Is It A Short-Memory, Long-Memory, or Permanently Granger-Causation Influence?

WEN-DEN CHEN*
Department and Graduate Institute of Economics,
Tung Hai University, Taiwan, ROC

ABSTRACT
Exploring the Granger-causation relationship is an important and interesting topic in the field of econometrics. In the traditional model we usually apply the short-memory style to exhibit the relationship, but in practice there could be other different influence patterns. Besides the short-memory relationship, Chen (2006) demonstrates a long-memory relationship, in which a useful approach is provided for estimation where the time series are not necessarily fractionally co-integrated. In that paper two different relationships (short-memory and long-memory relationship) are regarded whereby the influence flow is decayed by geometric, or cutting off, or harmonic sequences. However, it limits the model to the stationary relationship. This paper extends the influence flow to a non-stationary relationship where the limitation is on $-0.5 \leq \delta \leq 1.0$ and it can be used to detect whether the influence decays off ($-0.5 \leq \delta < 0.5$) or is permanent ($0.5 \leq \delta \leq 1.0$).

KEY WORDS Granger causality; non-stationary relationship; long-memory relationship; tapered cross-spectrum estimator; Féjer taper

INTRODUCTION
Ever since Granger (1969) proposed an idea about the ‘causality’ test, the Granger causality test has become an important topic in econometrics. The ‘Granger causality’ test is based on the ‘precedence’ concept as precisely discussed in Leamer (1985), who provided a useful idea to clarify the relationship among the variables. In the traditional model we usually regard the ‘Granger causality’ relationship as a short-memory pattern in which the Lagrange multiplier (LM) test is often used to test the effects. Chen (2006) provided a useful approach to capture the long-memory relationship (or harmonic influence relationship), which concentrates on the same random resource and validates the long-memory relationship between two observed variables. However, in practice we are more interested on how long (or profoundly) the influence component will last, which is not clearly discussed in Chen (2006). This article investigates this problem and develops an approach to evaluate how
long (or profoundly) the influential information stays. We extend the model to a non-stationary relationship, i.e., a permanent influence.

**PRE-WHITEN FOR FORWARD TRANSFER FUNCTION**

First, let us consider a model that includes an independent variable $x_t$ and a dependent variable $y_t$, where $x_t$ is derived from a data-generating mechanism that is combined by current and past shocks, i.e., $a_{t-k}$ and $k = 0, \ldots, \infty$. Following the conventional model, we assume $x_t$ is generated as follows.

**Assumption 1.** The status of $x_t$ is combined from the current and past impacts. We assume $x_t$ employs the following mechanism generator:

$$
\phi(B)x_t = (1 - B)^{-d_x}\theta(B)\alpha_t
$$

where $\phi^{-1}(B)(1 - B)^{-d_x}\theta(B)$ is a lag operator polynomial function, $\alpha_t$ is an independent noise process, and $\phi(B)$ and $\theta(B)$ are respectively the autoregressive and moving average polynomials whose roots lie outside the unit circle, $(1 - B)^{-d_x}$ is the integrated part which indicates $x_t$ is a short-memory ($d_x = 0$), or long-memory ($0 < d_x < 0.5$), or non-stationary process ($0.5 \leq d_x$).

If we want to capture how long the shocks that occur in $x_t$ affects $y_t$, we have to isolate the impulse of $x_t$ at time $t$, i.e., $\alpha_t$. If we apply the original raw data $x_t$, which include not only the current information, then they also contain past information. This easily causes confusion; for instance, if we have a correlative relation between $x_t$ and $y_t$, when the information in $x_t$ could be self-delivered, then we do not know whether the relationship comes from the current state $\alpha_t$, or the past impulse $a_{t-k}$, where $k > 0$, or both. Therefore, if we use $\alpha_t$ to instead of $x_t$ this will prompt us to measure how long the influential information lasts whereby $x_t$ affects $y_t$.

It is quite useful to isolate the impulses of $x_t$ that occur at time $t$, and we can use the whiten process, i.e.:

$$
\alpha_t = (1 - B)^d \theta_t(B) \phi_t(B) x_t
$$

(1)

Here, we should note that our purpose is to explore the type of influence in which $\alpha_t$ affects $y_t$, which is different from Chen (2006), who focuses on the relationship between $y_t$ and $x_t$. Therefore, we do not need to apply the filter on $y_t$, but rather we directly apply $\alpha_t$ as the explanatory variable, i.e.:

$$
y_t = v(B)\alpha_t + \epsilon_t \quad \text{or} \quad y_t = \sum_{\tau=-\infty}^{\infty} u_{x_{t-\tau}} + \epsilon_t
$$

(2)

This model contains three parts: ‘leading’, ‘contemporaneous’, and ‘succeeding’, i.e., $\tau > 0$, $\tau = 0$, and $\tau < 0$. In this article we focus on the information on $\alpha_t$ that affects $y_t$, which indicates we only need to concentrate on the part $\tau \geq 0$. Therefore, we rewrite (2) as follows:

$$
y_t = \sum_{\tau=0}^{\infty} u_{x_{t-\tau}} + N_t \quad \text{where} \quad N_t = \sum_{\tau=1}^{\infty} u_{x_{t+\tau}} + \epsilon_t
$$

(3)

Copyright © 2008 John Wiley & Sons, Ltd. 


DOI: 10.1002/for
where $\text{cov}(\alpha_t, N_t) = 0$. Here, we should note that if the model does not contain the ‘succeeding’ part, then $N_t$ reduces to an i.i.d. noise, i.e., $v_k = 0$ for $k < 0$.

According to (2) the lag operator polynomial function can be used to express the relationship between $\alpha_t$ and $y_t$. We apply a conventional ARIMA model now for the polynomial function, which is exhibited only in a forward direction, i.e., the power numbers are only the non-negative integers:

$$v(B) = \sum_{j=0}^{\infty} v_j B^j, \text{ or } v(B) = v_0\phi^{-1}(B)(1-B)^{-\delta}\theta(B) = v_0 g(B; \Theta)$$

where $\Theta = (\phi_1, \ldots, \phi_p, \delta, \theta_1, \ldots, \theta_q)'$. Model (3) can be expressed in an alternative form:

$$y_t = v_0 g(B; \Theta)\alpha_t + N_t$$

where $g(B; \Theta) = \phi^{-1}(B)(1-B)^{-\delta}\theta(B)$ indicates the pattern style of $\alpha_t$ affecting $y_t$, $\phi(B)$ and $\theta(B)$ respectively represent the autoregressive and moving average polynomial functions, i.e., $\phi(B) = 1 - \phi_1B - \ldots - \phi_pB^p$ and $\theta(B) = 1 - \theta_1B - \ldots - \theta_qB^q$, and their roots lie outside the unit circle.

We are interested in the part $(1-B)^{-\delta}$, which indicates how profoundly $\alpha_t$ affects $y_t$. Here, if $-0.5 \leq \delta \leq 0$, then the relationship is a short- or negative-memory relationship; the influential information is exhibited by a geometry decay or cut-off style. Otherwise, $0 < \delta < 0.5$ shows a long-memory influence and the pseudo spectrum at zero approaches infinity (see Chen, 2006). We can say that though the sequence $\{u_k\}$ is convergent, i.e., $\lim_{k \to \infty} u_k = 0$, the series $\sum_{k=0}^{\infty} v_k$ is divergent. The influence is decayed by a harmonic decay. Moreover, if $\delta \geq 0.5$, then the relationship is non-stationary and the spectrum of the first part in the right-hand side of (5) would lead to a non-integrable function. At least we cannot apply the model in the usual sense. For instance, if $\delta = 1$, then $\{u_k\}$ is a divergent sequence, i.e., $\lim_{k \to \infty} u_k \neq 0$. The influential coefficients $v_k$ no longer decay to zero and they obviously have a permanent influence.

**MEASURE STATIONARY RELATIONSHIP**

In the conventional approach if we want to analyze the relationship between $y_t$ and $\alpha_t$, then we usually focus on the cross-spectrum. The cross-spectrum of $y_t$ and $\alpha_t$ can be exhibited as follows:

$$f_{\alpha y}(\lambda; \Theta_0) = (2\pi)^{-1} \sum_{u=-\infty}^{\infty} c_{\alpha y}(u)e^{-i\lambda u} \quad \text{for } \lambda \in [-\pi, \pi] - \{0\},$$

where $\Theta_0$ is the true value of $\Theta$. Here, we shall note that if $y_t$ and $\alpha_t$ have a long-memory relationship, then the spectrum at zero frequency will approach infinity. In order to simplify our estimation, we exclude the zero part, which will not affect our result analysis.

In the traditional model the cross-spectrum is a complex function, because of $c_{\alpha y}(u) \neq c_{\alpha y}(-u)$, which includes three types of relationships: ‘leading’, ‘contemporaneous’, and ‘succeeding’ parts that can be seen in (2). According to (3), if we only focus on the information of $\alpha_t$ influencing $y_t$, i.e., ‘leading’ or ‘current’ parts, then we have to revise the cross-spectrum. We follow Chen (2006)
here to construct a pseudo cross-spectrum, which is constructed by the cross-covariance $c_{\alpha \tau}(k)$ for $k > 0$ that indicates $\alpha_t$ leading or being contemporaneous on $y_t$. We shall note that if $\delta \geq 0.5$, then the cross-spectrum does not exist. Therefore, we assume here that their relation is stationary, i.e., $-0.5 \leq \delta < 0.5$. Thus, the pseudo cross-spectrum can be expressed as

$$f^A_{\alpha \tau} (\lambda; \Theta_0) = \frac{1}{2\pi} \sum_{r=-\infty}^{\infty} c_{\alpha \tau}(|\tau|) e^{-i\lambda \tau} \quad \text{for } \lambda \in [-\pi, \pi] - \{0\} \tag{6}$$

where

$$c_{\alpha \tau}(|\tau|) = \int_{-\pi}^{\pi} f^A_{\alpha \tau} (\lambda; \Theta_0) e^{-i\lambda \tau} d\lambda$$

As $\{\alpha_t\}$ is a white noise process according to (3), we see $c_{\alpha \tau}(|\tau|) = \sigma^2_{\alpha} \delta_{|\tau|}$. The following relationship holds:

$$f^A_{\alpha \tau} (\lambda; \Theta_0) = \frac{\sigma^2_{\alpha}}{2\pi} \sum_{k=-\infty}^{\infty} \delta_{|k|} e^{-i\lambda k} = \frac{\sigma^2_{\alpha}}{2\pi} \left( \delta_{0} + 2 \sum_{k=1}^{\infty} \delta_{k} \cos \lambda k \right)$$

Applying (4), this also can be expressed by

$$f^A_{\alpha \tau} (\lambda; \Theta_0) = \frac{\sigma^2_{\alpha}}{2\pi} \left( u(e^{-i\lambda}) + u(e^{i\lambda}) - u_0 \right) = \frac{\sigma^2_{\alpha} u_0}{2\pi} \left( g(e^{-i\lambda}; \Theta_0) + g(e^{i\lambda}; \Theta_0) - 1 \right) = \frac{\sigma^2_{\alpha} u_0}{2\pi} \left( 2 \Re \left[ g(e^{-i\lambda}; \Theta_0) \right] - 1 \right)$$

where $g(e^{-i\lambda}; \Theta) = (1 - e^{-i\lambda})^{-\delta} (e^{-i\lambda})^{-\delta} (e^{i\lambda})^{-\delta}$. Here, we note that $u(e^{-i\lambda})$ is only on the forward direction part that has been discussed in (4). According to (6), the estimator for the pseudo spectrum can naturally be calculated by the estimated covariance, i.e.:

$$\hat{f}^A_{\alpha \tau} (\lambda) = \frac{1}{2\pi} \sum_{r=-\infty}^{\infty} \hat{c}_{\alpha \tau}(|\tau|) e^{-i\lambda \tau} \tag{7}$$

As $\{\alpha_t\}$ is a white noise process, then $E(\hat{c}_{\alpha \tau}(k)) = \sigma^2_{\alpha} \delta_{k}$, which indicates $E(\hat{f}^A_{\alpha \tau}(\lambda)) = f^A_{\alpha \tau}(\lambda; \Theta_0)$, where $\Theta_0$ is the true value of $\Theta$. However, in practice it is infeasible to apply the covariance estimator to an infinite-lagged number, i.e., $\tau \rightarrow \infty$. Therefore, a tapered cross-spectrum estimator is implemented. We use a simple Féjer tapered cross-spectrum estimator, which can be shown as follows:

$$\hat{f}^{(T)}_{\alpha \tau} (\lambda) = \frac{1}{2\pi} \sum_{r=-T}^{T-1} h(\tau) \hat{c}_{\alpha \tau}(|\tau|) e^{-i\lambda \tau} \tag{8}$$

where $T$ is time length, $h(\tau) = (T - |\tau|)/T$, $\hat{c}_{\alpha \tau} (\tau) = \frac{1}{T - \tau} \sum_{r=\tau}^{T-\tau} (y_{r+\tau} - \bar{y}) (\alpha_r - \bar{\alpha})$, $\bar{y} = \frac{1}{T - \tau} \sum_{j=\tau+1}^{T} y_j$, and $\bar{\alpha} = \frac{1}{T - \tau} \sum_{j=\tau+1}^{T} \alpha_j$, and therefore

$$h(\tau) \hat{c}_{\alpha \tau} (\tau) = \frac{1}{T} \sum_{r=1}^{T-\tau} (y_{r+\tau} - \bar{y}) (\alpha_r - \bar{\alpha})$$
Thus, we have the following property:

\[
\lim_{T \to \infty} E \left[ \hat{f}^{(T)}_{\lambda}(\lambda) \right] = \frac{1}{2\pi} \lim_{T \to \infty} \sum_{\tau=-T+1}^{T-1} h(\tau) E \left[ \hat{c}_{\lambda \alpha}(\tau) \right] e^{-i(\lambda - \lambda^*) \tau} = \frac{1}{2\pi} \int_{-\pi}^{\pi} f^{(\lambda)}_{\lambda \alpha}(\lambda^*; \Theta_0) \sum_{\tau=-T+1}^{T-1} h(\tau) e^{-i(\lambda - \lambda^*) \tau} \, d\lambda^* \]

\[
= \lim_{T \to \infty} \int_{-\pi}^{\pi} f^{(\lambda)}_{\lambda \alpha}(\lambda^*; \Theta_0) D_T(\lambda - \lambda^*) \, d\lambda^* = f^{(\lambda)}_{\lambda \alpha}(\lambda; \Theta_0)
\]

where \( D_T(\lambda) = \frac{1}{2\pi T} \left[ \sin T\lambda/2 \right]^2 \) and \( \int_{-\pi}^{\pi} D_T(\lambda) \, d\lambda = 1 \), which is a convergence factor. We shall note now that as the relationship between \( y_t \) and \( \alpha_t \) is stationary, we can naturally assume \( y_t \) is also a stationary process. (Of course, in a few cases \( y_t \) could be a non-stationary process. In order to simplify the calculations we exclude these situations). Therefore, when \( \hat{c}_{\alpha}(\tau) \) converges in probability to \( \sigma^2_{\alpha} \), \( \hat{c}_{\lambda \alpha}(\tau) \) \( \Rightarrow \sigma^2_{\alpha} v(\tau) \) and \( \lim_{T \to \infty} v(\tau) = 0 \) (for stationary relationship), which implies for a bounded estimator, \( f^{(T)}_{\alpha}(\lambda) \) for \( \lambda \neq 0 \), which employs \( \lim_{T \to \infty} f^{(T)}_{\alpha}(\lambda; \Theta_0) = f^{(\lambda)}_{\alpha}(\lambda; \Theta_0) \) for \( \lambda \in (0, \pi) \). It is useful for us to obtain a consistent estimator. If \( y_t \) is a non-stationary process, then of course these properties no longer exist. That will change the results. We shall then consider an alternative approach to obtain a consistent estimator, which can be seen in the next section.

After discussing the properties of the tapered cross-spectrum estimator, we know that the tapered estimator is consistent. If we want to estimate the parameters, then in the conventional approach we can minimize the sum of the squared residuals. However, we shall note that if we directly apply the model it will have heteroscedasticity and autocorrelation between the adjacent frequencies. Here, we can use a two-step estimation as follows.

In the first step we calculate the weighting matrix. Through the estimated weight matrix we can achieve an objective function:

\[
Q(\Theta) = \left[ f^{(T)}_{\alpha}(\lambda) - f^{(\lambda)}_{\alpha}(\lambda; \Theta) \right] W \left[ f^{(T)}_{\alpha}(\lambda) - f^{(\lambda)}_{\alpha}(\lambda; \Theta) \right] \]

\[
(10)
\]

where \( \lambda' = [\lambda_1, \lambda_2, \ldots, \lambda_M] \) and \( \lambda_j \) for \( j = 1, \ldots, M \) \( (M = T/2) \) are the harmonic frequencies. Term \( W \) is the weighting matrix, which can be derived from the residuals of the first-step estimation which can be seen in Greene (2003, p. 536) and White (1980). Through the optimization, we obtain the estimator \( \hat{\Theta} = \arg \min Q(\Theta, y, \alpha) \), and its corresponding covariance matrix for \( \hat{\Theta} \) can be regarded as

\[
\text{cov}\left( \sqrt{M} \left( \hat{\Theta} - \Theta_0 \right) \right) \approx \sigma^2 \left[ \frac{1}{M} \frac{\partial f^{(\lambda)}_{\alpha}(\lambda)}{\partial \Theta} \bigg|_{\Theta_0} W \frac{\partial f^{(\lambda)}_{\alpha}(\lambda)}{\partial \Theta'} \bigg|_{\Theta_0} \right]^{-1}
\]

\[
(11)
\]

As \( \hat{\Theta} \nrightarrow \Theta_0 \), we use \( \hat{\Theta} \) to replace the true value \( \Theta_0 \) of \( \Theta \). Here, we simply apply the first-order autoregressive and heteroscedastic approach to obtain the weight matrix. (The proofs and further details can be found in Appendix A.)
NON-STATIONARY RELATIONSHIP

In the previous section the relationship between $a_t$ and $y_t$ is assumed to be stationary, which indicates that the impact of $a_t$ on $y_t$ will decay, i.e., $-0.5 \leq \delta < 0.5$. An asymptotic unbiased estimator for the pseudo spectrum can be developed for $\lambda \in [-\pi, \pi] - \{0\}$. However, if the relationship between $a_t$ and $y_t$ is non-stationary, then this implies the influence is permanent, i.e., $\delta \geq 0.5$. Does the estimator have similar statistical properties? This section focuses on this problem and develops an approach to estimate.

According to (4), we know that if $d \geq 0.5$, then $\{u_k\}$ could be a divergence sequence, i.e., $\lim_{k \to \infty} u_k \neq v_0$, and the pseudo cross-spectrum in (6) will explode. In order to avoid this situation, naturally we apply a convergence factor (or a taper) on the pseudo spectrum estimator, which can be seen in (8), i.e.:

$$\hat{f}_{xy}(\lambda) = \frac{1}{2\pi} \sum_{\tau=-(T-1)}^{T-1} h(\tau) \hat{c}_{xy}(|\tau|) e^{-i\lambda \tau}$$

where $h(\tau) = \frac{T - |\tau|}{T}$, $\tau = 1, \ldots, T$.

As $\{y_t\}$ is non-stationary and $\lim_{T \to \infty} \text{var} [\hat{c}_{xy}(\lambda)] \neq 0$, the estimated cross-covariance will be diffused. There is no longer convergence in probability to $\nu_k$, i.e., $\lim_{T \to \infty} \hat{c}_{xy}(\lambda) \neq \sigma^2 v_t$, and (12) is not a consistent estimator.

In the traditional approach, if the time series data show a non-stationary process then we usually use the difference to obtain a stationary process. However, if we want to choose an appropriate difference we first need prior knowledge. Velasco and Robinson (2000) suggested a useful approach to detect the non-stationary process by using a taper method. In the following context, let us focus on the tapered pseudo cross-spectrum estimator and discuss its properties. Here, we limit the non-stationary relationship on $0.5 \leq \delta \leq 1$. If we apply the taper on the spectrum estimator, then we can achieve the following lemma.

**Lemma 1.** Assume $y_t = \nu_t \phi^{-1}(B)(1 - B)^{\delta} \theta(B) \alpha_t + \varepsilon_t$, where $0.5 \leq \delta \leq 1.0$, and the Féjer tapered cross-spectrum estimator is measured by

$$\hat{f}_{xy}^{(T)}(\lambda) = \frac{1}{2\pi} \sum_{\tau=-(T-1)}^{T-1} h(\tau) \hat{c}_{xy}(|\tau|) e^{-i\lambda \tau}$$

where $h(\tau) = \frac{T - |\tau|}{T}$, $\tau = 1, \ldots, T$.

Let $\lambda = \pi T^\beta$ and $-1 \leq \beta \leq 0$, i.e., $\lambda \in (0, \pi)$, and the tapered cross-spectrum has the following relationship (see the proof in Appendix B):

$$\lim_{T \to \infty} \hat{f}_{xy}^{(T)}(\lambda) = f^S(\lambda; \Theta_0) + O(T^{1-2\beta}),$$

where $f^S(\lambda; \Theta) = \frac{\sigma^2 \Delta \nu_0}{2\pi} (2 \text{Re} \left[ \phi^{-1}(e^{-i\lambda})(1 - e^{-i\lambda})^{-\delta} \theta(e^{-i\lambda}) \right] - 1) \text{ and } 0.5 \leq \delta \leq 1.0$.

According to Lemma 1, if $\beta > -0.5$ then $O(T^{1-2\beta}) \to 0$ as $T \to \infty$, i.e., $\lim_{T \to \infty} \hat{f}_{xy}^{(T)}(\lambda) = f^S(\lambda; \Theta_0)$. As $\beta \in (-0.5, 0)$ is an open set, in practice we need a closed set to apply the harmonic frequencies. Therefore, we choose a fixed-value $\beta^*$ that belongs to $(-0.5, 0)$ and let $\Omega_2 = \{ \lambda | \pi T^{1-2\beta^*} < \lambda \leq \pi \}$. We now have the following relationship:

$$\lim_{T \to \infty} \hat{f}_{xy}^{(T)}(\lambda) = f^S(\lambda; \Theta_0) \text{ for } \lambda \in \Omega_2$$
For frequency estimation we need to cover the whole frequency domain, i.e., \( \lambda \in (0, \pi] \). Thus, we can separate the frequencies into two parts, say:

\[
\Omega_1 = \{ \lambda | 0 < \lambda \leq \pi T^{1-2\beta} \} \quad \text{and} \quad \Omega_2 = \{ \lambda | \pi T^{1-2\beta} < \lambda \leq \pi \}
\]

In \( \Omega_1 \) we cannot obtain a consistent estimator through the tapered cross-spectrum estimator. Therefore, we use the differenced data instead. If the relationship is non-stationary, then we can use the difference on \( \{ y_t \} \) to obtain a stationary relationship. Thus, we obtain

\[
\lim_{T \to \infty} \hat{f}_{\lambda \alpha}^{(T)}(\lambda) = \frac{S_2}{2\pi} \left( 2 \text{Re} \left[ g(e^{-i\lambda}; \Theta^0) \right] - 1 \right) \quad \text{for} \quad \lambda \in \Omega_1
\]

and

\[
\lim_{T \to \infty} \hat{f}_{\lambda \alpha}^{(T)}(\lambda) = \frac{S_2}{2\pi} \left( 2 \text{Re} \left[ g(e^{-i\lambda}; \Theta^0) \right] - 1 \right) \quad \text{for} \quad \lambda \in \Omega_2
\]

where \( \Theta = (\phi_1, \ldots, \phi_p, \delta, \theta_1, \ldots, \theta_q)' \), \( \Theta^* = (\phi_1, \ldots, \phi_p, \delta - 1, \theta_1, \ldots, \theta_q)' \), \( S_2 = \sigma_u^2 \Delta \), and \( g(e^{-i\lambda}; \Theta) = \phi^*(e^{-i\lambda})(1-e^{-i\lambda})^{-\delta} \). A hybrid objective function can be used to detect the non-stationary relationship where the limitation is on \(-0.5 \leq \delta < 1.5\), i.e.:

\[
Q_2 = \frac{1}{T} \left[ \sum_{\lambda \in \Omega_1} \left( \hat{f}_{\lambda \alpha}^{(T)}(\lambda_j) - \frac{S_2}{2\pi} \left( 2 \text{Re} \left[ g(e^{-i\lambda}; \Theta^0) \right] - 1 \right) \right)^2 + \sum_{\lambda_j \in \Omega_2} \left( \hat{f}_{\lambda \alpha}^{(T)}(\lambda_j) - \frac{S_2}{2\pi} \left( 2 \text{Re} \left[ g(e^{-i\lambda}; \Theta^0) \right] - 1 \right) \right)^2 \right]
\]

We should note that the domain of \( \Omega_1 \) is \( 0 < \lambda \leq \pi T^{1-2\beta} \), \( T^{1-2\beta} \to 0^+ \) as \( T \to \infty \), and the ratio of the first part will decrease to zero as \( T \) increases. This could be used to detect whether it is a stationary or non-stationary relationship. If the relationship is non-stationary, then we apply the difference to obtain a stationary relationship, and then follow (10) and (11) for estimation and test.

**SIMULATION AND POWER TESTS**

This section applies the simulations to evaluate the asymptotic properties of the estimators. Two situations are considered: one is a stationary relationship, while the second is a non-stationary relationship. Through Monte Carlo experiments, we examine the properties of the estimators by growing the time length. In stationary relationship cases, (10) and (11) are used for power tests in which we can explore the asymptotic properties of the estimators.

In a non-stationary relationship, we examine the consistency of the estimator. Here, we shall note that the estimator for the non-stationary relationship is a detector, and we do not need to calculate its asymptotic covariance. If the relationship is acknowledged as a non-stationary relationship, then we use the difference to obtain a stationary relationship, and follow (10) and (11) to obtain its asymptotic variances. According to (4), we assume a data-generating mechanism as follows:

\[
y_t = (1 - \phi_1 B - \ldots - \phi_p B^p)^{-1}(1 - B)^{-\delta}(1 - \theta_1 B - \ldots - \theta_q B^q)\alpha_t + \varepsilon_t
\]

where \( \alpha_t \sim \text{NID}(0,1) \) and \( \varepsilon_t \sim \text{NID}(0,1) \).

In a stationary situation, we show two models: one is a fractional AR(1) model (\( \phi_1 = 0.5 \) and \( \delta = 0.2 \)) and the second is a fractional MA(1) model (\( \theta_1 = 0.3 \) and \( \delta = 0.2 \)). The results are seen in
Table I. Empirical size and the power test are based on 1000 replications for the fractional AR model, i.e., $y_t = (1 - B)^d(1 - \phi B)\alpha_t + \epsilon_t$ and $\alpha_t, \epsilon_t \sim$ NID(0, 1). The sizes are respectively 2000, 3000, and 4000, and the parameters are respectively $\phi = 0.5$ and $\delta = 0.2$. The power test is based on 1000 replications, the significance level is 5%, and the two-tailed standard normal distribution test is used. The corresponding pseudo spectrum is $f_{\alpha}^d(\lambda; \Theta) = s^2/\pi(2\Re\{1 - B)^d(1 - \phi B)\} - 1)$

<table>
<thead>
<tr>
<th>$H_0$: $\phi =$</th>
<th>0.20</th>
<th>0.30</th>
<th>0.40</th>
<th>0.50</th>
<th>0.60</th>
<th>0.70</th>
<th>0.80</th>
<th>0.90</th>
<th>$\hat{\phi}$</th>
<th>se($\hat{\phi}$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>2000</td>
<td>0.904</td>
<td>0.613</td>
<td>0.266</td>
<td>0.116</td>
<td>0.209</td>
<td>0.496</td>
<td>0.813</td>
<td>0.941</td>
<td>0.5248</td>
<td>0.1187</td>
</tr>
<tr>
<td>3000</td>
<td>0.972</td>
<td>0.782</td>
<td>0.315</td>
<td>0.103</td>
<td>0.253</td>
<td>0.669</td>
<td>0.927</td>
<td>0.988</td>
<td>0.5147</td>
<td>0.0907</td>
</tr>
<tr>
<td>4000</td>
<td>0.992</td>
<td>0.872</td>
<td>0.360</td>
<td>0.084</td>
<td>0.315</td>
<td>0.790</td>
<td>0.981</td>
<td>0.995</td>
<td>0.5090</td>
<td>0.0739</td>
</tr>
</tbody>
</table>

$H_0$: $\delta =$

<table>
<thead>
<tr>
<th>2000</th>
<th>-0.10</th>
<th>0.00</th>
<th>0.10</th>
<th>0.20</th>
<th>0.30</th>
<th>0.40</th>
<th>0.50</th>
<th>0.60</th>
<th>$\hat{\delta}$</th>
<th>se($\hat{\delta}$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>2000</td>
<td>0.816</td>
<td>0.545</td>
<td>0.236</td>
<td>0.115</td>
<td>0.296</td>
<td>0.694</td>
<td>0.952</td>
<td>0.988</td>
<td>0.1718</td>
<td>0.1265</td>
</tr>
<tr>
<td>3000</td>
<td>0.916</td>
<td>0.703</td>
<td>0.302</td>
<td>0.108</td>
<td>0.347</td>
<td>0.875</td>
<td>0.988</td>
<td>0.995</td>
<td>0.1854</td>
<td>0.1181</td>
</tr>
<tr>
<td>4000</td>
<td>0.975</td>
<td>0.822</td>
<td>0.371</td>
<td>0.079</td>
<td>0.423</td>
<td>0.929</td>
<td>0.996</td>
<td>0.997</td>
<td>0.1914</td>
<td>0.0795</td>
</tr>
</tbody>
</table>

$H_0$: $s_2 =$

<table>
<thead>
<tr>
<th>2000</th>
<th>0.70</th>
<th>0.80</th>
<th>0.90</th>
<th>1.00</th>
<th>1.10</th>
<th>1.20</th>
<th>1.30</th>
<th>1.40</th>
<th>$\hat{s}_2$</th>
<th>se($\hat{s}_2$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>2000</td>
<td>0.989</td>
<td>0.809</td>
<td>0.145</td>
<td>0.019</td>
<td>0.304</td>
<td>0.863</td>
<td>0.986</td>
<td>0.994</td>
<td>0.9890</td>
<td>0.0826</td>
</tr>
<tr>
<td>3000</td>
<td>0.999</td>
<td>0.967</td>
<td>0.284</td>
<td>0.015</td>
<td>0.440</td>
<td>0.958</td>
<td>0.997</td>
<td>0.998</td>
<td>0.9929</td>
<td>0.0656</td>
</tr>
<tr>
<td>4000</td>
<td>0.998</td>
<td>0.991</td>
<td>0.438</td>
<td>0.015</td>
<td>0.513</td>
<td>0.988</td>
<td>0.997</td>
<td>0.999</td>
<td>0.9966</td>
<td>0.0397</td>
</tr>
</tbody>
</table>

Table II. Empirical size and the power test are based on 1000 replications for the fractional MA model, i.e., $y_t = (1 - B)^d(1 - \theta B)\alpha_t + \epsilon_t$ and $\alpha_t, \epsilon_t \sim$ NID(0, 1). The sizes are respectively 1000, 2000, and 3000, and the parameters are respectively $\theta = 0.3$ and $\delta = 0.2$. The power test is based on 1000 replications, the significance level is 5%, and the two-tailed standard normal distribution test is used. The corresponding pseudo spectrum is $f_{\alpha}^d(\lambda; \Theta) = s^2/\pi(2\Re\{1 - B)^d(1 - \phi B)\} - 1)$

<table>
<thead>
<tr>
<th>$H_0$: $\theta =$</th>
<th>0.00</th>
<th>0.10</th>
<th>0.20</th>
<th>0.30</th>
<th>0.40</th>
<th>0.50</th>
<th>0.60</th>
<th>0.70</th>
<th>$\hat{\theta}$</th>
<th>se($\hat{\theta}$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1000</td>
<td>0.652</td>
<td>0.213</td>
<td>0.068</td>
<td>0.081</td>
<td>0.199</td>
<td>0.413</td>
<td>0.671</td>
<td>0.804</td>
<td>0.3121</td>
<td>0.1523</td>
</tr>
<tr>
<td>2000</td>
<td>0.957</td>
<td>0.572</td>
<td>0.117</td>
<td>0.050</td>
<td>0.280</td>
<td>0.679</td>
<td>0.911</td>
<td>0.976</td>
<td>0.2968</td>
<td>0.1062</td>
</tr>
<tr>
<td>3000</td>
<td>0.993</td>
<td>0.811</td>
<td>0.206</td>
<td>0.050</td>
<td>0.322</td>
<td>0.795</td>
<td>0.973</td>
<td>0.998</td>
<td>0.2982</td>
<td>0.0803</td>
</tr>
</tbody>
</table>

$H_0$: $\delta =$

<table>
<thead>
<tr>
<th>1000</th>
<th>-0.10</th>
<th>0.00</th>
<th>0.10</th>
<th>0.20</th>
<th>0.30</th>
<th>0.40</th>
<th>0.50</th>
<th>0.60</th>
<th>$\hat{\delta}$</th>
<th>se($\hat{\delta}$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1000</td>
<td>0.796</td>
<td>0.332</td>
<td>0.083</td>
<td>0.084</td>
<td>0.223</td>
<td>0.486</td>
<td>0.736</td>
<td>0.846</td>
<td>0.2161</td>
<td>0.1541</td>
</tr>
<tr>
<td>2000</td>
<td>0.990</td>
<td>0.765</td>
<td>0.166</td>
<td>0.067</td>
<td>0.340</td>
<td>0.765</td>
<td>0.934</td>
<td>0.977</td>
<td>0.2002</td>
<td>0.1013</td>
</tr>
<tr>
<td>3000</td>
<td>1.000</td>
<td>0.922</td>
<td>0.307</td>
<td>0.059</td>
<td>0.414</td>
<td>0.874</td>
<td>0.989</td>
<td>1.000</td>
<td>0.1988</td>
<td>0.0765</td>
</tr>
</tbody>
</table>

$H_0$: $s_2 =$

<table>
<thead>
<tr>
<th>1000</th>
<th>0.70</th>
<th>0.80</th>
<th>0.90</th>
<th>1.00</th>
<th>1.10</th>
<th>1.20</th>
<th>1.30</th>
<th>1.40</th>
<th>$\hat{s}_2$</th>
<th>se($\hat{s}_2$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1000</td>
<td>0.979</td>
<td>0.694</td>
<td>0.113</td>
<td>0.013</td>
<td>0.201</td>
<td>0.714</td>
<td>0.972</td>
<td>1.000</td>
<td>0.9951</td>
<td>0.0646</td>
</tr>
<tr>
<td>2000</td>
<td>1.000</td>
<td>0.966</td>
<td>0.317</td>
<td>0.014</td>
<td>0.405</td>
<td>0.949</td>
<td>1.000</td>
<td>1.000</td>
<td>0.9969</td>
<td>0.0464</td>
</tr>
<tr>
<td>3000</td>
<td>1.000</td>
<td>0.996</td>
<td>0.534</td>
<td>0.016</td>
<td>0.563</td>
<td>0.995</td>
<td>1.000</td>
<td>1.000</td>
<td>0.9980</td>
<td>0.0387</td>
</tr>
</tbody>
</table>

Tables I and II. Table I shows the power tests based on 1000 replications for the fractional AR(1) model when the time length increases from 2000 to 3000, and then to 4000. The rejected probability for the null hypothesis $\phi = 0.5$ decreases from 0.116 to 0.103, and then to 0.084. By contrast, it is obviously different from the fault null hypothesis of $\phi = 0.4$, whereby the rejected probability increases from 0.266 to 0.315, and then to 0.360. The means of the estimated coefficients for $\phi$ are from 0.5248 to 0.5147, and then to 0.5090. The mean is closer to the true value 0.5 as the sample size increases. Simultaneously, its corresponding standard deviations decrease from 0.1187 to 0.0907 and then to 0.0739. The likely properties are also found in the fractional MA(1) model, which can be seen in Table II. When the time length increases from 1000 to 2000 and then to 3000, the rejected
Short-Memory, Long-Memory, and Granger-Causation Influence

Table III. Empirical sizes are based on 1000 replications for the non-stationary fractional AR model, i.e., $y_t = (1 - B)^{-d}(1 - \phi B)^{-1}\alpha_t + \epsilon_t$ and $\alpha_t, \epsilon_t \sim \text{NID}(0, 1)$. The sizes are respectively 1000, 2000, and 3000, and the parameters are respectively $\phi = 0.5$, $\delta = 0.7$, and $\sigma^2 = 1.00$, where the corresponding pseudo spectrum is $f_{\omega}^*(\lambda; \Theta) = \sigma^2/(2\pi)(2\text{Re}[(1 - B)^{-d}(1 - \phi B)^{-1}])^{-1}$.

<table>
<thead>
<tr>
<th>True value</th>
<th>$\phi = 0.5$</th>
<th>$\delta = 0.7$</th>
<th>$\sigma^2 = 1.00$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Estimated coefficients and standard deviations</td>
<td>$\hat{\phi}$</td>
<td>se($\hat{\phi}$)</td>
<td>$\hat{\delta}$</td>
</tr>
<tr>
<td>Sample size = 1000</td>
<td>0.5096</td>
<td>0.1505</td>
<td>0.6720</td>
</tr>
<tr>
<td>Sample size = 2000</td>
<td>0.5158</td>
<td>0.1104</td>
<td>0.6755</td>
</tr>
<tr>
<td>Sample size = 3000</td>
<td>0.5074</td>
<td>0.0880</td>
<td>0.6869</td>
</tr>
</tbody>
</table>

Table IV. Empirical sizes are based on 1000 replications for the non-stationary fractional MA model, i.e., $y_t = (1 - B)^{-d}(1 - \theta B)\alpha_t + \epsilon_t$ and $\alpha_t, \epsilon_t \sim \text{NID}(0, 1)$. The sizes are respectively 1000, 2000, and 3000, and the parameters are respectively $q = 0.3$, $\delta = 0.7$, and $\sigma^2 = 1.00$ where the corresponding pseudo spectrum is $f_{\omega}^*(\lambda; \Theta) = \sigma^2/(2\pi)(2\text{Re}[(1 - B)^{-q}(1 - \phi B)])^{-1}$.

<table>
<thead>
<tr>
<th>True value</th>
<th>$q = 0.3$</th>
<th>$\delta = 0.7$</th>
<th>$\sigma^2 = 1.00$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Estimated coefficients and standard deviations</td>
<td>$\hat{q}$</td>
<td>se($\hat{q}$)</td>
<td>$\hat{\delta}$</td>
</tr>
<tr>
<td>Sample size = 1000</td>
<td>0.2933</td>
<td>0.1603</td>
<td>0.6888</td>
</tr>
<tr>
<td>Sample size = 2000</td>
<td>0.2937</td>
<td>0.1080</td>
<td>0.6919</td>
</tr>
<tr>
<td>Sample size = 3000</td>
<td>0.2908</td>
<td>0.0882</td>
<td>0.6917</td>
</tr>
</tbody>
</table>

probability for the null hypothesis $\theta = 0.3$ stays around 0.05. However, the rejected probability for the fault null hypothesis $\theta = 0.4$ increases from 0.199, to 0.280, and then to 0.322. Overall, this indicates that the estimator employs consistency properties and the larger the time length is, the more powerful the estimator will be.

Tables III and IV show the estimates for the non-stationary relationship model. According to the earlier section, ‘Measure stationary relationship’, if we want to obtain a consistent estimator that includes a non-stationary relationship, then we have to choose a fixed value $\beta \in (-0.5, 0)$. Here, we choose $\beta = -0.4$, i.e.: $$\Omega_1 = \{ \lambda_j | 0 < \lambda_j \leq \pi T^{-0.2} \}$$

and $$\Omega_2 = \{ \lambda_j | \pi T^{-0.2} \leq \lambda_j \leq \pi \}$$

Table III shows the non-stationary relationship with an AR(1) model where $\phi_t = 0.5$ and $\delta = 0.7$. From the table we see that the standard deviation of the estimated coefficients becomes smaller as the time length increases, and the means of the estimator are close to the true value. For instance, the means of the estimated coefficients for $\delta$ are from 0.6720, to 0.6755, and then to 0.6869. They are closer to the true value 0.7 as the sample size increases. Simultaneously, the corresponding standard deviations decrease from 0.1380 to 0.1015 and then to 0.0775. The same results are also found in Table IV for the non-stationary relationship with the MA(1) model. When the sample size
increases, the means of the estimated coefficients are closer to the true value and the corresponding standard deviations decrease. Tables III and IV show that the hybrid objective function for the non-stationary relationship is a consistent estimator for the pseudo cross-spectrum.

CONCLUDING REMARKS

This paper provides a useful methodology to capture the temporary and permanent effects under a ‘Granger causality’ sense, in which we follow Chen (2006) and extend to a non-stationary relationship where a hybrid objective function is used. By minimizing the sum of the squared residuals, the asymptotic properties are analyzed. In this paper our limitation is on the assumption of the influence parameter $-0.5 \leq \delta \leq 1$, $-0.5 \leq \delta < 0.5$ and indicates that the information flow will decay off in the long run; otherwise, $0.5 \leq \delta \leq 1$ indicates that the effect is permanent.

APPENDIX A

According to (10) we can obtain the objective function:

$$Q(\Theta) = [\hat{f}_{\alpha}^{TA}(\lambda) - f_{\alpha}^{A}(\lambda; \Theta)]'W[\hat{f}_{\alpha}^{TA}(\lambda) - f_{\alpha}^{A}(\lambda; \Theta)]$$

Following Greene (2003, p. 536) and White (1980), we simply apply an AR(1) and heteroscedastic model, and then $W$ is obtained:

$$W = PV^{-1}P$$

where

$$V^{-1} = \frac{1}{1 - \rho^2} \begin{bmatrix} 1 & -\rho & 0 & \cdots & 0 \\ -\rho & 1 + \rho^2 & \cdots & \cdots & \cdots \\ 0 & -\rho & \cdots & -\rho & 0 \\ \vdots & \vdots & \ddots & \ddots & \ddots \\ 0 & \cdots & \cdots & 1 + \rho^2 & -\rho \\ -\rho & \cdots & \cdots & -\rho & 1 \end{bmatrix}, \quad P = \begin{bmatrix} \sigma_1^{-1} & 0 & \cdots & 0 \\ 0 & \sigma_2^{-1} & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & \sigma_M^{-1} \end{bmatrix}$$

where $\rho$ is the first-order autoregressive coefficient and the heteroscedastic standard errors $\sigma_j$ for $j = 1, \ldots, M$ are obtained by White (1980) with the polynomial of the explanatory variable. The first-order derivatives of the objective function are then obtained as

$$\frac{\partial Q(\Theta)}{\partial \Theta} = -2 \frac{\partial f_{\alpha}^{A}(\lambda; \Theta)}{\partial \Theta}'W[\hat{f}_{\alpha}^{TA}(\lambda) - f_{\alpha}^{A}(\lambda; \Theta)]$$

At the same time, the second-order derivatives can be obtained as

$$\frac{\partial^2 Q(\Theta)}{\partial \Theta \partial \Theta'} = 2 \frac{\partial f_{\alpha}^{A}(\lambda; \Theta)'}{\partial \Theta} W \frac{\partial f_{\alpha}^{A}(\lambda; \Theta)}{\partial \Theta'} - 2 \sum_{j=1}^{M} \frac{\partial^2 f_{\alpha}^{A}(\lambda_j; \Theta)}{\partial \Theta \partial \Theta'} W[\hat{f}_{\alpha}^{TA}(\lambda) - f_{\alpha}^{A}(\lambda; \Theta)]$$

where $W_j$ is the $j$th row of $W$. The expectation of the second-order derivative is obtained:

$$E \left[ \frac{\partial^2 Q(\Theta)}{\partial \Theta \partial \Theta'} \right] = 2 \frac{\partial f_{\nu A}(\lambda; \Theta)'}{\partial \Theta} \frac{\partial f_{\nu A}(\lambda; \Theta)}{\partial \Theta'}$$

With a scoring method, the corresponding covariance matrix for the estimated parameters is now regarded as

$$\text{cov}(\sqrt{M} (\hat{\Theta} - \Theta_0)) = \left( E \left[ \frac{\partial^2 Q}{\partial \Theta \partial \Theta'} \right] \right)^{-1} \left( E \left[ \frac{\partial Q}{\partial \Theta} \frac{\partial Q}{\partial \Theta'} \right] \right)^{-1} \bigg|_{\Theta_0}$$

$$\approx \hat{\sigma}^2 \left[ \frac{1}{M} \frac{\partial f_{\nu A}(\lambda)}{\partial \Theta} \right] \left( \frac{\partial f_{\nu A}(\lambda)}{\partial \Theta'} \right)^{-1} \bigg|_{\Theta_0}$$

where $\hat{\sigma}^2 = \frac{1}{M} \sum_{j=1}^{M} \eta_j^2$, with $\eta = [\eta_1, \eta_2, \ldots, \eta_M]$, $\eta = \Gamma M (f_{\nu A}(\lambda) - f_{\nu A}(\lambda))$, and $\Gamma = V^{-1}$.

According to (7), the pseudo spectrum is

$$f_{\nu A}(\lambda; \Theta) = \frac{s_2}{2\pi} \left[ (1 - e^{-i\lambda t})^{-\delta} \theta (e^{-i\lambda t}) \phi^{-1} (e^{-i\lambda t}) + (1 - e^{i\lambda t})^{-\delta} \theta (e^{i\lambda t}) \phi^{-1} (e^{i\lambda t}) - 1 \right]$$

where $s_2 = \sigma_\theta^2 \nu_0$. The first-order derivative for the pseudo spectrum is next calculated by

$$\frac{\partial f_{\nu A}(\lambda; \Theta)}{\partial \phi_k} = \frac{s_2}{2\pi} \left[ (1 - e^{-i\lambda t})^{-\delta} \theta (e^{-i\lambda t}) \phi^{-1} (e^{-i\lambda t}) e^{-ik\lambda t} + (1 - e^{i\lambda t})^{-\delta} \theta (e^{i\lambda t}) \phi^{-1} (e^{i\lambda t}) e^{ik\lambda t} \right]$$

$$\frac{\partial f_{\nu A}(\lambda; \Theta)}{\partial \theta_k} = -\frac{s_2}{2\pi} \left[ (1 - e^{-i\lambda t})^{-\delta} \phi^{-1} (e^{-i\lambda t}) e^{-ik\lambda t} + (1 - e^{i\lambda t})^{-\delta} \phi^{-1} (e^{i\lambda t}) e^{ik\lambda t} \right]$$

$$\frac{\partial f_{\nu A}(\lambda; \Theta)}{\partial \delta} = -\frac{s_2}{2\pi} \left[ (1 - e^{-i\lambda t})^{-\delta} \theta (e^{-i\lambda t}) \phi^{-1} (e^{-i\lambda t}) \ln (1 - e^{-i\lambda t}) + (1 - e^{i\lambda t})^{-\delta} \theta (e^{i\lambda t}) \phi^{-1} (e^{i\lambda t}) \ln (1 - e^{i\lambda t}) \right]$$

and

$$\frac{\partial f_{\nu A}(\lambda; \Theta)}{\partial s_2} = \frac{1}{2\pi} \left[ (1 - e^{-i\lambda t})^{-\delta} \theta (e^{-i\lambda t}) \phi^{-1} (e^{-i\lambda t}) + (1 - e^{i\lambda t})^{-\delta} \theta (e^{i\lambda t}) \phi^{-1} (e^{i\lambda t}) - 1 \right]$$

Therefore (A.1), the asymptotic covariance, is achieved.
APPENDIX B

Let \( \Delta \hat{t}_t = \hat{t}_t - \hat{t}_{-1} \), and then we have the following relationship:

\[
(1 - e^{-i\lambda}) \sum_{\tau=0}^{T-1} \hat{\upsilon}_\tau e^{-i\lambda \tau} = \hat{\upsilon}_{-1} - \hat{\upsilon}_{T-1} e^{-i\lambda T} + \sum_{\tau=0}^{T-1} (\hat{\upsilon}_\tau - \hat{\upsilon}_{-1}) e^{-i\lambda \tau} \\
= \hat{\upsilon}_{-1} - \hat{\upsilon}_{T-1} e^{-i\lambda T} + \sum_{\tau=0}^{T-1} \Delta \hat{\upsilon}_\tau e^{-i\lambda \tau}.
\]

This is expressed by an alternative form:

\[
\sum_{\tau=0}^{T-1} \hat{\upsilon}_\tau e^{-i\lambda \tau} = \frac{\hat{\upsilon}_{-1}}{1 - e^{-i\lambda}} - \frac{\hat{\upsilon}_{T-1}}{1 - e^{-i\lambda}} e^{-i\lambda T} + \frac{1}{1 - e^{-i\lambda}} \sum_{\tau=0}^{T-1} \Delta \hat{\upsilon}_\tau e^{-i\lambda \tau} \quad (B.1)
\]

If we add a Féjer taper, \( h(\tau) = (T - \tau)/T \), then the above function has the following relationship (here we shall note that it is half of a taper, because in (B.1) there is only the positive index part):

\[
\sum_{\tau=0}^{T-1} h(\tau) \hat{\upsilon}_\tau e^{-i\lambda \tau} = \sum_{\tau=0}^{T-1} \left( \frac{T - \tau}{T} \right) \hat{\upsilon}_\tau e^{-i\lambda \tau}
\]

Putting (B.1) into the equation, we express it in an alternative form:

\[
\sum_{\tau=0}^{T-1} h(\tau) \hat{\upsilon}_\tau e^{-i\lambda \tau} = \frac{1}{(1 - e^{-i\lambda})} \sum_{\tau=0}^{T-1} h(\tau) \Delta \hat{\upsilon}_\tau e^{-i\lambda \tau} + \frac{\hat{\upsilon}_{-1}}{1 - e^{-i\lambda}} - \frac{1}{T (1 - e^{-i\lambda})} \sum_{\tau=0}^{T-1} \hat{\upsilon}_\tau e^{-i\lambda \tau}
\]

\[
= \frac{1}{(1 - e^{-i\lambda})} \sum_{\tau=0}^{T-1} h(\tau) \Delta \hat{\upsilon}_\tau e^{-i\lambda \tau} + \frac{\hat{\upsilon}_{-1}}{1 - e^{-i\lambda}} - \frac{1}{T (1 - e^{-i\lambda})} \left( \hat{\upsilon}_{-1} - \hat{\upsilon}_{T-1} e^{-i\lambda T} + \sum_{\tau=0}^{T-1} \Delta \hat{\upsilon}_\tau e^{-i\lambda \tau} \right)
\]

\[
= \frac{1}{(1 - e^{-i\lambda})} \sum_{\tau=0}^{T-1} h(\tau) \Delta \hat{\upsilon}_\tau e^{-i\lambda \tau} + \frac{\hat{\upsilon}_{-1}}{1 - e^{-i\lambda}} - \frac{1}{T (1 - e^{-i\lambda})} \left( \hat{\upsilon}_{-1} (1 + e^{-i\lambda T}) - \left( \sum_{\tau=0}^{T-1} \Delta \hat{\upsilon}_\tau \right) e^{-i\lambda T} + \sum_{\tau=0}^{T-1} \Delta \hat{\upsilon}_\tau e^{-i\lambda \tau} \right)
\]

If \( 0.5 \leq \delta \leq 1 \) indicates \( \{ \Delta \hat{\upsilon}_\tau \} \), which can be regarded as the coefficients of a short-memory or negative-memory relationship, then \( \sum_{\tau=0}^{\infty} \Delta \hat{\upsilon}_\tau \) and \( \sum_{\tau=0}^{\infty} \Delta \hat{\upsilon}_\tau e^{-i\lambda \tau} \), and their corresponding variances are bounded and have finite values (is similar to the spectrum at zero frequency for the short- or negative-memory process). Therefore, if \( \lambda = T^{\beta} \) and \( -1 < \beta < 0 \), which implies \( \lambda \to 0^+ \) as \( T \to \infty \), then the above equation is rewritten as

\[
\sum_{\tau=0}^{T-1} h(\tau) \hat{\upsilon}_\tau e^{-i\lambda \tau} = \frac{1}{(1 - e^{-i\lambda})} \sum_{\tau=0}^{T-1} h(\tau) \Delta \hat{\upsilon}_\tau e^{-i\lambda \tau} + \frac{\hat{\upsilon}_{-1}}{1 - e^{-i\lambda}} + O(T^{-1-2\beta})
\]

A tapered cross-spectrum estimator is next estimated as

\[
\hat{f}_{x\alpha}^{(T)A}(\lambda) = \frac{1}{2\pi} \sum_{\tau=-1}^{T-1} h(\tau) \hat{c}_{x\alpha}(\tau) e^{-i\lambda \tau}
\]

Here, in order to simplify the calculation we extract the variance of \( \alpha \). The tapered cross-spectrum estimator is now expressed as

Copyright © 2008 John Wiley & Sons, Ltd.


DOI: 10.1002/for
\[
\hat{f}_{\alpha}(\lambda) = \frac{\hat{\sigma}_\alpha}{2\pi} \left( \sum_{\tau=0}^{T-1} h(\tau) \hat{u}_\tau e^{-i\lambda \tau} + \sum_{\tau=0}^{T-1} h(\tau) \hat{u}_\tau e^{i\lambda \tau} - \hat{u}_0 \right)
\]

\[
= \frac{\hat{\sigma}_\alpha}{2\pi} \left( \frac{1}{1-e^{-i\lambda}} \sum_{\tau=0}^{T-1} h(\tau) \Delta \hat{u}_\tau e^{-i\lambda \tau} + \frac{1}{1-e^{i\lambda}} \sum_{\tau=0}^{T-1} h(\tau) \Delta \hat{u}_\tau e^{i\lambda \tau} + \hat{\Delta}_{-1} \hat{u}_\tau e^{i\lambda \tau} - \hat{u}_0 \right)
\]

\[+ \mathcal{O}(T^{-1-2\beta})\]

\[
= \frac{\hat{\sigma}_\alpha}{2\pi} \left( \frac{1}{1-e^{-i\lambda}} \sum_{\tau=0}^{T-1} h(\tau) \Delta \hat{u}_\tau e^{-i\lambda \tau} + \frac{1}{1-e^{i\lambda}} \sum_{\tau=0}^{T-1} h(\tau) \Delta \hat{u}_\tau e^{i\lambda \tau} - \Delta \hat{u}_0 \right) + \mathcal{O}(T^{-1-2\beta})
\]

\[
= \hat{f}(\lambda) + \mathcal{O}(T^{-1-2\beta}),
\]

where the pseudo spectrum is

\[
\hat{f}(\lambda) = \frac{\hat{\sigma}_\alpha^2}{2\pi} \left( \frac{1}{1-e^{-i\lambda}} \sum_{\tau=0}^{T-1} h(\tau) \Delta \hat{u}_\tau e^{-i\lambda \tau} + \frac{1}{1-e^{i\lambda}} \sum_{\tau=0}^{T-1} h(\tau) \Delta \hat{u}_\tau e^{i\lambda \tau} - \Delta \hat{u}_0 \right)
\]

If \(\beta > -0.5\) then \(\lim_{T \to \infty} \mathcal{O}(T^{-1-2\beta}) = 0\). Moreover, \(\Delta u_\tau\) is regarded as being the coefficients of the stationary relationship between \(\Delta y_t\) and \(\alpha_t\). According to the above section, ‘Pre-whiten for forward transfer function’, we know the stationary relationship:

\[
\lim_{T \to \infty} \sum_{\tau=0}^{T-1} h(\tau) \Delta u_\tau e^{-i\lambda \tau} = \Delta u_0 g_{\lambda,\alpha}(e^{-i\lambda})
\]

where \(g_{\lambda,\alpha}(e^{-i\lambda}) = \phi^{-1}(e^{-i\lambda})(1-e^{-i\lambda})^{-\delta} \theta(e^{-i\lambda})\). Thus, we obtain a useful result, i.e.:

\[
\lim_{T \to \infty} E[\hat{f}_{\alpha}(\lambda)] = f^S(\lambda; \Theta_b), \text{ for } \lambda \in \Omega_2 \text{ and } \Omega_2 = \{\lambda | \pi T^{-1-2\beta} < \lambda \leq \pi\}
\]

where \(f^S(\lambda; \Theta_b)\) is a pseudo spectrum for a non-stationary process and according to (4) and (9) has the following relationship:

\[
f^S(\lambda; \Theta) = \frac{\sigma^2 \Delta u_0}{2\pi} \left( \frac{1}{1-e^{-i\lambda}} g_{\lambda,\alpha}(e^{-i\lambda}) + \frac{1}{1-e^{i\lambda}} g_{\lambda,\alpha}(e^{i\lambda}) - 1 \right)
\]

or

\[
f^S(\lambda; \Theta) = \frac{\sigma^2 \Delta u_0}{2\pi} \left( 2 \Re \left[ \phi^{-1}(e^{-i\lambda})(1-e^{-i\lambda})^{-\delta} \theta(e^{-i\lambda}) \right] - 1 \right)
\]

We eventually know that \(\hat{f}_{\alpha}(\lambda)\) is an asymptotic unbiased estimator for \(f^S(\lambda; \Theta_b)\) when \(\lambda \in \Omega_2\).

REFERENCES


Copyright © 2008 John Wiley & Sons, Ltd.


DOI: 10.1002/for


Author's biography:

**Wen-Den Chen** is a Professor of the Economics Institute at Tung Hai University, Taiwan, ROC. He received a BA in Information Management and an MA in Industrial Economics from National Central University, and a PhD in Management Science from National Chiao Tung University. Chen’s research interests include stochastic processes to finance, time series analysis, and panel data analysis.

Author's address:

**Wen-Den Chen**, Tung Hai University, PO Box 5-0885, No. 181, Section 3, Taichung-Kan Road, Taichung, Taiwan 407.