$. Consider the following consistent estimator for V:

$$\widehat{V}_T \equiv \frac{\widehat{\sigma^2_{lrv,T}}}{\widehat{\sigma^2_{x,T}}}, \tag{18}$$

where $\widehat{\sigma^2_{x,T}}$ is the usual sample variance and $\widehat{\sigma^2_{lrv,T}}$ is given by (11). Here, again we restrict our attention only to the BT kernel. Then, the estimator $\widehat{V}_T$ is asymptotically equivalent to the variance ratio estimator proposed by Cochrane (1988)[10], as the aggregation value k, or S_T in our notation, grows with (but more slowly than) the sample size T. Since the sample variance is $\sqrt{T}$ consistent, by exactly the same argument as in the case of $\widehat{W}_T$, we can show that[11]

$$\sqrt{T/S_T}(\widehat{V_T} - V) \rightarrow_d N(0, (4/3)V^2), \tag{19}$$

where $S_T = O(T^g)$ with $1/3 < g < 1$. In particular, when $V = 1$, which implies that $\{y_t\}$ is a random walk process, the asymptotic distribution of $\sqrt{T/S_T}(\widehat{V_T} - 1)$ is $N(0, 4/3)$. This result is consistent with the result of Lo and MacKinlay (1988, p.47).[12] The upper 100α % confidence interval for $\widehat{V}_T$ is given by

$$\left[\frac{\widehat{V}_T}{1 + \ c_\alpha/\kappa_T}, \quad \infty\right), \tag{20}$$

where $\kappa_T \equiv \sqrt{(3T)/(4S_T)}$. With this $\widehat{V}_T$, there is no difficulty to estimate the upper bound ρ_{ub} if it exists; it can be consistently estimated by $\widehat{\rho}_{ub,T} \equiv -\sqrt{1 - \widehat{V}_T^{-1}}$ with $\widehat{V}_T > 1$.

Next, we consider the construction of confidence intervals for ρ_{ub}. When it comes to ρ_{ub}, we would be interested in how far it is away from zero. Therefore, it is more appropriate to consider lower one-sided confidence intervals. Also, since $\widehat{\rho}_{ub}$ is defined only when $\widehat{V}_T > 1$, it would be

[10]See the footnote 2.

[11]Here, we do not need the normality assumption on $\{u_t\}$, which was necessary to assure the $\sqrt{T}$ consistency of $\widehat{\sigma^2_{u,T}}$ in $\widehat{W}_T$.

[12]They showed that $\sqrt{T/q}(\widehat{M}_r(q) - 1) \sim_a N(0, 2(2q-1)(q-1)/3q^2)$, where $\widehat{M}_r(q)$ is their estimator for the ratio of the variances of q-th and the first differences of the series, which is asymptotically equivalent to $\widehat{V_T}$. Here q corresponds to S_t. When $q \rightarrow \infty$, the asymptotic variance reduces to 3/4. See also Lo and MacKinlay (1989).

reasonable to construct the confidence intervals only when $\widehat{V}_T > 1$. However, this brings an additional complication because then we have to deal with a conditional distribution conditioned on $\widehat{V}_T > 1$.

Let z be a random variable defined as $z \equiv \kappa_T V^{-1}(\widehat{V}_T - V)$. Note that the condition $\widehat{V}_T > 1$ is equivalent to $z > \kappa_T(V^{-1} - 1)$. First we will find a constant c_β such as $\beta = P(z < c_\beta \mid z > \kappa_T(V^{-1} - 1))$ for $0 < \beta < 1$. Noting that $z \sim_a N(0,1)$, this constant is given by $c_\beta = \Phi^{-1}(\beta\Phi(\kappa_T(1 - V^{-1}))$, where $\Phi(.)$ is the cumulative distribution function of the standard normal distribuion.[13] With this constant c_β, we have

$$\beta = P\left(z \leq c_\beta \,\middle|\, \widehat{V}_T > 1\right) = P\left(\widehat{V}_T \leq V + V\frac{c_\beta}{\kappa_T} \,\middle|\, \widehat{V}_T > 1\right). \tag{21}$$

Here, we use a crude approximation for c_β, namely, the first order Taylor expansion of c_β of V^{-1} around $\widehat{V}_T^{-1}$:

$$c_\beta \approx \widehat{c} - \frac{\beta\phi(\widehat{\nu})\kappa_T}{\phi(\widehat{c})}\left[\frac{1}{V} - \frac{1}{\widehat{V}_T}\right], \tag{22}$$

where $\widehat{c} \equiv \Phi^{-1}(\beta\Phi(\widehat{\nu}))$ and $\widehat{\nu} \equiv \kappa_T(1 - \widehat{V}_T^{-1})$. Substituting this into (21), we have

$$\begin{aligned}
&P\left(\widehat{V}_T \leq V + \tfrac{\widehat{c}}{\kappa_T}V - \tfrac{\beta\phi(\widehat{\nu})}{\phi(\widehat{c})} + \tfrac{\beta\phi(\widehat{\nu})}{\phi(\widehat{c})}\tfrac{V}{\widehat{V}_T} \,\middle|\, \widehat{V}_T > 1\right) \\
&= P\left(\widehat{V}_T + \tfrac{\beta\phi(\widehat{\nu})}{\phi(\widehat{c})} \leq \left[1 + \tfrac{\widehat{c}}{\kappa_T} + \tfrac{\beta\phi(\widehat{\nu})}{\phi(\widehat{c})}\tfrac{1}{\widehat{V}_T}\right]V \,\middle|\, \widehat{V}_T > 1\right). \\
&= P\left(\bar{c} \leq V \,\middle|\, \widehat{V}_T > 1\right) = P\left(\bar{c}^{-1} \geq V^{-1} \,\middle|\, \widehat{V}_T > 1\right) \\
&= P\left(1 - \bar{c}^{-1} \leq 1 - V^{-1} \,\middle|\, \widehat{V}_T > 1\right) \\
&= P\left(-\sqrt{1 - \bar{c}^{-1}} \geq -\sqrt{1 - V^{-1}} \,\middle|\, \widehat{V}_T > 1\right) \text{ or } P\left(-\sqrt{1 - \bar{c}^{-1}} \geq \rho_{ub} \,\middle|\, \widehat{V}_T > 1\right),
\end{aligned} \tag{23}$$

where

$$\bar{c} \equiv \frac{\widehat{V}_T + \frac{\beta\phi(\widehat{\nu})}{\phi(\widehat{c})}}{1 + \frac{\widehat{c}}{\kappa_T} + \frac{\beta\phi(\widehat{\nu})}{\phi(\widehat{c})}\frac{1}{\widehat{V}_T}}.$$

[13]Following the convention, we use $\Phi(.)$ for the cdf and $\phi(.)$ for the pdf of the standard normal distribution, although the same notations were used for the AR lag polynomials and the coefficients.

Note that the numerator and denominator of $\bar{c}$ are both always positive, which justifies the second and third equalities in (23). Based on this approximation, we suggest $(-1,\ -\sqrt{1-\bar{c}^{-1}}]$ as a β % lower confidence interval for ρ_{ub}. As it is obvious from the derivation, this confidence interval is valid only when ρ_{ub} exists; i.e., $V > 1$. For $1-\bar{c}^{-1}$ to be positive (otherwise, $\sqrt{1-\bar{c}^{-1}}$ becomes a complex number), $\bar{c}$ must be greater than 1. We can show that this is in fact always satisfied.[14]

To obtain this confidence interval, we have to calculate $\bar{c}$ at each time. This can be easily done numerically. The above argument relies on two approximations, (19) and (22). Although both of them become more and more accurate as the sample size increases, they may not be good approximations in a finite sample size. In the next section we examine the accuracy of this confidence interval by simulation.

4 Monte Carlo study

Using Monte Carlo experiments, we first check the finite sample properties of the proposed non-parametric estimators, namely, $\widehat{W}_T$ and $\widehat{V}_T$, and compare them with several parametric estimators derived from ARIMA models. Then, we examine the accuracy of our suggested method for constructing confidence intervals for W, V, and ρ_{ub}.

Throughout this section, we consider the following ARIMA(1,1,1) processes:

$$(1-\phi_1 L)(1-L)y_t = c + (1+\theta_1 L)u_t, \quad u_t \sim NID(0,\sigma_u^2), \quad t = 0, \pm 1, \pm 2, ... \tag{24}$$

Here, c and σ_u^2 are fixed at $c = 0.4431$ and $\sigma_u^2 = 0.9723$, which are the actual estimates of the ARIMA(1,1,1) model for the log of U.S. real GDP. The following four sets of ϕ_1 and θ_1 are

[14]First, we have $\widehat{c} \leq \kappa_T(1-\widehat{V}_T^{-1})$ since $\beta < 1$. Then, we have $1-\widehat{V}_T^{-1} \geq \widehat{c}/\kappa_T$. Since $(\widehat{V}_T-1)^2 > 0$, we have $\widehat{V}_T - 1 > 1-\widehat{V}_T^{-1}$ and thus $\widehat{V}_T - 1 > \widehat{c}/\kappa_T$ or $\widehat{V}_T > 1+\widehat{c}/\kappa_T$. Then, noting $\beta\phi(\widehat{\nu})/\phi(\widehat{c}) > 0$ and $\widehat{V}_T > 1$, we obtain the result.

examined:

$$(A)\quad (\phi_1,\,\theta_1) = \quad (0.4591,\,-0.1310) \qquad (W = 2.5811,\, V = 2.2713,\, \rho_{ub} = -0.7481),$$

$$(B)\quad (\phi_1,\,\theta_1) = \quad (0.3,\,-0.5) \qquad (W = 0.5102,\, V = 0.4887,\, \rho_{ub} \text{ does not exist}),$$

$$(C)\quad (\phi_1,\,\theta_1) = \quad (0.5,\,-0.3) \qquad (W = 1.9600,\, V = 1.8608,\, \rho_{ub} = -0.6801),$$

$$(D)\quad (\phi_1,\,\theta_1) = \quad (0.3,\,-0.29) \qquad (W = 1.0288,\, V = 1.0287,\, \rho_{ub} = -0.1669),$$

where the values inside of the parentheses are the true values of W, V, and ρ_{ub}. These values cover several important cases in which W is less than 1, close to 1 and much greater than 1.

The values of ϕ_1, θ_1 given in (A) (and c, σ_u^2) are the Gaussian ML estimates of the ARIMA(1,1,1) model for the log of U. S. quarterly real GDP data.[15] (B) and (C) are well identified cases. (D) is a near cancellation case; i.e., ϕ_1 and $-\theta$ are very close to each other, and it is known that the Gaussian MLE behaves very badly (see, e.g., Nelson and Startz, 2006). For (B), ρ may be positive or negative, and ρ_{ub} does not exist. For (A), (C) and (D), ρ must be negative and ρ_{up} exists.

We also compare our non-parametric estimators with parametric estimators computed from several ARIMA(p, 1, q) models. Specifically, we consider three ARIMA models: ARIMA(1,1,1), ARIMA(0,1,1), and ARIMA(2,1,2) models. The ARIMA(0,1,1) and ARIMA(2,1,2) models are under and over parameterized models, respectively.[16] They are used to address the concerns from Cochrane (1988) and Hauser *et al.* (1999).[17] Let $\widehat{W}_T(p,q)$ denote the parametric estimator from the ARIMA(p, 1, q) model.

[15]The quarterly real GDP data are the same as the data used in Morley, Nelson, and Zivot (2003) and Oh at al. (2006). Details of the data will be given in Section 4.

[16]Strictly speaking, if the true process is ARIMA(1,1,1) model, ARIMA(2,1,2) parameters are not uniquely identified and the parametric estimator of persistence with this model is not consistent. See Hauser *et al.* (1999) for more details on the effect of overparameterizing ARIMA model.

[17]These ARIMA models are estimated by exact MLE using the Kalman filter. The program was written by the Matlab programming language version 6.1. The program uses Matlab command

For the BT kernel, setting $g = 1/3$ in (13) is asymptotically optimal for estimating the long-run variance, however, as argued in Section 3.2, this could lead to inaccurate confidence intervals and even may not be optimal in finite samples. With this in mind, we examine several values for the growth rate of the bandwidth parameter; specifically, we include $g = 1/2$ and $g = 2/3$ [18] additionally. We denote the nonparametric estimator with the growth rate g by $\widehat{W}_{T,g}$; for example, when $g = 1/2$, it is denoted by $\widehat{W}_{T,1/2}$.

The Monte Carlo experiment is conducted as follows:

(1) Generate T samples from the ARIMA(1,1,1) processes in (24) with coefficients specified by (A)-(D).

(2) Estimate W and V using parametric and nonparametric estimators.

(3) Construct 100α % upper confidence intervals for V and W by the method suggested in Section 3.

(4) For Case (A), (C), and (D), if $\widehat{V}_{T,g} > 1$, then construct β % lower confidence intervals for ρ_{ub} by the method suggested in Section 3. If $\widehat{V}_{T,g} \leq 1$, regenerate T samples and estimate V. Repeat this until we can construct the confidence interval.

(5) Repeat (1)-(4) R times.

(6) Calculate the mean, bias, mean absolute error (MAE) of the estimators and the actual coverage probabilities of the confidence intervals.

Here, we set $T = 100, 200,$ and $R = 1000$ for the parametric estimators; $T = 100, 200, 1000,$

`fminunc` for the maximization, which is based on a quasi-Newton method with the numerical Hessian updated by the BFGS algorithm.

[18] Under our assumptions, (12) holds with these values of g. See Theorem 1(b) in Andrews (1991).

and $R = 10000$ for the nonparametric estimators; and $\alpha, \beta = 0.01, 0.05, 0.10$. Tables 1 and 2 report the mean, bias and mean absolute error (MAE) of the estimators for W and V, respectively; we report MAE instead of mean square error (MSE) since these two have essentially the same implications.

First, consider the results for parametric estimators of W. We can observe that the parametric estimators constructed from the misspecified ARIMA models, $\widehat{W}_T(0,1)$, and $\widehat{W}_T(2,2)$ tend to have larger biases than the one computed from the correctly specified ARIMA model, i.e., $\widehat{W}_T(1,1)$. Interestingly, in Case (D), $\widehat{W}_T(1,1)$ has the largest bias among the three. Also the biases of $\widehat{W}_T(0,1)$ and $\widehat{W}_T(2,2)$ are not always negative; they are so in three cases, but are positive in Case (B). In terms of MAE, the performance of $\widehat{W}_T(2,2)$ is usually inferior to the other two except in Case (D). Also the performances of $\widehat{W}_T(0,1)$ are better than those of $\widehat{W}_T(1,1)$ in Cases (B) and (D). The results for the estimator of V parallel those for the estimator of W. From these limited experiments, we may say that if our objective is to estimate W or V, under-parameterization does not pose a severe problem, though it would depend on the underlying true process, and over-parameterizing would be more problematic.

Next, consider the nonparametric estimators of W and V. We find that their performances are comparable to those of parametric ARIMA estimators; they indeed often perform better than the parametric ones, especially in Case (D). Choice of the growth rate does not seem to give much differences here, although $\widehat{W}_{T,1/3}$ and $\widehat{W}_{T,1/2}$ seem to perform slightly better than $\widehat{W}_{T,2/3}$, except in Case (B). We note that the models we examined here are very simple processes. For different processes, these estimators may perform very differently. Then the selection of the bandwidth parameters may become more important. The above comments apply to both $\widehat{W}_{T,g}$ and $\widehat{V}_{T,g}$.

Last, we check the accuracies of the confidence intervals for W, V and ρ_{ub}. Tables 3 and 4 report the empirical coverage probabilities of those confidence intervals. Again almost the same comments apply to both W and V. The empirical coverage probabilities tend to be higher than the nominal levels except in Case (B). Although as the sample size increases, the accuracies improve, the improvements are quite slow in some cases; confidence intervals constructed with $g = 1/3$ are quite inaccurate and the accuracy does not improve as the number of sample size increases, especially in Case (B). This is because for this process, the neglected mean is very large and thus the asymptotic approximation by the zero mean normal distribution does not work well. This result indicates that, at least when we use the asymptotic normal approximation for constructing confidence intervals, we have to be careful in choosing the growth rate of bandwidth parameter; it should be larger than the asymptotically optimal rate. Regarding the confidence intervals for ρ_{ub}, we find that their empirical coverage probabilities also have a tendency to be higher than the nominal level and they becomes quite inaccurate when the true value of ρ_{ub} is close to 0 or the true value of V is close to 1.

5 Empirical application to U. S. real GDP

In this section we apply our non-parametric estimators to the growth rate of US quarterly real GDP over the period 1947.1-1998.2 (sample size is 205). The data is the same as used in MNZ. Assuming real GDP is a difference stationary process defined in (1), we estimate the values of W, V, and ρ_{ub} by parametric and nonparametric approaches. We apply 6 estimators for each of W, V, and ρ_{ub}: $\widehat{W}_T(p,q)$ and $\widehat{V}_T(p,q)$ with $(p,q) = (0,1), (1,1)$, and $(2,2)$; $\widehat{W}_{T,g}$ and $\widehat{V}_{T,g}$ with $g = 1/3, 1/2$ and $2/3$. The estimation results are summarized in Table 5. For both W and V, all estimates are greater than 1, ranging from 1.3160 to 2.5811 for W and from 1.1315 to 2.22713 for V.[19] The values of W and V implied from the estimated parameters of ARIMA(2,1,2) reduced model in MNL are 1.6275 and 1.3878, which are in fact almost the same as the estimates of W and

[19]We select the bandwidth parameter following the formula in (13); it gives $S_T = 5.7720$ when $g = 1/3$, $S_T = 12.9612$ when $g = 1/2$, and $S_T = 29.1047$ when $g = 2/3$.

V by ARIMA(2,1,2) model here ($\widehat{W}_T(2,2) = 1.6278$ and $\widehat{V}_T(2,2) = 1.3880$). MNZ could estimate ρ since their UC model with AR(2) cycle is identified. The advantage of the nonparametric approach here for estimating the upper bound is that it does not require a parametric cycle nor that the UC model is identified. The estimates of ρ_{ub} range from -0.3409 to -0.7481.

Next, we construct 90, 95, and 99 % confidence intervals for W, V, and ρ_{ub}. The results are summarized in Table 6. The 95 % confidence intervals for W with $\widehat{W}_{T,1/3}$ and $\widehat{W}_{T,1/2}$ do not include $W = 1$. However, all of the the confidence intervals computed from $\widehat{W}_{T,2/3}$ include $W = 1$. Unfortunately, the conclusion about the true value of W depends on the value of the bandwidth parameter. The results for V are similar to those for W except that the 95% confidence interval with $\widehat{V}_{T,1/2}$ barely includes 1. The confidence intervals for ρ_{ub} indicate that the correlation is at least lower than -0.254, if it is negative. Note that the confidence intervals for ρ_{ub} are obtained under the assumption that it exists or $V > 1$, which may not be true.

6 Conclusion

In this paper we extended the theorem in Lippi and Reichlin (1992) to the UC model with correlated components. It was shown that the square of the impulse response measure has an important implication for the correlation between shocks in trend and cycle; if it is greater than or equal to one, then the correlation must be negative. Furthermore, we derived an upper bound for the correlation, which is a function of the variance ratio measure, and we suggested non-parametric estimators for the square of the impulse response measure and the upper bound.

A method for constructing confidence intervals of the proposed estimators was also discussed. It is based on an asymptotic normal approximation. We investigated properties of the estimators and the accuracies of the confidence intervals by Monte Carlo experiments. Our Monte Carlo experiments indicate that the choice of bandwidth parameter is important; in general, the bandwidth parameter must grow faster than its asymptotically optimal rate for constructing accurate confidence intervals.

The estimators are applied to U. S. quarterly real GDP data. It is found that the estimates by various methods, parametrically and non-parametrically, are all greater than 1, but one of the valid upper 95% confidence intervals does not include the value of 1. Hence, we find compelling evidence that the correlation between shocks in trend and cycle in the UC model is negative for U. S. quarterly real GDP data.

In the definition of the UC model, we assumed that the trend is a random walk. Several authors (e.g., Blanchard and Quah, 1989; Lippi and Reichlin, 1994; Quah, 1992) considered UC models in which the trend follows a general I(1) process. It would be interesting to extend our analysis to this larger class of UC models and see whether a similar result is obtained.

Appendix: Proofs

In this appendix, first we show that our assumptions actually satisfy the assumptions made in Hannan and Nicholls (1977). Second, we prove Proposition 1 and 2. Our proof of Proposition 1 is similar to that of the theorem in Lippi and Reichlin (1992).

A function $f(\omega)$ is said to belong to Λ_α, $0 < \alpha \leq 1$ if $\sup_\omega |f(\omega + d) - f(\omega)| \leq C|d|^\alpha$ with some constant $C > 0$ independent of d. In addition to normality of u_t, Hannan and Nicholls (1977) assumed that $f_x(\omega) \in \Lambda_\alpha$ with $\alpha > \frac{1}{2}$ to prove the asymptotic normality and $\sqrt{T}$ convergence of $\widehat{\sigma_u^2}$. Our assumption that $s(1) < \infty$ satisfies the above assumption as the following proposition shows.

Proposition 3 *If $s(q) < \infty$ for some $0 < q \leq 1$, where $s(q) \equiv \sum_{j=-\infty}^{\infty} |j|^q |\gamma_j|$, then $f_x(\omega) \in \Lambda_q$.*

Proof. From the definition of the spectral density function, we have

$$\begin{aligned}
&|f(\omega + d) - f(\omega)| = \tfrac{1}{2\pi} \left| \sum_{j=-\infty}^{\infty} \gamma_j \left[\cos(j\omega + jd) - i \sin(j\omega + jd) - \cos(j\omega) + i \sin(j\omega)\right] \right| \\
&= \tfrac{1}{2\pi} \left| \sum_{j=-\infty}^{\infty} \gamma_j \left[-2 \sin\left(\tfrac{2j\omega + jd}{2}\right) \sin\left(\tfrac{jd}{2}\right) + i2 \cos\left(\tfrac{2j\omega + jd}{2}\right) \sin\left(\tfrac{jd}{2}\right) \right] \right| \\
&\leq \tfrac{1}{2\pi} \sum_{j=-\infty}^{\infty} |\gamma_j| \left[4 \sin^2\left(\tfrac{2j\omega + jd}{2}\right) \sin^2\left(\tfrac{jd}{2}\right) + 4 \cos^2\left(\tfrac{2j\omega + jd}{2}\right) \sin^2\left(\tfrac{jd}{2}\right) \right]^{1/2} \\
&= \tfrac{2}{\pi} \sum_{j=-\infty}^{\infty} |\gamma_j| \left| \sin\left(\tfrac{jd}{2}\right) \right|,
\end{aligned}$$

where we used the well known formulas: $\cos(x) - \cos(y) = -2 \sin\left(\frac{x+y}{2}\right) \sin\left(\frac{x-y}{2}\right)$, and $\sin(x) - \sin(y) = -2 \cos\left(\frac{x+y}{2}\right) \sin\left(\frac{x-y}{2}\right)$. Noting that $|\sin(x)| \leq |x|^\alpha$ for any $0 \leq \alpha \leq 1$,[20] we have

$$|f(\omega + d) - f(\omega)| \leq \frac{2^{1-\alpha}}{\pi} |d|^\alpha \sum_{j=-\infty}^{\infty} |j|^\alpha |\gamma_j|$$

[20] This obviously holds for $|x| > 1$. It also holds for $|x| \leq 1$ because $\sin(x) \leq |x| \leq |x|^\alpha$ with $0 \leq \alpha \leq 1$.

Therefore, if $s(q) < \infty$ for some $0 < q \le 1$, then

$$|f(\omega + d) - f(\omega)| \le C|d|^q, \text{ where } \quad C = \frac{2^{1-q}}{\pi} s(q). \quad \text{Q.E.D.}$$

Next, we shall prove Proposition 1. Let Δy_t be rewritten as

$$\Delta y_t = \Delta(\tau_t + c_t) = \mu + \eta_t + C(L)\epsilon_t, \tag{25}$$

where $C(L) = \Delta B(L)$ and $C(L) = 1 + c_1 L + c_2 L^2 + \cdots$. It is easy to show that if coefficients of $B(L)$ are absolutely summable, so are coefficients of $C(L)$. Note that $c_j = b_j - b_{j-1}$ for j =1,... with $b_0 = 1$ and so $C(1) = 0$.

Here, we derive the spectral density function of Δy_t.

Lemma 1 *The spectral density function of Δy_t in (25) is given by*

$$f_{\Delta y}(\omega) = \frac{1}{2\pi}\left\{\sigma_\eta^2 + \sigma_{\eta\epsilon}[C(e^{i\omega}) + C(e^{-i\omega})] + \sigma_\epsilon^2 C(e^{i\omega})C(e^{-i\omega})\right\}. \tag{26}$$

Proof Rewrite Δy_t as $\Delta y_t = \mu + \mathbf{i}'\mathbf{v}_t$, where $\mathbf{i} = (1,1)'$ and $\mathbf{v}_t = (\eta_t, C(L)\epsilon_t)'$. $\mathbf{v}_t$ can be represented by the following vector MA(∞):

$$\begin{aligned}
\mathbf{v}_t \equiv \begin{bmatrix} \eta_t \\ C(L)\epsilon_t \end{bmatrix} &= \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}\begin{bmatrix} \eta_t \\ \epsilon_t \end{bmatrix} + \begin{bmatrix} 0 & 0 \\ 0 & c_1 \end{bmatrix}\begin{bmatrix} \eta_{t-1} \\ \epsilon_{t-1} \end{bmatrix} + \begin{bmatrix} 0 & 0 \\ 0 & c_2 \end{bmatrix}\begin{bmatrix} \eta_{t-2} \\ \epsilon_{t-2} \end{bmatrix} + \cdots \\
&= \mathbf{I}_2\mathbf{e}_t + \mathbf{C}_1\mathbf{e}_{t-1} + \mathbf{C}_2\mathbf{e}_{t-2} + \cdots \\
&= \mathbf{C}(L)\mathbf{e}_t,
\end{aligned}$$

where $\mathbf{e}_t = (\eta_t, \epsilon_t)'$ with variance-covariance matrix $\mathbf{\Sigma}$. The autocovariance generating function of Δy_t is defined as $g_{\Delta y}(z) = \sum_{j=-\infty}^{\infty} \gamma_j z^j$, where γ_j is the j-th order autocovariance of Δy_t. Noting that $\gamma_j = \mathrm{cov}(\Delta y_t, \Delta y_{t-j}) = E(\mathbf{i}'\mathbf{v}_t\mathbf{v}'_{t-j}\mathbf{i}) = \mathbf{i}'E(\mathbf{v}_t\mathbf{v}'_{t-j})\mathbf{i}$, we have

$$
\begin{aligned}
g_{\Delta y}(z) &= \textstyle\sum_{j=-\infty}^{\infty} \mathbf{i}'E(\mathbf{v}_t\mathbf{v}'_{t-j})\mathbf{i}z^j \\
&= \mathbf{i}'\left[\textstyle\sum_{j=-\infty}^{\infty} E(\mathbf{v}_t\mathbf{v}'_{t-j})z^j\right]\mathbf{i} \\
&= \mathbf{i}'\left[\mathbf{C}(z)\mathbf{\Sigma}\mathbf{C}(z^{-1})'\right]\mathbf{i} \\
&= \sigma_\eta^2 + \sigma_{\eta\epsilon}\left[C(z) + C(z^{-1})\right] + \sigma_\epsilon^2 C(z)C(z^{-1}).
\end{aligned} \tag{27}
$$

Thus, the spectral density function of Δy_t is given in (26). Q.E.D.

From the above lemma, we immediately have $A(1)^2\sigma_u^2 = 2\pi f_{\Delta y}(0) = \sigma_\eta^2$, the result already shown in Cochrane (1988).

Proof of Proposition 1. First we consider the case that $\rho\delta \neq -1$. Then, by Kolmogorov's formula[21]

$$
\begin{aligned}
\sigma_u^2 &= \exp\tfrac{1}{2\pi}\textstyle\int_{-\pi}^{\pi}\log 2\pi f_{\Delta y}(\omega)\mathrm{d}\omega \\
&= \exp\tfrac{1}{2\pi}\textstyle\int_{-\pi}^{\pi}\log\left[\sigma_\eta^2\left(1 + \frac{\sigma_{\epsilon\eta}}{\sigma_\eta^2}\left[C(e^{i\omega}) + C(e^{-i\omega})\right] + \frac{\sigma_\epsilon^2}{\sigma_\eta^2}C(e^{i\omega})C(e^{-i\omega})\right)\right]\mathrm{d}\omega \\
&= \sigma_\eta^2\exp\tfrac{1}{2\pi}\textstyle\int_{-\pi}^{\pi}\log\left\{1 + \rho\delta\left[C(e^{i\omega}) + C(e^{-i\omega})\right] + \delta^2 C(e^{i\omega})C(e^{-i\omega})\right\}\mathrm{d}\omega \\
&> \sigma_\eta^2\exp\tfrac{1}{2\pi}\textstyle\int_{-\pi}^{\pi}\log\left\{1 + \rho\delta\left[C(e^{i\omega}) + C(e^{-i\omega})\right] + (\rho\delta)^2 C(e^{i\omega})C(e^{-i\omega})\right\}\mathrm{d}\omega \\
&= \sigma_\eta^2\exp\tfrac{1}{2\pi}\textstyle\int_{-\pi}^{\pi}\log\left\{\left[1 + \rho\delta C(e^{i\omega})\right]\left[1 + \rho\delta C(e^{-i\omega})\right]\right\}\mathrm{d}\omega \\
&= \sigma_\eta^2\exp\tfrac{1}{2\pi}\textstyle\int_{-\pi}^{\pi}\log\left\{2\pi\tfrac{1}{2\pi}(1+\rho\delta)^2\left[1 + (1+\rho\delta)^{-1}\rho\delta c_1 e^{i\omega} + (1+\rho\delta)^{-1}\rho\delta c_2 e^{2i\omega} + \cdots\right]\right. \\
&\qquad\qquad \left.\times\left[1 + (1+\rho\delta)^{-1}\rho\delta c_1 e^{-i\omega} + (1+\rho\delta)^{-1}\rho\delta c_2 e^{-2i\omega} + \cdots\right]\right\}\mathrm{d}\omega \\
&= \sigma_\eta^2(1+\rho\delta)^2
\end{aligned} \tag{28}
$$

The inequality in the forth line comes from that $C(e^{i\omega})C(e^{-i\omega})$ is positive almost everywhere in $[-\pi, \pi]$ and $|\rho| < 1$.The last equality is obtained by regarding the inside of the logarithm as the spectral density of MA(∞) process with the prediction error variance $(1+\rho\delta)^2$ and coefficients $(1+\rho\delta)^{-1}\rho\delta c_j$ $j = 1, ...,$ which can be shown to be absolutely summable, and then applying

[21] Note that 2π in the formula cancels out the 2π of the spectral density function.

Kolmogorov's formula. Note that $1 + \rho\delta \neq 0$ since $\rho\delta \neq -1$ is assumed here. Substituting $\sigma_\eta^2 = A(1)^2\sigma_u^2$ into the above, we obtain the desired inequality. Care must be given when $\rho\delta = -1$. In this case, $1 + \rho\delta = 0$ and so the above argument does not apply. However a simple modification can still give a similar result.

Note that when $\rho\delta = -1$, we have $1 > 1/\delta$ since $\rho = -1/\delta$ and $|\rho| < 1$. From the third line in (28), we have

$$\begin{aligned}
\sigma_u^2 &= \sigma_\eta^2 \exp \tfrac{1}{2\pi} \textstyle\int_{-\pi}^{\pi} \log \left\{1 - \left[C(e^{i\omega}) + C(e^{-i\omega})\right] + \delta^2 C(e^{i\omega})C(e^{-i\omega})\right\} \mathrm{d}\omega \\
&> \sigma_\eta^2 \exp \tfrac{1}{2\pi} \textstyle\int_{-\pi}^{\pi} \log \left\{1/\delta^2 - \left[C(e^{i\omega}) + C(e^{-i\omega})\right] + \delta^2 C(e^{i\omega})C(e^{-i\omega})\right\} \mathrm{d}\omega \\
&= \sigma_\eta^2 \exp \tfrac{1}{2\pi} \textstyle\int_{-\pi}^{\pi} \log \left\{\left[1/\delta - \delta C(e^{i\omega})\right]\left[1/\delta - \delta C(e^{-i\omega})\right]\right\} \mathrm{d}\omega \\
&= \sigma_\eta^2 \exp \tfrac{1}{2\pi} \textstyle\int_{-\pi}^{\pi} \log \left\{[(1-\delta^2)^2/\delta^2]\left[1 - (1-\delta^2)^{-1}\delta^2 c_1 e^{i\omega} - (1+\delta^2)^{-1}\delta^2 c_2 e^{2i\omega} + \cdots\right]\right. \\
&\qquad\qquad \left. \times \left[1 - (1-\delta^2)^{-1}\delta^2 c_1 e^{-i\omega} - (1+\delta^2)^{-1}\delta^2 c_2 e^{-2i\omega} + \cdots\right]\right\} \mathrm{d}\omega \\
&= \sigma_\eta^2 (1-\delta^2)^2/\delta^2.
\end{aligned}$$

Again the last line is obtained by regarding the inside of logarithm as 2π times the spectral density function of MA(∞) process with prediction error variance $(1-\delta^2)^2/\delta^2$. Thus, we have $\delta^2/(1-\delta^2)^2 > A(1)^2$ when $\rho\delta = -1$. Q.E.D.

Proof of Proposition 2 Taking the variances of both sides in (25), we have $\sigma_{\Delta y}^2 = \sigma_{lrv}^2 + 2\sigma_{\eta\epsilon} + \omega\sigma_\epsilon^2$, or $V^{-1} = 1 + 2\rho\delta + \omega\delta^2$, where $\omega \equiv 1 + c_1^2 + c_2^2 + \cdots$. From the absolute summability of b_j $j = 1, 2, ...$, we have $\omega \neq 1$ and thus $\omega > 1$. Therefore, it follows that $V^{-1} > 1 + 2\rho\delta + \delta^2$. Solving this inequality for δ, we have $-\rho - \sqrt{\rho^2 - (1 - V^{-1})} < \delta < -\rho + \sqrt{\rho^2 - (1 - V^{-1})}$. For this inequality to have a solution for δ, it must be satisfied that $\rho^2 - 1 + V^{-1} > 0$. When $V < 1$, this is trivially satisfied. When $V > 1$, [22] which implies that $\rho < 0$, we have $\rho < -\sqrt{1 - V^{-1}}$. Q.E.D.

[22]We do not need to consider the case $V = 1$; we have $\delta > 0$ under the assumption $\sigma_\epsilon^2 > 0$, however if $V = 1$ then the above inequality for δ implies $\delta < 0$, which is a contradiction and

References

An H-Z. 1981. Asymptotic properties of an estimate of the prediction error variance. *Australian Journal of Statistics* **23**: 164-176.

Andrews DWK. 1991. Heterskedasticiy and autocorrelation consistent covariance matrix estimation. *Econometrica* **59**: 817-858.

Beveridge S, Nelson CR. 1981. A new approach to the decomposition of economic time series into permanent and transitory component with particular attention to measurement of the business cycle, *Journal of Monetary Economics* **7**: 151-154.

Bhansali RJ. 1974. Asymptotic properties of the Wiener Kolmogorov predictor I. *Journal of the Royal Statistical Society. Series B(Methodological)* **36**: 61-73.

Blanchard OJ, Quah D. 1989. The dynamic effects of demand and supply disturbances. *American Economic Review* **79**: 655-673.

Campbell JY, Mankiw NS. 1987a. Are output fluctuations transitory? *Quarterly Journal of Economics* **102**: 857-880.

Campbell JY, Mankiw NS. 1987b. Permanent and transitory components in macroeconomic fluctuations. *American Economic Review (Papers and Proceedings)* **77**: 111-117.

Chen Z-G, Hannan EJ. 1980. The distribution of periodogram ordinates. *Journal of Time Series Analysis* **1**: 73-82.

Cochrane JH. 1988. How big is the random walk component in the GNP? *Journal of Political Economy* **96**: 893-920.

---

therefore $V \neq 1$.

Christiano JL, Eichenbaum M. 1989. Unit roots in real GNP: do we know, and do we care? NBER working paper no. 3130.

Clark PK. 1987. The cyclical component of U.S. economic activity. *Quarterly Journal of Economics* **102**: 797-814.

Davis HT, and Jones RH. 1968. Estimation of the innovation variance of a stationary time series. *Journal of the American Statistical Association* **63**: 141-149.

Hamilton DJ. 1994. *Time series analysis*. Princeton University Press: Princeton, NJ.

Hannan EJ. 1970. *Multiple time series*. Wiley: New York, NY.

Hannan EJ, Nicholls DF. 1977. The estimation of the prediction error variance. *Journal of the American Statistical Assocation* **72**: 834-840.

Hannan EJ, Nicholls DF. 1979. Corrigenda: The estimation of the prediction error variance. *Journal of the American Statistical Assocation* **74**: 94.

Harvey AC, Koopman SJ. 2000. Signal extrction and the formulation of unobserved components models. *Econometrics Journal* **3**: 84-107.

Hauser AM, Pötscher BM, Reschenhofer E. 1999. Measuring persistence in aggregate output: ARMA models, fractionally integrated ARMA models and nonparametric procedures. *Empirical Economics* **24**: 243-269.

Janacek, G. 1975. Estimation of the minimum mean square error of prediction. *Biometrika* **62** 175-180.

Lippi M, Reichlin L. 1992. On persistence of shocks to economic variables: a common misconception. *Journal of Monetary Economics* **29**: 87-93.

Lippi M, Reichlin L. 1994. Diffusion of technical change and the decomposition of output into trend anc cycle. *The Review of Economic Studies* **61**: 19-30.

Morley JC, Nelson CR, Zivot E. 2003. Why are the Beveridge-Nelson and unobserved components decompositions of GDP so different? *Review of Economics and Statistics* **85**: 235-243.

Mohanty R, Pourahamadi M. 1996. Estimation of the generalized prediction error variance of a multiple time series. *Journal of the American Statistical Association* **91**: 294-299.

Lo AW, MacKinlay AC. 1988. Stock market prices do not follow random walks: evidence from a simple specification test. *The Review of Financial Studies* **1**: 41-66.

Lo AW, MacKinlay AC. 1989. The size and power of the variance ratio test in finite samples a Monte Carlo investigation. *Journal of Econometrics* **40**: 203-238.

Nelson CR, Startz R. 2006. The zero-information-limit condition and spurious inference in weakly identified models. forthcoming in *Journal of Econometrics*.

Newey WK, West KD. 1987. A simple, positive semi-definite, heteroskedasticity and autocorrelation consistent covarince matrix. *Econometrica* **55**: 703-708.

Newey WK, West KD. 1994. Automatic lag selection in covariance matrix estimation. *Review of Economic Studies* **61**: 631-53.

Oh K, Zivot E, Creal D. 2006. Identification and correlation in the Beveridge-Nelson decomposition. University of Washington, working paper.

Percival DB, Walden AT. 1993. *Spectral Anaslysis for Physical Applications: Multitapper and Conventional Univariate Techniques*. Cambridge University Press: Cambridge.

Priestley MB. 1981. *Spectral Analysis and Time Series*. Academic Press: New York.

Proiett T. 2006. Trend-cycle decompositions with correlated components. *Econometric Revies* **25**: 61-84.

Pukkila T, Nyquist H. 1985. On the frequency domain estimation of the innovation variance of a stationary univariate time series. *Biometrica* **72**: 317-323.

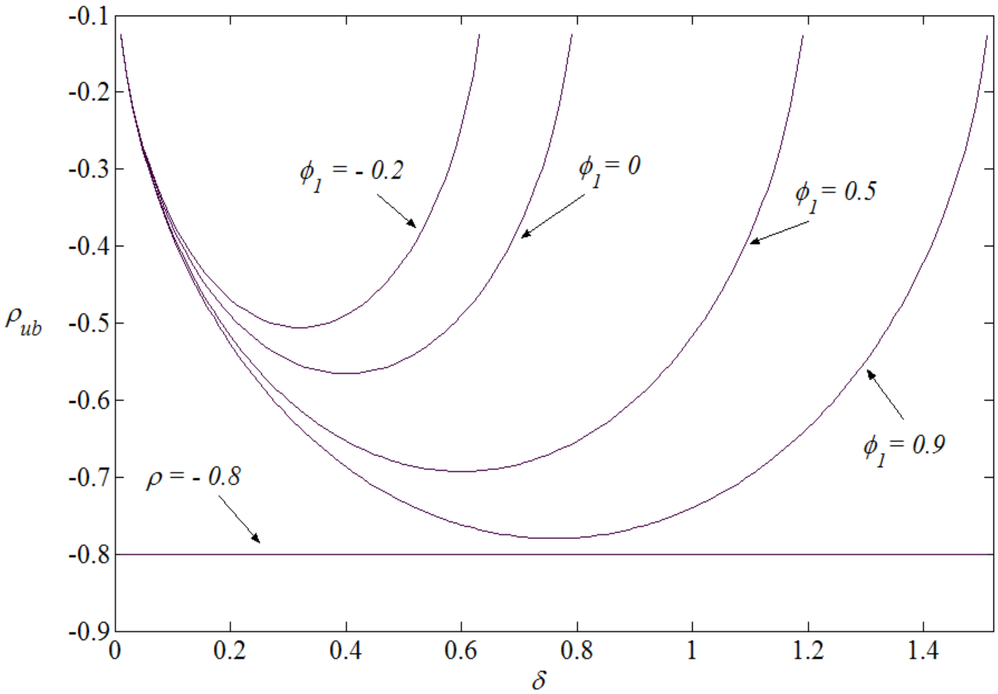
Quah D. 1992. The relative importance of permanent and transitory components: Identification and some theoritical bounds. *Econometrica* **60**: 107-118.

Taniguchi M. 1980. On estimation of the integrals of certain functions of spectral density. *Journal of Applied Probability* **17**: 73-83.

Walden AT. 1995. Multitaper estimation of the innovation variance of a stationary time series. *IEEE Transactions on Signal Processing 43*: 181-187.

Watson MW. 1986. Univariate deterending methods with stochastic trends. *Journal of Monetary Economics* **18**: 49-75.

Figure 1: Implied upper bound ρ_{ub} of $\rho = -0.8$ with various values of ϕ_1



Note: The figure plots the function $\rho_{ub} = -\sqrt{2\delta[\ |\rho| - \delta(1+\phi)^{-1}]}$ in (6) with various values of ϕ_1. Here the true value of the correlation is fixed at -0.8.

Table 1: Mean, Bias, and Mean absolute error (MAE) of Parametric and Nonparametric estimators for W

(a) Parameteric estimator ($\widehat{W}_T(p,q)$)

			$\widehat{W}_T(0,1)$			$\widehat{W}_T(1,1)$			$\widehat{W}_T(2,2)$		
ϕ_1	θ_1	T	Mean	Bias	MAE	Mean	Bias	MAE	Mean	Bias	MAE
0.4591	−0.1310	100	1.6722	−0.9089	0.9089	2.7774	0.1963	0.9436	1.8198	−0.7613	1.5855
(W = 2.5811)		200	1.6779	−0.9032	0.9032	2.6466	0.0655	0.5957	1.9569	−0.6242	1.4420
0.3	−0.5	100	0.6088	0.0986	0.1637	0.3693	−0.1409	0.2592	0.5888	0.0786	0.3650
(W=0.5102)		200	0.6143	0.1041	0.1286	0.4420	−0.0682	0.1631	0.5535	0.0433	0.3127
0.5	−0.3	100	1.4001	−0.5599	0.5621	2.1507	0.1907	0.7914	1.4321	−0.5280	1.0959
(W=1.96)		200	1.4050	−0.5550	0.5551	2.0621	0.1021	0.5200	1.5736	−0.3864	0.9641
0.3	−0.29	100	1.0094	−0.0193	0.1704	0.9430	−0.0858	0.4709	1.0128	−0.0160	0.3382
(W=1.0288)		200	1.0190	−0.0098	0.1145	0.9923	−0.0365	0.3403	0.9861	−0.0427	0.2230

(b) Non-parametric estimator ($\widehat{W}_{T,g}$)

			$\widehat{W}_{T,1/3}$			$\widehat{W}_{T,1/2}$			$\widehat{W}_{T,2/3}$		
ϕ_1	θ_1	T	Mean	Bias	MAE	Mean	Bias	MAE	Mean	Bias	MAE
0.4591	−0.1310	100	1.9161	−0.5935	0.7662	2.0460	−0.4636	0.8166	1.9202	−0.5894	0.9782
(W = 2.5811)		200	2.0550	−0.4546	0.5872	2.2031	−0.3065	0.6456	2.1001	−0.4090	0.8621
		1000	2.2856	−0.2240	0.3076	2.4115	−0.0981	0.4003	2.3292	−0.1804	0.6554
0.3	−0.5	100	0.7647	0.2545	0.2597	0.6870	0.1768	0.2119	0.6250	0.1148	0.2022
(W=0.5102)		200	0.7259	0.2157	0.2174	0.6394	0.1292	0.1545	0.5775	0.0673	0.1435
		1000	0.6375	0.1273	0.1276	0.5624	0.0522	0.0710	0.5217	0.0115	0.0813
0.5	−0.3	100	1.4224	−0.5376	0.6164	1.5293	−0.4307	0.6163	1.5430	−0.4170	0.6700
(W=1.96)		200	1.5038	−0.4562	0.5059	1.6378	−0.3222	0.4796	1.6702	−0.2898	0.5413
		1000	1.6863	−0.2737	0.2929	1.8212	−0.1388	0.2649	1.8367	−0.1233	0.3704
0.3	−0.29	100	0.9973	−0.0315	0.1361	0.9962	−0.0326	0.1503	0.9909	−0.0379	0.1642
(W=1.0288)		200	1.0033	−0.0255	0.0949	1.0044	−0.0243	0.1051	1.0027	−0.0260	0.1149
		1000	1.0104	−0.0183	0.0453	1.0129	−0.0158	0.0499	1.0146	−0.0142	0.0544

Note: The number of iterations is 1000 for the parametric estimators and 10000 for the nonparametric estimators. Given T samples, $(x_1, ..., x_T)$, Mean$= T^{-1}\sum_{t=1}^{T} x_t$; Bias = Mean−true value; MAE$= T^{-1}\sum_{t=1}^{T} |x_t - \text{true value}|$.

Table 2: Mean, Bias, and Mean absolute error (MAE) of Parametric and Nonparametric estimators for V

(a) Parameteric estimator ($\widehat{V}_T(p,q)$)

			$\widehat{V}_T(0,1)$			$\widehat{V}_T(1,1)$			$\widehat{V}_T(2,2)$		
ϕ_1	θ_1	T	Mean	Bias	MAE	Mean	Bias	MAE	Mean	Bias	MAE
0.4591	-0.1310	100	1.5230	-0.7483	0.7483	2.3639	0.0926	0.6904	1.4962	-0.7751	1.3518
(V=2.2713)		200	1.5356	-0.7357	0.7357	2.2975	0.0262	0.4443	1.6780	-0.5933	1.2309
0.3	-0.5	100	0.5810	0.0923	0.1717	0.3478	-0.1409	0.2548	0.5240	0.0353	0.3215
(V= 0.4887)		200	0.5869	0.0982	0.1319	0.4200	-0.0687	0.1617	0.5150	0.0263	0.2909
0.5	-0.3	100	1.3378	-0.5230	0.5230	1.9808	0.1200	0.6800	1.2600	-0.6008	1.0403
(V= 1.8608)		200	1.3507	-0.5101	0.5101	1.9317	0.0709	0.4499	1.4491	-0.4117	0.8988
0.3	-0.29	100	0.9984	-0.0303	0.1670	0.9226	-0.1061	0.4610	0.9251	-0.1036	0.3198
(V= 1.0287)		200	1.0137	-0.0150	0.1125	0.9819	-0.0468	0.3355	0.9419	-0.0868	0.2166

(b) Non-parametric estimator ($\widehat{V}_{T,g}$)

			$\widehat{V}_{T,1/3}$			$\widehat{V}_{T,1/2}$			$\widehat{V}_{T,2/3}$		
ϕ_1	θ_1	T	Mean	Bias	MAE	Mean	Bias	MAE	Mean	Bias	MAE
0.4591	-0.1310	100	1.7032	-0.5681	0.6181	1.8185	-0.4528	0.6445	1.7155	-0.5559	0.8157
(V=2.2713)		200	1.8184	-0.4529	0.4920	1.9492	-0.3221	0.5219	1.8623	-0.4090	0.7350
		1000	2.0127	-0.2586	0.2857	2.1236	-0.1477	0.3460	2.0518	-0.2195	0.5835
0.3	-0.5	100	0.7489	0.2602	0.2663	0.6746	0.1859	0.2205	0.6152	0.1265	0.2093
(V= 0.4887)		200	0.7038	0.2151	0.2174	0.6207	0.1320	0.1581	0.5612	0.0725	0.1458
		1000	0.6119	0.1232	0.1236	0.5399	0.0511	0.0702	0.5008	0.0121	0.0794
0.5	-0.3	100	1.3642	-0.4966	0.5245	1.4643	-0.3965	0.5112	1.4793	-0.3815	0.5668
(V= 1.8608)		200	1.4352	-0.4255	0.4438	1.5617	-0.2991	0.4033	1.5935	-0.2673	0.4671
		1000	1.6015	-0.2593	0.2675	1.7295	-0.1313	0.2332	1.7443	-0.1165	0.3381
0.3	-0.29	100	1.0119	-0.0168	0.0831	1.0108	-0.0179	0.0994	1.0055	-0.0232	0.1141
(V= 1.0287)		200	1.0111	-0.0176	0.0614	1.0122	-0.0164	0.0732	1.0105	-0.0181	0.0834
		1000	1.0114	-0.0173	0.0339	1.0139	-0.0148	0.0391	1.0155	-0.0132	0.0442

Note: The number of iterations is 1000 for the parametric estimators and 10000 for the nonparametric estimators. Given T samples, $(x_1, ..., x_T)$, Mean= $T^{-1}\sum_{t=1}^{T} x_t$; Bias = Mean$-$true value; MAE= $T^{-1}\sum_{t=1}^{T} |x_t - \text{true value}|$.

Table 3: Empirical coverage probabilities of confidence intervals for W and V

For W

			$\widehat{W}_{T,1/3}$			$\widehat{W}_{T,1/2}$			$\widehat{W}_{T,2/3}$		
			$1-\alpha$								
ϕ_1	θ_1	T	0.99	0.95	0.90	0.99	0.95	0.90	0.99	0.95	0.90
0.4591	-0.1310	100	0.9951	0.9861	0.9778	0.9925	0.9808	0.9673	0.9954	0.9862	0.9755
(W = 2.5811)		200	0.9958	0.9878	0.9799	0.9915	0.9759	0.9580	0.9924	0.9780	0.9617
		1000	0.9985	0.9926	0.9826	0.9911	0.9697	0.9450	0.9894	0.9686	0.9429
0.3	-0.5	100	0.4778	0.3475	0.2721	0.6689	0.5717	0.5040	0.7583	0.6895	0.6381
(W=0.5102)		200	0.4432	0.2892	0.2107	0.7423	0.6244	0.5471	0.8556	0.7821	0.7227
		1000	0.3728	0.1840	0.1144	0.9109	0.7894	0.6912	0.9750	0.9228	0.8702
0.5	-0.3	100	0.9960	0.9897	0.9821	0.9937	0.9817	0.9676	0.9928	0.9812	0.9699
(W=1.96)		200	0.9970	0.9924	0.9864	0.9925	0.9802	0.9663	0.9915	0.9765	0.9590
		1000	0.9991	0.9970	0.9933	0.9947	0.9762	0.9542	0.9893	0.9683	0.9434
0.3	-0.29	100	0.9724	0.9341	0.9004	0.9666	0.9310	0.8953	0.9631	0.9302	0.8979
(W=1.0288)		200	0.9809	0.9460	0.9127	0.9752	0.9416	0.9056	0.9722	0.9400	0.9033
		1000	0.9916	0.9651	0.9316	0.9870	0.9574	0.9240	0.9822	0.9539	0.9196

For V

			$\widehat{V}_{T,1/3}$			$\widehat{V}_{T,1/2}$			$\widehat{V}_{T,2/3}$		
			$1-\alpha$								
ϕ_1	θ_1	T	0.99	0.95	0.90	0.99	0.95	0.90	0.99	0.95	0.90
0.4591	-0.1310	100	0.9999	0.9990	0.9968	0.9983	0.9927	0.9838	0.9981	0.9922	0.9848
(V =2.2713)		200	0.9999	0.9985	0.9964	0.9981	0.9889	0.9754	0.9959	0.9856	0.9701
		1000	0.9997	0.9987	0.9958	0.9943	0.9770	0.9549	0.9914	0.9708	0.9469
0.3	-0.5	100	0.4336	0.3125	0.2466	0.6386	0.5384	0.4738	0.7347	0.6583	0.6078
(V= 0.4887)		200	0.4206	0.2729	0.2013	0.7199	0.5991	0.5191	0.8397	0.7630	0.6983
		1000	0.3700	0.1902	0.1201	0.9004	0.7728	0.6748	0.9714	0.9156	0.8632
0.5	-0.3	100	1.0000	0.9985	0.9970	0.9988	0.9935	0.9871	0.9968	0.9899	0.9800
(V= 1.8608)		200	0.9999	0.9985	0.9977	0.9983	0.9918	0.9813	0.9961	0.9837	0.9703
		1000	0.9997	0.9994	0.9980	0.9972	0.9847	0.9673	0.9908	0.9725	0.9486
0.3	-0.29	100	0.9980	0.9897	0.9743	0.9966	0.9841	0.9701	0.9960	0.9831	0.9697
(V= 1.0287)		200	0.9985	0.9901	0.9775	0.9973	0.9859	0.9710	0.9962	0.9853	0.9671
		1000	0.9990	0.9929	0.9800	0.9974	0.9853	0.9683	0.9963	0.9815	0.9626

Note: Here, $\widehat{W}_{T,g}$ and $\widehat{V}_{T,g}$ mean that we used these estimators in the formulas of the confidence intervals. T is the sample size. $1-\alpha$ is the nominal level of the confidence interval. The number of iterations is 10000.

Table 4: Actual coverage probabilities of confidence intervals for ρ_{ub}

			$\widehat{V}_{T,1/3}$			$\widehat{V}_{T,1/2}$			$\widehat{V}_{T,2/3}$		
						$1-\alpha$					
ϕ_1	θ_1	T	0.99	0.95	0.90	0.99	0.95	0.90	0.99	0.95	0.90
0.4591	-0.1310	100	0.9995	0.9980	0.9956	0.9832	0.9769	0.9655	0.9596	0.9510	0.9388
($V = 2.2713$)		200	0.9998	0.9991	0.9962	0.9877	0.9798	0.9666	0.9575	0.9430	0.9253
($\rho_{ub} = -0.7481$)		1000	0.9999	0.9986	0.9950	0.9946	0.9766	0.9536	0.9654	0.9487	0.9242
0.5	-0.3	100	0.9982	0.9973	0.9955	0.9849	0.9790	0.9694	0.9626	0.9563	0.9430
($V = 1.8608$)		200	0.9998	0.9985	0.9974	0.9907	0.9832	0.9743	0.9639	0.9538	0.9418
($\rho_{ub} = -0.6801$)		1000	0.9999	0.9996	0.9980	0.9968	0.9874	0.9704	0.9787	0.9606	0.9310
0.3	-0.29	100	0.6431	0.6176	0.5858	0.6304	0.6060	0.5767	0.6401	0.6152	0.5854
($V = 1.0287$)		200	0.6955	0.6743	0.6420	0.6817	0.6528	0.6219	0.6985	0.6648	0.6259
($\rho_{ub} = -0.1669$)		1000	0.8533	0.8388	0.8109	0.8102	0.7918	0.7660	0.7860	0.7610	0.7456

Note: Here, $\widehat{V}_{T,g}$ means that we used these estimators in the formulas of the confidence intervals for ρ_{ub}. T is ths sample size. $1-\alpha$ is the nominal level of the confidence interval. The number of iterations is 1000.

Table 5: Estimates of W, V and ρ_{ub} for U. S. GDP growth rate data

(a) Estimates of W

$\widehat{W_T}(0,1)$	$\widehat{W_T}(1,1)$	$\widehat{W_T}(2,2)$	$\widehat{W}_{T,1/3}$	$\widehat{W}_{T,1/2}$	$\widehat{W}_{T,2/3}$
1.6265	2.5811	1.6278	1.9893	1.6400	1.3160

(b) Estimates of V

$\widehat{V}_T(0,1)$	$\widehat{V}_T(1,1)$	$\widehat{V}_T(2,2)$	$\widehat{V}_{T,1/3}$	$\widehat{V}_{T,1/2}$	$\widehat{V}_{T,2/3}$
1.5119	2.2713	1.3880	1.7104	1.4101	1.1315

(c) Estimates of ρ_{ub}

$\widehat{\rho}_{ub,T}(0,1)$	$\widehat{\rho}_{ub,T}(1,1)$	$\widehat{\rho}_{ub,T}(2,2)$	$\widehat{\rho}_{ub,T,1/3}$	$\widehat{\rho}_{ub,T,1/2}$	$\widehat{\rho}_{ub,T,2/3}$
-0.5819	-0.7481	-0.5287	-0.6445	-0.5393	-0.3409

Table 6: Confidence intervals for W, V, and ρ_{ub} for U. S. GDP growth rate data

(a) CIs for W

α	$\widehat{W}_{T,1/3}$	$\widehat{W}_{T,1/2}$	$\widehat{W}_{T,2/3}$
0.90	$[\ 1.5940, \infty)$	$[\ 1.1957, \infty)$	$[\ 0.8453, \infty)$
0.95	$[\ 1.5090, \infty)$	$[\ 1.1110, \infty)$	$[\ 0.7680, \infty)$
0.99	$[\ 1.3706, \infty)$	$[\ 0.9782, \infty)$	$[\ 0.6535, \infty)$

(b) CIs for V

α	$\widehat{V}_{T,1/3}$	$\widehat{V}_{T,1/2}$	$\widehat{V}_{T,2/3}$
0.90	$[1.3705, \infty)$	$[1.0280, \infty)$	$[0.7268, \infty)$
0.95	$[1.2980, \infty)$	$[0.9553, \infty)$	$[0.6603, \infty)$
0.99	$[1.1784, \infty)$	$[0.8411, \infty)$	$[0.5619, \infty)$

(b) CIs for ρ_{ub}

β	$\widehat{V}_{T,1/3}$	$\widehat{V}_{T,1/2}$	$\widehat{V}_{T,2/3}$
0.90	$(-1, -0.5408]$	$(-1, -0.4410]$	$(-1, -0.3030]$
0.95	$(-1, -0.5192]$	$(-1, -0.4259]$	$(-1, -0.2763]$
0.99	$(-1, -0.5063]$	$(-1, -0.4149]$	$(-1, -0.2540]$

Note: Here, $\widehat{W}_{T,g}$ and $\widehat{V}_{T,g}$ means that we used these estimators in the formulas of the confidence intervals. $1-\alpha$ is the nominal level of the confidence interval. Note that the CIs for ρ_{ub} are obtained under the assumption that it exists or $V > 1$.