Empirical Analysis of the Causal Relations in Japanese and Chinese Macroeconomic Data[1]

Feng Yao

Abstract This paper aims to show the characterization of causal structure of the recent Japanese and Chinese macroeconomy. For this purpose, we first give an introduction to the one-way effect causal measure and its Wald test as well as their computational algorithm. In view of the causal measures (in frequency domain and in time domain) in cointegrated vector time series, the long-run and short-run economic relationships are showed. We can also see the processes of applying the one-way effect causality theory to the analysis of macroeconomy.

1. Introduction

To solve the problems of determining the direction of causality between a pair of time series and also of statistically testing the absence of feedback, Granger (1963, 69) introduced a celebrated definition of causality. His

---

1) The research is partially supported by the Special Research Expenses presented by the Faculty of Economics, the Project Expenses presented by the Department of Economics, Kagawa University.

concept of causality is a statistically testable criterion defined in terms of predictability based on the assumption that the cause chronologically precedes the effect and the future does not cause the past. As far as testing absence of feedback relation is concerned, the earlier representative studies are the Granger test of zero restriction of specific coefficients of an stationary autoregressive representation, and the Sims test of the zero restriction of some coefficients in moving-average representation of stationary bivariate processes.

As regards testing Granger's non-causality in levels of a nonstationary vector autoregressive (VAR) system, Sims, Stock and Watson (1990) dealt with trivariate VAR systems, to conclude that the Wald test statistic has a limiting χ^2 distribution if the time series are cointegrated and otherwise that it has a nonstandard limiting distribution. Lütkepohl and Leimers (1992), using the Wald test for Granger's non-causality in bivariate cointegrated finite order AR process, investigated the short and long-term interest rates in the U. S., whereas Toda and Phillips (1993) extended the results of Sims, Stock and Watson (1990). So far, the interest of the econometric literature seems mostly concerned with Granger's non-causality test.

For the purpose of quantitative characterization of the feedback relationship between two multivariate time series, Geweke (1982) introduced an early version of the measure of causality from one time series to another in the time domain as well as in the frequency domain. Developing Geweke's frequency-domain approach, Hosoya (1991) introduced three causal measures summarizing the interdependency between a pair of nondeterministic stationary processes. Granger and Lin (1995) gave an extended measure of one-way effect for an nonstationary bivariate cointegrated process. Yao and Hosoya (1995) showed the algorithms of numerical computations of the causal measures in cointegrated relations. Yao (1996b) discusses algorithm of the one-way effect measure applied to Japanese macroeconomy involving structural changes and also gives an emperical analysis to the financial and external trade of Japan. Hosoya (1997) extended all his causal measures to nonstationary reproducible processes.

For the purposes of testing causal relations in cointegrated processes and constructing their confidence-sets, Yao and Hosoya (1998) introduced the Wald statistics. In contract to the conventional tests of Granger's non-causality which amount to testing the hypothesis of zero restriction of a certain set of autoregressive coefficients, the approach enables us to examine a variety of causal characteristics between time-series ; it can test not only Granger's non-causality by means of testing the nullity of the overall measure of one-way effect (*OMO*), but also the strength of the one-way effect. Moreover by means of the integral of the frequency-wise measure of one-way effect (*FMO*) on specific frequency bands, the long-run and short-run causal relationships can also be tested.

In this paper, we apply the Wald test theory presented by Yao and Hosoya (1998) to Japanese and Chinese macroeconomic data over the span of the recent twenty years. Our empirical analysis of Japanese macroeconomic data shows that, at 0.05 significance level, the one-way effect in neither direction between money and income is significant, but at 0.1 significance level a weak causal effect from money to income is detected. Our investigation also shows that the one-way effects from interest rates to the other variables are notably strong in general. In contrast, the effects in the reverse direction are weak and not significant. For certain cases, even though a single series does not cause significantly a specific series, a multiple series including that series is observed to cause the other series, indicating that policy mix might be effective in those circumstances. During the period we analyzed, the Japanese economic growth can be thought caused by the exports in conformity with the common understanding. The emperical results also show that the economic relation between

Japan and China, at least in the meaning of international trade, is not competitional.

This paper is organized as follows: Section 2 shows a heuristic exposition of the *OMO* and the *FMO* for stationary and nonstationary processes. Based on an ECM, Section 3 summarizes Wald test statistic for testing the one-way effect causal measures, and exhibits relevant computational procedures. Section 4 is for a preliminary data analysis of Japanese and Chinese macroeconomic time-series in order to identify pertinent ECM's for the causal analysis. In that section, we apply Johansen's likelihood ratio test for cointegration rank identification and apply extensively the Hosking statistic and the Doornik-Hansen statistic for testing serial uncorrelation and Gaussianity of the residuals. Section 5 deals with empirical causal analysis of 7 economic time serieses in the recent twenty years. The estimates of the *FMO* for bivariate and trivariate as well as four-variate models are exhibited in the figures. The estimated cointegration rank, estimates of *OMO* and causal test statistics are also listed in the corresponding figures. For the cases where causality is statistically significant, the confidence intervals of the true *OMO* are also listed in the corresponding figures. Section 6 concludes the paper.

Throughout the paper, we use the following notations and symbols. The set of all integers and the set of positive integers are denoted by Z and Z^+ respectively. For a set of random variables $\{Z_i,\ i \in A\}$ with finite second moment, $H\{Z_i,\ i \in A\}$ implies the closure in mean square of the linear hull of $\{Z_i,\ i \in A\}$ in the Hilbert space of random variables with finite second moment. For a p-vector process $X(t)$ with finite covariance matrix and for S a set of integers, $H\{X(t),\ t \in S\}$ implies $H\{X_i(t),\ t \in S,\ i = 1, \cdots, p\}$. A^* indicates the conjugate transpose if A is a complex matrix and the simple transpose if A is a real matrix. The *vec* operator transforms a $m \times n$ matrix B into a vector by stacking the columns of the matrix one underneath the other, i. e. *vec* B is the $m\cdot n\times 1$ vector, whereas $v(C)$ denotes the $n(n+1)/2$ vector that is obtained from *vec* C by eliminating all supradiagnoal elements of a square $n\times n$ matrix C. In this way, for symmetric C, $v(C)$ contains only the generically distinct elements of C. For a random vector X or for a pair of random vectors X and Y, $Cov(X)$ and $Cov(X,\ Y)$ indicate the variance-covariance matrix of X and of *vec* $(X,\ Y)$ respectively. The trace of a square matrix C is denoted by $\mathrm{tr}C$ and the determinant is denoted by $detC$. The Kronecker product of any $m\times n$ matrix A and $p\times q$ matrix B is denoted by the $mp\times nq$ matrix $A\otimes B$, whereas the sum of two vector subspaces H_1 and H_2 is denoted by $H_1\oplus H_2$. The lag operator denoted by L so that $Lx_t = x_{t-1}$ and the difference operator is denoted by $\Delta = 1-L$.

2. The Causal Measure of One-way Effect

The section shows the measures *OMO* and *FMO* for nondeterministic stationary time-series and extentions to nonstationary time-series in cointegrated relations [see for details Hosoya (1991, 1997), Yao and Hosoya (1998)]. In the last part of this section, we discuss long-run and short-run relationships expressed by those one-way effect measures.

The construction of the causal measures, in particular the measures of one-way effect, is closely related to the prediction theory of stationary processes. Suppose that $\{U(t),\ V(t),\ t\in Z\}$ is a zero mean jointly covariance stationary process where the $U(t)$ and $V(t)$ are $p_1\times 1$ and $p_2\times 1$ real vectors respectively $(p = p_1+p_2)$. Suppose also that the process $\{U(t), V(t)\}$ is nondeterministic and has the $p\times p$ spectral density matrix

$$f(\lambda) = \begin{bmatrix} f_{11}(\lambda) & f_{12}(\lambda) \\ f_{21}(\lambda) & f_{22}(\lambda) \end{bmatrix},\ -\pi < \lambda \leq \pi,$$

where $f_{11}(\lambda)$ is the $p_1\times p_1$ spectral density of $\{U(t)\}$, and that $f(\lambda)$ satisfies

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

Under the condition (2.1), $f(\lambda)$ has a factorization such that

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

where $\Lambda(e^{-i\lambda})$ is the boundary value $\lim_{\mu \to 1-} \Lambda(\mu e^{-i\lambda})$ of a $p \times p$ matrix-valued function $\Lambda(z)$ which is analytic has no zeros inside the unit disc $\{z : |z| < 1\}$ of the complex plane. Such a factorization is said to be a canonical factorization in the sequel. Let Σ be the covariance of the one-step ahead linear prediction error of the process $\{U(t),\ V(t)\}$ by its own past ; then, we have

$$det\{\Lambda(0)\Lambda(0)^*\} = det\Sigma = (2\pi)^p \exp\left\{\frac{1}{2\pi}\int_{-\pi}^{\pi} \log\, det\, f(\lambda) d\lambda\right\}, \tag{2.3}$$

[see Rozanov (1967) pp. 71-7, for example]. The relationship (2.2) is the frequency domain version of the Wold decomposition

$$\begin{pmatrix} U(t) \\ V(t) \end{pmatrix} = \sum_{j=0}^{\infty} \tilde{\Lambda}(j) \epsilon(t-j),$$

where $\{\epsilon(t)\}$ is a white-noise process with $Var\{\epsilon(t)\} = I_p$ and the matrices $\tilde{\Lambda}(j)$ are the real-matrix coefficients in the expansion of the analytic function $\Lambda(z)$; namely $\Lambda(z) = \sum_{j=0}^{\infty} \tilde{\Lambda}(j) z^j$.

The one-way effect component of $V(t)$ is the component which causes $\{U(t)\}$ one-sidedly but suffers no feedback from it in the Granger sense. We can extract such component from $V(t)$ as the regression residual obtained by regressing $V(t)$ on $\{U(t+1-j),\ V(t-j),\ j \in Z^+\}$. Formally, let $V_{0,-1}(t)$ be the residual of orthogonal projection (with respect to the mean-square) of $V(t)$ onto $H\{U(t+1-j),\ V(t-j),\ j \in Z^+\}$. It turns out that $\{V_{0,1}(t)\}$ is a white noise process with covariance matrix $\Sigma_{22} - \Sigma_{21}\Sigma_{11}^{-1}\Sigma_{12}$.

In contrast to the Wold decomposition of $\{U(t),\ V(t)\}$ which is a decomposition into an orthogonal sum in the time domain, $\{U(t),\ V(t)\}$ is known to have the spectral representation

$$U(t) = \int_{-\pi}^{\pi} e^{i\lambda t} d\tilde{U}(\lambda) \text{ and } V(t) = \int_{-\pi}^{\pi} e^{i\lambda t} d\tilde{V}(\lambda)$$

where $\tilde{U}(\lambda)$ and $\tilde{V}(\lambda)$ are (frequency-wise orthogonal) random measures such that

$$Cov\{d\tilde{U}(\lambda),\ d\tilde{V}(\lambda)\} = f(\lambda);$$

namely, the processes $\{U(t)\}$ and $\{V(t)\}$ are interpreted as weighted sums of harmonic oscillations with orthogonal random weight for the respective frequency. Hence the prediction error formula (2.3) implies for instance that the one-step ahead prediction error of $U(t)$ measured in terms of the determinant of the prediction error covariance matrix is the geometric mean of the $detCov\{d\tilde{U}(\lambda)\}$ over the frequency domain $-\pi < \lambda \leq \pi$. In other words, the variability of $d\tilde{U}(\lambda)$ expresses the frequency-wise contribution to the one-step ahead prediction error of $U(t)$. In the case of the joint one-step ahead prediction of $\{U(t),\ V(t)\}$, a similar argument applies and the variability expressed by $det\ Cov\{d\tilde{U}(\lambda),\ d\tilde{V}(\lambda)\}$ indicates the contribution of the λ-frequency oscillation to the joint prediction error of $U(t)$ and $V(t)$.

Then in view of the Granger concept of causality, the questions to be asked are how much of the prediction error reduction in $U(t)$ is attributed to the other series $\{V(s),\ s \leq t-1\}$ when it is added for the prediction of $U(t)$ and which portion of the variability in the pair $\{d\tilde{U}(\lambda),\ i\tilde{V}(\lambda)\}$, which is correlated in general, is attributable to the series $\{U(t)\}$. The pairing $\{U(t),\ V_{0,-1}(t)\}$ instead of the original pair $\{U(t),\ V(t)\}$ helps us to deal with these questions. In view of the construction of $V_{0,-1}(t)$, the projection residual of $U(t)$ onto $H\{V_{0,-1}(s);\ s \leq t-1\}$ is given by

$$U'(t) = \int_{-\pi}^{\pi} e^{it\lambda}\{d\tilde{U}(\lambda) - \tilde{f}_{12}(\lambda)\tilde{f}_{22}^{-1}(\lambda)d\tilde{V}_{0,-1}(\lambda)\}, \tag{2.4}$$

where the spectral density matrix of the process $\{U(t),\ V_{0,-1}(t)\}$ is denoted by the p_1 to p_2 partitioned matrix

$$\tilde{f}(\lambda)=\begin{bmatrix}\tilde{f}_{11}(\lambda) & \tilde{f}_{12}(\lambda)\\ \tilde{f}_{21}(\lambda) & \tilde{f}_{222}(\lambda)\end{bmatrix}$$

and $\tilde{f}_{11}(\lambda)=f_{11}(\lambda)$, $\tilde{f}_{21}(\lambda)=\{-\Sigma_{21}\Sigma_{11}^{-1},\ I_{p_2}\}\Lambda(0)\Lambda(e^{-i\lambda})^{-1}f_{\cdot 1}(\lambda)$, where $f_{\cdot 1}(\lambda)$ is the matrix which consists of the first p_1 columns of $f(\lambda)$, $\tilde{f}_{22}(\lambda)=\dfrac{1}{2\pi}\{\Sigma_{22}-\Sigma_{21}\Sigma_{11}^{-1}\Sigma_{12}\}$ [see Hosoya (1991), pp. 432-3, and also see Whittle (1963) for the spectral regression (2.4)]. Since the one-step ahead prediction error of $U(t)$ on the basis of $U(s)$ and $V_{0,-1}(s)(s\leq t-1)$ is the same as that of $U'(t)$ on the basis of its own past, it follows that

$$det\Sigma'_{11}=(2\pi)^{p_1}\exp\left[\frac{1}{2\pi}\int_{-\pi}^{\pi}\log\ det\ Cov\{d\tilde{U}(\lambda)-\tilde{f}_{12}(\lambda)\tilde{f}_{22}^{-1}(\lambda)d\tilde{V}_{0,-1}(\lambda)\}d\lambda\right], \tag{2.5}$$

where Σ'_{11} denotes the covariance matrix of the one-step ahead prediction error of $U'(t)$; whereas as for the prediction of $U(t)$ by its own past values, we have the relation

$$det\Sigma_{11}=(2\pi)^{p_1}\exp\left[\frac{1}{2\pi}\int_{-\pi}^{\pi}\log\ det\ Cov\{d\tilde{U}(\lambda)\}d\lambda\right]. \tag{2.6}$$

The comparison of (2.5) and (2.6) implies that the prediction improvement by the additional information of $V_{0,-1}(t)$ is given by

$$M_{V\to U}=\log\{det\Sigma_{11}/det\Sigma'_{11}\} \tag{2.7}$$

and that the frequency-wise reduction of the variability from $d\tilde{U}(\lambda)$ to $d\tilde{U}'(\lambda)$ is given by

$$M_{V\to U}(\lambda)=\log[det\ Cov\{d\tilde{U}(\lambda)\}/det\ Cov\{d\tilde{U}(\lambda)-\tilde{f}_{12}(\lambda)\tilde{f}_{22}^{-1}(\lambda)d\tilde{V}_{0,-1}(\lambda)\}]. \tag{2.8}$$

It turns out that $\{V(t)\}$ does not cause $\{U(t)\}$ in the Granger sense if and only if $M_{V\to U}=0$. Consequently, in conformity to Granger's causality concept, we might call $M_{V\to U}$ the overall measure of one-way effect (OMO)

from V to U and $M_{V\to U}(\lambda)$ the frequency-wise measure of one-way effect (FMO). It is obvious that $M_{V\to U}(\lambda)$ in (2.8) can also be expressed by

$$M_{V\to U}(\lambda)=\log[\det f_{11}(\lambda)/\det\{f_{11}(\lambda)-\tilde{f}_{12}(\lambda)\tilde{f}_{22}^{-1}(\lambda)\tilde{f}_{21}(\lambda)\}]. \tag{2.9}$$

Then OMO from V to Ucan be expressed by

$$M_{V\to U}=\frac{1}{2\pi}\int_{-\pi}^{\pi}M_{V\to U}(\lambda)d\lambda. \tag{2.10}$$

In order to extend this causal analysis of nondeterministic stationary time-series to nonstationary processes, consider the process $\{X(t),\ Y(t)\}$ which is generated by

$$A(L)\begin{bmatrix}X(t)\\ Y(t)\end{bmatrix}=\begin{bmatrix}U(t)\\ V(t)\end{bmatrix},\ (t=1,\ 2,\ \cdots) \tag{2.11}$$

where $\{U(t),\ V(t),\ t\in Z\}$ is the stationary process defined as before, and the lag polynomial matrix $A(L)$ is a $p\times p$ matrix such that

$$A(L)=\begin{bmatrix}\sum_{j=0}^{l}A_{11,j}L^j & 0\\ 0 & \sum_{j=0}^{l}A_{22,j}L^j\end{bmatrix}$$

for some positive l where $A_{11,0}=I_{p1}$ and $A_{22,0}=I_{p2}$. Suppose in the sequel that $X(t)$ and $Y(t)$ for $t\le 0$ are random vectors which belong to $H\{U(t),\ t\le 0\}$ and $H\{V(t),\ t\le 0\}$ respectively. The process given by (2.11) has the characteristic that the one-step ahead prediction and the residual of $X(t)$ based on $H\{X(t-j),\ j\ge 1\}\oplus H\{U(t),\ t\le 0\}$ and $Y(t),\ t\ge 1$, based on $H\{Y(t-j),\ j\ge 1\}\oplus H\{V(t),\ t\le 0\}$ are the same as those of $U(t)$ and $V(t)$ based on $H\{U(t-j),\ j\ge 1\}$ and $H\{V(t-j),\ j\ge 1\}$ respectively where $\oplus$ denotes the sum of vector subspaces. Similarly, the joint prediction of $\{X(t),\ Y(t)\}$ based on $H\{X(t-j),\ Y(t-j),\ j\ge 1\}\oplus H\{U(t),\ V(t),\ t\le 0\}$ is the same as the prediction of $\{U(t),\ V(t)\}$ based on $H\{U(t-j),\ V(t-j),\ j\ge 1\}$. Therefore the predictional properties of the process $\{X(t),\ Y(t)\}$ for $t\ge 1$ are entirely determined by those of the generating stationary process $\{U(t),\ V(t)\}$. Since the one-way effect structure of

$\{X(t),\ Y(t)\}$ is determined only by its predictional properties, it follows that it is given by the corresponding structure of $\{U(t),\ V(t)\}$. Namely, the *OMO* and the *FMO* between $\{X(t)\}$ and $\{Y(t)\}$ can be equated with the corresponding measures between the generating processes $\{U(t)\}$ and $\{V(t)\}$. This is the basic idea for the extension of the definitions of *OMO* and *FMO* to nonstationary processes.

It should be noted, however, that the relationship (2.11) is not very well defined. Suppose that $B(L)$ is another block diagonal matrix given by

$$B(L)=\begin{bmatrix} B_{11}(L) & 0 \\ 0 & B_{22}(L) \end{bmatrix},$$

where $B_{11}(L)$ and $B_{22}(L)$ are lag polynomials such that $B_{11,0}=I_{p_1}$ and $B_{22,0}=I_{p_2}$. The left multiplication of $B(L)$ to each member of the equation (2.11) produces a different representation of the process $\{X(t),\ Y(t)\}$. Unless $B(L)=I_p$, the resulting generating process $\{B_{11}(L)U(t),\ B_{22}(L)V(t)\}$ might possibly possess a spectral structure different from that of $\{U(t),\ V(t)\}$. In order to retain invariance of the one-way effect structure under such a multiplication, a certain restriction on the generating mechanism (2.11) is required. Let $f_{11}(\lambda)=\frac{1}{2\pi}\Lambda^{(1)}(e^{-i\lambda})\Lambda^{(1)}(e^{-i\lambda})^*$ and $f_{22}(\lambda)=\frac{1}{2\pi}\{U(t),\ V_{0,-1}(t)\}\Lambda^{(2)}(e^{-i\lambda})\Lambda^{(2)}(e^{-i\lambda})^*$ be canonical factorizations respectively.

Assumption 2.1. The process (2.11) satisfies either

(i) the zeroes of det $A_{11}(z)$ and det $A_{22}(z)$ are all on or outside of the unit disc; or

(ii) There are no common zeroes between $detA_{11}(z)$ and $det\Lambda^{(1)}(z)$ and between $detA_{22}(z)$ and $det\Lambda^{(2)}(z)$.

The preceding consideration leads us to the following extended definitions of the Granger non-causality and of the measures *OMO* and *FMO*. Suppose the process $\{X(t),\ Y(t),\ t=1,\ 2,\ \cdots\}$ generated by (2.11) satisfies Assumption 2.1.

Definition 2.1 $\{Y(t)\}$ is said not to cause $\{X(t)\}$ if and only if the prediction error covariance matrices of $X(t)$ based on $H\{U(s),\ V(s),\ s \leq t-1\}$ and based on $H\{U(s),\ s \leq t-1\}$ are identical.

Definition 2.2 The *OMO* $M_{Y\to X}$ and the *FMO* $M_{Y\to X}(\lambda)$ are defined by

$$M_{Y\to X} \equiv M_{V\to U} \text{ and } M_{Y\to X}(\lambda) \equiv M_{V\to U}(\lambda)$$

respectively.

Remark 2.1. Note that we have $H\{U(s),\ V(s),\ s \leq t\} = H(X(s),\ Y(s),\ s \leq t-1;\ U(s),\ V(s),\ s \leq 0\}$ and $H\{U(s),\ s \leq t-1\} = H\{X(s),\ s \leq t-1;\ U(s),\ s \leq 0\}$, and also that $\{Y(t)\}$ does not cause $\{X(t)\}$ if and only if $\{V(t)\}$ does not cause $\{U(t)\}$.

Now consider the p-dimensional process $Z(t) = \{X(t)^*,\ Y(t)^*\}^*$ represented by a finite a-th order VAR model

$$Z(t) = \sum_{j=1}^{a}\Pi(j)Z(t-j)+\varepsilon(t) \quad (t = 0,\ 1,\ \cdots), \tag{2.12}$$

where the $\Pi(j)$'s are $p\times p$ matrices, $\{\varepsilon(t)\}$ is a p-dimensional white noise process such that $E(\varepsilon(t)) = 0$, $Cov(\varepsilon(t)) = \Sigma$, and $rank\Sigma = p$. Set $A(L) = I_p - \sum_{j=1}^{a}\Pi(j)L^j$, where the zeros of $det\ A(z)$ are assumed to be either on or outside of the unit disc. Denote by $C(L)$ the adjoint matrix of $A(L)$ so that

$$C(L)A(L) \equiv D(L),$$

where $D(L)$ is the diagonal matrix having $d(L) \equiv detA(L)$ as the common diagonal element, $d(L) = \sum_{j=0}^{b}d_jL^j$ be a lag polynomial with scalar coefficients such that $d_0 = 1$ and the zeros of $\sum_{j=0}^{b}d_jz^j$ are either on or outside the unit circle. Left-multiplying $C(L)$ to the members of the equation (2.12), we have

$$\left[\begin{array}{ccc|ccc} d(L) & & 0 & & & \\ & \ddots & & & 0 & \\ 0 & & d(L) & & & \\ \hline & & & d(L) & & 0 \\ & 0 & & & \ddots & \\ & & & 0 & & d(L) \end{array}\right]\left[\frac{X(t)}{Y(t)}\right] = C(L)\epsilon(t) \equiv \left[\frac{U(t)}{V(t)}\right], \tag{2.13}$$

where we set $W(t) \equiv [U(t)^*, \ V(t)^*]^*$ for p_1 and p_2 vectors $U(t)$ and $V(t)$. It follows from the above construction that $\{U(t), \ V(t)\}$ is a stationary MA process and that the process $\{X(t), \ Y(t)\}$ satisfies the condition (2.8) and Assumption 2.1 (i). Therefore, in view of Definition 2.2 above, all the measures of one-way effect for the possibly nonstationary processes $\{X(t), Y(t)\}$ are determined by the corresponding measures of the stationary processes $\{U(t), \ V(t)\}$.

Moreover, since the zeros of $detC(z)$ are either on or outside the unit circle, the covariance matrix of the one-step ahead prediction error of $W(t)$ is equal to Σ and if the spectral density matrix of $\{W(t)\}$ is denoted by $f(\lambda)$, it has a canonical factorization

$$f(\lambda) = \frac{1}{2\pi}\Lambda(e^{-i\lambda})\Lambda(e^{-i\lambda})^*, \tag{2.14}$$

where $\Lambda(e^{-i\lambda}) = C(e^{-i\lambda})\Sigma^{1/2}$ for the Cholesky factor $\Sigma^{1/2}$ of Σ such that $\Sigma = \Sigma^{1/2}\Sigma^{1/2}$. Then the causal measures can be calculated in view of (2.9) and (2.10) by means of the spectral density $f(t)$ and a feasible $\Lambda(e^{-i\lambda})$.

A variety of causal measures can be derived on the basis of the *OMO* and the *FMO* between $\{X(t)\}$ and $\{Y(t)\}$ for the purposes of the long-run or short-run characterization of their causal relationships. In case $M_{Y \to X} \neq 0$, for example, the contribution of a long-run effect in the overall one-way effect is given by

$$D_{Y \to X}(\varepsilon) = \frac{1}{2\pi}\int_{-\varepsilon}^{\varepsilon} M_{Y \to X}(\lambda)d\lambda / M_{Y \to X},$$

for a certain low frequency band $[-\varepsilon, \ \varepsilon]$. In some cases, one might be more interested in the contribution of the relative effect for a given period band $[t_1, t_2](2 \leq t_1 < t_2)$, which is defined by

$$D_{Y\to X}(t_1, t_2) = \frac{1}{\pi}\int_{2\pi/t_2}^{2\pi/t_1} M_{Y\to X}(\lambda)\,d\lambda / M_{Y\to X},$$

where we used the relation $t = 2\pi/\lambda$ between period t and frequency $\lambda(\lambda > 0)$. Since the measure of one-way effect is nonnegative, those causal measures $D_{Y\to X}(\varepsilon)$ and $D_{Y\to X}(t_1, t_2)(2 \leq t_1 < t_2)$ take values in the interval $[0, 1]$, if $M_{Y\to X} < \infty$.

The long-run effect may be measured in another way, for example, by the mean *FMO* which is given by

$$\bar{D}_{Y\to X}(\varepsilon) = \frac{1}{2\varepsilon}\int_{-\varepsilon}^{\varepsilon} M_{Y\to X}(\lambda)\,d\lambda,$$

where ε is a certain small positive number. In order to summarize the one-way effect in a period band $[t_1, t_2]$,

$$\bar{D}_{Y\to X}(t_1, t_2) = \frac{t_1 t_2}{2\pi(t_2 - t_1)}\int_{2\pi/t_2}^{2\pi/t_1} M_{Y\to X}(\lambda)\,d\lambda$$

may be more useful. A small value of $\bar{D}_{Y\to X}(\varepsilon)$ indicates no substantial long-run effect from Y to X, and small $\bar{D}_{Y\to X}(t_1, t_2)$ implies that there is no notable one-way effect from Y to X for the period band (t_1, t_2). In any case, in order to interpret those quantities based on empirical data, we need a statistical testing theory.

Remark 2.2. The existence of the Nyquist frequency seems often ignored in exclusively time-domain oriented causal analyses. The discernible highest frequency is $\lambda = \pi$, which corresponds with two periods ($t = 2\pi/\lambda = 2$); namely, half a year for quarterly data. The economic implication is that we cannot discern the one-way effect shorter than half a year for quarterly data.

3. Testing Causality in Cointegrated VAR Processes

This section considers the Wald tests for testing hypotheses on the measures of one-way effect based on the ECM given by (3.1) below, providing the computational procedure and also applying the test statistics to construction of confidence-sets of those measures.

Let $\{Z(t)\} = \{X(t)^*,\ Y(t)^*\}^*$ be generated by a cointegrated p-vector AR model which is represented in the error-correction form

$$\Delta Z(t) = \alpha\beta^* Z(t-1) + \sum_{j=1}^{a-1}\Gamma(j)\Delta Z(t-j) + \mu + \Phi P(t) + \varepsilon(t), \tag{3.1}$$

where α and β are $p\times r$ matrices $(r \leq p)$, and μ is a constant p-vector. Also in (3.1), $P(t)$ is a column (s_d-1)-vector of centered seasonal dummy variables, where s_d is the seasonal period so that for quarterly data, $s_d = 4$; suppose also that $\{\varepsilon(t)\}$ is a Gaussian white noise process with mean 0 and with positive definite non-degenerate variance-covariance matrix Σ. Let θ be a $(r\cdot p)\times 1$ vector consisting of the elements of β such $\theta = vec\beta^*$. Denoting $n_\psi = p\cdot(r+p\cdot(a-1))+p\cdot(p+1)/2$, let ψ be the $n_\psi\times 1$ vector which consists of the elements of α and $\Gamma(j)(j = 1, \cdots, a-1)$ and the elements in the lower triangular part of Σ ; namely $\psi = vec(vec(\alpha, \Gamma)^*, v(\Sigma))$, where $\Gamma = \{\Gamma(1), \cdots, \Gamma(a-1)\}$ and $v(\Sigma)$ denotes the $(p\cdot(p+1)/2)\times 1$ vector.

The spectral density matrix f and its canonical factor Λ derived for the joint process $\{Z(t)\}$ by the relation (2.15) are given respectively by

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

and

$$\Lambda(e^{-i\lambda}|\theta,\ \psi) = C(e^{-i\lambda}|\theta,\ \psi)\Sigma^{1/2},$$

where $C(e^{-i\lambda}|\theta,\ \psi)$ is the adjoint matrix of the complex-valued polynomial matrix

$$I_p - e^{-i\lambda}(I_p + \alpha\beta^*) - \sum_{j=1}^{a-1}\Gamma(j)(e^{-ij\lambda} - e^{-i(j+1)\lambda}).$$

It is important to note here that the Granger causality is defined only between non-deterministic time-series and there is no one-way effect between such deterministic components as the dummy variables and the intercept which appear in model (3.1); a deterministic component can be predicted exactly by its past values and there is no improvement in prediction if information of another series is added [see Hosoya (1977) for a formal proof for non-causality between deterministic processes].

By means of those f and Λ, we define $M_{Y\to X}(\lambda|\theta, \psi)$ the *FMO* from $\{Y(t)\}$ to $\{X(t)\}$, by (2.11) and the *OMO* by

$$G(\theta, \psi) = \frac{1}{2\pi}\int_{-\pi}^{\pi} M_{Y\to X}(\lambda/\theta, \psi)d\lambda. \tag{3.3}$$

Note that in these instances, $G(\theta, \psi)$ is differentiable functions with respect to (θ, ψ).

Johansen (1988, 1991) showed that, if (θ, ψ) is the true value and $(\hat{\theta}, \hat{\psi})$ is the ML estimate, $T(\hat{\theta}-\theta)$ tends to have a mixed multivariate normal distribution and $\sqrt{T}(\hat{\psi}-\psi)$ tends to have a multivariate normal distribution as $T\to\infty$, whence $G(\hat{\theta}, \hat{\psi})$ is a $\sqrt{T}$ consistent estimate of $G(\theta, \psi)$. By the stochastic expansion, we have

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

where $D_\psi G$ is a n_ψ-dimensional vector of the gradient of $G(\theta, \psi)$. It follows that $\sqrt{T}\{G(\hat{\theta}, \hat{\psi}) - G(\theta, \psi)\}$ is asymptotically normally distributed with mean 0 and variance

$$H(\theta, \psi) = D_\psi G(\theta, \psi)^*\Psi(\theta, \psi)D_\psi G(\theta, \psi), \tag{3.4}$$

where $\Psi(\theta, \psi)$ is the asymptotic variance-covariance matrix of $\sqrt{T}(\hat{\psi}-\psi)$. Note that the first-order asymptotic distribution of $G(\hat{\theta}, \hat{\psi})$ is completely determined by $\hat{\psi}$ and the nonstandard limiting distribution of $\hat{\theta}$ is not involved, the sampling error of $\hat{\theta}$ being negligible in comparison with that of $\hat{\phi}$. Consequently, the test for $G(\theta, \phi)$ and the confidence-set construction can be conducted based on the Wald statistic

$$W \equiv T\{G(\hat{\theta}, \hat{\phi})-G(\theta, \phi)\}^2/H(\hat{\theta}, \hat{\phi}), \tag{3.5}$$

which is asymptotically distributed as χ^2 distribution with one degree of freedom if (θ, ϕ) is the true value.

As regards evaluation of $D_\phi G$ at $\hat{\theta}$, $\hat{\phi}$, the numerical differentiation is practical in view of the complexity of the exact analytic expression. Specifically, the gradient of $G(\theta, \phi)$

$$D_\phi G = \left(\frac{\partial G}{\partial \phi_1}, \cdots, \frac{\partial G}{\partial \phi_{n_\phi}}\right)^*$$

is evaluated by

$$\frac{\partial G}{\partial \phi_i} \approx \{-G(\hat{\theta}, \hat{\phi}+h_i)-G(\hat{\theta}, \hat{\phi}-h_i)\}/(2h), \tag{3.6}$$

for sufficiently small positive h where h_i is the $n_\phi \times 1$ vector with the i-th element h and all the other elements zero; namely, $h_i=(0, \cdots, h, 0, \cdots, 0)^*$, $i=1, 2, \cdots, n_\phi$.

The numerical computation of $\Psi(\theta, \phi)$ in (3.4) can be conducted as follows. We set $\phi^{(1)}=vec\{\alpha, \Gamma\}$, $\phi^{(2)}=vec(\mu, \Phi)$ and $\phi^{(3)}=\nu(\Sigma)$, and also we set $\phi^{(12)}=vec(\phi^{(1)}, \phi^{(2)})$. Then the log-likelihood function of the parameter $\phi^{(12)}$ and $\phi^{(3)}$ based on observation $Z(1), \cdots, Z(T)$ can be given as

$$l_T(\phi^{(12)}, \phi^{(3)}|Z)=-\frac{T}{2}(p\log 2\pi+\log det\Sigma)-\frac{1}{2}tr\Sigma^{-1}V_T,$$

where

$$V_T=\sum_{t=1}^{T} V(t)V(t)^*,$$

and

$$V(t)=\Delta Z(t)-\alpha\beta^* Z(t-1)-\sum_{j=1}^{a-1}\Gamma(j)\Delta Z(t-j)-\mu-\Phi P(t).$$

Let D be the p^2 by $p(p+1)/2$ duplication matrix and let D^+ be the Moore-

Penrose inverse of matrix D [see Magnus and Neudecker (1988), p49]. Denote by $\hat{\psi}^{(12)}$ and $\hat{\psi}^{(3)}$ the ML estimators of $\psi^{(12)}$ and $\psi^{(3)}$ respectively, then the asymptotic variance-covariance matrix of $\sqrt{T}\{\hat{\psi}^{(12)}-\psi^{(12)}\}$ and $\sqrt{T}$ $(\hat{\psi}^{(3)}-\psi^{(3)})$ is equal to

$$\begin{pmatrix} \Sigma\otimes Q^{-1} & 0 \\ 0 & 2D^{+}(\Sigma\otimes\Sigma)D^{+*} \end{pmatrix}, \tag{3.7}$$

where $Q=\lim_{T\to\infty}(1/T)\sum_{t=1}^{T}S(t)S(t)^*$,

$$S(t)=vec(\beta^*Z(t-1),\ \Delta Z(t-1),\ \cdots,\ \Delta Z(t-a-1),\ 1_p,\ P(t))$$

[see Magnus and Neudecker (1988), p321]. The asymptotic covariance of $\sqrt{T}(\hat{\psi}^{(1)}-\psi^{(1)})$, which is denoted by $\Psi_{\psi(1)\psi(1)}$ is then constructed from $\Sigma\otimes Q^{-1}$ by eliminating the rows and columns corresponding to $\sqrt{T}(\hat{\psi}^{(2)}-\psi^{(2)})$. In fact we can write the symmetric $(p\cdot(r+p\cdot(a-1))+p\cdot s_d)$ dimensional matrix $\Sigma\otimes Q^{-1}$ into $p\times p$ partitioned matrix in the form of

$$\begin{pmatrix} \sigma_{11}Q^{-1} & \sigma_{12}Q^{-1} & \cdots & \sigma_{1p}Q^{-1} \\ \sigma_{21}Q^{-1} & \sigma_{22}Q^{-1} & \cdots & \sigma_{2p}Q^{-1} \\ \vdots & \vdots & \vdots & \vdots \\ \sigma_{p1}Q^{-1} & \sigma_{p2}Q^{-1} & \cdots & \sigma_{pp}Q^{-1} \end{pmatrix},$$

where all of the submatrices $\sigma_{ij}Q^{-1}(i,\ j=1,\ \cdots,\ p)$ are $(r+p\cdot(a-1)+s_d)$ dimensional squared matrix. The covariance matrix $\Psi\psi_{(1)}\psi_{(1)}$ is constructed by eliminating all the last s_d columns and the last s_d rows of the submatrices $\sigma_{ij}Q^{-1}$, $i,\ j=1,\ \cdots,\ p$.

As for the estimation of Σ and Q in (3.7), we set

$$\hat{\Sigma}=(1/T)\sum_{t=1}^{T}(\hat{V}(t)\hat{V}(t)^*), \tag{3.8}$$

$$\hat{Q}=(1/T)\sum_{t=1}^{T}\hat{S}(t)\hat{S}(t)^*, \tag{3.9}$$

where

$$\hat{V}(t)=\Delta Z(t)-\hat{\alpha}\hat{\beta}^*Z(t-1)-\sum_{j=1}^{a-1}\hat{\Gamma}(j)\Delta Z(t-j)-\hat{\mu}-\hat{\Phi}P(t),$$

and

$$\hat{S}(t) = vec(\hat{\beta}^* Z(t-1),\ \Delta Z(t-1),\ \cdots,\ \Delta Z(t-a-1),\ 1_p,\ P(t)).$$

In view of the consistency of $\hat{\phi}$ and $\hat{\theta}$, if $\hat{\Psi}\phi_{(1)}\phi_{(1)}$ denotes the variance-covariance matrix of $\sqrt{T}(\hat{\phi}^{(1)}-\phi^{(1)})$ evaluated at $(\hat{\theta},\ \hat{\phi})$, then

$$\Psi(\theta,\ \phi) = \begin{pmatrix} \hat{\Psi}\phi_{(1)}\phi_{(1)} & 0 \\ 0 & 2D^+(\hat{\Sigma}\otimes\hat{\Sigma})D^{+*} \end{pmatrix} + o_p(1). \tag{3.10}$$

Therefore we can use the first right-hand side member of (3.10) as a consistent estimate of $\Psi(\theta,\ \phi)$.

By (3.4) and (3.10), we then get a variance estimate $\hat{H} = H(\hat{\theta},\ \hat{\phi})$. Denote G_{01} the given scalar, for the purpose of testing the null hypothesis $G(\theta,\ \phi) = G_{01}$, we evaluate the test statistic W defined by (3.5). In order to test no-causality in Granger's sense, we set the null hypothesis as $G_{01} = 0$ and the test statistic is given by

$$W = T\{G(\hat{\theta},\ \hat{\phi})\}^2/H(\hat{\theta},\ \hat{\phi}). \tag{3.11}$$

If $W \geq \chi^2_\alpha(1)$, for $\chi^2_\alpha(1)$ the upper α quantile of the χ^2 distribution with one degree of freedom, we may reject the null hypothesis of non-cansality from Y to X. On the other hand, in view of (3.5), the $(1-\alpha)$ confidence interval of th causal measure $G(\theta,\ \phi)$ is provided by

$$(G(\hat{\theta},\ \hat{\phi})-H_\alpha,\ G(\hat{\theta},\ \hat{\phi})+H_\alpha), \tag{3.12}$$

Where $H_\alpha = \sqrt{(1/T)H(\hat{\theta},\ \hat{\phi})\chi^2_\alpha(1)}$

Remark 3.1. Note that our algorithm for evaluating the Wald statistic and the confidence set does not depend upon the kind of measures of one-way effect so that it applies also to $\bar{D}_{Y\to X}(\varepsilon)$ or $\bar{D}_{Y\to X}(t_1,\ t_2)$, given in Section 2.

4. Preliminary Analysis

The test theory of the causal measures developed in the foregoing sections is applied in this section to macroeconomic data of Japan, in order to examine the performance of our Wald test in practice. The data used are the quarterly observations of GDP (Y), M2+CD (M), Call Rates (R), Exports (Ex), Imports (Im) in Japan during the period of the first quarter of 1975 through the fourth quarter of 1994. For the same period, the exports to China (Ex-JC or Exports-JC) and the imports from China (Im-JC or Imports-JC) are also investigated. The data of GDP, M2+CD as well as Call Rates are based on 'Economic Statistics Monthly', by Research and Statistics Department, Bank of Japan. The Exports and Imports data are in U.S. Dollar and are based on 'Balance of Payments Monthly', by International Department, Bank of Japan. The Ex-JC and Im-JC are the sum of monthly data (which is originally based on The Summary Report on Trade of Japan MOF) from Nikkei NEEDS Macro Database. Both of the Ex-JC and the Im-JC are in U. S. Dollar. All the variables are nominal and, except for the Call Rates, are given in logarithmic scale. Figure

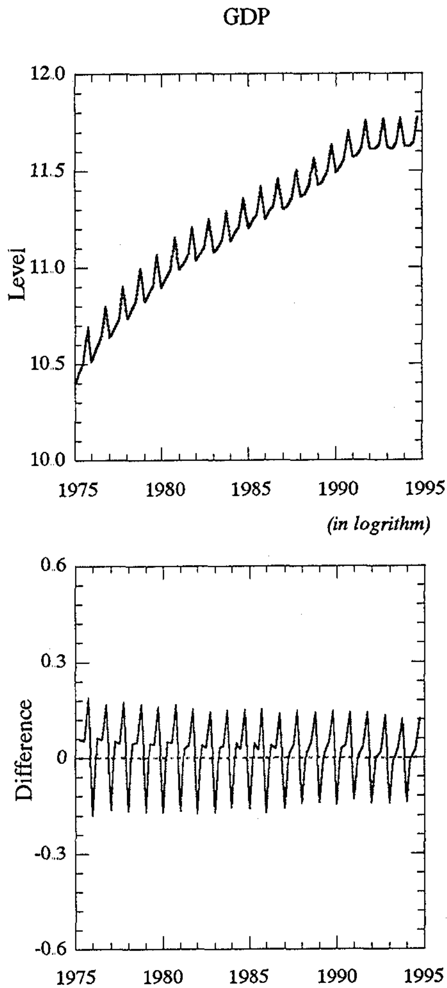


Figure 4.1 Japanese Macro Economic Data in Levels and Differences

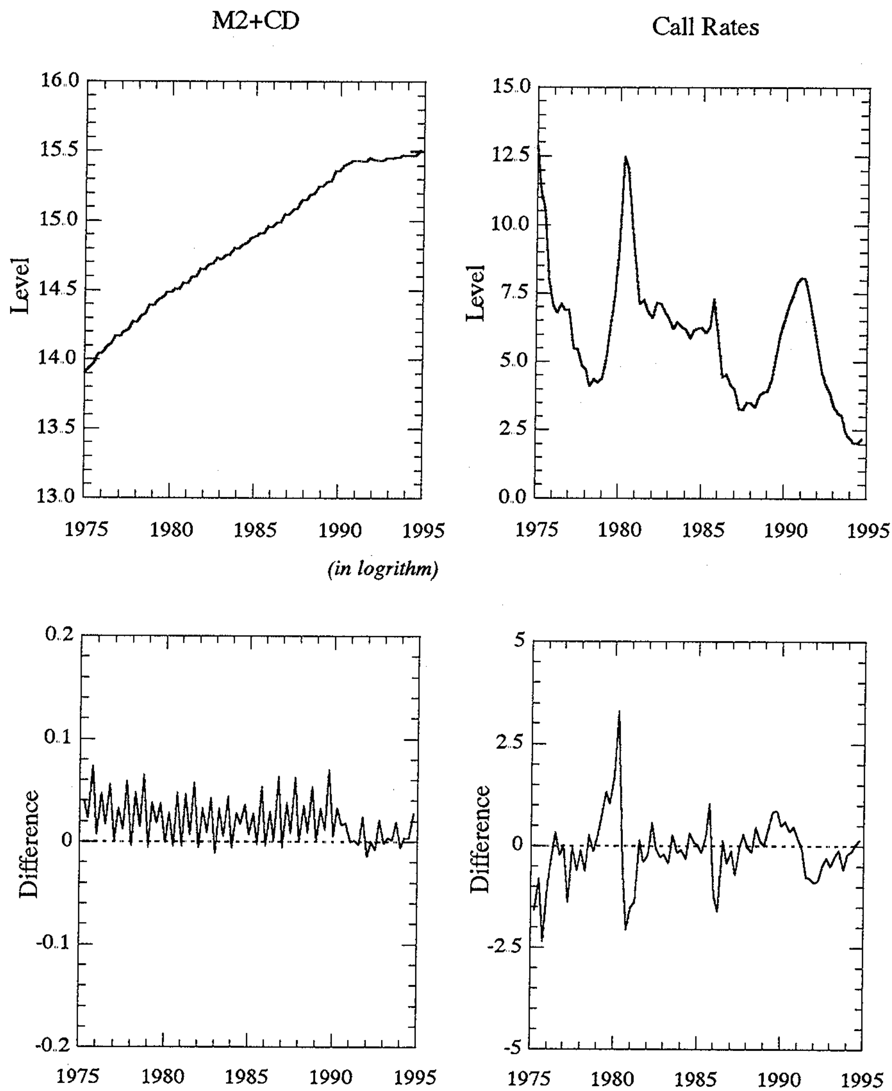


Figure 4.1 Continued

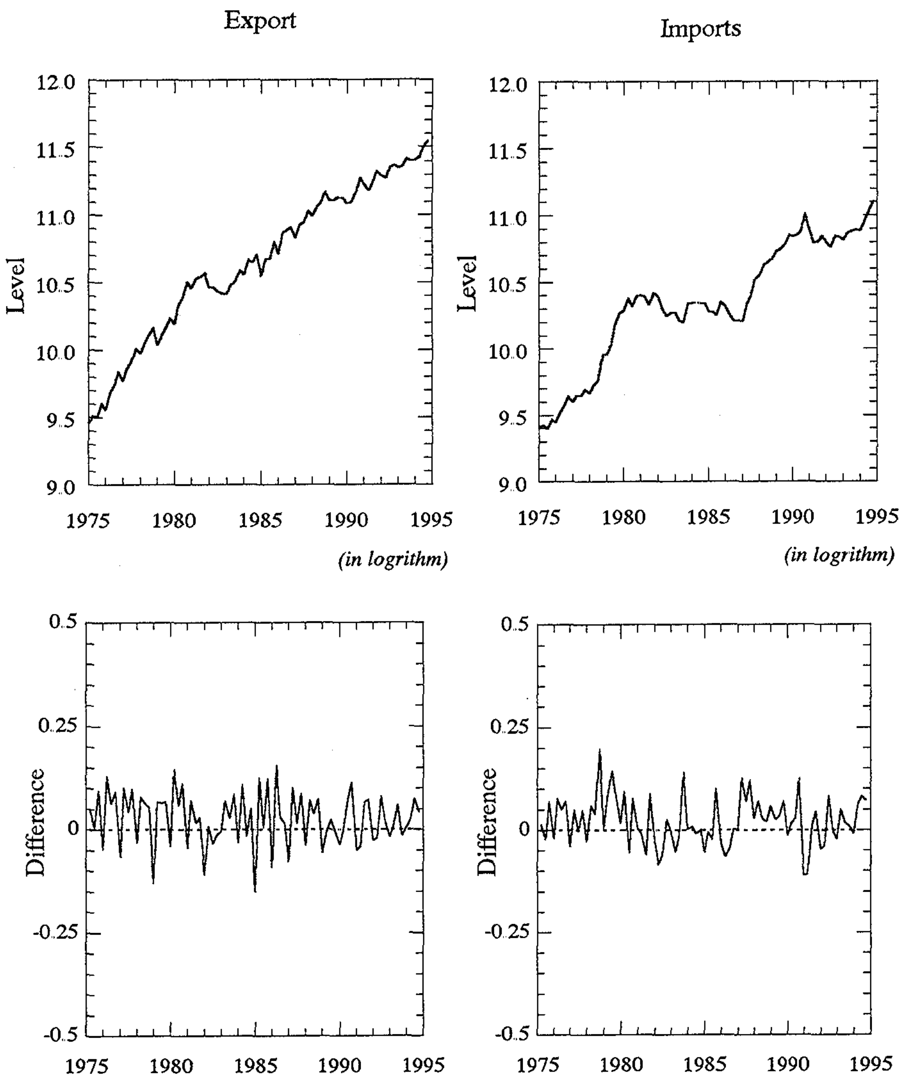


Figure 4.1 Continued

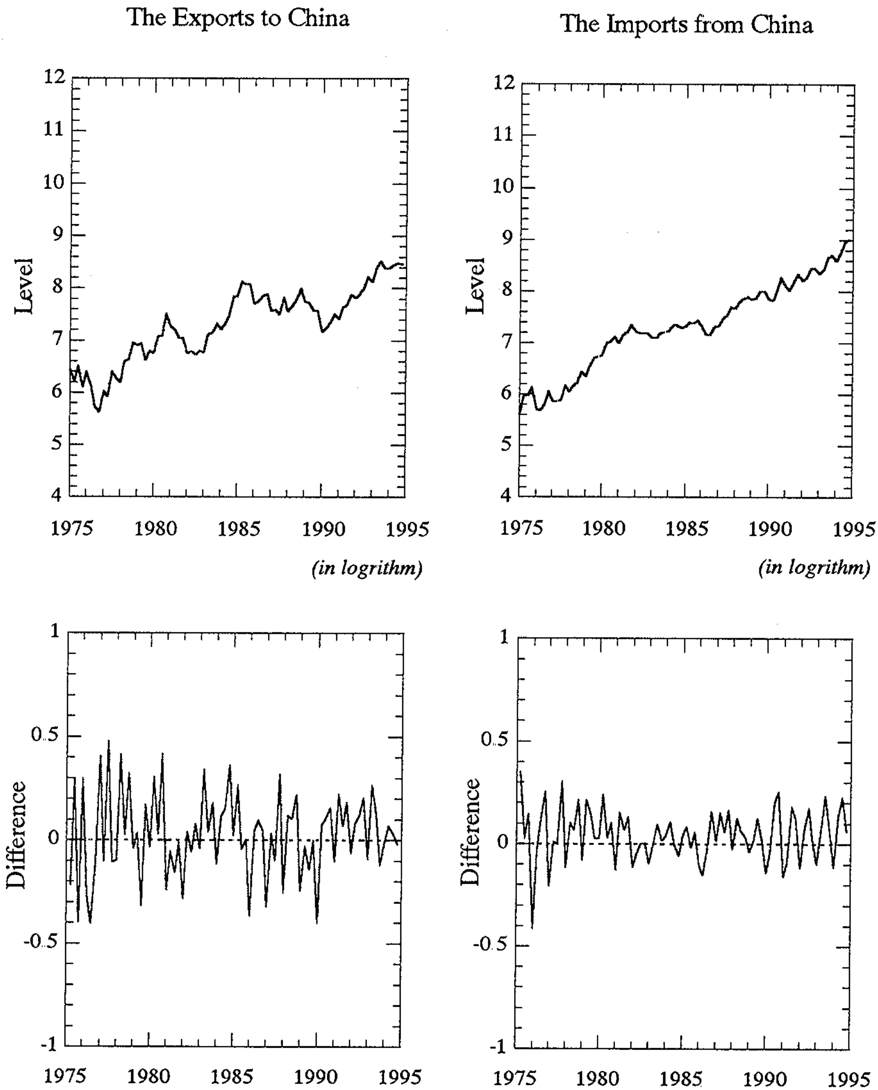


Figure 4.1 Continued

4.1 depicts the original data in levels and in differences. All the seven time series appear, to a reasonable extent, non-stationary with stationary differences.

In the following study, we apply the common lag-length $a=5$. The lag-length of autoregressive process delimits the range of possible configuration of the *FMO*. In order to avoid the lag-length playing a part in differentiation of the configuration, we do not use information criteria which are rather suited for identification of individual models. As is seen below, the uncorrelation and the Gaussianity hypotheses seem mostly supported for the residuals derived by fitting the lag-length $a=5$. The fitted model we used is the cointegrated p-dimensional AR(5) in ECM form represented by

$$\Delta Z(t) = \Pi Z(t-1) + \sum_{k=1}^{4} \Gamma(k) \Delta Z(t-k) + \mu + \Phi P(t) + \varepsilon(t), \qquad (4.1)$$

where $\varepsilon(t)$'s $(t=1, \cdots, T)$ are Gaussian white noise with mean 0 and variance-covariance matrix Σ, and we choose $P(t)$ the 3×1 vector of centered seasonal dummies so as not to produce seasonal trend effects in the level of $Z(t)$. The first 5 observations of $Z(t)$ are kept for initial values.

We summarize Johansen's ML test for cointegration rank [for the details see Johansen (1988, 1991, 1995)]. The hypothesis of independent r cointegration vectors is

$$H(r): \Pi = \alpha\beta^*, \qquad (4.2)$$

where α, β are $p\times r$ matrices $(r \leq p)$ such that rank $(\Pi)=r$. If $r=0$, (4.1) reduces to a full-rank unit root process. If $r=p$, then Π is full rank and the process $Z(t)$ is stationary. Denote by $R_0(t)$ and $R_1(t)$ the residuals obtained by regressing $\Delta Z(t)$ and $Z(t-1)$ on $\Delta Z(t-1), \cdots, \Delta Z(t-k+1)$, 1_p, $P(t)$ respectively. Define a $p\times p$ matrix S_{ij} by

$$S_{ij} = T^{-1}\sum_{t=1}^{T} R_i(t) R_j(t)^*, \ (i, \ j=0, \ 1). \qquad (4.3)$$

Under the hypothesis (4.2), the ML estimator of Π is found by the following procedure [see Johansen (1995), Theorem 6.1]:

(1) First solve the equation

$$|\lambda S_{11} - S_{10} S_{00}^{-1} S_{01}| = 0, \tag{4.4}$$

which produces the decreasing sequence of eigenvalues $1 > \hat{\lambda}_1 > \cdots > \hat{\lambda}_p > 0$ and the matrix constituted by the corresponding eigenvectors $\hat{V} = (\hat{V}_1, \cdots, \hat{V}_p)$, which is normalized so the $\hat{V}^* S_{11} \hat{V} = I$.

(2) Given r, the ML estimator of β is $\hat{\beta} = (\hat{V}_1, \cdots, \hat{V}_r)$, for which

$$L_{max}^{-2/T}(H(r)) = |S_{00}| \prod_{i=1}^{r} (1-\hat{\lambda}_i). \tag{4.5}$$

The likelihood ratio test statistic for the hypothesis $H(r)$ against $H(p)$ is given by the 'trace' statistic $\tau_{trace}(r)$ (abbreviated as $\tau(r)$):

$$\tau(r) = -T \sum_{i=r+1}^{p} 1n(1-\hat{\lambda}_i). \tag{4.6}$$

The asymptotic distribution of $\tau(t)$ is nonstandard and quantile tables are given by Osterwald-Lenum (1992) based upon Monte Carlo simulations. In the case of there is no or little prior information about r, we might estimate r as follows: Denote by $\tau(i|1-\alpha)$ the $(1-\alpha)$ quantile of $\tau(i)$ and by $\hat{\tau}(i)$ be the observation of $\tau(i)$. If $\hat{\tau}(0) < \tau(o|1-\alpha)$, we chose $\hat{\tau} = 0$. For $r = 1, \cdots, p-1$, let $\hat{r}$ be the first r such that

$$\hat{\tau}(r-1) > \tau(r-1|1-\alpha), \text{ and } \hat{\tau}(r) < \tau(r|1-\alpha),$$

and if there is no such r, then set $\hat{r} = p$.

The estimates of the other parameters are obtained by OLS by setting $\Pi = \alpha\hat{\beta}^*$ in the equation (4.1). A variety of aspects of the identification problem are discussed by Johansen (1995), but we choose in our analysis the least restrictive model specification.

Remark 4.1. The numerical computations of the paper were conducted by FORTRAN programs [see Yao (1996a)]. By applying those programs to the seven macroeconomic series, we investigated bivariate, trivariate as well as four-variate models. Since the size of twenty-year quarterly data cannot be regarded as large, to be conservative, we use $T-n_{\phi}$ instead of the sample size T in (3.8), (3.9) and (3.11).

The estimated eigenvalues and the corresponding eigenvectors of the bivariate and trivariate as well as four-variate in ECM are given in tables 4.1.1, 4.1.2, 4.2, 4.3, respectively. The variables of the models are indicated in the tables. The observed trace statistics are also listed in the tables. The 90 and 95 percent quantiles in Table 4.3 for cointegrating rank $r=1$, 2, 3, 4 are from Table 1 in Osterwald-Lenum (1992). We estimate r in this paper based not only on the $\tau(r)$ statistic but also on the consideration of other aspects of data and the corresponding model. Consider for example the process of determining the cointegration rank r for the case of four-variate model where the necessary quantiles are listed in Table 4.3. It shows that $\hat{\tau}(0)=48.73>43.95=\tau(0|0.9)$ and $\hat{\tau}(1)=25.38<26.79=\tau(1|0.9)$. According to the above procedure we select $\hat{r}=1$, which is listed in figures 5.1 (c1) and (c2). Consider for another example the determination of the cointegration rank r for bivariate model $Z=(Y,\ R)^*$ where the necessary quantiles are listed in Table 4.1.2. Even though the observed test statistics indicate two cointegrated relations, considering the obvious nonstationary nature of the nominal GDP, we chose $\hat{r}=1$. The parameters α and $\Gamma(k)$, which will be used in the following causality analysis, are then estimated by the OLS method and denoted by $\hat{\alpha}$, $\hat{\Gamma}(k)$, $(k=1,\ 2,\ 3,\ 4)$, respectively.

A criterion for the lag length selection is that the resulting residuals are uncorrelated to a reasonable degree. This is checked by Portmanteau tests. In this paper, in stead of using the Ljung-Box test statistic [Ljung and Box (1978)] which is given by

$$LB(s)=T(T+2)\sum_{j=1}^{s}\frac{1}{T-j}tr\{\hat{C}_{0j}\hat{C}_{00}^{-1}\hat{C}_{0j}{}^{*}\hat{C}_{00}^{-1}\},$$

we use the following modified form given by Hosking (1980), which has

Table 4.1.1 The Eigenvalues and the Eigenvectors and the Trace Statistics for Bivariate Models

	Eigenvalues					Eigenvalues		
	(0.077 0.028)					(0.113 0.000)		
	Eigenvectors	$p-r$	$\hat{\tau}$			Eigenvectors	$p-r$	$\hat{\tau}$
M	0.583 -0.319	1	2.13		R	0.601 -0.105	1	0.01
Ic	-0.812 0.948	2	8.14		Ic	0.799 0.994	2	8.97
	Edgenvalues					Eigenvalues		
	(0.094 0.004)					(0.145 0.013)		
	Eigenvectors	$p-r$	$\hat{\tau}$			Eigenvectors	$p-r$	$\hat{\tau}$
Ec	7.405 7.691	1	0.29		R	0.902 0.991	1	1.01
Ic	0.286 0.286	2	7.69		Ec	0.432 -0.137	2	12.78
	Eigenvalues					Eigenvalues		
	(0.183 0.047)					(0.152 0.049)		
	Eigenvectors	$p-r$	$\hat{\tau}$			Eigenvectors	$p-r$	$\hat{\tau}$
Y	-0.766 0.993	1	3.64		M	0.840 0.968	1	3.80
Ec	0.643 -0.115	2	18.78		Ec	-0.543 -0.251	2	16.12

1. Y : GDP, M : M2+CD, R : Call Rates, Ex : Exports, Im : Imports, Ec : the Exports to China, Ic : the Imports from China.
2. r is the cointegration rank.
3. The notations are also used for the following tables.

better performance for small sample size :

$$Hg(s) = T^2 \sum_{j=1}^{s} \frac{1}{T-j} tr\{\hat{C}_{0j}\hat{C}_{00}^{-1}\hat{C}_{0j}^{*}\hat{C}_{00}^{-1}\}, \tag{4.7}$$

when

$$\hat{C}_{0j} = T^{-1} \sum_{t=j+1}^{T} \hat{\varepsilon}_t \hat{\varepsilon}_{t-j}^{*}.$$

Under the null hypothesis of uncorrelation, the distribution of this test statistic is approximated for large T and for $s > a$ by χ_2 distribution with degrees of freedom $f = p^2(s-a)$ where a is the lag length of the model. For our cases of bivariate, trivariate and four-variate models, we choose $s =$ 18 and the observed statistics are listed in Table 4.4. The results in Table 4.5 support that all the residuals in the models are reasonably uncorrelated.

Table 4.1.2 The Eigenvalues and the Eigenvectors and the Trace Statistics for Bivariate Models

	Eigenvalues					Eigenvalues			
	(0.146	0.035)				(0.112	0.055)		
	Eigenvectors		$p-r$	$\hat{\tau}$		Eigenvectors		$p-r$	$\hat{\tau}$
Y	0.834	-0.713	1	2.66	Y	0.986	0.997	1	4.28
M	-0.551	0.702	2	14.46	R	0.166	-0.075	2	13.16
	Eigenvalues					Eigenvalues			
	(0.152	0.049)				(0.119	0.034)		
	Eigenvectors		$p-r$	$\hat{\tau}$		Eigenvectors		$p-r$	$\hat{\tau}$
Y	0.840	0.968	1	3.80	Y	-0.620	0.902	1	2.60
Ex	-0.543	-0.251	2	16.12	Im	0.784	-0.432	2	12.12
	Eigenvalues					Eigenvalues			
	(0.152	0.051)				(0.204	0.036)		
	Eigenvectors		$p-r$	$\hat{\tau}$		Eigenvectors		$p-r$	$\hat{\tau}$
M	0.911	0.999	1	3.95	M	-0.735	0.909	1	2.77
R	0.413	0.018	2	16.33	Ex	0.678	-0.417	2	19.90
	Eigenvalues					Eigenvalues			
	(0.145	0.032)				(0.126	0.015)		
	Eigenvectors		$p-r$	$\hat{\tau}$		Eigenvectors		$p-r$	$\hat{\tau}$
R	0.262	0.259	1	2.45	R	0.316	-0.077	1	1.13
Ex	0.965	-0.966	2	14.19	Im	0.949	0.997	2	11.27
	Eigenvalues								
	(0.146	0.036)							
	Eigenvectors		$p-r$	$\hat{\tau}$					
Ex	-0.574	0.866	1	2.71					
Im	0.819	-0.499	2	14.53					

The Gaussian assumption of the disturbance term is checked by applying the omnibus test for multivariate normality given by Doornik and Hansen (1994) to the residuals of the estimated models [see, for technical details, Shenton and Bowman (1977) and Doornik and Hansen (1994)]. Let R_T^* be the $p\times T$ matrix of the residuals with sample covariance matrix $F = (f_{ij})$. Create a matrix D with the reciprocals of the standard deviations on the diagonal,

Table 4.2　The Eigenvalues and the Eigenvectors and the Trace Statistics for Trivariate Models

	The Eigenvalues (0.303 0.128 0.020)						The Eigenvalues (0.242 0.141 0.033)				
	The Eigenvectors			$p-r$	$\hat{\tau}$		The Eigenvectors			$p-r$	$\hat{\tau}$
Y	0.829	0.886	0.788	1	1.51	M	-0.746	-0.591	0.845	1	2.52
M	-0.560	-0.462	-0.615	2	11.75	R	-0.012	0.069	-0.017	2	13.93
R	-0.007	0.021	0.010	3	38.83	Ex	0.665	0.804	-0.535	3	34.70
	The Eigenvalues (0.201 0.131 0.031)						The Eigenvalues (0.258 0.134 0.031)				
	The Eigenvectors			$p-r$	$\hat{\tau}$		The Eigenvectors			$p-r$	$\hat{\tau}$
M	0.749	-0.508	0.820	1	2.37	Y	-0.794	-0.844	0.932	1	2.39
R	0.111	0.004	0.016	2	12.93	Ex	0.596	0.346	-0.147	2	13.15
Im	-0.653	0.861	-0.572	3	29.77	Im	-0.120	0.409	-0.331	3	35.53
	The Eigenvalues (0.164 0.106 0.034)						The Eigenvalues (0.169 0.115 0.032)				
	The Eigenvectors			$p-r$	$\hat{\tau}$		The Eigenvectors			$p-r$	$\hat{\tau}$
M	0.934	-0.797	0.919	1	2.59	M	0.749	-0.508	0.820	1	2.44
R	0.358	-0.018	0.012	2	10.99	Ec	-0.388	-0.365	0.069	2	11.58
Ic	0.021	0.603	-0.394	3	24.42	Ic	0.331	0.773	-0.443	3	25.48

$$D = diag(f_{11}^{-1/2}, \cdots, f_{pp}^{-1/2}), \tag{4.8}$$

and form the correlation matrix $C = DFD$. Define the transformed matrix of R_r by

$$R_c = HL^{-1/2}H^*DR_r^*, \tag{4.9}$$

where L is the diagonal matrix with the eigenvalues of C on the diagonal. The columns of H are the corresponding eigenvectors, such that $H^*H = I_p$ and $L = H^*CH$. Then we compute univariate skewness $\sqrt{b_{1i}}$ and kurtosis b_{2i} of each vector of the transformed R_c^*, $i = 1, \cdots, p$, where we follow the notations by Doornik and Hansen (1994). Under the null hypothesis of multivariate normal distribution of the residuals, the test statistic is asymptotically distributed as:

$$E_p = Z_1^*Z_1 + Z_2^*Z_2 \sim \chi^2(2p), \tag{4.10}$$

Table 4.3 The Eigenvalues and the Eigenvectors and the Trace Statistics for Four-Variate Models

		The Eigenvalues				
	(0.268	0.163	0.124	0.028)		
		The Eigenvectors			$p-r$	$\hat{\tau}$
M	0.735	0.083	−0.382	0.864	1	2.13
R	0.015	−0.021	0.047	0.006	2	12.07
Ex	−0.677	−0.584	0.781	−0.096	3	25.38
Im	0.041	0.808	−0.492	−0.495	4	48.73
		The Eigenvalues				
	(0.241	0.162	0.093	0.032)		
		The Eigenvectors			$p-r$	$\hat{\tau}$
M	0.369	−0.477	−0.518	0.110	1	2.42
R	0.566	−0.069	0.781	−0.508	2	9.73
Ec	−0.731	0.871	−0.342	0.854	3	22.97
Ic	0.097	0.097	−0.074	0.021	4	43.65

	Trace Statistics: τ-statistic		
$p-r$	80%	90%	95%
1	1.66	2.69	3.76
2	11.07	13.33	15.41
3	23.64	26.79	29.68
4	40.15	43.95	47.21

1. *The Trace statistic quantiles are from table 1 in Osterwald-Lenum (1992).*
2. *These quantities are also used for Tables 4.1.1, 4.1.2, 4.2.*

where $Z_1^* = (z_{11}, \cdots, z_{1p})$ and $Z_2^* = (z_{21}, \cdots, z_{2p})$ are determined by (4.11) and (4.12) given below in Remark 4.2. The observed test statistics E_p for all the models used in this paper are listed in Table 4.5. Those test statistics seem to indicate that there is no significant departure from Gaussianity. The results in tables 4.4 and 4.5 ensure us that we may proceed to the discussions on the one-way effect measurement on the basis of the proposed ECM's.

Remark 4.2. (i) For $i = 1, \cdots, p$, the transformation for the skewness $\sqrt{b_{1i}}$ into z_{1i} is due to D'Agostino (1970):

Table 4.4 The Hg-Statistics and the *p*-values

	Hg-Stat.	*p*-value		Hg-Stat.	*p*-value
Y&M	55.8605	0.3319	R&Im	51.3962	0.4976
Y&R	56.9648	0.2956	Ex&Im	68.1974	0.0653
Y&Ex	62.9925	0.1413	Y&Ex&Im	141.7412	0.0596
Y&Im	58.3853	0.2524	Y&M&R	129.4265	0.2037
M&R	42.1847	0.8325	M&R&Im	116.0028	0.5087
M&Ex	61.1332	0.1808	M&R&Ex	133.3932	0.1427
R&Ex	59.6746	0.2168	M&R&Ex&Im	229.2855	0.1492
Y&Ec	60.6232	0.1928	Ec&Ic	62.4273	0.1525
M&Ec	50.4564	0.5348	M&Ec&Ic	125.2771	0.2836
M&Ic	58.1883	0.2581	M&R&Ic	122.7860	0.3388
R&Ec	63.3306	0.1349	M&R&Ec&Ic	250.1354	0.0242
R&Ic	58.1436	0.2595			

1. Hg-statistic is defined by (4.7).
2. The degree of freedom of the Hg-Statistic is 4, 6 or 8 for bivariate model, trivariate model and four-variates models respectively. This is also true for the next Table 4.5.

Table 4.5 Testing Normality of Residuals

	Ep-Stat.	*p*-value		Ep-Stat.	*p*-value
Y&M	0.0940	0.9989	R&Im	0.1067	0.9986
Y&R	1.1102	0.8927	Ex&Im	0.8692	0.9289
Y&Ex	2.2849	0.6835	Y&Ex&Im	0.5600	0.9970
Y&Im	0.6071	0.9623	Y&M&R	5.0317	0.5398
M&R	5.6457	0.2272	M&R&Im	10.3059	0.1123
M&Ex	2.7026	0.6088	M&R&Ex	8.9454	0.1767
R&Ex	7.6296	0.1061	M&R&Ex&Im	12.5157	0.1296
Y&Ec	2.6474	0.6184	Ec&Ic	0.5454	0.9689
M&Ec	2.5871	0.6291	M&Ec&Ic	3.5588	0.7361
M&Ic	1.1489	0.8864	M&R&Ic	1.4793	0.9609
R&Ec	7.5628	0.1090	M&R&Ec&Ic	0.5714	0.9998
R&Ic	0.1205	0.9983			

1. Ep-statistic is defined by (4.10)

$$
\begin{aligned}
\beta &= 3(T^2+27T-70)(T+1)(T+3)/(T-2)(T+5)(T+7)(T+9),\\
\omega^2 &= -1+\{2(\beta-1)\}^{1/2},\\
\delta &= 1/\{\log\sqrt{\omega^2}\}^{1/2},\\
y &= \sqrt{b_{1i}}[(\omega^2-1)(T+1)(T+3)/\{12(T-2)\}]^{1/2},\\
z_{1i} &= \delta\log\{y+(y^2+1)^{1/2}\}.
\end{aligned}
\tag{4.11}
$$

(ii) For $i=1, \cdots, p$, the kurtosis b_{2i} is transformed from a gamma distribution to χ^2, and then transformed into standard normal z_{2i} using the Wilson-Hilferty cubed root transformation:

$$
\begin{aligned}
\delta &= (T-3)(T+1)(T^2+15T-4),\\
a &= (T-2)(T+5)(T+7)(T^2+27T-70)/6\delta,\\
c &= (T-7)(T+5)(T+7)(T^2+2T-5)/6\delta,\\
k &= (T+5)(T+7)(T^3+37T^2+11T-313)/12\delta,\\
\alpha &= a+b_{1i}c,\\
\chi &= (b_{2i}-1-b_{1i})2k,\\
z_{2i} &= \{(\chi/2\alpha)^{1/3}-1+(1/9\alpha)\}\{9\alpha\}^{1/2}.
\end{aligned}
\tag{4.12}
$$

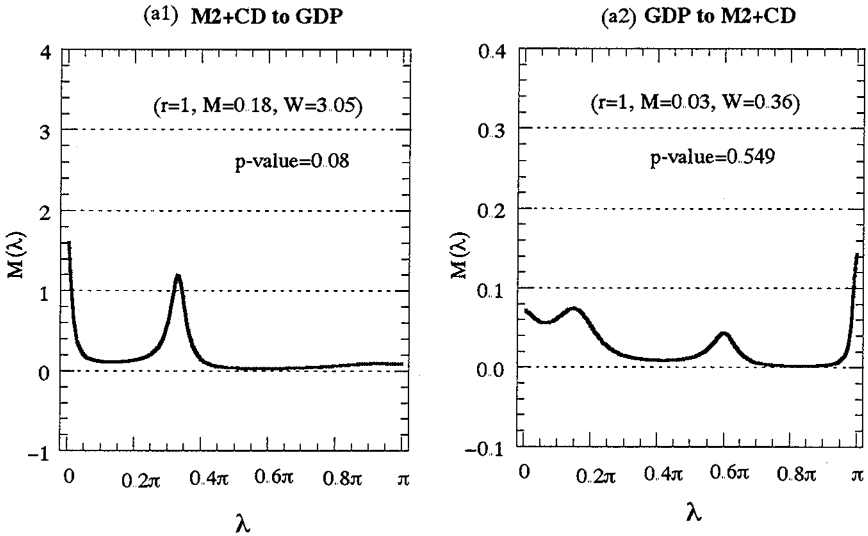
5. Empirical Measurement of One-way Effect

In order to illustrate the performance of the Wald test, we apply the Wald test in this section to the study of the causal relationships of Japanese macroeconomic time series. The following analyses are conducted on the model (4.1). The empirical examples would characterize the recent Japanese macroeconomy in view of the one-way causality.

For the model (4.1), let $\hat{C}(e^{-i\lambda})$ be the adjoint of the matrix

$$I_p-(I_p+\hat{\alpha}\hat{\beta}^*)e^{-i\lambda}-\sum_{j=1}^{4}\hat{\Gamma}(j)(e^{-ij\lambda}-e^{-i(j+1)\lambda})$$

as given in Section 3. Then the measures of one-way effect from Y to X are estimated on the basis of the frequency response estimate $\hat{\Lambda}(e^{-i\lambda}) = \hat{C}(e-^{i\lambda})\hat{\Sigma}^{1/2}$ and the spectral density estimate $\hat{f}(\lambda) = \frac{1}{2\pi}\hat{\Lambda}(e^{-i\lambda})\hat{\Lambda}(e^{-i\lambda})^*$. As for the numerical evaluation of $D_\phi G$ in (3.6), we choose $h = 0.0001$;



1. *r : Identified cointegration rank ;*
2. *W : Wald statistic given by (3.11)*
3. *CI : The 95% confidence interval of OMO (in case the non-causality null hypothesis is rejected).*
4. *Exports-JC means the exports to China, Imports-JC means the imports from China.*

Figure 5.1 Estimated measures of one-way effect, identified cointegration ranks, Wald statistics and confidence intervals

after having conducted evaluation of the Jacobian matrix for numerous choice including smaller h, we found that the results were sufficiently stable for $h = 0.0001$.

Figure 5.1 lists 26 plots of the estimated *FMO*. There, plots (a1) through (a15) show bivariate cases, and plots (b1) through (b7) show trivariate cases, while plots (c1) to (c4) are for four-variate. The estimates of cointegrating rank (r) and the *OMO* (M) as well as the Wald test statistic

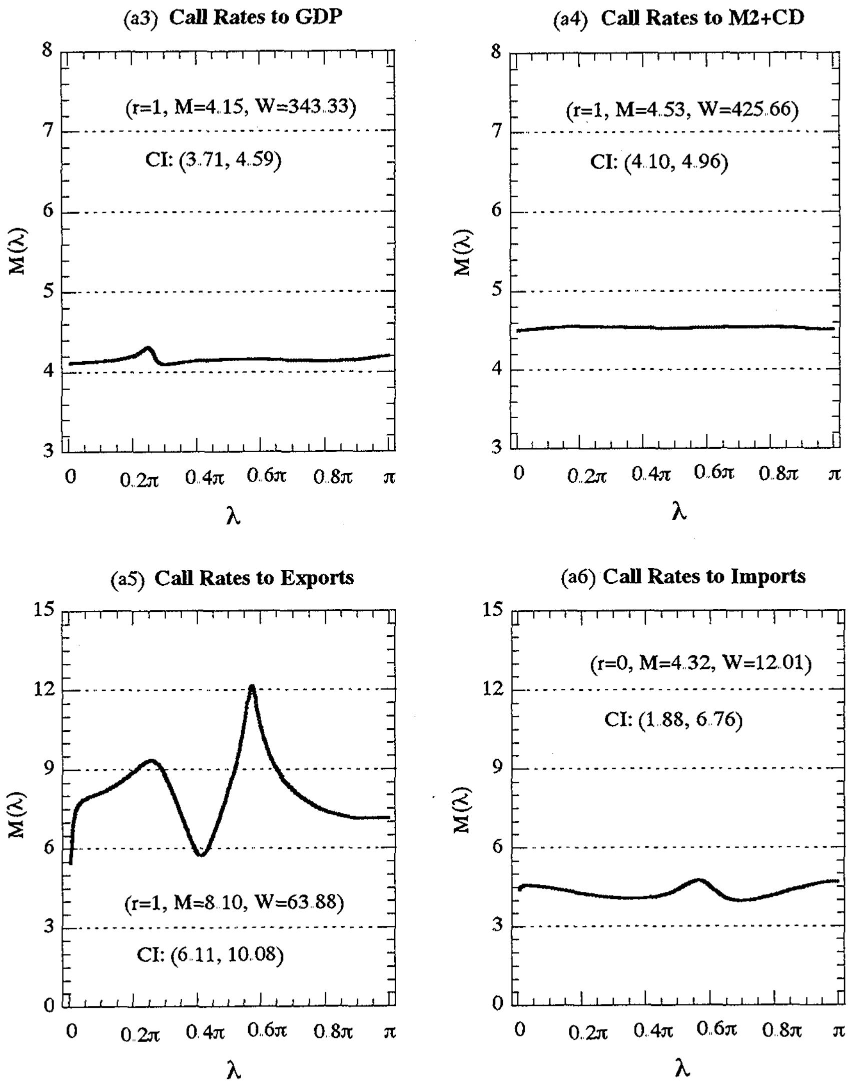


Figure 5.1 Continued

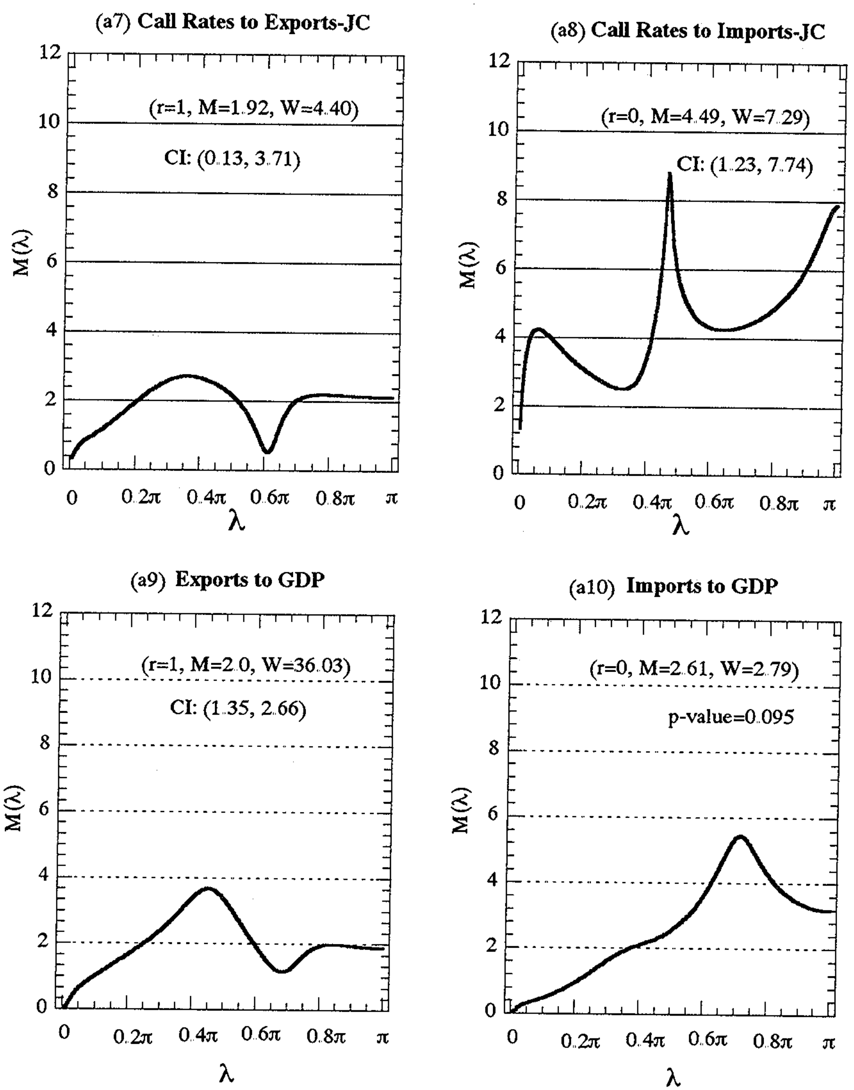
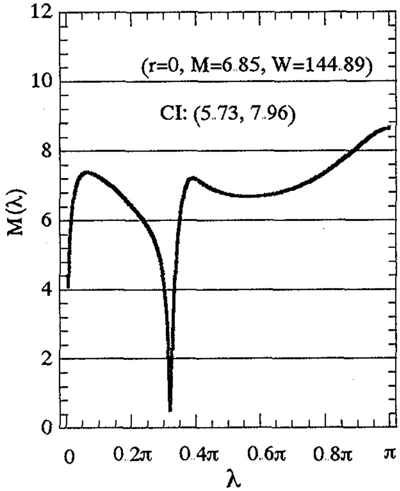
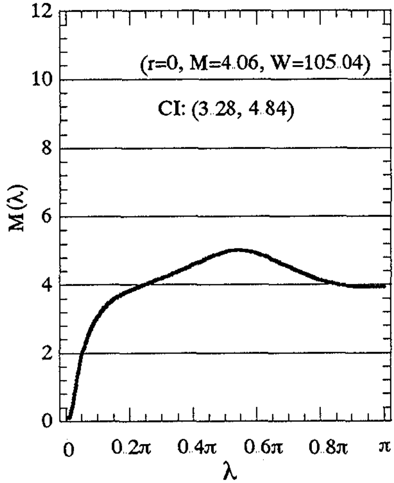


Figure 5.1 Continued

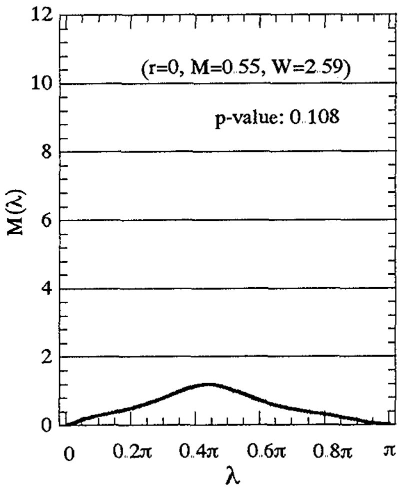
(a11) **Exports-JC to M2+CD**



(a12) **Imports-JC to M2+CD**



(a13) **Exports-JC to Imports-JC**



(a14) **Imports-JC to Exports-JC**

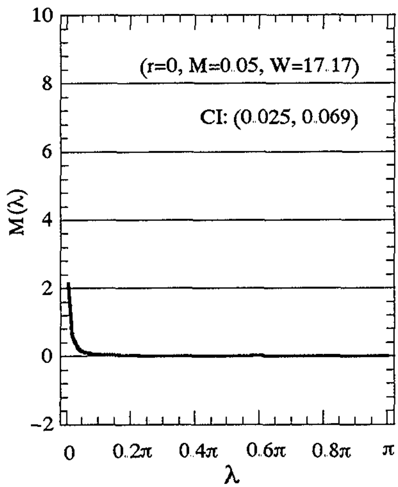


Figure 5.1 Continued

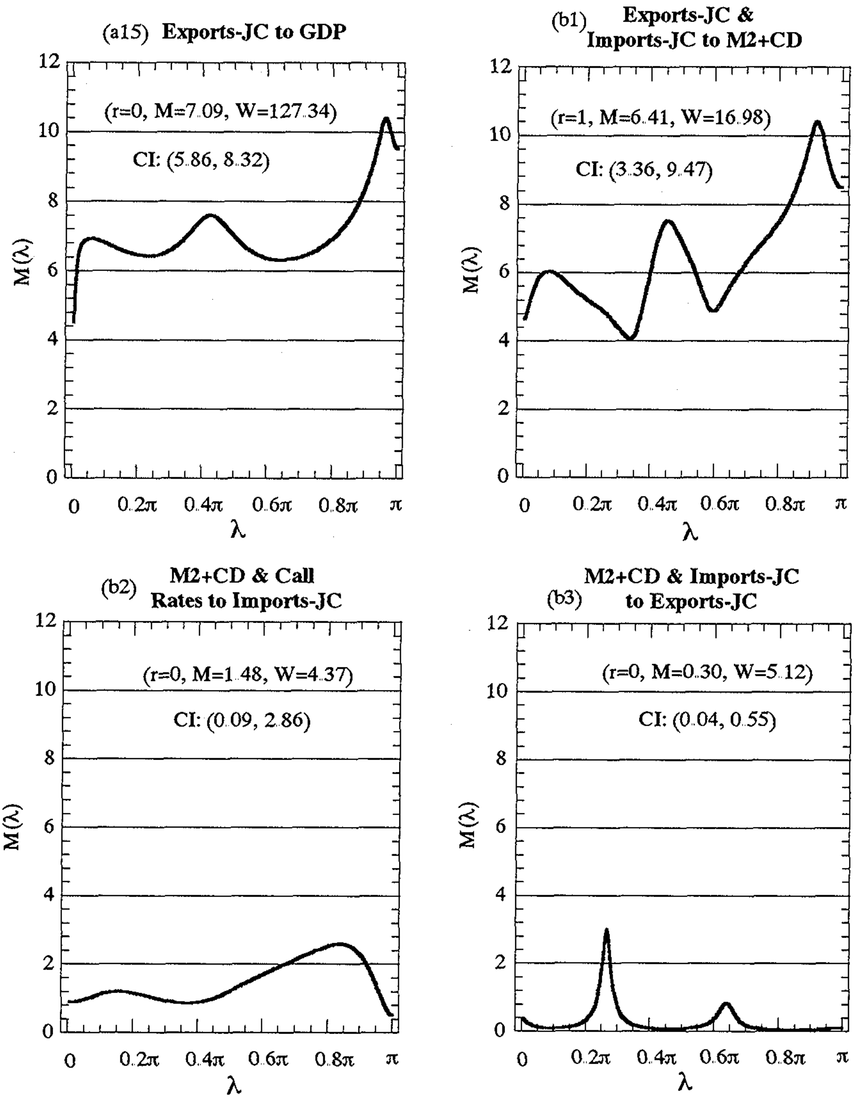


Figure 5.1　Continued

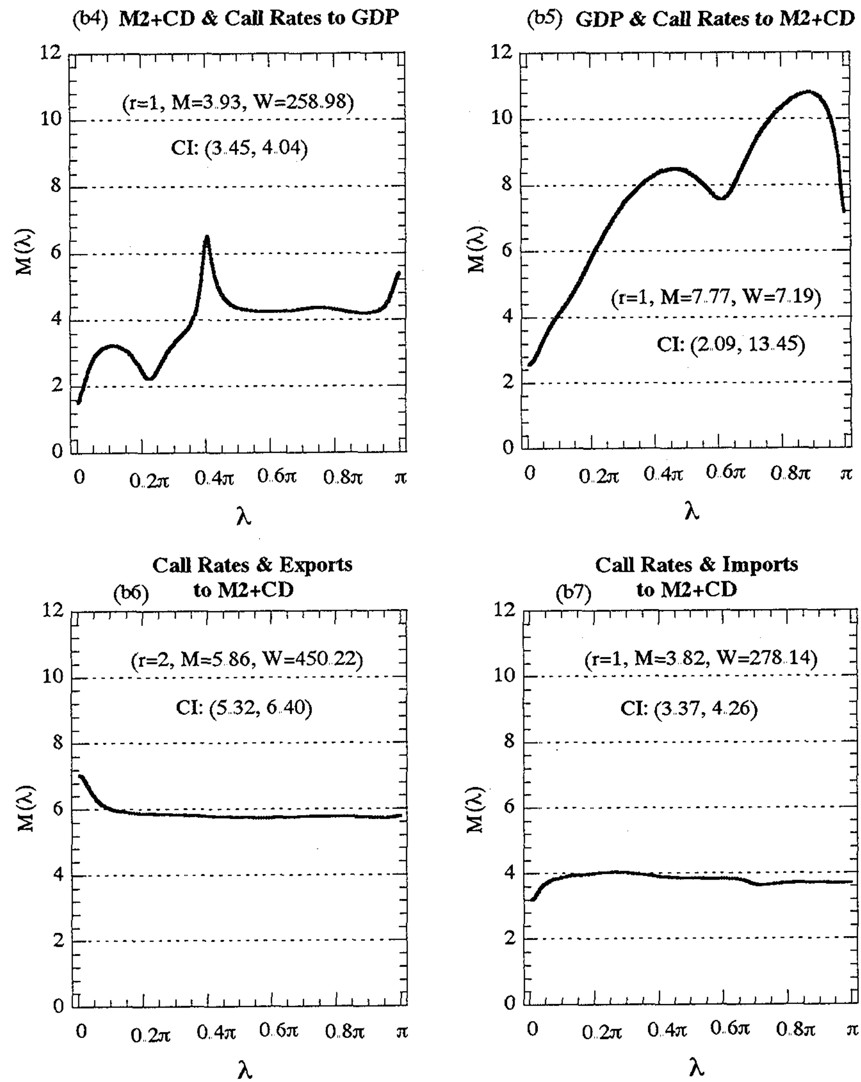


Figure 5.1 Continued

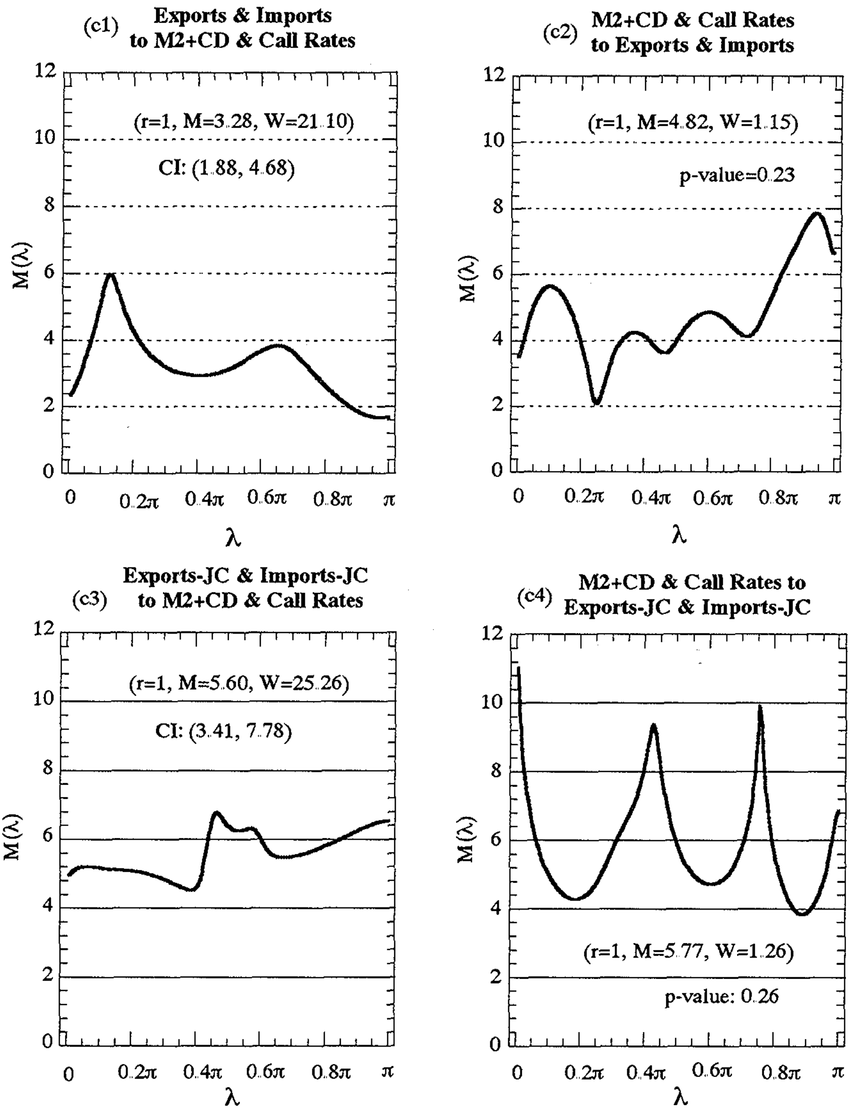


Figure 5.1 Continued

W defined by (3.11) are also presented in the figures. The 95 per cent confidence intervals of the *OMO*, in case the null hypothesis of non-causality is rejected, are also listed in the corresponding figures. The *OMO* estimates are obtained by numerical integration of the estimated *FMO*'s by dividing $[0, \pi]$ into 200 equal intervals. For each of the models we calculate *FMO* for frequency points $\lambda_i = i\pi/200$, $i = 1, 2, \cdots, 200$. As for the number of division of $[0, \pi]$, we checked many cases of interval division up to 1200, and we found that the 200 equal-division of the interval $[0, \pi]$ is fine enough.

Although similar computations were conducted on possible combinations and pairing of the seven variables, only a few are exhibited in the paper to save the space. In view of Figure 5.1, notable findings are as follows:

- The estimated *OMO* from M2+CD to nominal GDP is about three times of that in the reverse direction, but both of the two measures are not significant at 0.05 critical level [see plot (a1)]. Since the p-value of the Wald test for testing the one-way effect from money supply to GDP is 0.08, even though the effect is small, but significant at 10 per cent significance level. Plot (a1) also shows that there is no one-way effect in the frequency band $[0.4\pi, \pi]$, or in a period band shorter than one year and a quarter. The estimate of the *FMO* from money to GDP has a peak in the interval of $[0.25\pi, 0.4\pi]$, suggesting that it possibly takes about one year and a quarter for the effect to appear. The second peak near the origin indicates the existence of long-run effect.
- In general, Call Rate has conspicuous one-way effects to the other variables. In contrast, the effects in the reverse direction are small and not significant. [see plots (a3) to (a8) in Figure 5.1]. The one-way effects of Call Rates to GDP, to M2+CD and to Imports are very steady in all the frequency region. The plots (a5) and (a8) show that the effects

of Call Rates to Exports and to Imports-JC are not long-run. The one-way effect in frequency domain of Call Rates to Exports has a peak at frequency 0.55π [see plot (a5)], implying that the highest effect comes at the 3rd quarter period. The presence of one-way effect from interest rates would seem rather conformable to the conventional understanding of macroeconomic activities. Also this role of interest rates seems to be consistent with what Sims (1980) found in the macroeconomic data of U. S.

- The plot (a9) shows that the one-way effect from Exports to GDP is significant. The exports to China has one-way effect to GDP and the causal measure is comparatively large [$OMO=7.09$, see plot (a15)]. The corresponding one-way effect in the frequency domain shows that both of the two effects are not long-run. The estimate of the OMO from Imports to GDP is 2.61 and W=2.79 with a p-value 0.095 [see plot (a10)]. It is only significant at 0.1 critical level. The one-way effect from Exports and Imports to GDP is 3.82 and W=3.27 with a p-value 0.07. The effects from GDP to Exports and to Imports are not significant. The significant one-way effects from GDP to the exports to China and to the imports from China are not obseved. On the whole, it shows that during the period we analyzed, the Japanese economic growth can be thought derived by the external trades.
- The estimated one-way effect from Exports-JC to M2+CD is significant [see plot (a11), $OMO=6.85$, with a 95 percent confidence region (5.73, 7.96)]. The FMO is very low around the frequency 0.32π. The OMO of Imports-JC to M2+CD is 4.06 [Wald-statistic $W = 105.04$, see plot (a12)]. The effect is comparativelly stedy in the lower frequency band and is not long-run. Plot (b1) shows that the one-way effect of Exports-JC and Imports-JC to M2+CD is significant ($OMO=6.41$, $W=16.98$). The three plots (a11), (a12) and (b1) show that Japanese money supply is

partially effected by the trade between Japan and China.

- The *OMO* from Imports-JC to Exports-JC is significant but very small [$OMO=0.05$, $W=17.17$, see plot (a14)]. The one-way effect is only long-run. The reverse of the *OMO*, the one-way effect from Imports-JC to Exports-JC, is comparativelly large but not significant. Plot (a13) implies that, in a short frequency band including the frequency 0.5π, the *OMO* may be significant (the work of statistical test will be left for the next paper). The emperical results show that at 95 confidence level there is no significant one-way effects between Exports and Imports. A further investigation can tell us that at a compararively large critical value, there exists a comparatively weak one-way effect from Exports to Imports, and the one-way effect may be only long-run.
- The one-way effect of M2+CD and Call Rates to Imports-JC is significant and comparatively short-runt [see plot (b2)]. The *OMO* of M2+CD and Imports-JC to Exports-JC is 0.3 with a 95 confidence region (0.04, 0.55). Plot (b3) shows that the one-way effect are only concentrated to two short frequency bands including 0.25π and 0.65π. The one-way effect from M2+CD and Call Rates to GDP is significant and the corresponding one-way effect in the frequency domain has a peak at frequency 0.4π [see plot (b4)], implying that the highest effect comes from about one year and a quarter period. Both of the one-way effects from M2+CD and Call Rates to Exports and to Imports are not significant at 0.05 significance level. These findings seem to indicate that money supply is ineffective to the Japanese external trades in this period of the floating exchange-rate of Japanese Yen.
- Our Wald test shows that both of the effect from interest rates and exports, and that from interest rates and imports to money supply are significant at 0.05 critical value. The former effect is greater than that of the latter [see plots (b6), (b7)]. The one-way effect of interest rates

and imports to money supply is very stable in all the frequency region $[0, \pi]$. The evidence also indicates that the one-way effect of exports to money is not significant and we have

$$M_{Y\to M} = 0.03,\ M_{R\to M} = 4.53,\ M_{Ex\to M} = 1.89$$

whereas

$$M_{Y+R\to M} = 7.77,\ M_{R+Ex\to M} = 5.86.$$

These results imply that for some cases a policy mix is needed and is perhaps more effective than pursuing a single policy objective.

- The effect from Exports and Imports to M2+CD and Call Rates is strong and significant [see plot (c1)]. The effect in the reverse direction, the effect of M2+CD and Call Rates to Exports and Imports [see plot (c2)], is comparatively large in value ($OMO=4.82$) but not significant ($W=1.15$ with a p-value$=0.23$). The effect from Exports-JC and Imports-JC to M2+CD and Call Rates is strong and significant [see plot (c3), $OMO=5.60$, $W=25.26$]. The effect of M2+CD and Call Rates to Exports-JC and Imports-JC [see plot (c4)], is comparatively large in value ($OMO=5.77$) but not significant ($W=1.26$ with a p-value$=0.26$). The magnitude of the estimated OMO itself does not tell us whether a one-way effect is statistically significant or not, and a test is needed in judging the significance. As a whole, the findings imply that in the recent Japanese economy, the external trades have a significant one-way effect on the monetary side of the Japanese economy.

To summarize, the above empirical analyses show that there is no significant one-way effect from income to money, but the reverse effect is significant at size 0.1 but not significant at size 0.05. The interest rates in general cause the other variables but not the other way around. In general, the external trade causes monetary economy but not in the other direction. Even so, the monetary economy causes the imports from China. As for the effects of external trades to Japanese economic growth in the period we dealt with, the cause is mainly from exports but it is not long-run. The emperical result of the imports from China does not affect Japanese economic growth. This may support the comman understanding that the economy of Japan and that of China are cooperative, especially in the field of external trade. The empirical results also indicate the cases for which policy mix might be more effective.

6. Concluding Remarks

In this paper we show the one-way effect causal measure for cointegrated relations and show that causality hypotheses can be tested by standard asymptotically χ^2 distributed Wald statistics. Not only testing causality in cointegrated relationships by overall one-way effect measure, we also discussed inference on the long-run and short-run effects between pairs of vector-valued time series in the frequency domain. The proposed method includes testing Granger's non-causality as an instance of its multiple applications.

We presented how the theory of the one-way effect is put into practice and how to interpret empirical evidence in view of the theory. The empirical analyses were conducted indetail for seven quarterly macroeconomic series for the period of the first quarter of 1975 through the fourth quarter of 1994 in Japan. Our Wald test shows that money causes income mildly but not vice versa [see also Morimune and Zhao (1997), where mainly the standard F test and the Wald test presented by Toda and Phillips (1993) are used]. The one-way effects from interest rates to the other variables are comparatively strong and significant in general but not in the reverse directions. Our findings seem to indicate that monetary policies is ineffective to the external trades of Japan and that the growth of Japanese economy is driven by exports. The emperical results also support the comman understanding that, in the period we discussed, the relationship of international trade between Japan and China is cooperative.

In this paper, we did not pursue sophistication with respect to model specification and inference procedures. Although the cointegration model and the accompanying inference method of the paper is mainly based on Johansen's, they are not essential to our causal analysis at all and could be relaxed in many directions. As for other extensions of the paper, a model which allows breaks in the deterministic trend might be more realistic. By means of the integral of the *FMO* on specific frequency bands, the long-run and short-run causal relationships should also be tested. Moreover, although the analysis of this paper relies entirely upon "simple" causal relations, ignoring interaction with a third series, "partial" causal measures, which explicitly take into account the presence of a third series effect and its elimination, might be more desirable if we start from a well-defined full model of a macroeconomy. The problem of eliminating a third-series effect has been discussed in Granger (1969), Geweke (1984), Hosoya (1998), and recently in Hosoya and Yao (1999). Statistical inference and empirical studies based on these approach will be dealt with in the forthcoming papers.

References

Doornik, J. A. and H. Hansen, 1994, An omunibus test for univariate and multivariate normality, mimeo, Nuffield College, Oxford.

D'Agostino, R.B., 1970, Transformation to normality of the null distribution of g_1, Biometrika, vol. 57, pp. 679-81.

Engle, R. F. and C. W. J. Granger, 1987, Co-integration and error correction: representation, estimation and testing, Econometrica, vol. 55, no. 2, pp. 251-76.

Geweke, J., 1982, Measurement of linear dependence and feedback between multiple time series, Journal of the American Statistical

Association, vol. 77, no. 378, pp.304-13.

Geweke, J., 1984, Measures of conditional linear dependence and feedback between time series, Journal of the American Statistical Association, vol. 79, no. 388, pp.907-15.

Granger, C.W.J., 1963, Economic process involving feedback. Information and Control, vol.6, no.1, pp. 28-48.

Granger, C.W.J., 1969, Investigating causal relations by cross-spectrum methods, Econometrica, vol. 39, no. 3, pp. 424-38.

Granger, C.W.J. and J.L. Lin, 1995, Causality in the long run, Econometric Theory, vol. 11, no. 3, pp. 530-36.

Hosking, J. R. M., 1980, The multivariate portmanteau statistics, Journal of the American Statistical Association, vol. 75, pp. 602-8.

Hosoya, Y., 1977, On the Granger condition for non-causality, Econometrica, vol. 45, no. 7, pp. 1735-6.

Hosoya, Y., 1991, The decomposition and measurement of the interdependency between second-order stationary processes, Probability Theory and Related Fields, vol. 88, pp.429-44.

Hosoya, Y., 1997, 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, eds D. M. Kreps and K. F. Wallis, Cambridge : Cambridge University press, pp. 1-33.

Hosoya, Y., 1998, Elimination of a third-series effect in statistical causal analysis, Annal Report of the Economic Society, Tohoku University, vol. 59, no. 4, pp.136-55.

Hosoya, Y. and Yao, F., 1999, Statistical causal analysis and its application to economic time-series, manuscript presented to 1999 NBER/NSF Time Serier Confference, Taipei, Taiwan.

Johansen, S., 1988, Statistical analysis of cointegration vectors, Journal of

Economic Dynamic and Control, vol. 12, no. 213, pp. 231-54.

Johansen, S., 1991, Estimation and hypothesis testing of cointegration vectors in Gaussian vector autoregressive models, Econometrica, vol. 59, no. 6, pp. 1551-80.

Johansen, S., 1995, Likelihood-based Inference in Cointegrated Auto-regressive Models, Oxford University Press, Oxford.

Lütkepohl, H. and H. E. Reimers, 1992, Granger causality in cointegrated VAR processes, Economics Letters, vol. 40, no. 3, pp. 263-8.

Ljung, G. M. and G. E. P. Box, 1978, On a measure of lack of fit in time series models, Biometrika, vol. 65, pp. 297-303.

Magnus J. R. and H. Neudecker, 1988, Matrix Differential Calculus with Applications in Statistics and Econometrics, John Wiley & Sons, New York.

Morimune, K. and G. Q. Zhao, 1997, Unit root analysis of the causality between Japanese money and income, The Japanese Economic Review, vol. 48, no. 3, pp. 343-67.

Osterwald-Lenum, M., 1992, A note with quantiles of the asymptotic distribution of the maximum likelihood cointegration rank test, Oxford Bulletin of Economics and Statistics, Vol. 54, pp.461-71.

Shenton, L. R. and K. O. Bowman, 1977, A bivariate model for the distribution of $\sqrt{b_1}$ and b_2, Journal of the American Statistical Association, vol. 72, pp. 206-11.

Sims, C.A., 1980, Macroeconomics and reality, Econometrica, vol. 48, pp. 1-48.

Sims, C. A. and J. H. Stock and M. W. Watson, 1990, Inference in linear time series models with some unit roots, Econometrica, vol. 58, pp. 113-44.

Toda, H. and P. C. B. Phillips, 1993, Vector autoregressions and causality, Econometrica, vol. 61, no. 6, pp. 1367-93.

Whittle, P., 1963, Prediction and Regulation by Linear Least-Square Methods, D. Van Nostrand, Princeton.

Yao, F., 1996a, Econometric Analysis of Nonlinear and Nonstationary Relationships : Inference and Computational Methods, Ph. D. thesis, The Graduate School of Tohoku University, Sendai.

Yao, F., 1996b, Causal analysis of Japanese Money and International Trade, Statistical Analysis of Time Series : Theory and Application, The Institute of Statistical Mathematics Cooperative Research Report, no. 90, pp. 119-30.

Yao, F. and Y. Hosoya, 1995, Empirical causality analysis of Japanese macro economic data, Statistical Analysis of Time Series : Theory and Application, The Institute of Statistical Mathematics Cooperative Research Report, no. 79, pp. 85-96.

Yao, F. and Y. Hosoya, 1998, Inference on One-Way Effect and Evidence in Japanese Macroeconomic Data, Discussion Paper no. 145, Faculty of Economics, Tohoku University.