Measuring the Partial Causality in the Frequency Domain

Hosoya, Yuzo
Department of Economics, Meisei University

Takimoto, Taro
Faculty of Economics, Kyushu University

Discussion Paper Series

Discussion Paper No.2012-4

Measuring the Partial Causality in the Frequency Domain

Yuzo Hosoya・Taro Takimoto
Kyushu University
December 2012

Faculty of Economics
Kyushu University
Hakozaki, Higashi-ku, Fukuoka, 812-8581, Japan

Measuring the Partial Causality in the Frequency Domain[*1]

Yuzo Hosoya[*2] and Taro Takimoto[*3]

December, 2012

Abstract

The paper is about a statistical analysis aiming at quantitative characterization in the frequency domain the strength of partial causality of one series effecting on another in the presence of a third series. The paper provides an estimation procedure of the proposed partial measures based on a direct factorization method of spectral densities and also proposes a Monte Carlo Wald tests for allied causal measures.

JEL classification: C12, C13, C32, C53
Key Words: Canonical factorization; Granger non-causality; Measure of one-way effect; Measure of reciprocity; Partial causality; Prediction error; Vector ARMA model; Wald statistics.

1 Introduction

The paper presents a numerically feasible method of constructing partial causal measures between a pair of time series in the presence of a third series, providing an estimation and testing procedure of such measures, which quantitatively characterize the causal aspects in the presence of a third series. The presented estimation method is relied on the direct factorization of the spectral density matrix developed by Hosoya and Takimoto (2010). Also the paper proposes Monte Carlo Wald tests for the purposes of testing the strength

[*1]The research is partially supported by JSPS Grant-in-Aid for Scientific Research (C)22530211and JSPS Grant-in-Aid for Young Scientists (B) 22730180, respectively.
[*2]Department of Economics, Meisei University.
E-mail: yhosoya@econ.meisei-u.ac.jp.
[*3]Faculty of Economics, Kyushu University, Hakozaki, Higashi-ku, Fukuoka, 812-8581 Japan.
E-mail: takimoto@en.kyushu-u.ac.jp.

and direction of the partial causality and allied statistical inference.

Detecting such properties between a pair of time series as causal directions and the extent of their effects and also testing non-existence of feedback relation between them constitute major focal points in multivariate time-series analysis since Granger (1963, 69) introduced the celebrated definition of causality in view of prediction improvement. The Granger non-causality test in the time-domain is originally introduced for stationary processes, but the recent literature is rich in respect to the extension to non-stationary vector autoregressive system; see, for example, by Sims, Stock and Watson (1990), Mosconi and Giannini (1992), Lütkepohl and Reimers (1992), Toda and Phillips (1993) Yamamoto and Kurozumi (2006).

As regards causal analysis in the frequency domain there are studies by Gel'fand and Yaglom (1959), Granger (1969) and Geweke (1982, 1984). Also Hosoya (1991) proposed a system of causal measures which are decomposable in the frequency domain, improving on the Gel'fand-Yaglom and the Geweke measures, whereas Granger and Lin (1995) extended the one-way effect measure to bivariate cointegrated processes. Hosoya's one-way effect measure from a series $\{y(t)\}$ to another $\{x(t)\}$ is basically defined in terms of prediction improvement of $\{x(t)\}$ due to the addition of the past values of the one-way effect component of $\{y(t)\}$, in contrast to Geweke's feedback measure which is defined in terms of the improvement due to addition of the past values of $\{y(t)\}$ as a whole. Hosoya's decomposition has the merit that the equivalence relationship is established between the overall causal measure of $\{y(t)\}$ to $\{x(t)\}$ and the integral of its associating frequency-wise causal measure, whereas Geweke obtained only the inequality relation between them and, for the equality to hold, certain additional conditions are required.

The frequency-domain approach seems more informative than time-domain counterparts, since it enables us not only to conduct significance testing of the Granger non-causality, but also to measure frequency-wise as well as overall causal strength and to construct a variety of confidence intervals of those measures; see Hosoya (1997a), Yao and Hosoya (2000) and Hosoya, Yao and Takimoto (2005) for large-sample Wald tests of those measures and the allied confidence set construction. A related approach is provided by Breitung and Candelon (2006) who propose an F-test to test the frequency-wise Granger non-causality for the bivariate VAR model, in which the null hypothesis is shown representable as a set of linear hypotheses on the autoregressive parameters; see also Gronwald

(2009) for a related analysis.

Canonical factorization of spectral density matrix constitutes a crucial step in the construction of predictors and the evaluation of the prediction errors, which have a variety of applications in time series analysis and control. Consequently, it also constitutes an essential step in constructing causal measures, since the Granger concept is framed on the basis of the prediction error evaluation. Rational spectrum estimation based on a set of finite observations has a wide application in time-series analysis and it is often conducted based on a time-domain representation of the data generating process; see for typical examples Hannan-Rissanen (1982) and Hannan-Kavalieris (1984). The ARMA models fitting in the time domain automatically estimates the transfer function (and desirably a canonical response function) of the data generating process. In case the spectral density to be used does not correspond to the direct observation process but to a derived one, however, a certain factorization algorithm for rational spectra is required. In particular, the construction of measures introduced by Hosoya (1991, 97a) requires factorization of spectrum which is not necessarily obtained directly from the observation process. Another use of canonical factorization algorithm is the modification of MA coefficients with the MA spectrum kept invariant so that the resulting MA representation is canonical. This paper derives the partial causal measures by means of the second method of two methods to evaluate the frequency-wise measure of one-way effect suggested by Hosoya (1991, p.434). This approach is motivated by Breitung and Candelon (2006)'s paper.

Section 2.1 of the paper describes how third-series effect elimination is carried out and introduces such partial causality concepts as partial one-way effect measures and reciprocal measures. Although those causal measures are primarily defined for vector second-order stationary processes, they are extensible to non-stationary cointegrated processes with the aid of the reproducibility assumption as shown in Hosoya (1997a, 2001). Furthermore, Section 2.2 discusses in detail the partial spectral density construction and estimation of the partial one-way effect measure for the vector ARMA processes. Also the subsection shows that the canonical factorization of the general ARMA model for construction of the partial measures is reducible to that of a vector finite-order MA spectrum. Section 3 provides inferential procedure for several types of the partial causal measures for stationary ARMA processes and allied confidence set construction. In particular, the section proposes Monte Carlo Wald tests to detect the causal strength and direction.

Appendix A.1 is for proofs of the theorems, Appendix A.2 provides an explicit representation of the spectral density of the reciprocal components of a pair of series in interest.

The paper uses the following notations and terminology: The sets of all integers and nonnegative integers are denoted respectively by $\mathbb{Z}$, $\mathbb{Z}^{0+}$. For $\{z_i, i \in \mathbb{A}\}$ a set of real-valued random variables possessing finite second-order moment, $H\{z_i, i \in \mathbb{A}\}$ indicates the closure in mean-square of the linear hull of $\{z_i, i \in \mathbb{A}\}$ in the Hilbert space, defined over the real-number field, of random variables of finite second-order moment. Suppose that a p-vector stochastic process $x(t)$ has finite covariance matrix and that $\mathbb{S}$ denotes a set of certain integers, then $H\{x(t), t \in \mathbb{S}\}$ implies in the sequel the subspace $H\{x_i(t), t \in \mathbb{S}, i = 1, \cdots, p\}$. For brevity, $H\{x(t_1 - j), y(t_2 - j), z(t_3 - j);\ j \in \mathbb{Z}^{0+}\}$ is written as $H\{x(t_1), y(t_2), z(t_3)\}$ and $H\{x(j); j \in \mathbb{Z}\}$ is written as $H\{x(\infty)\}$. For a random vector $x(t)$ indexed by t, $\{x(t)\}$ denotes the process $\{x(t); t \in \mathbb{Z}\}$ unless otherwise specified. The identity matrix of order p is written as I_p. Given a matrix A, A^* denote the transpose if A is a real matrix and conjugate transpose if A is a complex matrix. For a random-vector x or a pair of random vectors x and y, $Cov(x)$ and $Cov(x, y)$ denote respectively the covariance matrices of x and $vec(x, y)$.The determinant of a square matrix C is written as $\det C$. Definition is indicated by $\equiv$. Suppose that a real sequence $c[j],\ j = -a, \cdots, a$ satisfies the condition $c[j] = c[-j],\ c[0] > 0$ and $c(z) = \sum_{j=-a}^{a} c[j]z^j$ is nonnegative for $z = e^{-i\lambda}$ $(-\pi \leq \lambda < \pi)$. Then there exists a real sequence $b[j]$ $(j = 0, \cdots, a)$ such that $b(z) = \sum_{j=0}^{a} b[j]z^j$ does not have zeros inside the unit circle and the relation $c(z) = \frac{1}{2\pi} b(z)b(z^{-1})$ holds. Such a factorization is said to be canonical and $b(z)$ is said to be a canonical factor of $c(z)$. If $b_0 > 0$, the factorization is unique.

2 Partial causality

2.1 Elimination of a third-series effect

This section describes elimination of a third series effect and the construction of the overall as well as frequency-wise partial causal measures in a general set-up. Suppose that $\{x(t), y(t), z(t); t \in \mathbb{Z}\}$ is a real vector-valued second-order stationary process. Denote by H the Hilbert space defined over the real-number field which is the closure of the linear hull of the union $\{x_j(t);\ t \in \mathbb{Z},\ j = 1, \cdots, p_1\} \cup \{y_k(t);\ t \in \mathbb{Z},\ k = 1, \cdots, p_2\} \cup$

$\{z_l(t);\ t \in \mathbb{Z},\ l = 1, \cdots, p_3\}$, where $x_j(t)$ denotes the j-th element of the vector $x(t)$. The projection of a random vector $w = \{w_j;\ j = 1, \cdots, s\}$ to a closed subspace $H(\cdot)$ of H implies the element-wise orthogonal projection. Namely, if $\bar{w}_j$ is the orthogonal projection of w_j onto $H(\cdot)$, the projection implies $\bar{w}$ whose j-th element is $\bar{w}_j$. In the three series system $\{x(t), y(t), z(t)\}$, the one-way effect component of $z(t)$ implies the projection residual (the perpendicular) of $z(t)$ when it is projected onto the closed linear subspace $H\{x(t), y(t), z(t-1)\}$ which is spanned by the set of the vector components $\{x(s), y(s), z(s-1),\ -\infty < s \leq t\}$ and the residual (the perpendicular) is denoted by $z_{0,0,-1}(t)$.

Although interpretation of the Granger causality concept does not accompany difficulty as long as it is focused on a pair of series, it incurs certain difficulty in the presence of a third confounding series, since the third series may produce such anomalous phenomena as spurious or indirect causality; see Granger (1980) and Hsiao (1982).

It is a standard method in time-series analysis to characterize the dependency between a pair $\{x(t), y(t)\}$ of series in the presence of a third series $\{z(t)\}$ by means of the partial serial correlation coefficient and the partial cospectrum. They are the kind of dependencies defined between $\{x_{\cdot,\cdot,\infty}(t)\}$ and $\{y_{\cdot,\cdot,\infty}(t)\}$, where $x_{\cdot,\cdot,\infty}(t)$ is the projection of $x(t)$ onto $H\{z(\infty)\}$, the closed linear space generated by $\{z(t), t \in \mathbb{Z}\}$, and $y_{\cdot,\cdot,\infty}(t)$ is defined similarly. Such total-effect elimination of a third variable is essentially the direct extension of the partial concepts in multivariate statistical analysis where temporal order of observations is not taken into account, but it may possibly distort the temporal dependence relation between $\{x(t)\}$ and $\{y(t)\}$ due to their feedback relation with $\{z(t)\}$. To deal with the difficulty, Hosoya (2001) proposed to define the partial relations between $\{x(t)\}$ and $\{y(t)\}$ in the presence of a third series by the corresponding simple relations between $\{u(t)\}$ and $\{v(t)\}$ which are respectively defined as the projection residuals of $x(t)$ and $y(t)$ onto $H\{z_{0,0,-1}(\infty)\}$ which is equivalent to the projection residuals onto $H\{z_{0,0,-1}(t-1)\}$. For example, the partial measure of one-way effect of $\{x(t)\}$ to $\{y(t)\}$ is defined to be the simple one-way effect of $\{u(t)\}$ to $\{v(t)\}$. To distinguish the ordinary Granger causality which focuses on a pair of processes alone from the partial version which takes account of a third series, the former causality is said simple in the paper.

Suppose that $x(t), y(t), z(t)$ are respectively p_1, p_2, p_3-vectors and let $f(\lambda)$ be the joint spectral density of the second-order process $w(t) = (x(t)^*, y(t)^*, z(t)^*)^*, t \in \mathbb{Z}$. Suppose that $f(\lambda)$ satisfies the Szegö condition

$$\int_{-\pi}^{\pi} \log\det f(\lambda) d\lambda > -\infty, \tag{2.1}$$

then the density has the factorization

$$f(\lambda) = \frac{1}{2\pi}\Lambda(e^{-i\lambda})\Lambda(e^{-i\lambda})^* \tag{2.2}$$

by means of a $(p_1+p_2+p_3)\times(p_1+p_2+p_3)$ matrix $\Lambda(z)$ which is analytic and of full rank inside the unit disc. Namely, in (2.2) $\Lambda(e^{-i\lambda})$ is the boundary value of the analytic function

$$\Lambda(z) = \sum_{j=0}^{\infty} \Lambda[j] z^j$$

with the real matrix coefficients $\Lambda[j]$. Such a factorization is said to be canonical in the sequel; see Rozanov (1967, pp.71-77) and Hannan (1970, pp.157-163). Let $\varepsilon(t) \equiv (\varepsilon_1(t)^*, \varepsilon_2(t)^*, \varepsilon_3(t)^*)^* \equiv w_{-1}(t) \equiv (x_{-1,-1,-1}(t)^*, y_{-1,-1,-1}(t)^*, z_{-1,-1,-1}(t)^*)^*$ be the one-step ahead prediction-error of the process $w(t)$. Denote the covariance matrix of $\varepsilon(t)$ by $\Sigma^\dagger$ and denote the partition matrix as

$$\Sigma^\dagger = \begin{bmatrix} \Sigma^\dagger_{..} & \Sigma^\dagger_{.3} \\ \Sigma^\dagger_{3.} & \Sigma^\dagger_{33} \end{bmatrix}.$$

Then the residual of the projection of $\varepsilon_3(t)$ on the linear space spanned by $\varepsilon_.(t) \equiv (\varepsilon_1(t)^*, \varepsilon_2(t)^*)^*$ is given by $\varepsilon^\dagger_3(t) = \varepsilon_3(t) - \Sigma^\dagger_{3.}\Sigma^{\dagger\,-1}_{..}\varepsilon_.(t)$ and it constitutes the one-way effect component of $z(t)$. Define

$$\begin{bmatrix} \varepsilon^\dagger_.(t) \\ \varepsilon^\dagger_3(t) \end{bmatrix} = \begin{bmatrix} \Sigma^{\dagger\,-1/2}_{..} & 0 \\ 0 & (\Sigma^\dagger_{33:.})^{-1/2} \end{bmatrix} \begin{bmatrix} I_{p_1+p_2} & 0 \\ -\Sigma^\dagger_{3.}\Sigma^{\dagger\,-1}_{..} & I_{p_3} \end{bmatrix} \begin{bmatrix} \varepsilon_.(t) \\ \varepsilon_3(t) \end{bmatrix}; \tag{2.3}$$

and define Π a $(p_1+p_2+p_3)\times(p_1+p_2+p_3)$ matrix by

$$\Pi = \begin{bmatrix} \Sigma^{\dagger\,-1/2}_{..} & 0 \\ 0 & (\Sigma^\dagger_{33:.})^{-1/2} \end{bmatrix} \begin{bmatrix} I_{p_1+p_2} & 0 \\ -\Sigma^\dagger_{3.}\Sigma^{\dagger\,-1}_{..} & I_{p_3} \end{bmatrix};$$

where $\Sigma^\dagger_{33:.} \equiv \Sigma^\dagger_{33} - \Sigma^\dagger_{3.}\Sigma^{\dagger\,-1}_{..}\Sigma^\dagger_{.3}$, so that Π is a lower triangular block matrix

$$\Pi = \begin{bmatrix} \Pi_{..} & 0 \\ \Pi_{3.} & \Pi_{33} \end{bmatrix}. \tag{2.4}$$

Set $\tilde{\Lambda}(L) = \Lambda(L)\Lambda(0)^{-1}\Pi^{-1}$ and set its partition as

$$\tilde{\Lambda}(z) = \left[\begin{array}{cc} \tilde{\Lambda}_{\cdot\cdot}(z) & \tilde{\Lambda}_{\cdot 3}(z) \\ \tilde{\Lambda}_{3\cdot}(z) & \tilde{\Lambda}_{33}(z) \end{array} \right].$$

Then it follows from the relationships

$$w(t) = \Lambda(L)\Lambda(0)^{-1}\varepsilon(t) = \Lambda(L)\Lambda(0)^{-1}\Pi^{-1}\Pi\varepsilon(t) = \tilde{\Lambda}(L)\varepsilon^{\dagger}(t),$$

that

$$\left[\begin{array}{c} x(t) \\ y(t) \end{array} \right] = \tilde{\Lambda}_{\cdot\cdot}(L)\varepsilon^{\dagger}_{\cdot}(t) + \tilde{\Lambda}_{\cdot 3}(L)\varepsilon^{\dagger}_{3}(t). \tag{2.5}$$

Since then $\{\varepsilon^{\dagger}_{3}(t)\}$ subordinates to the one-way effect process from $\{z(t)\}$ to $\{x(t), y(t)\}$, the spectral density of $\{x(t), y(t)\}$ is given in view of (2.5) by

$$f_{\cdot\cdot}(\lambda) = \frac{1}{2\pi}\{\tilde{\Lambda}_{\cdot\cdot}(e^{-i\lambda})\tilde{\Lambda}_{\cdot\cdot}(e^{-i\lambda})^{*} + \tilde{\Lambda}_{\cdot 3}(e^{-i\lambda})\tilde{\Lambda}_{\cdot 3}(e^{-i\lambda})^{*}\}.$$

Denote by $\{u(t), v(t)\}$ the residuals of the projection of $x(t)$ and $y(t)$ onto $H\{z_{0,0,-1}(\infty)\} = H\{\varepsilon^{\dagger}_{3}(\infty)\}$; then it is given in view of (2.5) by

$$\left[\begin{array}{c} u(t) \\ v(t) \end{array} \right] = \tilde{\Lambda}_{\cdot\cdot}\varepsilon^{\dagger}_{\cdot}(t), \tag{2.6}$$

whence the spectral density of $\{u(t), v(t)\}$ is represented by

$$h(\lambda) = \frac{1}{2\pi}\tilde{\Lambda}_{\cdot\cdot}(e^{-i\lambda})\tilde{\Lambda}_{\cdot\cdot}(e^{-i\lambda})^{*} \tag{2.7}$$

Note that although $\tilde{\Lambda}(z)$ is a canonical factor if $\Lambda(z)$ is canonical, its square diagonal block $\tilde{\Lambda}_{\cdot\cdot}(z)$ in (2.7) is not warranted to be so. When the factorization given in (2.7) is not canonical, a certain factorization procedure must be implemented to construct partial causal measures between $\{x(t)\}$ and $\{y(t)\}$, since all the partial casusal measures are based on the knowledge of a canonical factor of $h(\lambda)$.

Breitung and Candelon (2006) proposes, for the simple non-causality test of $\{z(t)\}$ not causing $\{x(t), y(t)\}$, for instance, testing the null hypothesis which is given in terms of (2.5) by

$$\tilde{\Lambda}_{\cdot 3}(e^{-i\lambda}) = 0. \tag{2.8}$$

Although an F-test might be applied at least to the case Breitung and Candelon deal with, the hypothesis (2.8) generally imposes non-linear restrictions on the model parameters so that a kind of either the likelihood ratio test or the Wald test rather than the F-test would be pertinent.

Remark 2.1. Suppose that a matrix $\tilde{\Lambda}(z) = \{\tilde{\Lambda}_{ij}(z),\ i,j = 1,2,3\}$ is given by

$$\tilde{\Lambda}(z) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} - \begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 1 \\ 0 & 3 & 1/2 \end{bmatrix} z;$$

then all the zeros of $|\tilde{\Lambda}(z)|$ are either on or outside of the unit circle. On the other hand, $|\tilde{\Lambda}_{\cdot\cdot}(z)| = |(1-z)(1-2z)|$ has one zero inside the unit circle, where $\tilde{\Lambda}_{\cdot\cdot}(z)$ denotes the upper 2×2 diagonal block of $\tilde{\Lambda}(z)$. Consequently, when a partial spectral density is derived as in (2.7), the factor $\tilde{\Lambda}_{\cdot\cdot}(e^{-i\lambda})$ on the right-hand side is not guaranteed to be canonical. □

Remark 2.2. To deal with the third-series presence problem, Breitung and Candelon (2006, p.369) propose a way to eliminate a third series effect by a time-domain regression which is to use the estimated relation between a pair of series in interest to construct the partial causal measures. Specifically in the case of a three-variate AR model, they fit

$$x_t = \sum_{j=1}^{p} \alpha_j x_{t-j} + \sum_{j=1}^{p} \beta_j y_{t-j} + \sum_{j=1}^{p} \gamma_j v_{t-j} + \varepsilon_t$$

and test the null hypothesis $\beta_1 = \cdots = \beta_p = 0$ by F-statistic, where v_t is the residual obtained by regressing z_t on x_t, y_t and $w_{t-1}, \cdots, w_{t-p}$ where $w = (x,y,z)^*$. While the method suggests a way to avoid the spectral canonical factorization problem, it may not produce the exactly same partial measures this paper proposes. By contrast, the approach we propose have the following merits:

- The MA part can be included in the basic model so that the partial causal analysis can be extended to the ARMA model as shown in this paper.
- The dimensions p_1, p_2, p_3 can be more than 1.
- Without assuming such a specific parametric model as the ARMA model, causal measures are able to be constructed as long as a canonical factorization of partial spectral density is available.
- Partial measures of reciprocity and association can be dealt with. □

2.2 Defining the partial measures

The partial causal measures between $\{x(t)\}$ and $\{y(t)\}$ in the presence of $\{z(t)\}$ is defined in terms of $\{u(t)\}$ and $\{v(t)\}$ as given in (2.6).

D1. The partial OMO (overall measure of one-way effect) from $\{y(t)\}$ to $\{x(t)\}$ is defined by

$$PM_{y\to x:z} \equiv M_{v\to u} = \log \frac{\det Cov\{u_{-1,\cdot}(t)\}}{\det Cov\{u'_{-1,-1}(t)\}},$$

where $u_{-1,\cdot}(t)$ and $u'_{-1,-1}(t)$ are the projection residuals of $u(t)$ onto $H\{u(t-1)\}$ and onto $H\{u(t-1), v_{0,1}(t-1)\}$ respectively.

Suppose that the spectral density $h(\lambda)$ satisfies the Szegö condition (2.1) so that $h(\lambda)$ has a canonical factorization

$$h(\lambda) = \frac{1}{2\pi}\Gamma(e^{-i\lambda})\Gamma(e^{-i\lambda})^*. \tag{2.9}$$

The equality implies the following MA representation of the series $\{u(t), v(t)\}$ in terms of the one-step ahead prediction error $\epsilon(t) \equiv (\epsilon_1(t)^*, \epsilon_2(t)^*)^* \equiv (u_{-1,-1}(t)^*, v_{-1,-1}(t)^*)^*$ holds in the time domain:

$$\begin{bmatrix} u(t) \\ v(t) \end{bmatrix} = \Gamma(L)\Gamma(0)^{-1}\begin{bmatrix} \epsilon_1(t) \\ \epsilon_2(t) \end{bmatrix}, \tag{2.10}$$

where $E\{\epsilon(t)\} = 0$ and $E\{\epsilon(t)\epsilon(t)^*\} = \Gamma(0)\Gamma(0)^* = \Sigma$. Set, as in (2.3),

$$\begin{aligned} \begin{bmatrix} \epsilon_1^\dagger(t) \\ \epsilon_2^\dagger(t) \end{bmatrix} &\equiv \begin{bmatrix} \Sigma_{11}^{-1/2} & 0 \\ 0 & \Sigma_{22:1}^{1/2} \end{bmatrix} \begin{bmatrix} I_{p_1} & 0 \\ -\Sigma_{21}\Sigma_{11}^{-1} & I_{p_2} \end{bmatrix} \begin{bmatrix} \epsilon_1(t) \\ \epsilon_2(t) \end{bmatrix} \\ &\equiv \Xi\epsilon(t), \end{aligned} \tag{2.11}$$

whence $E\{\epsilon^\dagger(t)\epsilon^\dagger(t)^*\} = I_{p_1+p_2}$. Then we have

$$\begin{aligned} \begin{bmatrix} u(t) \\ v(t) \end{bmatrix} &= \Gamma(L)\Gamma(0)^{-1}\Xi^{-1}\Xi\epsilon(t) \\ &\equiv \Gamma^\dagger(L)\epsilon^\dagger(t) \\ &\equiv \begin{bmatrix} \Gamma_{11}^\dagger(L) & \Gamma_{12}^\dagger(L) \\ \Gamma_{21}^\dagger(L) & \Gamma_{22}^\dagger(L) \end{bmatrix} \begin{bmatrix} \epsilon_1^\dagger(t) \\ \epsilon_2^\dagger(t) \end{bmatrix}, \end{aligned} \tag{2.12}$$

where $\{\epsilon_2^\dagger(t)\}$ is the normalized one-way effect component of $v(t)$ to $u(t)$.

D2. The partial FMO (frequency-wise measure of one-way effect) in terms of the frequency response function $\Gamma^{\dagger}(e^{-i\lambda})$ is defined by

$$\begin{aligned}
PM_{y\to x:z}(\lambda) &\equiv M_{v\to u}(\lambda) \\
&= \log \frac{\det\{\Gamma^{\dagger}_{11}(e^{-i\lambda})\Gamma^{\dagger}_{11}(e^{-i\lambda})^{*} + \Gamma^{\dagger}_{12}(e^{-i\lambda})\Gamma^{\dagger}_{12}(e^{-i\lambda})^{*}\}}{\det\{\Gamma^{\dagger}_{11}(e^{-i\lambda})\Gamma^{\dagger}_{11}(e^{-i\lambda})^{*}\}} \\
&= \log\det\{I_{p_1} + \Gamma^{\dagger}_{11}(e^{-i\lambda})^{-1}\Gamma^{\dagger}_{12}(e^{-i\lambda})(\Gamma^{\dagger}_{11}(e^{-i\lambda})^{-1}\Gamma^{\dagger}_{12}(e^{-i\lambda}))^{*}\}, \quad (2.13)
\end{aligned}$$

whereas $PM_{x\to y:z}(\lambda)$ is given in a similar way; see for another representation Hosoya (2001). Theorem 2.1 below asserts that the partial overall measure is equal to the integration of the partial FMO over the whole frequency domain. Since the proof proceeds in a manner paralleling to Hosoya (1991, p.433), it is omitted.

Theorem 2.1. The following equality holds between the partial OMO and FMO:

$$PM_{y\to x:z} = \frac{1}{2\pi}\int_{-\pi}^{\pi} PM_{y\to x:z}(\lambda)d\lambda.$$

To define the partial measures of association and reciprocity, denote by $\ddot{u}_{\cdot,\infty}(t)$ and $\ddot{v}_{\infty,\cdot}(t)$ the projection residuals of $u(t)$ onto $H\{v_{0,-1}(\infty)\}$ and $v(t)$ onto $H\{u_{0,-1}(\infty)\}$ respectively, and set their joint spectral density matrix as

$$\ddot{h}(\lambda) = \begin{bmatrix} \ddot{h}_{11}(\lambda) & \ddot{h}_{12}(\lambda) \\ \ddot{h}_{21}(\lambda) & \ddot{h}_{22}(\lambda) \end{bmatrix}.$$

Appendix A.2 exhibits a representation of $\ddot{h}(\lambda)$.

D3. The partial measure of reciprocity at frequency λ and the corresponding overall measure between $x(t)$ and $y(t)$ are defined by:

$$PM_{x.y:z}(\lambda) \equiv M_{u.v}(\lambda) = \log\left[\frac{\det \ddot{h}_{11}(\lambda)\det \ddot{h}_{22}(\lambda)}{\det \ddot{h}(\lambda)}\right],$$

and

$$PM_{x.y:z} = \frac{1}{2\pi}\int_{-\pi}^{\pi} PM_{x.y:x}(\lambda)d\lambda.$$

Theorem 2.2. The partial frequency-wise measure of reciprocity (FMR) is a constant over the whole frequency domain. Namely, set $\ddot{\sigma}^2 = \det\Sigma_{11}\det\Sigma_{22}/\det\Sigma$; then we have

$$PM_{x.y:z}(\lambda) \equiv M_{u,v}(\lambda) = \log\ddot{\sigma}^2. \quad (2.14)$$

Evidently $\ddot{\sigma}^2$ is a constant not less than 1 which Geweke (1982) calls the measure of instantaneous feedback.

The partial measure of association at frequency λ and the corresponding overall measure between $x(t)$ and $y(t)$ are defined respectively by:

$$PM_{x,y:z}(\lambda) \equiv M_{u,v}(\lambda) = \log\left[\frac{\det h_{11}(\lambda)\det h_{22}(\lambda)}{\det \ddot{h}(\lambda)}\right],$$

$$PM_{x,y:z} = \frac{1}{2\pi}\int_{-\pi}^{\pi} PM_{x,y:z}(\lambda)d\lambda.$$

The following relationships are straightforward consequences of the definitions of the respective terms; see Hosoya (1991).

Theorem 2.3.

$$\begin{aligned} PM_{x,y:z}(\lambda) &= PM_{x\to y:z}(\lambda) - PM_{x.y:z}(\lambda) + PM_{y\to x:z}(\lambda), \\ PM_{x,y:z} &= PM_{x\to y:z} + PM_{x.y:z} + PM_{y\to x:z}. \end{aligned}$$

There is an important case for application in which the spectral density matrix $k(\lambda)$ of the process $\{u(t), v(t)\}$ is expressed as

$$k(\lambda) = |\gamma(e^{-i\lambda})|^2 h(\lambda). \tag{2.15}$$

Suppose that $h(\lambda) = \Gamma(e^{-i\lambda})\Gamma(e^{-i\lambda})^*$ for a canonical factor $\Gamma(z)$, so that

$$k(\lambda) = \gamma(e^{-i\lambda})\Gamma(e^{-i\lambda})\{\gamma(e^{-i\lambda})\Gamma(e^{-i\lambda})\}^*. \tag{2.16}$$

Moreover suppose that $\gamma(z)$ is a scalar-valued function defined on the complex plane such that $c(0) = 1$ and that it is a analytic function with real coefficients and has no zeros inside the unit circle. In this specific circumstance, we have the following theorem:

Theorem 2.4. Suppose $\{u(t), v(t)\}$ has the spectral density $k(\lambda)$ given in (2.15), for which the canonical factorization (2.16) holds. Then the $M_{v\to u}(\lambda), M_{u\to v}(\lambda)$ and $M_{u.v}(\lambda)$ are the same as the corresponding measures which are given if the spectral density is $h(\lambda)$.

Remark 2.3. Breitung and Candelon (2006, p.364) directly derive $\epsilon^{\dagger}(t)$ in (2.11) by multiplying the Cholesky factor matrix of the inverse of the covariance matrix of $\varepsilon(t)$.

Although the procedure makes $\epsilon_1^\dagger(t)$ and $\epsilon_2^\dagger(t)$ orthogonal, it does not necessarily produce the one-way effect component of $v(t)$, namely, $\epsilon_2^\dagger(t)$. In the case of their bivariate model where $u(t)$ and $v(t)$ are scalar-valued and the orthogonalization is done by the lower triangle Cholesky matrix, the one-way effect component is automatically derived, since then orthogonalization is conducted by eliminating the effect of $\varepsilon_1(t)$ from $\varepsilon_2(t)$ via the projection. In general, however, when $u(t)$ and $v(t)$ are vector-valued, arbitrary orthogonalization of $\varepsilon_1(t)$ and $\varepsilon_2(t)$ does not necessarily produce the one-way effect measure. □

Remark 2.4. The Sims' version of non-causality in the presence of a third series $\{z(t)\}$ is described as this: A necessary and sufficient condition for $\{y(t)\}$ not to cause partially $\{x(t)\}$ is that $y(t)$ is expressed as

$$y(t) = y^{(1)}(t) + y^{(2)}(t),$$

where $y^{(1)}(t)$ is the projection of $y(t)$ onto $H\{x(t), z_{0,0,-1}(t)\}$ and $y^{(2)}(t)$ is orthogonal to $H\{x(\infty), z_{0,0,-1}(\infty)\}$. Moreover, a necessary and sufficient condition for $\{y(t)\}$ not to cause partially $\{x(t)\}$ is $PM_{y\to x:z} = 0$; see for allied studies Sims (1972), Hosoya (1977) and Hosoya (2001). □

3 Inference based on the ARMA model

Focused specifically on the stationary vector ARMA process, Section 3.1 shows how the partial causal measures defined in Section 2 are evaluated. Section 3.2 discusses statistical inference on those measures.

3.1 The stationary ARMA model

Suppose that the process $\{x(t), y(t), z(t)\}$ is a stationary multivariate ARMA process which is generated by

$$A(L)\begin{bmatrix} x(t) \\ y(t) \\ z(t) \end{bmatrix} = B(L)\varepsilon(t), \quad t \in \mathbb{Z}, \tag{3.1}$$

where $x(t), y(t), z(t)$ are respectively p_1, p_2, p_3-vectors, $A(L)$ and $B(L)$ are a-th and b-th order polynomials of the lag operator L and $A[0] = B[0] = I_{p_1+p_2+p_3}$; namely, we have $A(L) = \sum_{j=0}^{a} A[j]L^j$ and $B(L) = \sum_{j=0}^{b} B[j]L^j$.

Moreover suppose that the zeros of $\det A(z)$ are all outside of the unit circle, $\det B(z)$ has the zeros either on or outside of the unit circle and does not share any common zeros. We assume that the innovation $\{\varepsilon(t)\}$ is a white noise process with mean 0 and covariance matrix $\Sigma^\dagger$. Because of the zero conditions of $A(z)$ and $B(z)$, the joint spectral density $f(\lambda)$ of the process satisfies the Szegö condition

$$\int_{-\pi}^{\pi} \log\det f(\lambda) d\lambda > -\infty,$$

whence it has a canonical factorization

$$f(\lambda) = \frac{1}{2\pi}\Lambda(e^{-i\lambda})\Lambda(e^{-i\lambda})^*. \tag{3.2}$$

In view of the zero conditions of $A(z)$ and $B(z)$, a version of the canonical factor $\Lambda(z)$ is given by

$$\Lambda(z) = A(z)^{-1}B(z)\Sigma^{\dagger\frac{1}{2}} = (\det A(z))^{-1}A^\sharp(z)B(z)\Sigma^{\dagger\frac{1}{2}} \equiv (\det A(z))^{-1}C(z),$$

where $\Sigma^{\dagger\frac{1}{2}}$ is the Cholesky factor of $\Sigma^\dagger$ satisfying $\Sigma^\dagger = \Sigma^{\dagger\frac{1}{2}}(\Sigma^{\dagger\frac{1}{2}})^*$; $A^\sharp(z)$ denotes the transposed cofactor matrix of $A(z)$ and $C(z)(\equiv A^\sharp(z)B(z)\Sigma^{\dagger\frac{1}{2}})$ is a finite-order real matrix-coefficient polynomial such that

$$C(z) = \sum_{j=0}^{\bar{a}} C[j]z^j,\ 0 \le \bar{a} \le (p_1 + p_2 + p_3 - 1)a + b.$$

As in the previous section, denote the projection residuals of $x(t)$ and $y(t)$ onto $H\{z_{0,0,-1}(\infty)\}$ respectively by $u(t)$ and $v(t)$, and denote the joint spectral density matrix of $\{u(t), v(t)\}$ by $h(\lambda)$. Now set

$$\tilde{\Lambda}(z) = (\det A(z))^{-1}A^\sharp(z)B(z)\Sigma^{\dagger\frac{1}{2}} \begin{bmatrix} \Sigma^{-1/2} & 0 \\ 0 & (\Sigma^\dagger_{33:\cdot})^{-1/2} \end{bmatrix} \begin{bmatrix} I_{p_1-p_2} & 0 \\ -\Sigma^\dagger_{3\cdot}\Sigma^{\dagger\,-1}_{\cdot\cdot} & I_{p_3} \end{bmatrix} \tag{3.3}$$

and let $\tilde{\Lambda}_{\cdot\cdot}(z)$ be the $(p_1 + p_2) \times (p_1 + p_2)$ upper diagonal block of $\tilde{\Lambda}(z)$. It follows from (2.7) that the spectral density $h(\lambda)$ of $\{u(t), v(t)\}$ is given by

$$h(\lambda) = \frac{1}{2\pi}|\det A(e^{-i\lambda})|^{-2}\tilde{\Lambda}_{\cdot\cdot}(e^{-i\lambda})\tilde{\Lambda}_{\cdot\cdot}(e^{-i\lambda})^*.$$

In view of Theorem 2.4 the causal measures between $\{u(t)\}$ and $\{v(t)\}$ are derived assuming that the joint spectral density is given as $k(\lambda) = \frac{1}{2\pi}\tilde{\Lambda}_{\cdot\cdot}(e^{-i\lambda})\tilde{\Lambda}_{\cdot\cdot}(e^{-i\lambda})^*$. Although

$\tilde{\Lambda}_{..}(e^{-i\lambda})$ is not necessarily canonical, since $k(\lambda)$ is a MA spectrum, Hosoya and Takimoto (2010)'s algorithm is exploited to derive a canonical factor $\Gamma(e^{-i\lambda})$ such that

$$k(\lambda) = \frac{1}{2\pi}\tilde{\Lambda}_{..}(e^{-i\lambda})\tilde{\Lambda}_{..}(e^{-i\lambda})^* = \frac{1}{2\pi}\Gamma(e^{-\lambda})\Gamma(e^{-i\lambda})^*. \tag{3.4}$$

Consequently, the causal measures introduced in Section 2 are able to be computed according to the definitions D.2 through D.4 using the factor $\Gamma(z)$ given in (3.4).

3.2 Inferential procedures

Based on a finite set of observations $\{x(t), y(t), z(t);\ t = 1, \cdots, T\}$ and the VARMA modeling (3.1) of the data generating process, we are able to conduct statistical inference on the partial causal measures introduced in Section 2. Denote the whole model parameter by θ ; namely, set

$$\theta \equiv vec\{A[1], \cdots, A[a], B[1], \cdots, B[b], v(\Sigma^\dagger)\}$$

where $v(\Sigma^\dagger)$ denotes the $(p_1 + p_2) \times (p_1 + p_2 + 1)/2$ vector obtained from $vec(\Sigma^\dagger)$ by eliminating all the supradiagonal elements of the $(p_1 + p_2) \times (p_1 + p_2)$ matrix $\Sigma^\dagger$. Let $G(\theta)$ be an m-vector whose components are respectively certain quantities allied to the partial causal measures. Then the estimator $G(\hat{\theta})$ enable not only testing the Granger non-causality but also making confidence statements on $G(\theta)$.

Takimoto and Hosoya (2004, 2006) provide a relevant parameter estimation procedure, in contrast to conventional nonrestrictive estimation procedures for the VARMA model parameter which do not necessarily produce estimates satisfying the zero conditions of $\det A(z)$ and $\det B(z)$. Modifying the maximum Whittle likelihood estimation, Takimoto and Hosoya's forgoing papers provide a three-step root-modification procedure which produces coefficient estimates warranting the stationarity and invertibility conditions. The procedure is essentially carried out as follows:

Step 1. By fitting a sufficiently higher order VAR process and applying the ordinary least-square method, obtain an estimate of the unobservable disturbance terms as the regression residual series. In the case the DGP is VAR process, this step is skipped.

Step 2. Substituting the disturbances in the MA part by the corresponding residuals obtained in Step 1, estimate VARMA model by the least square method, selecting the lag-orders of the model by means of an information criterion.

Step 3. Determine the estimate $\hat{\theta}$ of the model parameter by maximizing the Whittle likelihood endowed with a penalty function of zero conditions by means of a quasi-Newton iteration method, using the parameter values obtained in Step 2 as the initial value of the iteration.

To deal with inferential issues, consider first the case in which the null hypothesis does not involve the Granger non-causality hypothesis. Since the partial measures $M_{v\to u}$ and $M_{v\to u}(\lambda)$ are non negative, testing them being equal to zero constitutes a boundary value test. For such tests, the direct use of the stochastic expansion of the estimates is not pertinent since the Jacobian matrix is not of full rank. Suppose specifically that $G_i(\theta), i = 1, \cdots, m$, are different kinds of scalar-valued measures and let $G(\theta)$ be a m-vector such that $G(\theta) = (G_1(\theta), \ldots, G_m(\theta))^*$. By the stochastic expansion, we have

$$\sqrt{T}\{G(\hat{\theta}) - G(\theta)\} = (D_\theta G)\sqrt{T}(\hat{\theta} - \theta) + o_p(1),$$

where $D_\theta G$ is the $m \times n_\theta$ Jacobian matrix of $G(\theta)$ where n_θ denotes the size of the vector θ. Suppose that $\sqrt{T}(\hat{\theta} - \theta)$ is asymptotically normally distributed with mean 0 and covariance matrix $\Psi(\theta)$. Then $\sqrt{T}\{G(\hat{\theta} - G(\theta)\}$ is asymptotically normally distributed with mean 0 and the $m \times m$ asymptotic covariance matrix

$$H(\theta) = D_\theta G(\theta)\Psi(\theta)D_\theta G(\theta)^*. \tag{3.5}$$

Assume that the vector G of causal measures is chosen so that rank $H(\theta) = m$ in a neighborhood of the true θ; then the Wald statistic

$$W^{(m)} \equiv T\{G(\hat{\theta}) - G(\theta)\}^* H(\hat{\theta})^{-1}\{G(\hat{\theta}) - G(\theta)\} \tag{3.6}$$

is asymptotically χ^2-distributed with m degrees of freedom if θ is the true value. Let G_0 be a given m vector, then the null hypothesis $G(\theta) = G_0$ is tested by the test statistic

$$W^{(m)} \equiv T\{G(\hat{\theta}) - G_0\}^* H(\hat{\theta})^{-1}\{G(\hat{\theta}) - G_0\}.$$

Also a confidence set for $G(\theta)$ is able to constructed by means of the statistic $W^{(m)}$.

There are several alternative procedures available to estimate the asymptotic covariance matrix $H(\theta)$. For example, we might use the asymptotic covariance matrix formula given by Yao and Hosoya (2000) which is based on the numerical differentiation for $D_\theta G$

and evaluation of $\Psi(\hat{\theta})$ in the case of the cointegrated VAR model, but the formula becomes much more complex computationally in the ARMA model set-up. An alternative simpler approach is to use the Monte Carlo Wald test procedure which is conducted as follows:

Step 1. Estimate θ, all the parameter involved in the model (3.1), and evaluate the vector $G(\hat{\theta})$.

Step 2. Generate the data series $\{x(t)^\dagger, y(t)^\dagger, z(t)^\dagger; t = 1, \cdots, T\}$ by the model (3.1) using the parameter estimate $\hat{\theta}$ obtained in Step 1 and simulated independently normally distributed random vectors $\{\varepsilon(t)\}$ with mean 0 and the estimated variance-covariance matrix $\hat{\Sigma}^\dagger$ in Step 1.

Step 3. Estimate $G(\theta)$ from the simulated series $\{x(t)^\dagger, y(t)^\dagger, z(t)^\dagger; t = 1, \cdots, T\}$, where the estimate is denoted by $G(\theta^\dagger)$.

Step 4. Iterate Steps 2 and 3 N times, and produce $G(\theta_n^\dagger); n = 1, \cdots, N$, and estimate the covariance matrix $H(\theta)$, denoted as $\bar{H}(\hat{\theta})$, as the Monte Carlo sample covariance matrix of $G(\theta_n^\dagger)$; namely,

$$\bar{H}(\hat{\theta}) = \frac{T}{N} \sum_{n=1}^{N} \left(G(\theta_n^\dagger) - \bar{G}(\theta^\dagger) \right) \left(G(\theta_n^\dagger) - \bar{G}(\theta^\dagger) \right)^*, \tag{3.7}$$

where

$$\bar{G}(\theta^\dagger) = \frac{1}{N} \sum_{n=1}^{N} G(\theta_n^\dagger).$$

As we have alluded above, the foregoing approach is not used for testing the frequency-wise non-causality, and so we must look for other statistics. For that purpose, Breitung and Candelon (2006), based on the stationary bivariate VAR model, propose an F-test for a set of linear hypotheses on the autoregressive parameters. To deal with a wider class of stationary models, however, we need a somewhat more general approach. Since the measure $PM_{y \to x:z}(\lambda)$ is not determined by $\Gamma_{12}^\dagger(e^{-i\lambda})$ alone, but by the ratio $\Gamma_{11}^\dagger(e^{-i\lambda})^{-1}\Gamma_{12}^\dagger(e^{-i\lambda})$ in view of the formula (2.13), we may appropriately conduct the test of the null hypothesis of v not causing u by testing

$$\Gamma_{11}^\dagger(e^{-i\lambda})^{-1}\Gamma_{12}^\dagger(e^{-i\lambda}) = 0,$$

rather than testing the hypothesis $\Gamma_{12}^\dagger(e^{-i\lambda}) = 0$, where $\Gamma_{ij}^\dagger(e^{-i\lambda}) \equiv \sum_{j=0}^{\bar{a}} \Gamma_{ij}^\dagger[j] e^{-ij\lambda}$ and $\Gamma_{ij}^\dagger[j]$ is the j-th coefficient matrix of the polynomial $\Gamma_{ij}^\dagger(z)$, where $i, j = 1, 2$ and

$\bar{a} \equiv a(p_1 + p_2 + p_3 - 1) + b$. Define

$$\psi(\theta, \lambda) \equiv \text{vec}\{Re\Gamma_{11}^{\dagger}(e^{-i\lambda})^{-1}\Gamma_{12}^{\dagger}(e^{-i\lambda}), Im\Gamma_{11}^{\dagger}(e^{-i\lambda})^{-1}\Gamma_{12}^{\dagger}(e^{-i\lambda})\}$$

and let the stochastic expansion of $\psi(\hat{\theta}, \lambda)$ be

$$\sqrt{T}(\psi(\hat{\theta}, \lambda) - \psi(\theta, \lambda)) = (D_\theta\psi(\theta, \lambda))\sqrt{T}(\hat{\theta} - \theta) + o_p(1),$$

where $D_\theta\psi(\theta, \lambda)$ is the Jacobian matrix of $\psi(\theta, \lambda)$; then we have asymptotically

$$\sqrt{T}(\psi(\hat{\theta}, \lambda) - \psi(\theta, \lambda)) \overset{a}{\sim} N(0, H(\theta, \lambda)),$$

where $H(\theta, \lambda) = D_\theta\psi(\theta, \lambda)^*\Psi(\theta)D_\theta\psi(\theta, \lambda)$. The Wald statistic for the null hypothesis that $y(t)$ does not cause $x(t)$ in the presence of $z(t)$ is given by

$$W^{(n)} = T(\psi(\hat{\theta}, \lambda))^* H(\hat{\theta}, \lambda)^{-1}\psi(\hat{\theta}, \lambda),$$

where $W^{(n)}$ is, under the null hypothesis, asymptotically distribution as χ^2 distribution with degrees of freedom which is equal to the dimension of the vector ψ. Another approach of evaluating the Wald statistics is to estimate the covariance matrix $H(\hat{\theta}, \lambda)$ by applying the four-step Monte Carlo procedure presented above, in particular the formula (3.7).

Lastly consider the test of the overall OMO $M_{v\to u} = 0$. The component $\Gamma_{12}^{\dagger}(z)$ in (2.12) has a finite-order MA expression such that $\Gamma_{12}^{\dagger}(z) = \sum_{j=0}^{l_1}\Gamma_{12}^{\dagger}[j]z^j$ where the coefficients $\Gamma_{12}^{\dagger}[j]$ are in general nonlinear functions of θ; namely, $\Gamma_{12}^{\dagger}[j] = \Gamma_{12}^{\dagger}[j, \theta], j = 0, \cdots, l_1$. One way of testing the hypothesis $M_{v\to u} = 0$, namely $\{v(t)\}$ not causing $\{u(t)\}$, is to test $vec(\Gamma_{12}^{\dagger}[j, \theta], j = 0, \cdots, l_1) = 0$, which does not constitute a boundary-value test. Another method to test the null overall OMO is to test $\Gamma_{11}^{\dagger}(z)^\sharp\Gamma_{12}^{\dagger}(z) \equiv \sum_{j=0}^{l_2}\Delta[j]z^j = 0$, where $\Gamma_{11}^{\dagger}(z)^\sharp$ denotes the transposed cofactor matrix of $\Gamma_{11}^{\dagger}(z)$. The test is reduced to the test of $vec(\Delta[j, \theta], j = 0, \cdots, l_2) = 0$. For those tests, we can apply the Wald test approach, using the Whittle estimator $\hat{\theta}$ and its pertinent covariance matrix estimate.

4 Concluding remarks

By means of the cointegrated VAR model fitted to Japanese macroeconomic data, Hosoya (1997), Yao and Hosoya (2000) and Hosoya, Yao and Takimoto (2005) investigated the empirical one-way effect structure for a variety of pairs of variables. But the studies were limited to the simple one-way effects, whereas this paper presented a numerically practicable method which enables estimating and testing the partial causal measures introduced in Hosoya (2001) which explicitly takes confounding third series into account. In contrast to the simple causal measures, the numerical construction of the partial causal measures needs an explicit knowledge of canonical factor of a spectral density matrix involved. By implementing the numerical factorization procedure of Hosoya and Takimoto (2010), which is an improved version of the Rozanov (1967)'s factorization method, this paper provided a numerical procedure to evaluate the partial causal measures for stationary VARMA model. The paper shows that the evaluation of the measures is reducible to the one for a finite-order MA spectral density matrix. The paper presented a parametric statistical inference approach which consists of estimation based on the Whittle likelihood asymptotic theory, and testing and confidence-set construction relied on the standard limiting theory of the Wald statistics.

There remain some open problems. First of all, the paper has left it untouched to scrutinize numerically the performance of the proposed theory; the authors' research is in progress on these issues. To improve the performance of the Wald test in small-sample circumstances and the feasibility in application, employment of a certain time-series bootstrap method for probability evaluation and/or introduction of nonlinear transformation as proposed by Hosoya and Terasaka (2009) might be useful. Although Section 3 deals only with the stationary ARMA process mainly for the sake of expositional simplicity, extension to a wider class of processes is necessary for applications to empirical economic analyses. By utilizing the asymptotic covariance-matrix formula provided by Hosoya (1997), our statistical inference procedure can be extended to more general time-series models in which the disturbance series is possibly non-Gaussian. Hosoya, Yao and Takimoto (2005) took trend-breaks explicitly into account for testing the simple one-way effect measures in a cointegrated VAR set-up. The extension of the partial causal measures in that direction as well as the extension to nonlinear processes might be also important. But the most important open issue above all would be to develop a testing theory of the Granger causality which is more conformable to out-sample prediction, and thus to find a way to identify predictors endowed with substantial out-sample prediction ability; see Granger (1999) who emphasized the importance of this kind of research.

An enormous amount of empirical economic studies has dealt with predictive ability of the term structure and other asset-price characteristics for the future growth rate of economic activities and inflation rates. Stock and Watson (2003) and Wheelock and Wohar (2009) respectively give wide-ranging reviews of the literature; see also Hamilton and Kim (2002) and Assenmacher-Wesche, Gerlach and Sekine (2008) for example. A common understanding seems to be that the prediction ability of the term structure has fallen since the middle of 1980's in the U.S. economy and also that the predictive content of the original as well as the Freedman version of the Phillips curve is rather meager; see Staiger, Stock and Watson (1997) for the latter aspect.

Stock and Watson (2003) assert that in- sample tests of significance for Granger causality are, in general, poor guides for identification of potent predictors, providing little assurance that the identified predictive relations are stable. Although focusing not on the causality issue itself but on predictability, Wheelock and Wohar (2009) also note considerable variation of prediction ability of the term spread across countries and over time as far as prediction of a variety of economic-activity changes is concerned.

To be specific, Stock and Watson (2003) argue the problem of prediction ability, relying mainly on the single equation autoregressive-distributive lag model of the form

$$x_t = \sum_{j=1}^{a} \alpha_j x_{t-j} + \sum_{k=1}^{b} \beta_j y_{t-j} + \varepsilon_t. \tag{4.1}$$

The Granger test result in itself does not bring into question how much the prediction is improved by inclusion of the sum $\sum_{k=1}^{b} \beta_j y_{t-j}$ in case the null hypothesis $(\beta_1, \cdots, \beta_b) = 0$ is rejected. The problem is not indigenous to the Granger non-causality test. If a relation changes over time, it is natural to expect that in-sample observation is not extrapolated for out-sample prediction. Characteristically, while giving negative assessment to the Granger causality test in respect of prediction ability, Stock and Watson does not question directly the use of the Granger test when the stability of the relation (4.1) extends over a certain out-sample range; namely, they do not ask whether the rejection of non-causality indicates the usefulness of the corresponding variable over such a time interval of relative stability.

In case causal relation is stable over time, the relation between statistical and practical significance is reduced to the general dictum that a significant test result does not measure the practical significance. Even if the estimates of the β_j ' s are small in magnitude, they can be well significant when the corresponding standard errors are small and the Granger non-causality hypothesis is rejected, but it does not necessarily imply the notable prediction improvement by inclusion of those predictors. In contrast to statistical significance, confidence statements seem fit to represent the strength of effects. The one-way effect measures proposed in the paper are a way of quantifying prediction improvement and the suggested confidence sets would provide information the Granger causality test does not cover; see also Yao and Hosoya (2000) who suggested an approach of confidence-set construction of the OMO.

ACKNOWLEDGMENT

To conduct this research, the first author is partially supported by JSPS Grant-in-Aid for Scientific Research (C) 22530211 and the second author by JSPS Grant-in-Aid for Young Scientists (B) 22730180. Parts of the paper were presented in sessions of the Spring Meeting of Japan Association of Applied Economics (Fukuoka University, June 2012) and the Annual Meeting of Japan Statistical Society (Hokkaido University, September 2012). The authors are thankful to the comments given to the reports by the participants.

A Appendix

A.1 Proofs of Theorems

Proof of Theorem 2.2. It follows from the representation (2.12), the reciprocal component $\ddot{u}_{\cdot,\infty}(t)$ of $u(t)$ is given by

$$\ddot{u}_{\cdot,\infty}(t) = \Gamma^{\dagger}_{11}(L)\Sigma_{11}^{-1/2}\epsilon_1(t). \tag{A.1}$$

Similarly, setting

$$\Psi = \begin{bmatrix} \Sigma_{11:2} & 0 \\ 0 & \Sigma_{22}^{-1/2} \end{bmatrix} \begin{bmatrix} I_{p_1} & -\Sigma_{12}\Sigma_{22}^{-1} \\ 0 & I_{p_2} \end{bmatrix} \quad \text{and} \quad \xi(t) = \Psi\epsilon(t),$$

we have

$$\begin{aligned} \begin{bmatrix} u(t) \\ v(t) \end{bmatrix} &= \Gamma(L)\Gamma(0)^{-1}\Psi^{-1}\Psi\epsilon(t) \\ &\equiv \check{\Gamma}(L)\xi(t) \\ &\equiv \begin{bmatrix} \check{\Gamma}_{11}(L) & \check{\Gamma}_{12}(L) \\ \check{\Gamma}_{21}(L) & \check{\Gamma}_{22}(L) \end{bmatrix} \begin{bmatrix} \xi_1(t) \\ \xi_2(t) \end{bmatrix}. \end{aligned} \tag{A.2}$$

In view of the construction, $\{\xi_1(t)\}$ is the one-way effect componet process of $\{u(t)\}$ to $\{v(t)\}$. It follows from the representations (A.1) and (A.2), the reciprocal components of $u(t)$ and $v(t)$ are respectively given by

$$\ddot{u}_{\cdot,\infty}(t) = \Gamma_{11}^{\dagger}(L)\Sigma_{11}^{-1/2}\epsilon_1(t) \quad \text{and} \quad \ddot{v}_{\infty,\cdot}(t) = \check{\Gamma}_{22}(L)\Sigma_{22}^{-1/2}\epsilon_2(t).$$

Consequently, the joint spectral density matrix $\ddot{h}(\lambda)$ of the process $\{\ddot{u}_{\cdot,\infty}(t), \ddot{v}_{\infty,\cdot}(t)\}$ is given by

$$\begin{aligned}
\ddot{h}(\lambda) &= \frac{1}{2\pi}\begin{bmatrix} \Gamma_{11}^{\dagger}(e^{-i\lambda})\Sigma_{11}^{-1/2} & 0 \\ 0 & \check{\Gamma}_{22}(e^{-i\lambda})\Sigma_{22}^{-1/2} \end{bmatrix}\begin{bmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{bmatrix} \\
&\quad \cdot \begin{bmatrix} \Sigma_{11}^{-1/2}\Gamma_{11}^{\dagger}(e^{-i\lambda})^* & 0 \\ 0 & \Sigma_{22}^{-1/2}\check{\Gamma}_{22}(e^{-i\lambda})^* \end{bmatrix} \\
&= \frac{1}{2\pi}\begin{bmatrix} \Gamma_{11}^{\dagger}(e^{-i\lambda})\Gamma_{11}^{\dagger}(e^{-i\lambda})^* & \Gamma_{11}^{\dagger}(e^{-i\lambda})\Sigma_{11}^{-1/2}\Sigma_{12}\Sigma_{22}^{-1/2}\check{\Gamma}_{22}(e^{-i\lambda})^* \\ \check{\Gamma}_{22}(e^{-i\lambda})\Sigma_{22}^{-1/2}\Sigma_{21}\Sigma_{11}^{-1/2}\Gamma_{11}^{\dagger}(e^{-i\lambda})^* & \check{\Gamma}_{22}(e^{-i\lambda})\check{\Gamma}_{22}(e^{-i\lambda})^* \end{bmatrix} \\
&= \frac{1}{2\pi}\begin{bmatrix} \Gamma_{11}^{\dagger}(e^{-i\lambda}) & 0 \\ 0 & \check{\Gamma}_{22}(e^{i\lambda}) \end{bmatrix}\begin{bmatrix} I_{p_1} & \Sigma_{11}^{-1/2}\Sigma_{12}\Sigma_{22}^{-1/2} \\ \Sigma_{22}^{-1/2}\Sigma_{21}\Sigma_{11}^{-1/2} & I_{p_2} \end{bmatrix} \\
&\quad \cdot \begin{bmatrix} \Gamma_{11}^{\dagger}(e^{-i\lambda})^* & 0 \\ 0 & \check{\Gamma}_{22}(e^{-i\lambda})^* \end{bmatrix}.
\end{aligned} \tag{A.3}$$

Then it follows from (A.3) and the definition of $\ddot{\sigma}^2$ that

$$\det \ddot{h}_{11}(\lambda) \det \ddot{h}_{22}(\lambda) / \det \ddot{h}(\lambda) = \ddot{\sigma}^2. \quad \square$$

Proof of Theorem 2.4. Let $\epsilon_i^{\dagger}(t)$ and $\Gamma_{ij}^{\dagger}(L), i, j = 1, 2$, be defined as in (2.12) based on the factorization $h(\lambda) = (2\pi)^{-1}\Gamma(e^{-i\lambda})\Gamma(e^{-i\lambda})^*$. If the spectral density $k(\lambda)$ has the canonical factorization (2.16), we have the time domain representation, in parallel to (2.10),

$$\begin{aligned}
\begin{bmatrix} u(t) \\ v(t) \end{bmatrix} &= \gamma(L)\Gamma(L)\Gamma(0)^{-1}\begin{bmatrix} \epsilon_1(t) \\ \epsilon_2(t) \end{bmatrix} \\
&= \begin{bmatrix} \Gamma_{11}^{\dagger\dagger}(L) & \Gamma_{12}^{\dagger\dagger}(L) \\ \Gamma_{21}^{\dagger\dagger}(L) & \Gamma_{22}^{\dagger\dagger}(L) \end{bmatrix}\begin{bmatrix} \epsilon_1^{\dagger}(t) \\ \epsilon_2^{\dagger}(t) \end{bmatrix}
\end{aligned} \tag{A.4}$$

where

$$\Gamma_{ij}^{\dagger\dagger} = \gamma(L)\Gamma_{ij}^{\dagger}(L), i = 1, 2.$$

Hence we have

$$\begin{aligned} M_{v\to u}(\lambda) &\equiv \log \frac{\det\{\Gamma_{11}^{\dagger\dagger}(e^{-i\lambda})\Gamma_{11}^{\dagger\dagger}(e^{-i\lambda})^* + \Gamma_{12}^{\dagger\dagger}(e^{-i\lambda})\Gamma_{12}^{\dagger\dagger}(e^{-i\lambda})^*\}}{\det\{\Gamma_{11}^{\dagger\dagger}(e^{-i\lambda})\Gamma_{11}^{\dagger\dagger}(e^{-i\lambda})^*\}} \\ &\equiv \log \frac{\det\{\Gamma_{11}^{\dagger}(e^{-i\lambda})\Gamma_{11}^{\dagger}(e^{-i\lambda})^* + \Gamma_{12}^{\dagger}(e^{-i\lambda})\Gamma_{12}^{\dagger}(e^{-i\lambda})^*\}}{\det\{\Gamma_{11}^{\dagger}(e^{-i\lambda})\Gamma_{11}^{\dagger}(e^{-i\lambda})^*\}}. \qquad (A.5) \end{aligned}$$

Namely the right-hand side member of (A.5) implies that the FMO base on $k(\lambda)$ is the same one for the spectral density $h(\lambda) = \frac{1}{2\pi}\Gamma(e^{-i\lambda})\Gamma(e^{-i\lambda})^*$. In the same way, for the process given by (A.4), the joint spectral density matrix $\ddot{k}(\lambda)$ of the reciprocal-component process $\{\ddot{u}_{\cdot,\infty}(t), \ddot{v}_{\infty,\cdot}(t)\}$ is equal to $|\gamma(e^{-i\lambda})|^2\ddot{h}(\lambda)$ where $\ddot{h}(\lambda)$ the density given by (A.3). Therefore the frequency-wise measure of reciprocity is given by

$$\begin{aligned} M_{u.v}(\lambda) &= \log[\det\{|\gamma(e^{-i\lambda})|^2\ddot{h}_{11}(\lambda)\}\det\{|\gamma(e^{-i\lambda})|^2\ddot{h}_{22}(\lambda)\}/\det\{|\gamma(e^{-i\lambda})|^2\ddot{h}(\lambda)\} \\ &= \log\ddot{\sigma}^2. \quad \square \end{aligned}$$

A.2 The joint spectral density of the reciprocal components

A representation of the joint spectral density of the reciprocal components is given in (A.3), whereas this subsection presents another representation and some errata contained in Hosoya (1991) are corrected. Suppose that the joint process $\{u(t), v(t)\}$ introduced in Section 2 has the spectral representation with respect to a random spectral measure:

$$r(t) \equiv \begin{bmatrix} u(t) \\ v(t) \end{bmatrix} = \int_{-\pi}^{\pi} e^{i\lambda t} \begin{bmatrix} \Phi_u(d\lambda) \\ \Phi_v(d\lambda) \end{bmatrix} \equiv \int_{-\pi}^{\pi} e^{i\lambda t}\Phi_r(d\lambda) \quad t \in \mathbb{Z}.$$

Denote by $h(\lambda)$ the spectral density matrix of the process $\{r(t)\}$. Let $\tilde{h}$ and $\breve{h}$ be respectively the spectral densities of the reciprocal-component processes $\{u(t), v_{0,-1}(t)\}$ and $\{u_{-1,0}(t), v(t)\}$ and let the partitions of them be given by

$$\tilde{h}(\lambda) = \begin{pmatrix} \tilde{h}_{11}(\lambda) & \tilde{h}_{12}(\lambda) \\ \tilde{h}_{21}(\lambda) & \tilde{h}_{22}(\lambda) \end{pmatrix} \quad \text{and} \quad \breve{h}(\lambda) = \begin{pmatrix} \breve{h}_{11}(\lambda) & \breve{h}_{12}(\lambda) \\ \breve{h}_{21}(\lambda) & \breve{h}_{22}(\lambda) \end{pmatrix}.$$

Also denote the partition of the spectral density matrix $\ddot{h}(\lambda)$ of the joint process $\{\ddot{u}_{\cdot,\infty}(t), \ddot{v}_{\infty,\cdot}(t)\}$ by

$$\ddot{h}(\lambda) = \begin{bmatrix} \ddot{h}_{11}(\lambda) & \ddot{h}_{12}(\lambda) \\ \ddot{h}_{21}(\lambda) & \ddot{h}_{22}(\lambda) \end{bmatrix}.$$

Set $A = (-\Sigma_{21}\Sigma_{11}^{-1}, I_{p_2})$; $\quad B = (I_{p_1}, -\Sigma_{12}\Sigma_{22}^{-1})$, where

$$\Sigma = \begin{bmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{bmatrix}$$

is the covariance of the one-step ahead prediction error of the process $\{u(t), v(t)\}$.

Theorem A. The spectral density $\ddot{h}(\lambda)$ is represented as follows:

$$\begin{aligned}
\ddot{h}_{11}(\lambda) &= h_{11}(\lambda) - 2\pi h_{1\cdot}(\lambda)\Gamma(e^{-i\lambda})^{-1*}\Gamma(0)^* A^* \Sigma_{22:1}^{-1} A\Gamma(0)\Gamma(e^{-i\lambda})^{-1} h_{\cdot 1}(\lambda), && \text{(A.6)}\\
\ddot{h}_{22}(\lambda) &= h_{22}(\lambda) - 2\pi h_{2\cdot}(\lambda)\Gamma(e^{-i\lambda})^{-1*}\Gamma(0)^* B^* \Sigma_{11:2}^{-1} B\Gamma(0)\Gamma(e^{-i\lambda})^{-1} h_{\cdot 2}(\lambda), && \text{(A.7)}\\
\ddot{h}_{12}(\lambda) &= h_{12}(\lambda) - 2\pi h_{1\cdot}(\lambda)\Gamma(e^{-i\lambda})^{-1*}\Gamma(0)^*(A^*\Sigma_{22:1}^{-1}A + B^*\Sigma_{11:2}^{-1}B)\Gamma(0)\Gamma(e^{-i\lambda})^{-1} h_{\cdot 2}(\lambda)\\
&\quad -2\pi \tilde{h}_{12}(\lambda)\Sigma_{22:1}^{-1}(-\Sigma_{21} + \Sigma_{21}\Sigma_{11}^{-1}\Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21})\Sigma_{11:2}^{-1}\breve{h}_{12}(\lambda), && \text{(A.8)}
\end{aligned}$$

where $\Sigma_{11:2} = \Sigma_{11} - \Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21}$ and $\Sigma_{22:1} = \Sigma_{22} - \Sigma_{21}\Sigma_{11}^{-1}\Sigma_{12}$.

Proof. Since the proofs of the three components proceed in parallel ways, only the proof for (A.8) is given below. It follows from the definition of $\ddot{u}_{\cdot,\infty}(t)$ and $\ddot{v}_{\infty,\cdot}(t)$ which proceeds to D.2 of Section 2 that

$$\Phi_{\ddot{u}_{\cdot,\infty}}(d\lambda) - \Phi_u(d\lambda) - \tilde{h}_{12}(\lambda)\tilde{h}_{22}^{-1}(\lambda)\Phi_{v_{0,-1}}(d\lambda),$$

$$\Phi_{\ddot{v}_{\infty,\cdot}}(d\lambda) = \Phi_v(d\lambda) - \breve{h}_{21}(\lambda)\breve{h}_{11}^{-1}(\lambda)\Phi_{u_{-1,0}}(d\lambda), \tag{A.9}$$

whereas it follows from the definition of the one-way effect components that

$$\Phi_{v_{0,-1}}(d\lambda) - A\Gamma(0)\Gamma(e^{-i\lambda})^{-1}\Phi_r(d\lambda),$$

$$\Phi_{u_{-1,0}}(d\lambda) = B\Gamma(0)\Gamma(e^{-i\lambda})^{-1}\Phi_r(d\lambda). \tag{A.10}$$

Now in view of (A.9) the submatrix $\ddot{h}_{12}(\lambda)$ is given by

$$\begin{aligned}
\ddot{h}_{12}(\lambda) &= E\{\Phi_u(d\lambda) - \tilde{h}_{12}(\lambda)\tilde{h}_{22}^{-1}(\lambda)\Phi_{v_{0,-1}}(d\lambda)\}\{\Phi_v^*(d\lambda) - \Phi_{u_{-1,0}}^*(d\lambda)\breve{h}_{11}^{-1}(\lambda)\breve{h}_{12}(\lambda)\}\\
&= E\{\Phi_u(d\lambda)\Phi_v^*(d\lambda)\} - \tilde{h}_{12}(\lambda)\tilde{h}_{22}^{-1}(\lambda)E\{\Phi_{v_{0,-1}}(\lambda)\Phi_v^*(d\lambda)\}\\
&\quad -E\{\Phi_u(d\lambda)\Phi_{u_{-1,0}}^*(d\lambda)\}\breve{h}_{11}^{-1}(\lambda)\breve{h}_{12}(\lambda)\\
&\quad +\tilde{h}_{12}(\lambda)\tilde{h}_{22}^{-1}(\lambda)E\{\Phi_{v_{0,-1}}(\lambda)\Phi_{u_{-1,0}}^*(d\lambda)\}\breve{h}_{11}^{-1}(\lambda)\breve{h}_{12}(\lambda)\\
&= h_{12}(\lambda) - \tilde{h}_{12}(\lambda)\tilde{h}_{22}^{-1}(\lambda)A\Gamma(0)\Gamma(e^{-i\lambda})^{-1}h_{\cdot 2}(\lambda)\\
&\quad -h_{1\cdot}(\lambda)\Gamma(e^{-i\lambda})^{-1*}\Gamma(0)^*B^*\breve{h}_{11}^{-1}(\lambda)\breve{h}_{12}(\lambda)\\
&\quad +\tilde{h}_{12}(\lambda)\tilde{h}_{22}^{-1}(\lambda)A\Gamma(0)\Gamma(e^{-i\lambda})^{-1}h(\lambda)\Gamma(e^{-i\lambda})^{-1*}\Gamma(0)^*B^*\breve{h}_{11}^{-1}(\lambda)\breve{h}_{12}(\lambda)\\
&\equiv C1 - C2 - C3 + C4. && \text{(A.11)}
\end{aligned}$$

Since $h(\lambda)$ has the canonical factorization

$$h(\lambda) = \frac{1}{2\pi}\Gamma(e^{-i\lambda})\Gamma^*(e^{-i\lambda}),$$

the last member C4 on the right hand side of (A.11) is equal to

$$\tilde{h}_{12}(\lambda)\tilde{h}_{22}^{-1}(\lambda)A\frac{1}{2\pi}\Gamma(0)\Gamma(0)^*B^*\breve{h}_{11}^{-1}(\lambda)\breve{h}_{12}(\lambda)$$

$$= \frac{1}{2\pi}\tilde{h}_{12}(\lambda)\tilde{h}_{22}^{-1}(\lambda)A\Sigma B^*\breve{h}_{11}^{-1}(\lambda)\breve{h}_{12}(\lambda). \tag{A.12}$$

On the other hand, since

$$\begin{aligned} A\Sigma B^* &= (-\Sigma_{21}\Sigma_{11}^{-1}, I_{p_2})\begin{bmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{bmatrix}\begin{bmatrix} I_{p_1} \\ -\Sigma_{22}^{-1}\Sigma_{21} \end{bmatrix} \\ &= [0, -\Sigma_{21}\Sigma_{11}^{-1}\Sigma_{12} + \Sigma_{22}]\begin{bmatrix} I_{p_1} \\ -\Sigma_{22}^{-1}\Sigma_{21} \end{bmatrix} \\ &= \Sigma_{21}\Sigma_{11}^{-1}\Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21} - \Sigma_{21}, \end{aligned}$$

the last member in (A.12) is expressed as

$$\begin{aligned} &\frac{1}{2\pi}\tilde{h}_{12}(\lambda)\tilde{h}_{22}^{-1}(\lambda)(-\Sigma_{21} + \Sigma_{21}\Sigma_{11}^{-1}\Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21})\breve{h}_{11}^{-1}(\lambda)\breve{h}_{12}(\lambda) \\ &= 2\pi\tilde{h}_{12}(\lambda)\Sigma_{22:1}^{-1}(-\Sigma_{21} + \Sigma_{21}\Sigma_{11}^{-1}\Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21})\Sigma_{11:2}^{-1}\breve{h}_{12}(\lambda). \end{aligned}$$

It follows from (A.10) that

$$\begin{aligned} \tilde{h}_{12}(\lambda) &= E[\Phi_u(d\lambda)\Phi_{v_{0,-1}}^*(\lambda)] = h_{1\cdot}\Gamma(e^{-i\lambda})^{-1*}\Gamma(0)^*A^*, \\ \breve{h}_{12}(\lambda) &= E[\Phi_{u_{-1,0}}(d\lambda)\Phi_v^*(\lambda)] = B\Gamma(0)\Gamma(e^{-i\lambda})^{-1}h_{\cdot 2}(\lambda). \end{aligned} \tag{A.13}$$

Also, it follows from (A.13) that the second member C2 on the right-hand side of (A.11) is expressed as

$$\begin{aligned} &\tilde{h}_{12}(\lambda)\tilde{h}_{22}^{-1}(\lambda)A\Gamma(0)\Gamma(e^{-i\lambda})^{-1}h_{\cdot 2}(\lambda) \\ &= 2\pi h_{1\cdot}(\lambda)\Gamma(e^{-i\lambda})^{-1*}\Gamma(0)^*A^*\Sigma_{22:1}^{-1}A\Gamma(0)\Gamma(e^{-i\lambda})^{-1}h_{\cdot 2}(\lambda) \end{aligned} \tag{A.14}$$

whereas the third member C3 is given by

$$h_{1\cdot}(\lambda)\Gamma(e^{-i\lambda})^{-1*}\Gamma(0)^*B^*\breve{h}_{11}^{-1}(\lambda)\breve{h}_{12}(\lambda)$$

$$= 2\pi h_{1\cdot}(\lambda)\Gamma(e^{-i\lambda})^{-1*}\Gamma(0)^* B^* \Sigma_{11:2}^{-1} B\Gamma(0)\Gamma(e^{-i\lambda})^{-1} h_{\cdot 2}(\lambda).$$

Therefore the sum of the second and third members C2, C3 is equal to

$$2\pi h_{1\cdot}(\lambda)\Gamma(e^{-i\lambda})^{-1*}\Gamma(0)^*(A^*\Sigma_{22:1}^{-1}A + B^*\Sigma_{11:2}^{-1}B)\Gamma(0)\Gamma(e^{-i\lambda})^{-1}h_{\cdot 2}(\lambda). \qquad (A.15)$$

Hence (A.8) follows from (A.14) and (A.15). It provides a representation of the $(1, 2)$-th block of the joint spectral density of $\{\ddot{u}_{\cdot,\infty}(t), \ddot{v}_{\infty,\cdot}(t)\}$. By means of parallel arguments, we have (A.6) and (A.7). □

References.

Assenmacher-Wesche, K., Gerlach, S. and Sekine, T. (2008) "Monetary factors and inflation in Japan", *Journal of the Japanese and International Economies*, vol.22, pp.343-363.

Breitung, J. and Candelon, B. (2006) "Testing for short- and long-run causality: A frequency-domain approach", *Journal of Econometrics*, vol.132, pp.363–378.

Geweke, J. (1982) "Measurement of linear dependence and feedback between multiple time series", *Journal of the American Statistical Association*, vol.77, pp.304–313.

Geweke, J. (1984) "Measures of conditional linear dependence and feedback between time series", *Journal of the American Statistical Association*, vol.79, pp.907–915.

Granger, C.W.J. (1963) "Economic process involving feedback", *Information and Control*, vol.6, pp.28–48.

Granger, C.W.J. (1969) "Investigating causal relations by cross-spectrum methods", *Econometrica*, vol.39, pp.424–438.

Granger, C.W.J. (1980) "Testing for causality: a personal view-point", *Journal of Economic Dynamics and Control*, vol.2, pp.329–352.

Granger, C.W.J. (1999) *Empirical Modeling in Economics: Specification and Evaluation*, Cambridge: Cambridge University Press,

Granger, C.W.J. and Lin, J.L. (1995) "Causality in the long run", *Econometric Theory*, vol.11, pp.530–536.

Gronwald, M. (2009) "Reconsidering the macroeconomics of the oil price in Germany: testing for causality in the frequency domain, *Empirical Economics*, vol.36, pp.441–453.

Hamilton, J.D. and Kim, D.H. (2002) "A reexamination of the predictability of economic activity using the yield spread", *Journal of Money, Credit and Banking*, vol.34, pp.340–360.

Hannan, E.J. (1970) *Multiple Time Series*, Wiley, New York.

Hannan, E.J. and Rissanen, J. (1982) "Recursive estimation of mixed autoregressive-moving average order", *Biometrika*, vol.69, pp.81–94.

Hannan, E.J. and Kavalieris, L. (1984) "A method for autoregressive-moving average estimation", *Biometrika*, vol.71, pp.273–280.

Hosoya, Y. (1977) "On the Granger condition for non-causality", *Econometrica*, vol.45, pp.1735–1736.

Hosoya, Y. (1991) "The decomposition and measurement of the interdependency between second-order stationary processes", *Probability Theory and Related Fields*, vol.88, pp.429–444.

Hosoya, Y. (1997a) "Causal analysis and statistical inference on possibly non-stationary time series", in *Advances in Economics and Econometrics: theory and application, Seventh World Congress Vol.III, Chapter 1, pp.1–33*, eds D.M. Kreps and K.F. Wallis, Cambridge: Cambridge University press.

Hosoya, Y. (1997b) "A limit theory for long-range dependence and statistical inference on related models", *The Annals of Statistics*, vol.25, pp.105–137.

Hosoya, Y. (2001) "Elimination of third-series effect and defining partial measures of causality", *Journal of Time Series Analysis*, vol.22, pp.537–554.

Hosoya, T. and Takimoto, T. (2010) "A numerical method for factorizing the rational spectral density matrix", *Journal of Time Series Analysis*, vol.31, pp.229–240.

Hosoya, Y. and Terasaka, T. (2009) " Inference on transformed stationary time series", *Journal of Econometrics*, vol.151, pp.129–139.

Hosoya, Y., Yao, F. and Takimoto, T. (2005) "Testing the one-way effect in the presence of trend breaks", *The Japanese Economic Review*, vol.56, pp.107–126.

Hsiao, C. (1982) "Time series modelling and causal ordering of Canadian money, income and interest rates", in *Time Series Analysis: Theory and Practice I*, ed. O.D. Anderson, pp.671–98. Amsterdam: North-Holland.

Lütkepohl, H. and Reimers, H.E. (1992) "Granger causality in cointegrated VAR process", *Economics Letters*, vol.40, pp.263–268.

Mosconi, R. and Giannini, C. (1992) "Non-causality in cointegrated systems: Representation, estimation and testing", *Oxford Bulletin of Economics and Statistics*, vol.54, pp.399–417.

Rozanov, Y.A. (1967) *Stationary Random Processes*, San Francisco: Holden-day.

Sims, C.A. (1972) "Money, income and causality", *American Economic Review*, vol.62, pp. 540-552.

Staiger, D., Stock, J.H. and Watson, W.W. (1997) "The NAIRU, unemployment and monetary policy", *Journal of Economic Perspective*, vol. 11, pp.33–49.

Stock, J.H. and Watson, M.W. (2003) "Forecasting output and inflation: The role of asset prices", *Journal of Economic Literature*, XLI, pp.788–829.

Takimoto, T. and Hosoya, Y. (2004) "A three-step procedure for estimating and testing cointegrated ARMAX models", *The Japanese Economic Review*, vol.55, pp.418–450.

Takimoto, T. and Hosoya, Y. (2006) "Inference on the cointegration rank and a procedure for VARMA root-modification", *Journal of Japan Statistical Society*, vol.36, pp.149–171.

Toda, H.Y. and Phillips, P.C.B. (1993) "Vector autoregression and causality", *Econometrica*, vol.59, pp.229–255.

Yamamoto, T. and Kurozumi, E. (2006) "Tests for long-run Granger non-causality in cointegrated systems", *Journal of Time Series Analysis*, vol.27, pp.703–723.

Yao, F. and Hosoya, Y. (2000) "Inference on one-way effect and evidence in Japanese macroeconomic data", *Journal of Econometrics*, vol.98, pp.225–255.

Wheelock, D.C. and Wohar, M.E. (2009) "Can the term spread predict output growth and recession?", *Federal Reserve Bank of st. Louis Review*, September/October, part1, pp.419–440.

Discussion Paper Series

Number	Author	Title	Date
2000-1	Seiichi Iwamoto	Nearest Route Problem	2000/ 5
2000-2	Hitoshi Osaka	Productivity Analysis for the Selected Asian Countries : Krugman Critique Revisited	2000/ 5
2000-3	徳賀　芳弘	資産負債中心観への変化の検討 －会計観の変化と会計処理－	2000/ 7
2000-4	堀江　康熙	地域金融機関の不良債権問題	2000/ 7
2000-5	Hitoshi Osaka	Economic Development and Income Distribution : Survey and Regional Empirical Analysis	2000/10
2001-1	Yoshihiko Maesono	Nonparametric confidence intervals based on asymptotic expansions	2001/ 2
2001-2	堀江　康熙 川向　　肇	大都市所在信用金庫の営業地盤	2001/ 3
2001-3	Akinori Isogai	The Increasing Fluidity of Employment Re-examined	2001/ 4
2001-4	Hitoshi Osaka	Empirical Analysis on the Economic Effects of Foreign Aid	2001/ 5
2001-5	Tōru Nakai	Learning Procedure for a Partially Observable Markov Process and its Applications	2001/ 6
2001-6	Isao Miura	Secret Collusion and Collusion-proof Mechanism in Public Bidding	2001/ 8
2001-7	堀江　康熙	金融政策の有効性と貸出行動	2001/11
2001-8	大坂　　仁	環境クズネッツ曲線の検証： 国際データによるクロスカントリー分析	2001/11
2001-9	堀江　康熙 川向　　肇	信用金庫の営業地盤分析	2001/12

Number	Author	Title	Date
2002-1	Horie Yasuhiro	Economic Analysis of the "Credit Crunch" in the late 1990s	2002/ 3
2002-2	大坂　　仁	日本のODA政策と経済効果：民主主義と経済発展におけるアジア地域とサブサハラ・アフリカ地域の比較分析	2002/ 6
2002-3	Hirofumi Ito	Can the Local Allocation Tax Break Free of the Doldrums? - Japan's Development of and Difficulties with Fiscal Equalization	2002/ 9
2002-4	堀江　康熙	信用格付を用いた不良債権規模の推計	2002/ 10
2003-1	三浦　　功	長期公共契約の経済分析 —コミットメント，ラチェット効果および再交渉の問題—	2003/ 2
2003-2	Toshiyuki Fujita	Design of International Environmental Agreements under Uncertainty	2003/ 3
2003-3	Tōru Nakai	Some Thoughts on a Job Search Problem on a Partially Observable Markov Chain	2003/ 3
2003-4	Horie Yasuhiro	Monetary Policy and Problem Loans	2003/ 7
2003-5	磯谷　明徳	企業組織への契約論アプローチと能力論アプローチ —知識・制度・組織能力—	2003/ 8
2003-6	Horie Yasuhiro	Credit Rating and Nonperforming Loans	2003/ 9
2003-7	磯谷　明徳	制度経済学のエッセンスは何か	2003/ 11
2004-1	磯谷　明徳	制度とは何か	2004/ 2
2004-2	大坂　　仁	日本ODAの再考：国際資本フローと主要援助国の動向に関するデータからの考察	2004/ 2
2004-3	Toshiyuki Fujita	Game of Pollution Reduction Investment under Uncertainty	2004/ 10

Number	Author	Title	Date
2005-1	大坂　　仁	東アジアの所得配分と平等性の再検証	2005/ 3
2005-2	大坂　　仁	東アジアにおける成長会計分析の再考	2005/ 3
2005-3	佐伯　親良 福井　昭吾	産業連関分析―IOMetrics の開発―	2005/10
2005-4	Koichi Matsumoto	Optimal Growth Rate with Liquidity Risk	2005/11
2006-1	三浦　　功 川﨑　晃央	ネットワーク外部性下での逐次的価格競争と 最適特許戦略	2006/ 3
2006-2	石田　　修	市場の階層化と貿易構造	2006/ 3
2006-3	Koichi Matsumoto	Portfolio Insurance with Liquidity Risk	2006/ 4
2006-4	Kazushi Shimizu	The First East Asia Summit (EAS) and Intra-ASEAN Economic Cooperation	2006/ 7
2006-5	Yuzo Hosoya Taro Takimoto	A numerical method for factorizing the rational spectral density matrix	2006/ 8
2006-6	三浦　　功	公共入札における総合評価落札方式	2006/12
2007-1	佐伯　親良 福井　昭吾 森田　　充	所得分布と不平等度尺度の計量分析 ―PPID の開発―	2007/ 3
2007-2	Koichi Matsumoto	Mean-Variance Hedging in Random Discrete Trade Time	2007/ 4

Number	Author	Title	Date
2007-3	清水　一史	東アジアの地域経済協力とFTA —ASEAN域内経済協力の深化と東アジアへの拡大—	2007/ 6
2007-4	Kazushi Shimizu	East Asian Regional Economic Cooperation and FTA: Deepening of Intra-ASEAN Economic Cooperation and Expansion into East Asia	2007/ 7
2008-1	Naoya Katayama	Portmanteau Likelihood Ratio Tests for Model Selection	2008/ 1
2008-2	三浦　　功 大野　正久	ソフトな予算制約とスピルオーバー効果	2008/ 1
2008-3	Koichi Matsumoto	Dynamic Programming and Mean-Variance Hedging with Partial Execution Risk	2008/ 3
2008-4	Naoya Katayama	On Multiple Portmanteau Tests	2008/ 5
2008-5	Kazushi Shimizu	The ASEAN Charter and Regional Economic Cooperation	2008/ 7
2008-6	Noriyuki Tsunogaya Hiromasa Okada	Boundaries between Economic and Accounting Perspectives	2008/11
2009-1	Noriyuki Tsunogaya	Four Forms of Present Value Method: From the Standpoint of Income Measurement	2009/ 2
2009-2	日野　道啓	市場的手段の効果と環境イノベーションに関する一考察	2009/ 3
2009-3	松本　浩一 坪田　健吾	アメリカンオプション価格の上方境界の改善	2009/ 3

Number	Author	Title	Date
2009-4	Naoya Katayama	Simulation Studies of Multiple Portmanteau Tests	2009/ 4
2009-5	北澤　満	両大戦間期における三池炭の販売動向	2009/ 5
2009-6	Koichi Matsumoto	Option Replication in Discrete Time with Illiquidity	2009/ 6
2009-7	Naoya Katayama	合理的バブルの検定の検出力について	2009/ 7
2009-8	Mika Fujii Koichi Matsumoto Kengo Tsubota	Simple Improvement Method for Upper Bound of American Option	2009/ 7
2009-9	Kazushi SHIMIZU	ASEAN and the Structural Change of the World Economy	2009/ 9
2010-1	Tadahisa Ohno Akio Kawasaki	Who should decide the corporation tax rate?	2010/ 2
2010-2	大野　正久	環境税の分権的政策決定と民営化	2010/ 2
2010-3	Yuta Katsuki Koichi Matsumoto	Tail VaR Measures in a Multi-period Setting	2010/ 3
2010-4	清水　一史	ASEAN域内経済協力と生産ネットワーク —ASEAN自動車部品補完とIMVプロジェクトを中心に—	2010/ 6
2011-1	三浦　功	市場化テストの競争促進効果	2011/ 1

Number	Author	Title	Date
2011-2	Fujita Toshiyuki	Realization of a self-enforcing international environmental agreement by matching schemes	2011/ 2
2011-3	Yasuhisa Hirakata	British Health Policy and the Major Government	2011/ 2
2011-4	西釜　義勝 藤田　敏之	組織能力の構築メカニズムとリーダーシップの役割 －インドにおけるスズキの国際戦略を事例として－	2011/ 5
2011-5	Noriyuki Tsunogaya Chris Patel	The Accounting Ecology and Change Frameworks: The Case of Japan	2011/ 6
2011-6	三浦　　功 前田　隆二	医療機関の競争と最適リスク調整 ：Jack（2006）モデルの再検討	2011/ 7
2011-7	瀧本　太郎 坂本　直樹	国・都道府県レベルにおける歳入・歳出構造について	2011/ 8
2011-8	Kunio Urakawa Yusuke Kinari	Impact of the financial crisis on household perception - The case of Japan and the United States -	2011/10
2011-9	Koichi Matsumoto	Hedging Derivatives with Model Risk	2011/10
2011-10	三浦　　功	PFIを活用した公立病院の経営改革に関する経済分析 ：医療・介護の連携にシナジー効果が存在するケース	2011/10
2011-11	Yusuke Kinari	Time Series Properties of Expectation Biases	2011/11
2011-12	西釜　義勝 藤田　敏之	企業活性化に向けたイノベーションの検討 －自動車の環境技術開発の事例より－	2011/11

Number	Author	Title	Date
2012-1	阪田　和哉 瀧本　太郎 中嶌　一憲 生川　雅紀 坂本　直樹 阿部　雅浩	「心拍再開」の内生性を考慮したウツタイン統計 データによる救命曲線の推定	2012/ 9
2012-2	西釜　義勝 藤田　敏之	経営戦略論における資源アプローチの理論研究 －経営資源・能力論の展開－	2012/10
2012-3	Masaharu Kuhara	Employment Issues Involving Japanese Banks: A Case Study of Shinsei Bank	2012/10
2012-4	Yuzo Hosoya Taro Takimoto	Measuring the Partial Causality in the Frequency Domain	2012/12