SSE-DP-2025-5

Akaike's Relative Power Contribution: A Revisit

Naoto Kunitomo

(Institute of Statistical Mathematics)

and

Xue Yujie

(Institute of Statistical Mathematics)

May 2025

Akaike's Relative Power Contribution: A Revisit *

Naoto Kunitomo †
and
Xue Yujie ‡

May 29, 2025

Abstract

Statistical analysis of inter-variable relationships in multivariate time series was initiated by Akaike (1968, 1971) at the Institute of Statistical Mathematics, where methods such as RPC (Relative Power Contribution) were developed for engineering applications. In the field of econometrics, vector autoregressive (VAR) analysis has evolved since the seminal works of Granger (1969) and Sims (1980), with further developments including the decomposition proposed by Pesaran and Shin (1998), Diebold and Yilmaz (2012, 2014), and Barunik and Krehlik (2018). This paper sheds some new lights on the limitations of existing methods regarding correlations among innovation variables, and proposes the use of decomposition of the predictive spectral density matrix with finite prediction horizon. The practical utility of this approach is discussed.

Key Words

VAR, Impulse Response, Akaike's RPC, Granger Causality, Pesaran-Shin Decomposition, predictive spectral density

*2025-5-29. This memorandum was inspired by the topic discussed in the Statistical Expert Group Training at the Institute of Statistical Mathematics and summarizes content that may be useful for statisticians and econometricians. We would like to express our gratitude to Wentei Zhang for introducing recent literature in financial econometrics, and to Genshiro Kitagawa for providing the data used in his work. The research has been supported by JSPS-Grant JP17H02513.

†The Institute of Statistical Mathematics, Midoricho 10-3, Tachikawa-shi, Tokyo 190-8562, JAPAN, kunitomo@ism.ac.jp

‡The Institute of Statistical Mathematics, Midoricho 10-3, Tachikawa-shi, Tokyo 190-8562, JAPAN, Xue@ism.ac.jp

1. Introduction

Statistical analysis of multivariate time series data has been conducted for quite some time as an application of statistics. In engineering, statistical methods based on Multivariate Autoregressive (MAR) models were developed, notably by Akaike (1968, 1971). At the Institute of Statistical Mathematics, Hirotsugu Akaike not only developed the TIMSAC software for statistical computation of multivariate time series with feedback, but also achieved applied success in addressing significant engineering problems at the time in Japan, such as controlling cement kilns and ship trajectories [1]. Statistical techniques including the Relative Power Contribution (RPC) method played an important role in these developments. Some of these methods are discussed, for example, in Kitagawa (2021).

In econometrics, the contribution of Sims (1980) is particularly influential. He proposed using vector autoregressive (VAR) models for macroeconomic analysis as an alternative to the then-dominant structural econometric models. Since then, a large body of research has emerged. In applied econometric analysis, decompositions based on impulse responses (hereafter IR) of forecast error variance have become especially important. In this context, Pesaran and Shin (1998) proposed the Generalized Impulse Response (GIR), which was later used in applied econometric studies developed by Diebold and Yilmaz (2012, 2014), Barunik and Krehlik (2018), and Zhang and Hamori (2021) among others, for analyzing spillover effects of financial volatilities due to innovation shocks.

The concept of Granger causality (G-causality), introduced by Granger (1969), has become central in the statistical analysis of multivariate economic time series. A number of statistical tests of G-causality have been developed. Debates and developments surrounding G-causality in econometrics are discussed in works such as Geweke (1982, 1984), Dufour and Taamouti (2010), and Hosoya et al. (2017), which explore various theoretical aspects of causality measures. For foundational references on statistical prediction theory and spectral decomposition in stationary multivariate time series, we refer to Hannan (1970) and Brockwell and Davis (1990); for time series econometrics, Hamilton (1994).

This paper first investigates the statistical properties of the Pesaran-Shin decomposition (hereafter referred to as PS decomposition), which has been applied in some econometric analysis. We also discuss the methodology employed by Diebold and Yilmaz (2012, 2014). Furthermore, we explore Akaike ’s RPC decomposition (1968), which was implemented in TIMSAC, to assess its relation to the PS decomposition and its applicability and limitations. We consider how this relates to prior discussions surrounding G-causality.

[1]TIMSAC is available at `https://www.ism.ac.jp/computer_system/jpn/software/software01.html` or a package at `https://cran.r-project.org/web/packages/timsac/index.html`

We aim to provide a unified interpretation of various forms of innovation-based decomposition in multivariate time series: including IR, GIR, PS decomposition, spectral decompositions (e.g., Akaike's RPC), and forecast error variance-covariance decompositions (causality measures). Moreover, we propose a new concept termed predictive spectral density, derived from the Fourier transform of forecast errors with finite horizon. This concept not only helps organize the previously fragmented discussions but also opens up new possibilities including statistical analysis of co-integrated processes for practical applications.

The structure of this paper is as follows: Section 2 introduces the multivariate stationary AR (or VAR) models. Section 3 describes the Pesaran-Shin decomposition, followed by Section 4 on the spillover effects defined by Diebold and Yilmaz. Section 5 explains a general error decomposition approach via spectral decomposition of multivariate AR models. Section 6 discusses causality measures in the sense of Granger causality, including those discussed by Hosoya et al. (2017) in a systematic way. Then, Section 7 define the predictive spectral density and its use in multivariate AR modeling. Section 8 provides numerical illustrations and empirical data analysis (and the related decomposition figures are given in the appendix). Finally, Section 9 offers concluding remarks and provisional observations on key issues discussed throughout the paper.

2 Multivariate Stationary AR Processes and Forecast Error Decomposition

We consider an m-dimensional (discrete-time) stochastic process $\mathbf{x}_t = (x_{kt})$ for $k = 1, \dots, m$ and $t = 1, \dots, T$, with $m \geq 2$, under the assumption of weak stationarity. Throughout this section, we follow the notation of Pesaran and Shin (1998). Suppose that the process $\mathbf{x}_t$ satisfies the vector autoregressive (VAR) model:

$$\mathbf{x}_t = \sum_{s=1}^{p} \mathbf{\Phi}_s \mathbf{x}_{t-s} + \boldsymbol{\epsilon}_t, \tag{2.1}$$

where the innovation term (or noise term) $\boldsymbol{\epsilon}_t$ is an i.i.d. sequence of random vectors such that $\mathbf{E}(\boldsymbol{\epsilon}_t) = \mathbf{0}$ and $\mathbf{E}(\boldsymbol{\epsilon}_t \boldsymbol{\epsilon}_t') = \mathbf{\Sigma}$. We assume that $\mathbf{\Sigma}$ is positive definite and the stochastic process is of full rank.

If the weakly stationary process $\mathbf{x}_t$ satisfies the following condition : **(A)** all roots of the characteristic equation $|\mathbf{I}_m \lambda^p - \sum_{s=1}^{p} \mathbf{\Phi}_s \lambda^{p-s}| = 0$ have modulus strictly less than one, then $\mathbf{x}_t$ admits the following causal moving average (MA) representation:

$$\mathbf{x}_t = \sum_{s=0}^{\infty} \mathbf{A}_s \boldsymbol{\epsilon}_{t-s}, \tag{2.2}$$

where $\mathbf{A}_0 = \mathbf{I}_m$ and, due to weak stationarity, $\sum_{s=0}^{\infty} \|\mathbf{A}_s\|^2 < \infty$.
Given the information available up to time t, the optimal H-step ahead forecast

$\mathbf{x}_{t+H|t}$ (for $H \geq 1$) is given by the conditional expectation

$$\mathbf{x}_{t+H|t} = \mathbf{E}[\mathbf{x}_{t+H}|\mathbf{x}_t, \mathbf{x}_{t-1}, \ldots] = \sum_{s=H}^{\infty} \mathbf{A}_s \boldsymbol{\epsilon}_{t+H-s}.$$

Hence, the mean squared forecast error is

$$\mathbf{E}\left[(\mathbf{x}_{t+H} - \mathbf{x}_{t+H|t})(\mathbf{x}_{t+H} - \mathbf{x}_{t+H|t})'\right] = \sum_{s=0}^{H-1} \mathbf{A}_s \boldsymbol{\Sigma} \mathbf{A}_s'. \tag{2.3}$$

Letting $H \to \infty$, we obtain the unconditional variance-covariance matrix [2]

$$\boldsymbol{\Gamma}(0) = \mathbf{E}[\mathbf{x}_t \mathbf{x}_t'] = \sum_{s=0}^{\infty} \mathbf{A}_s \boldsymbol{\Sigma} \mathbf{A}_s'. \tag{2.4}$$

Let $\mathbf{P}$ be an $m \times m$ matrix (typically a triangular matrix) such that $\mathbf{P}^{-1}\boldsymbol{\Sigma}\mathbf{P}^{-1\prime} = \mathbf{I}_m$, equivalently, $\boldsymbol{\Sigma} = \mathbf{P}\mathbf{P}'$. Define the transformed noise process $\mathbf{u}_t = \mathbf{P}^{-1}\boldsymbol{\epsilon}_t$, then $\mathbf{E}(\mathbf{u}_t\mathbf{u}_t') = \mathbf{I}_m$.

In the traditional (econometric) VAR analysis, from

$$\mathbf{x}_t = \sum_{s=0}^{\infty} \mathbf{A}_s \mathbf{P} \left(\mathbf{P}^{-1}\boldsymbol{\epsilon}_{t-s}\right),$$

the variance of the k-th component ($k = 1, \ldots, m$) is expressed as

$$\gamma_{kk} = \sum_{s=0}^{\infty} \mathbf{e}_k' \mathbf{A}_s \mathbf{P} \left(\sum_{j=1}^{m} \mathbf{e}_j \mathbf{e}_j'\right) \mathbf{P}' \mathbf{A}_s' \mathbf{e}_k = \sum_{j=1}^{m} \sum_{s=0}^{\infty} \left(\mathbf{e}_k' \mathbf{A}_s \mathbf{P} \mathbf{e}_j\right)^2, \tag{2.5}$$

where $\mathbf{e}_j$ denotes the $m \times 1$ vector with 1 in the j-th entry and zeros elsewhere. Each term on the right-hand side, $\mathbf{e}_k'\mathbf{A}_s\mathbf{P}\mathbf{e}_j$, depends in general on the ordering of variables. Therefore, the variance decomposition and the IR measure are sensitive to variable ordering selected in the statistical analysis, which is often subjective, unless $\boldsymbol{\Sigma}$ is diagonal [3].

---

[2] Let $\mathbf{f}(\lambda)$ denote the $m \times m$ spectral density matrix in (5.19). Then,

$$\boldsymbol{\Gamma}(0) = \int_{-\pi}^{\pi} \mathbf{f}(\lambda)\, d\lambda, \quad \mathbf{f}(0) = \frac{1}{2\pi} \left[\sum_{s=0}^{\infty} \mathbf{A}_s\right] \boldsymbol{\Sigma} \left[\sum_{s'=0}^{\infty} \mathbf{A}_{s'}'\right].$$

In particular, the variance of the k-th variable (the (k,k)-th element of $\boldsymbol{\Gamma}(0)$) can be expressed as the integral of the spectral density function $f_{kk}(\lambda)$, i.e., as the sum of power across frequencies. (See Section 7 for further discussion.)

[3] This issue is closely related to discussions surrounding structural equations, simultaneous equation systems, and structural VAR models in economic time series analysis. For classical treatments, see Chapter 11 of Hamilton (1994) or Chapter 1 of Hosoya et al. (2017).

3 Pesaran-Shin Decomposition

Let the regression coefficient with respect to the $k-$th ($k = 1, \cdots, m$) innovation term be defined as $\boldsymbol{\beta}_k = \mathbf{E}(\boldsymbol{\epsilon}_t | \epsilon_{kt} = \delta_k)$. Pesaran-Shin (1998) developed the GIR and their decomposition based on the difference of two conditional expectations. Since the multivariate AR model is linear in the noise terms, it follows from the conditional expectation that

$$\text{(3.6)} \qquad \mathbf{E}[\mathbf{x}_{t+H} | \epsilon_{kt} = \delta_k, \mathbf{x}_{t-1}, \mathbf{x}_{t-2}, \cdots] - \mathbf{E}[\mathbf{x}_{t+H} | \mathbf{x}_{t-1}, \mathbf{x}_{t-2}, \cdots] = \mathbf{A}_H \boldsymbol{\beta}_k.$$

In particular, if $\boldsymbol{\epsilon}_t \sim N_m(\mathbf{0}, \boldsymbol{\Sigma})$, then $\boldsymbol{\beta}_k = [\boldsymbol{\Sigma}\mathbf{e}_k / \sigma_{kk}]\delta_k$.
We have the standard representation as

$$\text{(3.7)} \qquad \begin{aligned} \mathbf{x}_t &= \sum_{s=0}^{\infty} \mathbf{A}_s \boldsymbol{\Sigma} (\sum_{j=1}^{m} \mathbf{e}_j \mathbf{e}_j^{'}) \boldsymbol{\Sigma}^{-1} \boldsymbol{\epsilon}_{t-s} \\ &= \sum_{j=1}^{m} \left\{ \sum_{s=0}^{\infty} \mathbf{A}_s \boldsymbol{\Sigma} \mathbf{e}_j \mathbf{e}_j^{'} \mathbf{P}^{'-1} (\mathbf{P}^{-1} \boldsymbol{\epsilon}_{t-s}) \right\} \end{aligned}$$

and

$$\text{(3.8)} \qquad \mathbf{E}[\mathbf{x}_t \mathbf{x}_t^{'}] = \sum_{s=0}^{\infty} \mathbf{A}_s \boldsymbol{\Sigma} \sum_{j=1}^{m} (\mathbf{e}_j \mathbf{e}_j^{'}) \boldsymbol{A}_s^{'}.$$

The variance of the k-th component is then decomposed as

$$\text{(3.9)} \qquad \begin{aligned} \gamma_{kk} &= \sum_{j=1}^{m} \sum_{s=0}^{\infty} \mathbf{e}_k^{'} \mathbf{A}_s \boldsymbol{\Sigma} \mathbf{e}_j (\mathbf{e}_j^{'} \mathbf{A}_s^{'} \mathbf{e}_k) \\ &= \sum_{j=1}^{m} \sum_{s=0}^{\infty} \left[\frac{\mathbf{e}_k^{'} \mathbf{A}_s \boldsymbol{\Sigma} \mathbf{e}_j}{\sqrt{\sigma_{jj}}} \right] \left[\mathbf{e}_k^{'} \mathbf{A}_s \mathbf{e}_j \sqrt{\sigma_{jj}} \right] \\ &= \sum_{j=1}^{m} \sum_{s=0}^{\infty} \left[\frac{\mathbf{e}_k^{'} \mathbf{A}_s \boldsymbol{\Sigma} \mathbf{e}_j}{\sqrt{\sigma_{jj}}} \right]^2 \left[\frac{\mathbf{e}_k^{'} \mathbf{A}_s \mathbf{e}_j \sqrt{\sigma_{jj}}}{\frac{\mathbf{e}_k^{'} \mathbf{A}_s \boldsymbol{\Sigma} \mathbf{e}_j}{\sqrt{\sigma_{jj}}}} \right] \end{aligned}$$

The terms on the right-hand side are invariant to variable ordering, but they are not necessarily non-negative.

In particular, when $\boldsymbol{\Sigma} = \sigma^2 \mathbf{I}_m$, we have

$$\text{(3.10)} \qquad \gamma_{kk}^{*} = \sum_{j=1}^{m} \sum_{s=0}^{\infty} \left[\frac{\mathbf{e}_k^{'} \mathbf{A}_s \boldsymbol{\Sigma} \mathbf{e}_j}{\sqrt{\sigma_{jj}}} \right]^2 .$$

More generally, for $\boldsymbol{\Sigma} = \sigma^2 \mathbf{I}_m + o(1)$, it follows that $\gamma_{kk} - \gamma_{kk}^{*} = o(1)$, leading to Proposition 1 below.

Let $\mathbf{x}_t(H) = \sum_{s=0}^{H-1} \mathbf{A}_s \boldsymbol{\epsilon}_{t-s}$. Then, insted of (2.2), the corresponding expressions for the variance-covariance matrix of the H-step ahead forecast error at time $t - H$.

Since the forecast error variance depends on H, we denote it by $\gamma_{kk}(H)$ and $\gamma^*_{kk}(H)$. In the $\mathcal{L}^2$ sense, as $H \to \infty$, $\|\mathbf{x}_t(H) - \mathbf{x}_t\| \to 0$, thus $\gamma_{kk}(H) \to \gamma_{kk}$ and $\gamma^*_{kk}(H) \to \gamma^*_{kk}$ for $k = 1, \cdots, m$.

Next, for any k ($k = 1, \cdots, m$), define $\boldsymbol{\beta}^*_k = \boldsymbol{\Sigma}\mathbf{e}_k/\sigma_{kk}$, and consider the decomposition of the innovation at time t:

$$\boldsymbol{\epsilon}_t = \boldsymbol{\beta}^*_k \boldsymbol{e}^{'}_k \boldsymbol{\epsilon}_t + [\mathbf{I}_m - \boldsymbol{\beta}^*_k \mathbf{e}^{'}_k]\boldsymbol{\epsilon}_t. \tag{3.11}$$

The two terms on the right-hand side are uncorrelated since

$$\mathbf{E}\left[\boldsymbol{\beta}^*_k \boldsymbol{e}^{'}_k \boldsymbol{\epsilon}_t \boldsymbol{\epsilon}^{'}_t (\mathbf{I}_m - \mathbf{e}_k \boldsymbol{\beta}^{*'}_k)\right] = \boldsymbol{\beta}^*_k \boldsymbol{e}^{'}_k \boldsymbol{\Sigma}(\mathbf{I}_m - \mathbf{e}_k \boldsymbol{\beta}^{*'}_k) = \mathbf{O}.$$

Thus, the covariance matrix $\boldsymbol{\Sigma}$ can be decomposed as $\boldsymbol{\Sigma} = \boldsymbol{\beta}^*_k \boldsymbol{\beta}^{*'}_k \sigma_{kk} + \left[\boldsymbol{\Sigma} - \boldsymbol{\beta}^*_k \boldsymbol{\beta}^{*'}_k \sigma_{kk}\right]$.

Now consider the use of the non-negative definite matrix

$$\boldsymbol{\Sigma}^* = \sum_{j=1}^{m} \boldsymbol{\beta}^*_j \boldsymbol{\beta}^{*'}_j \sigma_{jj}. \tag{3.12}$$

Note, however, that this matrix is not generally the covariance matrix of the random vector $\mathbf{u}^{(m)}_t = \sum_{j=1}^m \boldsymbol{\beta}^*_j \mathbf{e}^{'}_j \boldsymbol{\epsilon}_t$. For the random vector $\mathbf{u}^{(q)}_t = \sum_{j=1}^q \boldsymbol{\beta}^*_j \mathbf{e}^{'}_j \boldsymbol{\epsilon}_t$ $(1 \le q \le m)$, its covariance matrix is

$$\begin{aligned}
\mathbf{E}[\mathbf{u}^{(q)}_t \mathbf{u}^{(q)'}_t] &= \boldsymbol{\Sigma}\left[\sum_{j=1}^{q} \frac{\mathbf{e}_j \mathbf{e}^{'}_j}{\mathbf{e}^{'}_j \boldsymbol{\Sigma} \mathbf{e}_j}\right] \boldsymbol{\Sigma} \left[\sum_{j^{'}=1}^{q} \frac{\mathbf{e}_{j^{'}} \mathbf{e}^{'}_{j^{'}}}{\mathbf{e}^{'}_{j^{'}} \boldsymbol{\Sigma} \mathbf{e}_{j^{'}}}\right] \boldsymbol{\Sigma} \\
&= \boldsymbol{\Sigma}\left[\sum_{j=1, j^{'}=1}^{q} \frac{\mathbf{e}^{'}_j \boldsymbol{\Sigma} \mathbf{e}_{j^{'}} \mathbf{e}_j \mathbf{e}^{'}_{j^{'}}}{\mathbf{e}^{'}_j \boldsymbol{\Sigma} \mathbf{e}_j \mathbf{e}^{'}_{j^{'}} \boldsymbol{\Sigma} \mathbf{e}_{j^{'}}}\right] \boldsymbol{\Sigma} \\
&= \boldsymbol{\Sigma}\left[\sum_{j=1, j^{'}=1}^{q} \frac{\rho_{j,j^{'}}}{\sqrt{\mathbf{e}^{'}_j \boldsymbol{\Sigma} \mathbf{e}_j}\sqrt{\mathbf{e}^{'}_{j^{'}} \boldsymbol{\Sigma} \mathbf{e}_{j^{'}}}} \mathbf{e}_j \mathbf{e}^{'}_{j^{'}}\right] \boldsymbol{\Sigma},
\end{aligned} \tag{3.13}$$

where $\rho_{j,j^{'}} = \dfrac{\mathbf{e}^{'}_j \boldsymbol{\Sigma} \mathbf{e}_{j^{'}}}{\sqrt{\mathbf{e}^{'}_j \boldsymbol{\Sigma} \mathbf{e}_j}\sqrt{\mathbf{e}^{'}_{j^{'}} \boldsymbol{\Sigma} \mathbf{e}_{j^{'}}}}$.

When $q = 1$, choosing $j = k$ gives the first term of the decomposition

$$\boldsymbol{\Sigma} = \boldsymbol{\beta}^*_k \boldsymbol{\beta}^{*'}_k \sigma_{kk} + \left[\boldsymbol{\Sigma} - \boldsymbol{\beta}^*_k \boldsymbol{\beta}^{*'}_k \sigma_{kk}\right].$$

However, for $q \ge 2$, the vectors $\boldsymbol{\Sigma}\mathbf{e}_{j^{'}}$ are not necessarily orthogonal, so this does not correspond to an orthogonal decomposition of the covariance matrix. In particular, when $q = m$ and $\rho_{j,j^{'}} = 0$ for all $j \ne j^{'}$, we have $\mathbf{E}[\mathbf{u}^{(m)}_t \mathbf{u}^{(m)'}_t] = \boldsymbol{\Sigma}$. Next, by expressing the covariance matrix formally as $\boldsymbol{\Sigma} = \boldsymbol{\Sigma}^* + (\boldsymbol{\Sigma} - \boldsymbol{\Sigma}^*)$, and using the formula $\mathbf{E}[\mathbf{x}_t \mathbf{x}^{'}_t] = \sum_{s=0}^{\infty} \mathbf{A}_s \boldsymbol{\Sigma} \mathbf{A}^{'}_s$, we obtain

$$\mathbf{E}[\mathbf{x}_t \mathbf{x}^{'}_t] = \sum_{s=0}^{\infty} \mathbf{A}_s \left[\sum_{j=1}^{m} \boldsymbol{\beta}^*_j \sigma_{jj} \boldsymbol{\beta}^{*'}_j\right] \mathbf{A}^{'}_s + \sum_{s=0}^{\infty} \mathbf{A}_s \left[\boldsymbol{\Sigma} - \sum_{j=1}^{m} \boldsymbol{\beta}^*_j \sigma_{jj} \boldsymbol{\beta}^{*'}_j\right] \mathbf{A}^{'}_s.$$

Hence, the variance of the k-th component becomes

$$\gamma_{kk} = \sum_{j=1}^{m}\sum_{s=0}^{\infty} \mathbf{e}_k^{'}\mathbf{A}_s\boldsymbol{\beta}_j^*\sigma_{jj}\beta_j^{*'}\mathbf{A}_s^{'}\mathbf{e}_k + \sum_{s=0}^{\infty}\mathbf{e}_k^{'}\mathbf{A}_s\left[\boldsymbol{\Sigma} - \sum_{j=1}^{m}\boldsymbol{\beta}_j^*\sigma_{jj}\boldsymbol{\beta}_j^{*'}\right]\mathbf{A}_s^{'}\mathbf{e}_k.$$

The first term on the right-hand side corresponds to γ_{kk}^*. Let us denote the second term by γ_{kk}^{**}, then we can write

$$\gamma_{kk} = \gamma_{kk}^* + \gamma_{kk}^{**}. \tag{3.14}$$

The first term corresponds to the Pesaran-Shin decomposition. Although the second term is not necessarily non-negative, it vanishes when the innovation terms are uncorrelated. All the parameters appearing here are computable from data.

The relative contribution of the j-th innovation to the k-th variable in the Pesaran-Shin sense is given by

$$\begin{aligned}\theta_{kj} &= \frac{\left[\sum_{s=0}^{\infty}\mathbf{e}_k^{'}\mathbf{A}_s\boldsymbol{\Sigma}\mathbf{e}_j/\sqrt{\sigma_{jj}}\right]^2}{\sum_{j=1}^{m}\left[\sum_{s=0}^{\infty}\mathbf{e}_k^{'}\mathbf{A}_s\boldsymbol{\Sigma}\mathbf{e}_j/\sqrt{\sigma_{jj}}\right]^2} \times \frac{\gamma_{kk}^*}{\gamma_{kk}} \\ &= \frac{\left[\sum_{s=0}^{\infty}\mathbf{e}_k^{'}\mathbf{A}_s\boldsymbol{\Sigma}\mathbf{e}_j/\sqrt{\sigma_{jj}}\right]^2}{\sum_{s=0}^{\infty}\mathbf{e}_k^{'}\mathbf{A}_s\boldsymbol{\Sigma}\mathbf{A}_s^{'}\mathbf{e}_k}, \quad (j = 1, \cdots, m)\end{aligned} \tag{3.15}$$

which generally does not satisfy $\sum_{j=1}^{m}\theta_{kj} = 1$ for each $k = 1, \cdots, m$. (From a practical viewpoint, it may be preferable to compute each term explicitly for data analysis.)

Now, let us consider the eigenvalue decomposition $\boldsymbol{\Sigma} = \sum_{j=1}^{m}\mathbf{p}_j\mathbf{p}_j^{'}\lambda_j$ (assuming $\lambda_j > 0$). Focusing on the second term, we have

$$\boldsymbol{\Sigma} - \boldsymbol{\Sigma}^* = \boldsymbol{\Sigma}\left[\boldsymbol{\Sigma}^{-1} - \sum_{j=1}^{m}\frac{\mathbf{e}_j\mathbf{e}_j^{'}}{\mathbf{e}_j^{'}\boldsymbol{\Sigma}\mathbf{e}_j}\right]\boldsymbol{\Sigma} = \boldsymbol{\Sigma}\sum_{j=1}^{m}\left[\mathbf{p}_j\mathbf{p}_j^{'}\frac{1}{\lambda_j} - \mathbf{e}_j\mathbf{e}_j^{'}\frac{1}{\mathbf{e}_j^{'}\boldsymbol{\Sigma}\mathbf{e}_j}\right]\boldsymbol{\Sigma}.$$

Thus, if there exists a constant $c > 0$ such that $\lambda_j = c\sigma_{jj}$ and $\mathbf{e}_j = c\boldsymbol{\beta}_j^*$ for all $j = 1, \cdots, m$, then the second term vanishes. When all cross-correlations are small, it holds approximately that $\mathbf{e}_j \approx c\boldsymbol{\beta}_j^*$.

4 Spillover Effects

As an application of the previous section, we now consider the econometric method of spillover effects in asset price volatility using the forecast-error decomposition as proposed by Diebold and Yilmaz (2012).

Given information up to time t, the mean squared forecast error of the H-step-ahead forecast at time t is

$$\sum_{s=0}^{H-1}\mathbf{A}_s\boldsymbol{\Sigma}\mathbf{A}_s^{'}.$$

Let the diagonal elements of this matrix be denoted as

$$\gamma_{kk}(H) = \gamma^*_{kk}(H) + \gamma^{**}_{kk}(H) \quad (k = 1, \cdots, m).$$

Replacing ∞ with $H-1$ in the discussion from the previous section, we obtain the adjusted Pesaran-Shin contribution ratio for the k-th variable as

$$\tilde{\theta}_{kj}(H) = \frac{\left[\sum_{s=0}^{H-1} \mathbf{e}_k^{\prime} \mathbf{A}_s \mathbf{\Sigma} \mathbf{e}_j / \sqrt{\sigma_{jj}}\right]^2}{\sum_{j=1}^{m} \left[\sum_{s=0}^{H-1} \mathbf{e}_k^{\prime} \mathbf{A}_s \mathbf{\Sigma} \mathbf{e}_j / \sqrt{\sigma_{jj}}\right]^2} \quad (j = 1, \cdots, m). \tag{4.16}$$

In this case, each term is non-negative and satisfies

$$\sum_{j=1}^{m} \tilde{\theta}_{kj}(H) = 1.$$

Let us further decompose the first term of the variance of the k-th variable ’ s H-step-ahead forecast, $\gamma^*_{kk}(H)$, as follows:

$$\gamma^*_{kk}(H) = \sum_{j=1}^{m} \gamma^*_{j \to k}(H) = \sum_{j=1, j \neq k}^{m} \gamma^*_{j \to k}(H) + \gamma^*_{k \to k}(H), \tag{4.17}$$

where

$$\gamma^*_{j \to k}(H) = \frac{\sum_{s=0}^{H-1} \mathbf{e}_k^{\prime} \mathbf{A}_s \mathbf{\Sigma} \mathbf{e}_j}{\sigma_{jj}}. \tag{4.18}$$

Then, the adjusted contribution ratio can be written as

$$\tilde{\theta}_{kj}(H) = \frac{\gamma^*_{j \to k}(H)}{\sum_{j=1}^{m} \gamma^*_{j \to k}(H)}.$$

Consequently, since $\sum_{k,j=1}^{m} \tilde{\theta}_{kj}(H) = m$, the total spillover effects proposed by Diebold and Yilmaz (2012) can be defined as

$$S_{k.}(H) = \frac{100}{m} \frac{\sum_{j=1, j \neq k}^{m} \gamma^*_{j \to k}(H)}{\sum_{j=1}^{m} \gamma^*_{j \to k}(H)},$$

and

$$S_{.k}(H) = \frac{100}{m} \frac{\sum_{j=1, j \neq k}^{m} \gamma^*_{k \to j}(H)}{\sum_{j=1}^{m} \gamma^*_{k \to j}(H)}.$$

These measures are considered valid when $\gamma^{**}_{kk}(H) = 0$ (i.e., when $\gamma_{kk}(H) = \gamma^*_{kk}(H)$), or when the second term is sufficiently small. Whether this condition empirically holds in the measurement of volatility in multivariate financial time series is an important and interesting issue.

5 Spectral Decomposition

For weak-stationary processes, the prediction and spctrum theory of vector dicrete time series using the Hilbert space device have been well-establiehed and see Chapter III of Hannan (1970), for instance. The multivariate AR model of (2.1) and (2.2) with full rank $\mathbf{\Sigma}$ is a special case of purely non-deterministic processes. A decomposition of the vector time series in the frequency domain can be conducted. From the standard representation $\mathbf{x}_t = \sum_{s=0}^{\infty} \mathbf{A}_s \mathbf{P}(\mathbf{P}^{-1}\boldsymbol{\epsilon}_{t-s})$, the spectral density matrix (of size $m \times m$) is given by

$$
\begin{aligned}
(5.19) \qquad \mathbf{f}(\lambda) &= \frac{1}{2\pi}\left[\sum_{s=0}^{\infty} \mathbf{A}_s \mathbf{P} e^{-is\lambda}\right]\left[\sum_{s'=0}^{\infty} (\mathbf{A}_{s'}\mathbf{P})' e^{is'\lambda}\right] \\
&= \frac{1}{2\pi}\left[\sum_{s=0}^{\infty} \mathbf{A}_s e^{-is\lambda}\right] \mathbf{\Sigma} \left[\sum_{s'=0}^{\infty} \mathbf{A}_{s'}' e^{is'\lambda}\right],
\end{aligned}
$$

where $i^2 = -1$.

The (k,k)-th element of $\mathbf{f}(\lambda)$ $(k = 1, \cdots, m)$ is given by

$$
(5.20) \qquad f_{kk}(\lambda) = \frac{1}{2\pi}\sum_{j=1}^{m}\left|\sum_{s=0}^{\infty} \mathbf{e}_k' \mathbf{A}_s \mathbf{e}_j e^{-is\lambda} \mathbf{P} \mathbf{e}_j\right|^2 .
$$

Note that this decomposition depends on the choice of $\mathbf{P}$, i.e., the ordering of variables. In particular, when $\mathbf{\Sigma}$ (and hence $\mathbf{P}$) is a diagonal matrix, i.e., $\mathbf{\Sigma} = \mathrm{diag}(\sigma_{jj})$, the (k,k)-th element becomes

$$
f_{kk}(\lambda) = \frac{1}{2\pi}\sum_{j=1}^{m}\left|\sum_{s=0}^{\infty} \mathbf{e}_k' \mathbf{A}_s \mathbf{e}_j e^{-is\lambda}\right|^2 \sigma_{jj},
$$

where each term on the right-hand side is non-negative.
When the noise covariance matrix $\mathbf{\Sigma}$ is not diagonal, the Pesaran-Shin decomposition allows us to write the spectral density as

$$
(5.21) \qquad f_{kk}(\lambda) = f_{kk}^{*}(\lambda) + f_{kk}^{**}(\lambda),
$$

where

$$
(5.22) \qquad f_{kk}^{*}(\lambda) = \sum_{j=1}^{m} \frac{1}{2\pi}\left\| \mathbf{e}_k' \sum_{s=0}^{\infty} \mathbf{A}_s \mathbf{\Sigma} e^{-is\lambda} \mathbf{e}_j / \sqrt{\sigma_{jj}} \right\|^2 ,
$$

though the second term is not necessarily non-negative.
If we define

$$
f_{kk}^{*}(\lambda) = \sum_{j=1}^{m} f_{j\to kk}^{*}(\lambda),
$$

then the relative contribution of each frequency component based on the Pesaran-Shin decomposition is given by

$$R^{PS}_{j\to k}(\lambda) = \frac{f^*_{j\to kk}(\lambda)}{f_{kk}(\lambda)}. \tag{5.23}$$

In addition, the following measures provide useful spectral domain information:

$$RR^{PS*}_{j\to k}(\lambda) = \frac{f^*_{j\to kk}(\lambda)}{f^*_{kk}(\lambda)}, \quad RR^{PS}_{k}(\lambda) = \frac{f^*_{kk}(\lambda)}{f_{kk}(\lambda)}.$$

Let us now consider connections to Akaike's Relative Power Contribution (RPC, 1968) and the extensions proposed by Kitagawa et al. (2023). In general, the spectral density of the k-th variable is expressed as

$$\begin{aligned} f_{kk}(\lambda) &= \frac{1}{2\pi}\sum_{g,h=1}^{m}\Bigg[\sum_{s=0}^{\infty}(\mathbf{A}_s\cos s\lambda)_{kg}\sum_{s'=0}^{\infty}(\mathbf{A}_{s'}\cos s'\lambda)_{kh} \\ &\qquad + \sum_{s=0}^{\infty}(\mathbf{A}_s\sin s\lambda)_{kg}\sum_{s'=0}^{\infty}(\mathbf{A}_{s'}\sin s'\lambda)_{kh}\Bigg]\sigma_{gh}. \end{aligned} \tag{5.24}$$

When $g = h$, the term in the parenthesis reduces to

$$\sum_{s,s'=0}^{\infty}(\mathbf{A}_s)_{kg}(\mathbf{A}_{s'})_{kg}\cos((s'-s)\lambda) = \left|\sum_{s=0}^{\infty}\mathbf{e}'_k\boldsymbol{A}_s\mathbf{e}_g e^{-is\lambda}\right|^2,$$

which matches the standard spectral form.
Akaike (1968) defines the RPC as

$$RR^{A}_{g\to k}(\lambda) = \frac{\frac{1}{2\pi}\left|\sum_{s=0}^{\infty}\mathbf{e}'_k\mathbf{A}_s\mathbf{e}_g e^{-is\lambda}\right|^2\sigma_{gg}}{\sum_{j=1}^{m}\frac{1}{2\pi}\left|\sum_{s=0}^{\infty}\mathbf{e}'_k\mathbf{A}_s\mathbf{e}_j e^{-is\lambda}\right|^2\sigma_{jj}}. \tag{5.25}$$

Kitagawa, Tanokura, and Sato (2023) proposed the extended RPCs as

$$RR^{KTS}_{g\to k}(\lambda) = \frac{\frac{1}{2\pi}\left|\sum_{s=0}^{\infty}\mathbf{e}'_k\mathbf{A}_s\mathbf{e}_g e^{-is\lambda}\right|^2\sigma_{gg}}{f_{kk}(\lambda)}, \tag{5.26}$$

and for $g \neq h$,

$$RR^{KTS}_{gh\to k}(\lambda) = \frac{\frac{1}{2\pi}\sum_{s,s'=0}^{\infty}(\mathbf{A}_s(\lambda))_{kg}(\mathbf{A}_{s'}(\lambda))_{kh}\cos((s'-s)\lambda)\sigma_{gh}}{f_{kk}(\lambda)}. \tag{5.27}$$

Note that the above notation is slightly different from theirs. Because for $g \neq h$, $RR^{KTS}_{gh\to k}(\lambda)$ is not necessarily non-negative, it may cause some difficulty in the

interpretation of decomposition when it is not negligible.
Then, for $g, k \in \{1, \cdots, m\}$,

$$\frac{R^{PS}_{g\to k}(\lambda)}{RR^{KTS}_{g\to k}(\lambda)} = \frac{\left|\sum_{s=0}^{\infty} \mathbf{e}_k' \mathbf{A}_s \mathbf{\Sigma} e^{-is\lambda} \mathbf{e}_g / \sqrt{\sigma_{gg}}\right|^2}{\left|\sum_{s=0}^{\infty} \mathbf{e}_k' \mathbf{A}_s \mathbf{e}_g e^{-is\lambda}\right|^2 \sigma_{gg}}.$$

Next, we consider the case where the correlations among the innovations are small —namely, the so-called *small-correlation asymptotics*. For any covariance matrix, let

$$\sigma_{gh}(\rho) = \sigma_{gh}\delta_{gh} + \rho\omega_{gh} \quad (g, h = 1, \cdots, m), \tag{5.28}$$

where $\delta_{gh} = 1$ if $g = h$ and $\delta_{gh} = 0$ if $g \neq h$.

Under this setting, the following result holds.

Proposition 1: Under the above assumptions, as $\rho \to 0$, for any $g, k = 1, \cdots, m$,

$$\lim_{\rho\to 0} R^{PS}_{g\to k}(\lambda) = \lim_{\rho\to 0} RR^{KTS}_{g\to k}(\lambda) = RR^{A}_{g\to k}(\lambda). \tag{5.29}$$

In multivariate time series analysis, the contribution of each innovation and the strength of its spillover effects are not necessarily uniform across frequencies. For instance, one may need to separately analyze short-term, medium-term, and long-term frequencies to fully capture the nature of spillover or causality.
Therefore, it is crucial to interpret spectral-domain measures with care. In particular, when investigating the empirical validity of Proposition 1, the measure

$$RR^{PS}_{k}(\lambda) = \frac{f^*_{kk}(\lambda)}{f_{kk}(\lambda)} \quad (k = 1, \cdots, m)$$

can serve as a useful diagnostic index.

6 Causality Measures

Several measures of causality related to G-causality have been proposed in the literature, including those by Geweke (1982, 1984) and Dufour and Taanouti (2010). Here, we focus on the framework developed by Hosoya et al. (2017) as an example. Hosoya et al. (2017) propose a one-way causality measure (which we shall refer to as the Hosoya causality measure), although they do not mention the method of Pesaran and Shin (1998). In Chapters 2 and 3 of their work, they discuss a variety of causality measures developed in both the time and frequency domains. In this section, we examine the connections between the measures introduced so far and those of Hosoya et al. (2017).

To simplify the discussion, we consider the case $m = 2$ (in the notation of Hosoya et al. (2017), $p_1 = p_2 = 1$, $p = 2$). Using the Pesaran-Shin framework with $m = p = 2$, we have:

$$\boldsymbol{\beta}_1^* = \boldsymbol{\Sigma}\mathbf{e}_1/\sigma_{11} = \begin{bmatrix} 1 \\ \sigma_{21}/\sigma_{11} \end{bmatrix}, \quad \boldsymbol{\beta}_1^*\boldsymbol{\beta}_1^{*'}\sigma_{11} = \begin{bmatrix} \sigma_{11} & \sigma_{12} \\ \sigma_{12} & \sigma_{12}^2/\sigma_{11} \end{bmatrix}.$$

Then, $\boldsymbol{\Sigma}$ can be decomposed as: $\boldsymbol{\Sigma} = \boldsymbol{\beta}_1^*\boldsymbol{\beta}_1^{*'}\sigma_{11} + \left(\boldsymbol{\Sigma} - \boldsymbol{\beta}_1^*\boldsymbol{\beta}_1^{*'}\sigma_{11}\right)$. The spectral density of the first variable is:

$$f_{11}(\lambda) = \frac{1}{2\pi}\left[\sum_{s=0}^{\infty}\mathbf{e}_1^{'}\mathbf{A}_s e^{-is\lambda}\right]\boldsymbol{\Sigma}\left[\sum_{s'=0}^{\infty}\mathbf{A}_{s'}^{'}\mathbf{e}_1 e^{is'\lambda}\right]. \tag{6.30}$$

Conditioning on the first innovation variable, we define:

$$f_{11.1}(\lambda) = \frac{1}{2\pi}\left[\sum_{s=0}^{\infty}\mathbf{e}_1^{'}\mathbf{A}_s e^{-is\lambda}\right]\left[\boldsymbol{\beta}_1^*\boldsymbol{\beta}_1^{*'}\sigma_{11}\right]\left[\sum_{s'=0}^{\infty}\mathbf{A}_{s'}^{'}\mathbf{e}_1 e^{is'\lambda}\right]. \tag{6.31}$$

Remark 1: In the more general case where $p_1 \geq 1$ and $p_2 \geq 1$ (see Chapter 2 of Hosoya et al. (2017)), we consider the factorization of the spectral density matrix:

$$\tilde{\boldsymbol{\Gamma}}(e^{-i\lambda}) = \sum_{s=0}^{\infty}\mathbf{A}_s e^{-is\lambda}\boldsymbol{\Sigma}^{1/2}, \quad \mathbf{f}(\lambda) = \frac{1}{2\pi}\tilde{\boldsymbol{\Gamma}}(e^{-i\lambda})\tilde{\boldsymbol{\Gamma}}(e^{-i\lambda})^*,$$

where $\tilde{\boldsymbol{\Gamma}}(\cdot)^*$ stands for the complex conjugate. (We usethe notation $\tilde{\boldsymbol{\Gamma}}$, which is slightly differerent from Hosoya et al. (2017). It may be straightforward to discuss their method in the general case, but we omit the details here.)

Remark 2: The conditioning on the specific innovation variable serves to remove the effect of correlations among innovations, thereby trying to isolate the dynamic influence between variables. The idea is similar to the PS decomposition when $m = 2$.

Now define the one-way causal effect (Hosoya-measure) as:

$$\tilde{M}_{2\to1}(\lambda) = \log\left[\frac{f_{11}(\lambda)}{f_{11.1}(\lambda)}\right],$$

which corresponds to equation (2.26) in Hosoya et al. (2017). The overall causality measure from series 2 to series 1 is then:

$$M_{2\to1} = \frac{1}{2\pi}\int_{-\pi}^{\pi}\tilde{M}_{2\to1}(\lambda)\,d\lambda,$$

matching equation (2.17) in Hosoya et al. (2017). They also proposed related measures such as the measure of association and measure of reciprocity.

Remark 3: Rather than directly using $M_{2\to 1}(\lambda)$, it may be more intuitive to work with a modified relative power contribution (MRPC) ratio:

$$\mathrm{MRPC}^*_{11.1}(\lambda) = \frac{f_{11.1}(\lambda)}{f_{11}(\lambda)},$$

as this directly reflects the variance decomposition of forecast errors.

By defining

$$\gamma_{11} = \int_{-\pi}^{\pi} f_{11}(\lambda)d\lambda, \quad \gamma_{11.1} = \int_{-\pi}^{\pi} f_{11.1}(\lambda)d\lambda,$$

we obtain the ratio of predictive variances:

$$RV_{11.1} = \gamma_{11.1}/\gamma_{11}, \quad (0 \leq RV_{11.1} \leq 1),$$

ensuring that $\mathrm{MRPC}^*_1(\lambda) \leq 1$.
When $m = 2$ and $\sigma_{12} = 0$, the expression reduces to:

$$f_{11.1}(\lambda) = \frac{1}{2\pi}\left[\sum_{s=0}^{\infty} \mathbf{e}_1^{'}\mathbf{A}_s e^{-is\lambda}\mathbf{e}_1\right]^2 \sigma_{11},$$

which matches the relative power contribution (RPC).

Now consider

$$(6.32) \qquad f_{12.1}(\lambda) = \frac{1}{2\pi}\left[\sum_{s=0}^{\infty} \mathbf{e}_1^{'}\mathbf{A}_s e^{-is\lambda}\right]\left[\boldsymbol{\beta}_2^*\boldsymbol{\beta}_2^{*'}\sigma_{22}\right]\left[\sum_{s'=0}^{\infty} \mathbf{A}_{s'}^{'}\mathbf{e}_1 e^{is'\lambda}\right].$$

In the case of correlated innovations in a multivariate time series model, we can express the noise covariance matrix as:

$$\boldsymbol{\Sigma} = \boldsymbol{\beta}_1^*\boldsymbol{\beta}_1^{*'}\sigma_{11} + \begin{bmatrix} 0 & 0 \\ 0 & \sigma_{22} - \sigma_{12}^2/\sigma_{11} \end{bmatrix} = \boldsymbol{\beta}_2^*\boldsymbol{\beta}_2^{*'}\sigma_{22} + \begin{bmatrix} \sigma_{11} - \sigma_{12}^2/\sigma_{22} & 0 \\ 0 & 0 \end{bmatrix}.$$

Also,

$$\boldsymbol{\beta}_1^*\boldsymbol{\beta}_1^{*'}\sigma_{11} + \boldsymbol{\beta}_2^*\boldsymbol{\beta}_2^{*'}\sigma_{22} = \begin{bmatrix} \sigma_{11} + \sigma_{12}^2/\sigma_{22} & 2\sigma_{12} \\ 2\sigma_{12} & \sigma_{22} + \sigma_{12}^2/\sigma_{11} \end{bmatrix},$$

which equals $\boldsymbol{\Sigma}$ only when the innovations are uncorrelated. Therefore, in general, the sum

$$\sum_{j=1}^{m} \mathrm{MRPC}^*_j(\lambda) \neq 1.$$

Now define

$$(6.33) \qquad f^*_{11.2}(\lambda) = \frac{1}{2\pi}\left[\sum_{s=0}^{\infty} \mathbf{e}_1^{'}\mathbf{A}_s e^{-is\lambda}\right]\left[\boldsymbol{\Sigma} - \boldsymbol{\beta}_1^*\boldsymbol{\beta}_1^{*'}\sigma_{11}\right]\left[\sum_{s'=0}^{\infty} \mathbf{A}_{s'}^{'}\mathbf{e}_1 e^{is'\lambda}\right],$$

so that

$$f_{11}(\lambda) = f_{11.1}(\lambda) + f^*_{11.2}(\lambda).$$

Then define

$$\mathrm{MRPC}^*_{12.1}(\lambda) = \frac{f_{12.1}(\lambda)}{f_{11}(\lambda)},$$

so that

$$\mathrm{MRPC}^*_{11.1}(\lambda) + \mathrm{MRPC}^*_{12.1}(\lambda) = 1,$$

which gives a consistent decomposition.

Our interpretation—unlike Kitagawa et al. (2023)—unifies the innovation-based decomposition (Pesaran—Shin), spectral decomposition (RPC *á la* Akaike), and forecast-error variance decomposition (causality measures), enabling consistent analysis across these frameworks.

We summarize discussions on multivariate stochastic processes relevant. For a discrete-time, m-dimensional stationary time series process that is purely non-deterministic (i.e., regular), there exists a relationship (an extension of Kolmogorov ' s formula to the multivariate case) between the innovation covariance matrix $\mathbf{\Sigma}$ and the spectral density matrix $\mathbf{f}(\lambda)$:

$$\det(\mathbf{\Sigma}) = \exp\left[\frac{1}{2\pi}\int_{-\pi}^{\pi} \log\det(2\pi\mathbf{f}(\lambda))\, d\lambda\right]. \tag{6.34}$$

This fundamental result is discussed, for example, in Hannan (1970, p.162), in the context of Wold decomposition for weakly stationary processes.

Remark 4: Suppose the process is given by

$$\mathbf{x}_t = \sum_{s=0}^{\infty} \mathbf{A}_s \boldsymbol{\epsilon}_{t-s}, \quad \mathbf{E}(\boldsymbol{\epsilon}_t) = \mathbf{0}, \quad \mathbf{E}(\boldsymbol{\epsilon}_t\boldsymbol{\epsilon}_t^{'}) = \mathbf{\Sigma}, \quad \mathbf{E}(\boldsymbol{\epsilon}_t\boldsymbol{\epsilon}_s^{'}) = \mathbf{O}\ (t \neq s).$$

Then the spectral density is given by (5.19). The optimal one-step-ahead forecast given information up to time t is $\mathbf{x}_{t+1|t} = \sum_{s=1}^{\infty} \mathbf{A}_s \boldsymbol{\epsilon}_{t+1-s}$, and the forecast error is $\mathbf{x}_{t+1} - \mathbf{x}_{t+1|t}$. Normalizing by $\mathbf{A}_0 = \mathbf{I}_m$, the variance-covariance matrix of the forecast error equals $\mathbf{\Sigma}$.

7 Use of Predictive Spectral Density

We now introduce the concept of the predictive spectral density, which is based on the Fourier transform of forecast errors with finite horisons. For a forecast horizon $H \geq 1$ and a multivariate process of dimension $m \geq 2$, define the predictive spectral density matrix $\mathbf{f}^H(\lambda)$ as:

$$\mathbf{f}^H(\lambda) = \frac{1}{2\pi}\left[\sum_{s=0}^{H-1} \mathbf{A}_s e^{-is\lambda}\right] \mathbf{\Sigma} \left[\sum_{s'=0}^{H-1} \mathbf{A}_{s'}^{'} e^{is'\lambda}\right]. \tag{7.35}$$

This complex-valued matrix has real-valued diagonal entries:

$$f_{kk}^{H}(\lambda) = \mathbf{e}_k^{'}\mathbf{f}^{H}(\lambda)\mathbf{e}_k \quad (k = 1, \ldots, m).$$

This matrix corresponds to the spectral density of a vector MA process of order $H-1$. Moreover, for any non-zero complex vector $\mathbf{c}$, we have $\mathbf{c}^{'}\mathbf{f}^{H}(\lambda)\mathbf{c}^{*} \geq 0$. Intuitively, as $H \to \infty$, we have $\|\mathbf{f}^{H}(\lambda) - \mathbf{f}(\lambda)\|^2 \to 0$, and hence $\mathbf{f}^{H}(\lambda) \to \mathbf{f}(\lambda)$ in L^2-sense for each $\lambda \in [-\pi, \pi]$.

In particular, $\mathbf{f}^{H}(0)$ is given by

$$\mathbf{f}^{H}(0) = \frac{1}{2\pi}\left[\sum_{s=0}^{H-1}\mathbf{A}_s\right]\mathbf{\Sigma}\left[\sum_{s'=0}^{H-1}\mathbf{A}_{s'}^{'}\right]. \tag{7.36}$$

When $H = 1$, the predictive spectral density reduces to the noise covariance matrix: $\mathbf{f}^{1}(0) = \mathbf{\Sigma}$.

The corresponding forecast error covariance matrix is

$$\mathbf{\Gamma}^{H}(0) = \int_{-\pi}^{\pi}\mathbf{f}^{H}(\lambda)\,d\lambda = \sum_{s=0}^{H-1}\mathbf{A}_s\mathbf{\Sigma}\mathbf{A}_s^{'}\,. \tag{7.37}$$

Then, $\mathbf{O} < \mathbf{\Sigma} = \mathbf{\Gamma}^{1}(0) \leq \mathbf{\Gamma}^{2}(0) \leq \cdots \leq \mathbf{\Gamma}^{H}(0) \leq \mathbf{\Gamma}(0)$ for any finite H (≥ 1) in the sense of positive or non-negative definiteness.
For the k-th component of the process, the H-step-ahead forecast error variance is the (k,k)-th element of $\mathbf{\Gamma}^{H}(0)$ and is given by:

$$\gamma_{kk}(H) = \int_{-\pi}^{\pi} f_{kk}^{H}(\lambda)\,d\lambda. \tag{7.38}$$

As $H \to \infty$, $\mathbf{\Gamma}^{H}(0) \to \mathbf{\Gamma}(0)$ and $\gamma_{kk}(H) \to \gamma_{kk}$ for any k $(k = 1, \cdots, m)$, where γ_{kk} is the unconditional variance of the k-th component under the assumption of weak-stationarity of stochastic processes.

Thus, forecast error variance is expressible as the integral of predictive spectral density over frequency. This formulation provides a natural framework to organize short-run, medium-run, and long-run spectral analyses in both the time and frequency domains.

The arguments from Sections 5 and 6 naturally extend to predictive spectral density. The (k,k)-th element of $\mathbf{f}^{H}(\lambda)$ is given by:

$$f_{kk}^{H}(\lambda) = \frac{1}{2\pi}\sum_{j=1}^{m}\left|\sum_{s=0}^{H-1}\mathbf{e}_k^{'}\mathbf{A}_s\mathbf{e}_j e^{-is\lambda}\mathbf{P}\mathbf{e}_j\right|^2, \tag{7.39}$$

which still depends on the choice of $\mathbf{P}$, i.e., the ordering of variables.

Using the Pesaran—Shin decomposition, the predictive spectral density can be expressed as:

(7.40) $$f_{kk}^{H}(\lambda) = f_{kk}^{*,H}(\lambda) + f_{kk}^{**,H}(\lambda),$$

shortrun where

$$f_{kk}^{*,H}(\lambda) = \sum_{j=1}^{m} \frac{1}{2\pi} \left\| \mathbf{e}_k^{'} \sum_{s=0}^{H-1} \mathbf{A}_s \boldsymbol{\Sigma} e^{-is\lambda} \mathbf{e}_j / \sqrt{\sigma_{jj}} \right\|^2 .$$

Although the second term of (7.40) is not necessarily non-negative, this decomposition remains valid.

By redefining the following relative power decompositions for finite H(prediction horizon) instead of (5.23), (5.25), and (5.26) as :

$$R_{g\to k}^{PS,H}(\lambda), \quad RR_{g\to k}^{KTS,H}(\lambda), \quad RR_{g\to k}^{A,H}(\lambda),$$

we obtain the following result.

Proposition 2: Under the small-correlation asymptotic framework, as $\rho \to 0$, for all $g, k = 1, \ldots, m$,

(7.41) $$\lim_{\rho\to 0} R_{g\to k}^{PS,H}(\lambda) = \lim_{\rho\to 0} RR_{g\to k}^{KTS,H}(\lambda) = RR_{g\to k}^{A,H}(\lambda).$$

In practice, Propositions 1 and 2 may not hold exactly for real-world multivariate time series. Thus, it is advisable to utilize finite-horizon ($H \geq 1$) decompositions—such as the Pesaran—Shin innovation decomposition, variance ratios, and modified RPC-based predictive spectral measures—as informative and interpretable diagnostics.

Note that Sections 2 through 6 assume that the multivariate process (2.1) satisfies the weak-stationarity condition (**A**). However, the framework developed in Section 7 still holds under a relaxed assumption:

(**Condition A′**): All roots of the characteristic equation

(7.42) $$\left| \mathbf{I}_m \lambda^p - \sum_{s=1}^{p} \boldsymbol{\Phi}_s \lambda^{p-s} \right| = 0$$

are either equal to 1 or lie strictly inside the unit circle.

This includes, for example, the cointegrated processes proposed by Engle and Granger (1987). Although unit roots cause the forecast variances to diverge as $H \to \infty$, the predictive spectral density of (7.35) and the prediction variances for finite H remain meaningful. Therefore, even in the presence of unit roots in multivariate discrete-time AR models, the statistical analyses remain valid in terms of predictive spectral density instead of the standard spectral density.

Also, in financial market analysis, continuous-time diffusion processes or stochastic differential equations are frequently used. The frequency-domain methods discussed here may also prove important for the study of nonstationary processes and high-frequency financial data.

8 Numerical Illustrations and Empirical Analysis

To empirically support the theoretical results discussed in the previous sections, we provide both simulation-based numerical illustrations and real data analyses. Three spectral decomposition approaches are compared:

- Akaike's Relative Power Contribution (RPC), defined in equation (5.25),
- Kitagawa's Extended RPC (ERPC), proposed by Kitagawa et al. (2023), defined in equation (5.26), and
- Pesaran—Shin RPC (PSRPC), defined in equation (5.23).

Note that we use RPC, ERPC, PSRPC with finite horizon H as we discussed in Section 7.

8.1 Simulation Study

We first consider a simulated 3-dimensional stationary VAR(2) process of the form:

$$\boldsymbol{y}_n = A_1 \boldsymbol{y}_{n-1} + A_2 \boldsymbol{y}_{n-2} + \boldsymbol{\varepsilon}_n, \tag{8.43}$$

where the coefficient matrices are given by:

$$A_1 = \begin{bmatrix} 0.5 & 0.2 & 0 \\ 0.1 & 0.4 & 0 \\ 0 & 0.3 & 0.5 \end{bmatrix}, \quad A_2 = \begin{bmatrix} 0.1 & 0 & 0 \\ 0 & 0.1 & 0.1 \\ 0 & 0 & 0.1 \end{bmatrix},$$

and the innovation term $\boldsymbol{\varepsilon}_n \sim \mathcal{N}(\mathbf{0}, \Sigma)$ is a white noise process. The absolute values of all roots of $|\lambda^2 \boldsymbol{I}_3 - \lambda \boldsymbol{A}_1 - \boldsymbol{A}_2| = 0$ are less than one (-0.1534 $\pm$ 0.0885i, -0.1531, 0.3908, 0.6531, 0.8160). Two scenarios for the covariance matrix Σ innovations are considered:

- **Scenario 1 (Uncorrelated Innovations)**:

$$\Sigma_1 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

- **Scenario 2 (Correlated Innovations)**:

$$\Sigma_2 = \begin{bmatrix} 1 & 0.5 & 0.2 \\ 0.5 & 1 & 0 \\ 0.2 & 0 & 1 \end{bmatrix}$$

To conserve space, we report the results for the first component of the process only. We take the prediction horizon as $H = 2$, $H = 11$, and $H = +\infty$ (the last case corresponds to the spectral decomposition in the stationary process). Figures 1–3 illustrate the frequency-domain decompositions under Scenario 1, and Figures 4–6 under Scenario 2. In Scenario 1, all three decomposition methods yield identical results, which validates the theoretical equivalence under independent innovations. In contrast, under Scenario 2, the results of ERPC and PSRPC differ substantially, particularly when the forecast horizon H is finite. In Scenario 2, the PS decomposition results often differ significantly from the decompotion of ERPC sigificantly in low and high frequencies. It may be of some interest to find the effects of finite prediction horizons as Figure 6.

We also similated the multivariate AR model given by (8.43), where the coefficient matrices are given by:

$$A_1 = \begin{bmatrix} 0.6 & 0.2 & 0.1 \\ -0.4 & 0.5 & 0.2 \\ 0.1 & 0.3 & 0.4 \end{bmatrix}, \quad A_2 = \begin{bmatrix} 0.2 & 0.1 & -0.1 \\ 0.1 & -0.2 & 0.3 \\ 0.2 & 0.4 & 0.1 \end{bmatrix},$$

and the innovation term $\boldsymbol{\varepsilon}_n \sim \mathcal{N}(\mathbf{0}, \Sigma)$ is a white noise process. The absolute values of all roots of $|\lambda^2 \boldsymbol{I}_3 - \lambda \boldsymbol{A}_1 - \boldsymbol{A}_2| = 0$ are less than one or one (-0.2902 $\pm$ 0.1309i, 0.1758 $\pm$ 0.6234i,, 0.7289, 1.00). Thus, the simulated process is a cointedrated process. We only report the simulation results on the second scenarios for the covariance matrix Σ innovations are considered. In Figures 7–9, we give the spectral decomposition results of the first variable and compare three types of RPCs. In the first row, we show the true decompositions from the true model while in the second row we show the estimated decomposisions used by the ordinary least squares (OLS) estimation method. The estimated deomposition is quite similar to the true decomposition in this case.

As disucssed at the end of Section 7, the spectral decompositions give useful information even if the multivariate time series process has a unit root. The general observations of three types of RPC in this case are not different from those for the stationary processes. Although the prediction errors diverge as $H \to +\infty$ in this case, we found that the spectral decomposition results are rather stable numerically.

8.2 Empirical Examples

We next analyze two real-world datasets used in Kitagawa, Tanokura and Sato (2023) as typical examples. These examples reflect contrasting levels of estimated innovation correlation:

- **Example 1 (Ship Control Data)**: A 7-dimensional control system dataset from a ship named Hakusan, where the innovations are estimated to be weakly correlated.
- **Example 2 (Japanese Macroeconomic Data)**: A 4-dimensional macroeconomic dataset, in which the cyclical components were extracted using the DECOMP method. The innovations in this case exhibit strong mutual correlations.

Tables 1 and 2 report the estimated innovation correlations for these datasets, obtained via TSSS estimation [4]. We used the command $marfit$ and minimum AIC criteria of TSSS in our computation.

Table 1: Estimated Innovation Correlations : Hakusan (Ship Control) Data

	V1	V2	V3	V4	V5	V6	V7
V1	1.0000	0.0102	-0.0383	-0.1172	-0.0412	-0.0341	-0.0590
V2	0.0102	1.0000	0.0909	0.0713	-0.0367	0.0794	0.0234
V3	-0.0383	0.0909	1.0000	-0.0412	-0.0352	0.0448	-0.2758
V4	-0.1172	0.0713	-0.0412	1.0000	-0.0656	0.1154	0.0221
V5	-0.0412	-0.0367	-0.0352	-0.0656	1.0000	0.0010	-0.0165
V6	-0.0341	0.0794	0.0448	0.1154	0.0010	1.0000	-0.0002
V7	-0.0590	0.0234	-0.2758	0.0221	-0.0165	-0.0002	1.0000

Table 2: Estimated Innovation Correlations : Japanese Macroeconomic Data

	V1	V2	V3	V4
V1	1.0000	0.8412	0.7461	0.2659
V2	0.8412	1.0000	0.6487	0.1988
V3	0.7461	0.6487	1.0000	-0.1679
V4	0.2659	0.1988	-0.1679	1.0000

Figures 10–12 present the frequency-domain decomposition results for the Hakusan data, and Figures 13–15 for the macroeconomic data with three prediction horizons. In the former case, the differences among RPC, ERPC, and PSRPC are negligible,

[4]TSSS including Decomp is available as a package in R. See Kitagawa (2021) for the detail.

while in the latter, substantial differences are observed depending on the forecast horizon. In the macroeconomic data, the PS decomposition results often differ from the decompotion of ERPC sigificantly in low and high frequencies. This corresponds to Scinario 2 in our simulation since the absolute values of estimated correlations are large in comparison with Hakusan data. The results of estimated ERPC and PSRPC from data often differ substantially. On the other hand, the spectral characteristics of the Hakusan dataset correspond closely to those assumed in Scenario 1 of the simulation study.

The most important finding from our numerical and data analysis is that we should be careful whether the absolute values of the estimated correlations among inovations are large significantly or not before interpreting the estimated results. The decomposition of predictive spectral density depends on the frequency as well as the prediction horizon H. The inovation variances and covariances may have different effects with some horizon on the taeget variable in low and high frequencies. In applications, we are often interested in its prediction and the prediction hrizon may have an important implication. It is often the case for macroeconomic data in particular.

9 Concluding Remarks

In recent applied econometric studies, the decomposition method proposed by Pesaran and Shin (1998) has been employed within the context of VAR analysis. Since the introduction of the VAR model by Sims (1980), the analysis of multivariate time series in macroeconometrics has long been troubled by the issue that forecast error variance decompositions depend on the ordering of variables. The Pesaran-Shin decomposition is regarded by some economists as a welcome contribution, as it yields results that are invariant to variable ordering in multivariate time series analysis. In applications such as Diebold and Yilmaz (2012) concerning volatility in financial data, it is implicitly assumed that the correlation among innovations is negligible.

However, as discussed in this paper, the methodology of Pesaran and Shin (1996) contains fundamental limitations. In general, the Pesaran-Shin decomposition does not precisely correspond to the true forecast error decomposition, resulting in approximation errors. Upon closer examination, this issue is related to problems already recognized in the works of Akaike (1968) and Sims (1980). In cases where the correlations among innovations are small, the Pesaran-Shin decomposition coincides with the RPC (Relative Power Contribution) decomposition developed by Akaike (1968) (see Proposition 1). As shown in this paper, when the cross-correlations among innovations are negligible, the results obtained using the method of Pesaran and Shin (1996) are approximately consistent with those derived from Akaike's methodology (1968, 1971).

Nevertheless, when such correlations cannot be ignored, the analysis should be carried out with care. In such cases, it would be appropriate to consider alternative methods such as the ERPC (Extended RPC) and modified RPC proposed by Kitagawa, Tanokura, and Sato (2023). As they pointed out, particular caution is warranted in the case of macro-economic data. The influence of innovation correlations on the forecast error decomposition can be assessed via the ratio $r_k(H) = \gamma^*_{kk}(H)/\gamma_{kk}(H)$ $(k = 1, \cdots, m)$, for instance. The recent issues in econometric analysis are closely related to the long-standing debate on G-causality since Granger (1969), and it is essential to understand the connection between forecast error decompositions and causality measures.

This paper has proposed the use of the forecast spectral density matrix based on the Fourier transform of forecast errors with finite prediction horizon, and its practical utility has been highlighted. We visualize the decomposition of predictive spectral density in Section 8 for illustrations. The discussion presented here demonstrates that innovation-based decompositions (IR, GIR, Pesaran-Shin), spectral decompositions (Akaike's RPC, ERPC), and forecast error covariance decompositions (Causality Measures) can all be interpreted within a coherent framework. Moreover, these ideas may potentially be extended to applications in high-frequency financial markets based on non-stationary stochastic processes or continuous-time diffusion processes.

Although the practical implications of the theoretical arguments in this paper were only preliminarily validated through simulations and limited empirical data, the results suggest a wide range of possible applications. Further methodological development is anticipated, particularly for analyses involving high-dimensional setting. While a considerable amount of time has passed since the seminal works of Akaike (1968, 1971), we believe that there remains significant potential for advancing applications of multivariate time series analysis.

References

[1] Akaike, H. (1968),' 'On the Use of a Linear Model for the Identification of Feedback Systems," *Annals of Institute of Statistical Mathematics*, 20-1, 425-439.

[2] Akaike, H. (1971),"Autoregeressive model fitting for control," *Annals of Institute of Statistical Mathematics*, 23, 163-180.

[3] Barunık, J. and Krehlik, T. (2018),"Measuring the Frequency Dynamics of Financial Connectedness and Systemic Risk," *Journal of Financial Econometrics*, 2018, Vol. 16, No. 2, 271–296.

[4] Brockwell, P. and R. Davis (1990), *Time Series : Theory and Methods*, Springer.

[5] Diebold, F. and Yilmaz, K. (2012), "Better to give than recieve : Predictive

directional measurement of volatility spillovers," *International Journal of Forecasting*, 28, 57-66.

[6] Diebold, F. and Yilmaz, K. (2014), "On the network topology of variance decompositions ; Measuring the connectedness of financial firms," *Journal of Econometrics*182, 119-134.

[7] Dufour, J.M. and Taamouti, A. (2010), "Short and long run causality measures: Theory and inference," *Journal of Econometrics*, 2010, vol. 154, issue 1, 42-58.

[8] Engle, R. and Granger, C. (1987), "Co-integration and Error Correction : Representation, Estimation, and Testing," *Econometrica*, vol. 55-2, 251-276.

[9] Geweke, J. (1982), "Measurement of Linear Dependence and Feedback Betwveen Multiple Time Series," *Journal of the American Statistical Association*, 77-378, 304-313.

[10] Geweke, J. (1984), "Measures of Conditional Linear Dependence and Feedback Betvveen Time Series," *Journal of the American Statistical Association*, 79-388, 907-915.

[11] Granger, C. (1969), "Investigating causal relationships by econometric methods and cross-spectral methods," *Econometrica*, 37, 424-438.

[12] Hamilton, J. (1994), *Time Series Analysis*, Princeton University Press.

[13] Hannan, E.J. (1970), *Multiple Time Series*, Wiley.

[14] Hosoya, Y., Oya,K., Takimoto, T. and Kinoshita, R. (2017), *Characterizing Indterdependencies of Multiple Time Series: Theory and Practice*, Springer.

[15] Kitagawa, G. (2021), *Introduction to Time Series Modeling with Applications in R*, CRC Press, 2nd Edition.

[16] Kitagawa G., Tanokura, Y. and Sato, S. (2023) "Extended Relative Power Contribution that Allows to Evaluate the Effects of Correlated Noise," Unpublished Manuscript.

[17] Pesaran, H.H. and Y. Shin (1998), "Generalized impulse response analysis in linear multivariate models," *Economics Letters*, 58, 19-29.

[18] Sims, C. (1980), "Macroeconomics and Reality," *Econometrica*, 48, 1-48.

[19] Zhang, W., and Hamori, S. (2021). Crude oil market and stock markets during the COVID-19 pandemic: Evidence from the US, Japan, and Germany. International Review of Financial Analysis, 74, 101702

APPENDIX : Figures of Spectral Decomposition

In this Appendix, we present some figures discussed in Section 8. In each figure, we draw the estimated decompotion of spectral density for each innovations. Because there are many figures with similar characteristics in simulations, we have omitted some figures.

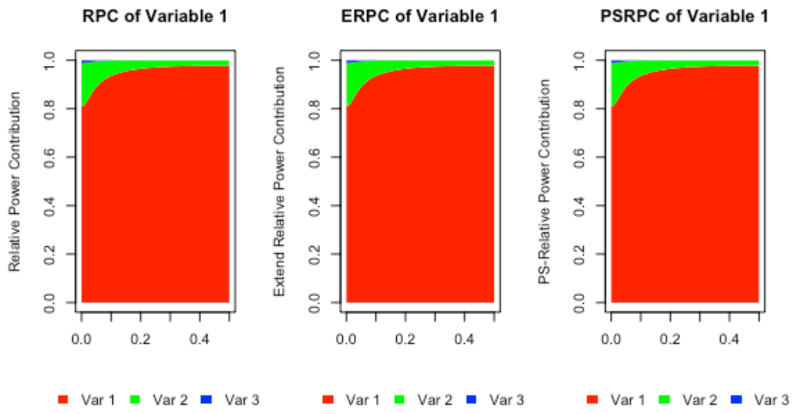

Figure 1: independent noises (Stationary Case, Sinario 1) $H = \infty$

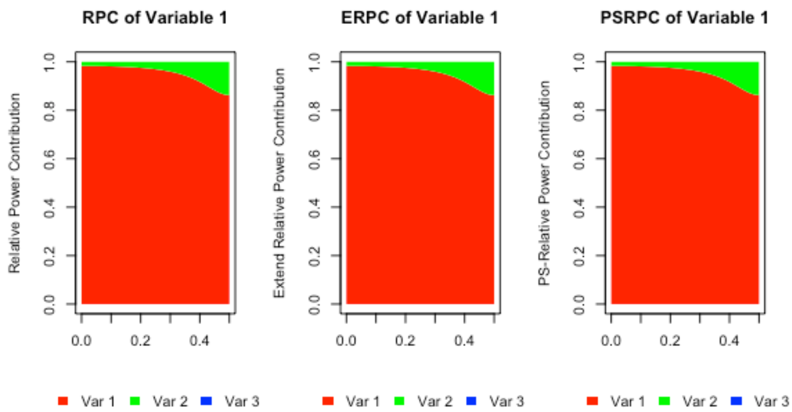

Figure 2: independent noises (Stationary Case, Sinario 1) $H = 2$

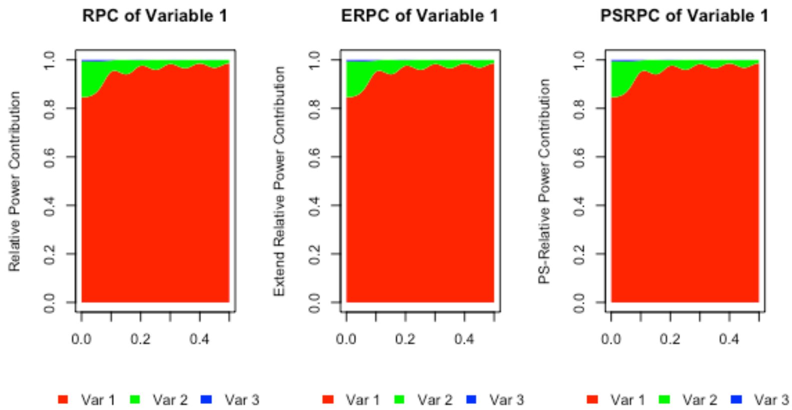

Figure 3: independent noises (Stationary Case, Scinario 1) $H = 11$

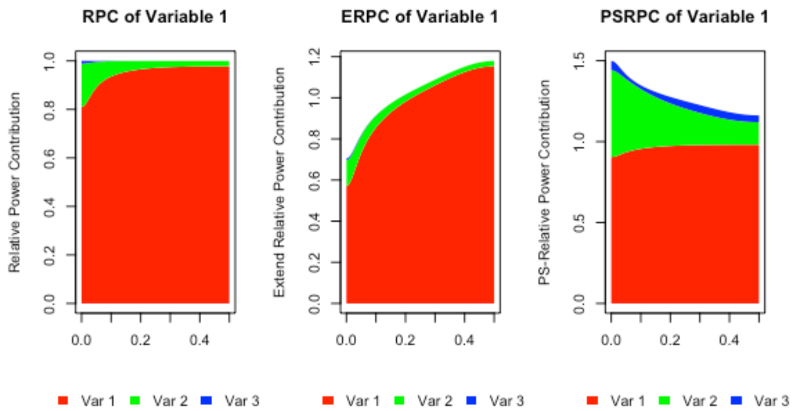

Figure 4: dependent noises (Stationary Case, Scinario 2) $H = \infty$

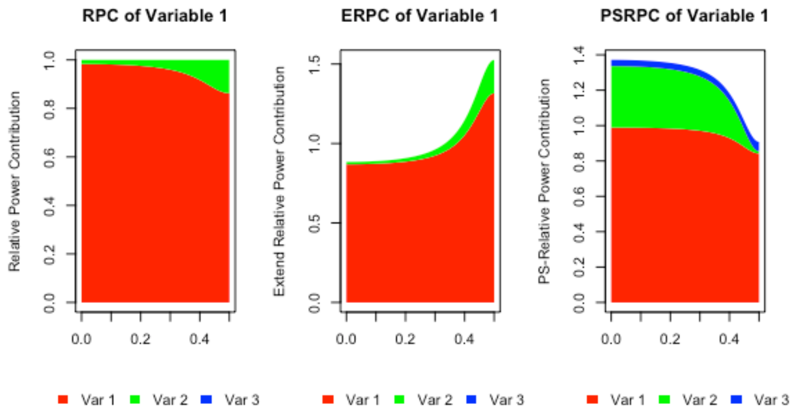

Figure 5: dependent noises (Stationary Case, Scinario 2) $H = 2$

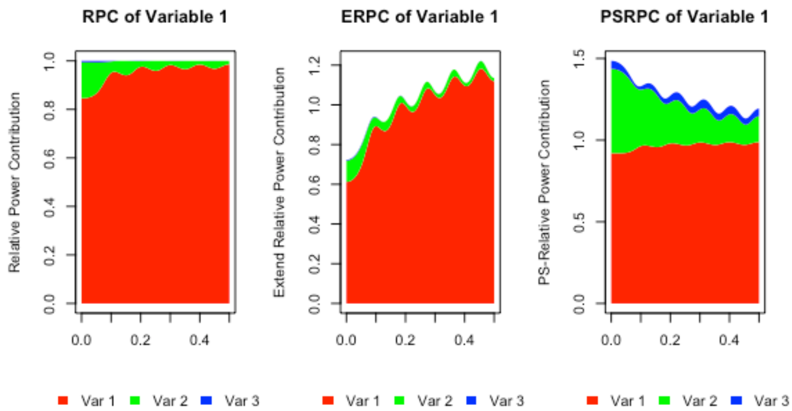

Figure 6: dependent noises (Stationary Case, Scinario 2) $H = 11$

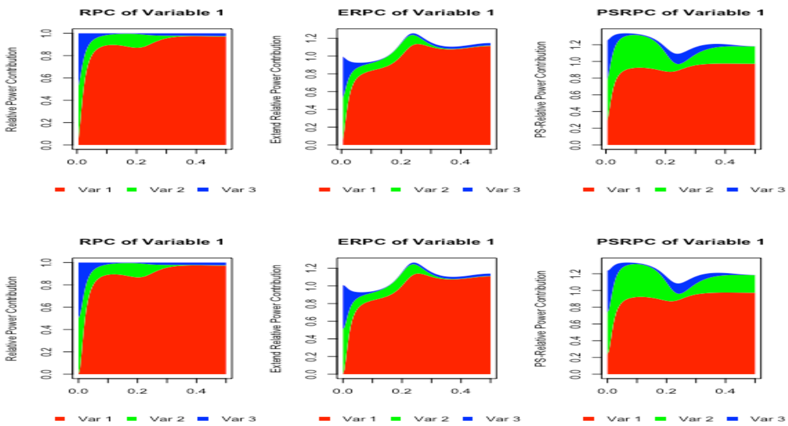

Figure 7: In the case of dependent noises (Co-integrated Case, Scinario 2) $H = \infty$, the first row shows the results by true coefficients while the second row shows the results by OLS estimation.

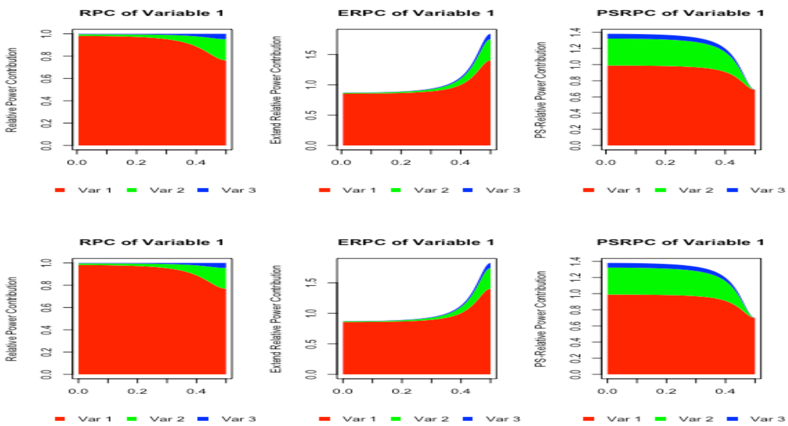

Figure 8: In the case of dependent noises (Co-integrated Case, Scinario 2) $H = 2$, the first row shows the results by true coefficients while the second row shows the results by OLS estimatation.

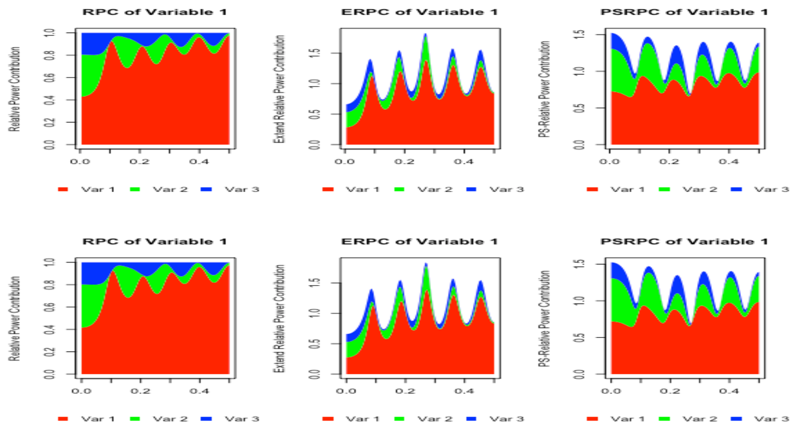

Figure 9: In the case of dependent noises (Co-integrated Case, Scinario 2) $H = 11$, the first row shows the results by true coefficients while the second row shows the results by OLS estimatation.

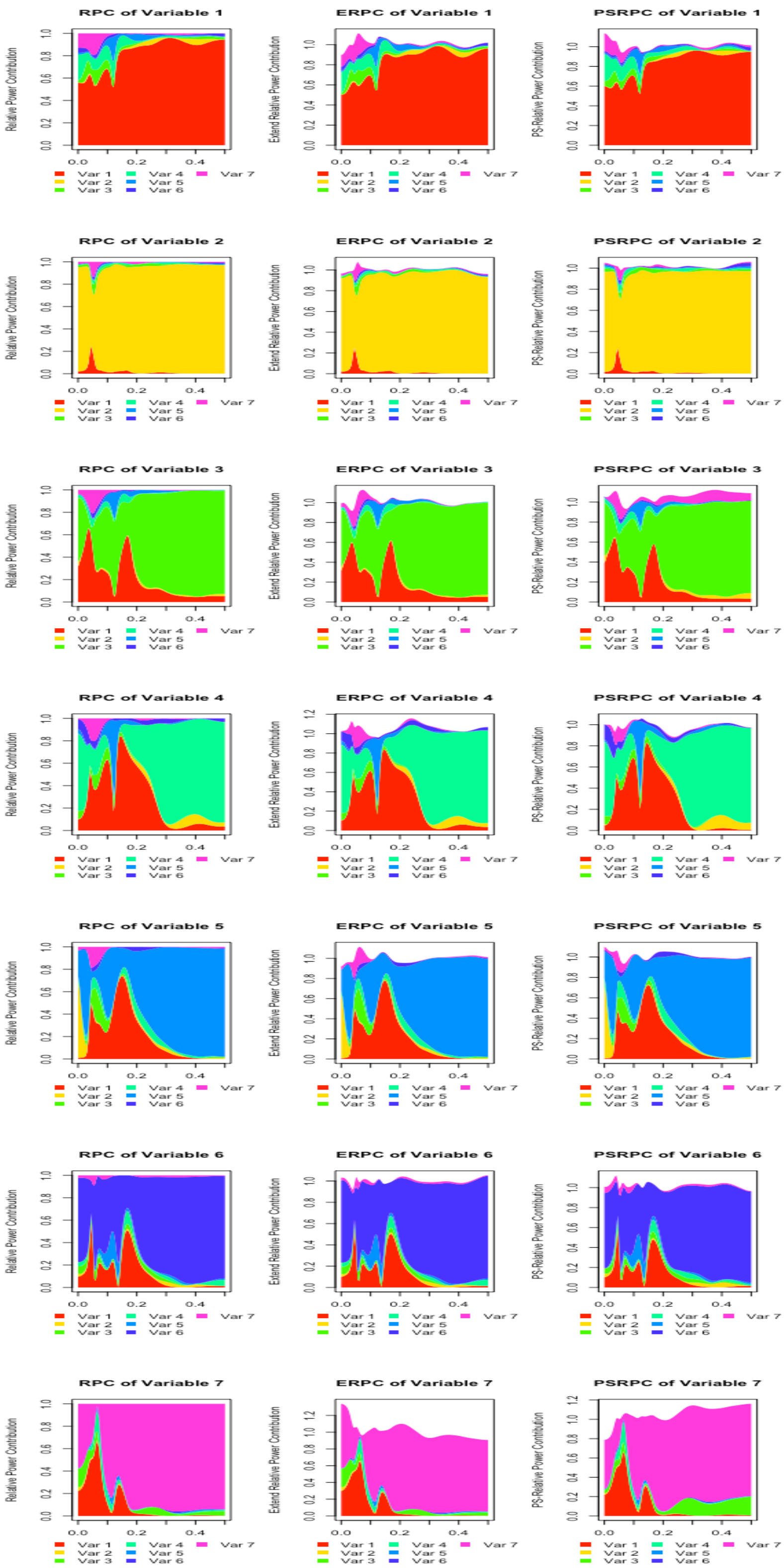

Figure 10: Hakusan data $H = \infty$

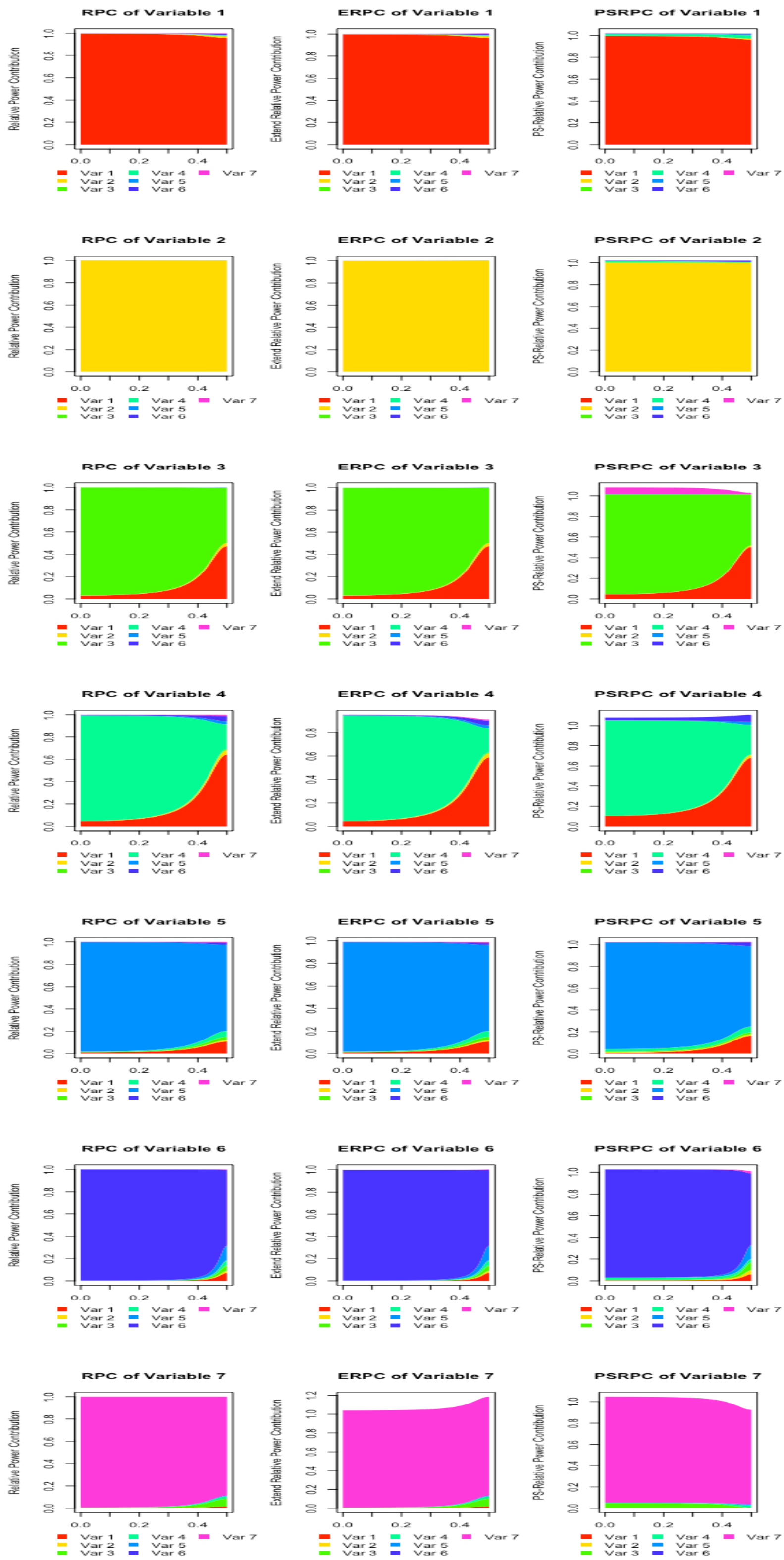

Figure 11: Hakusan data $H = 2$

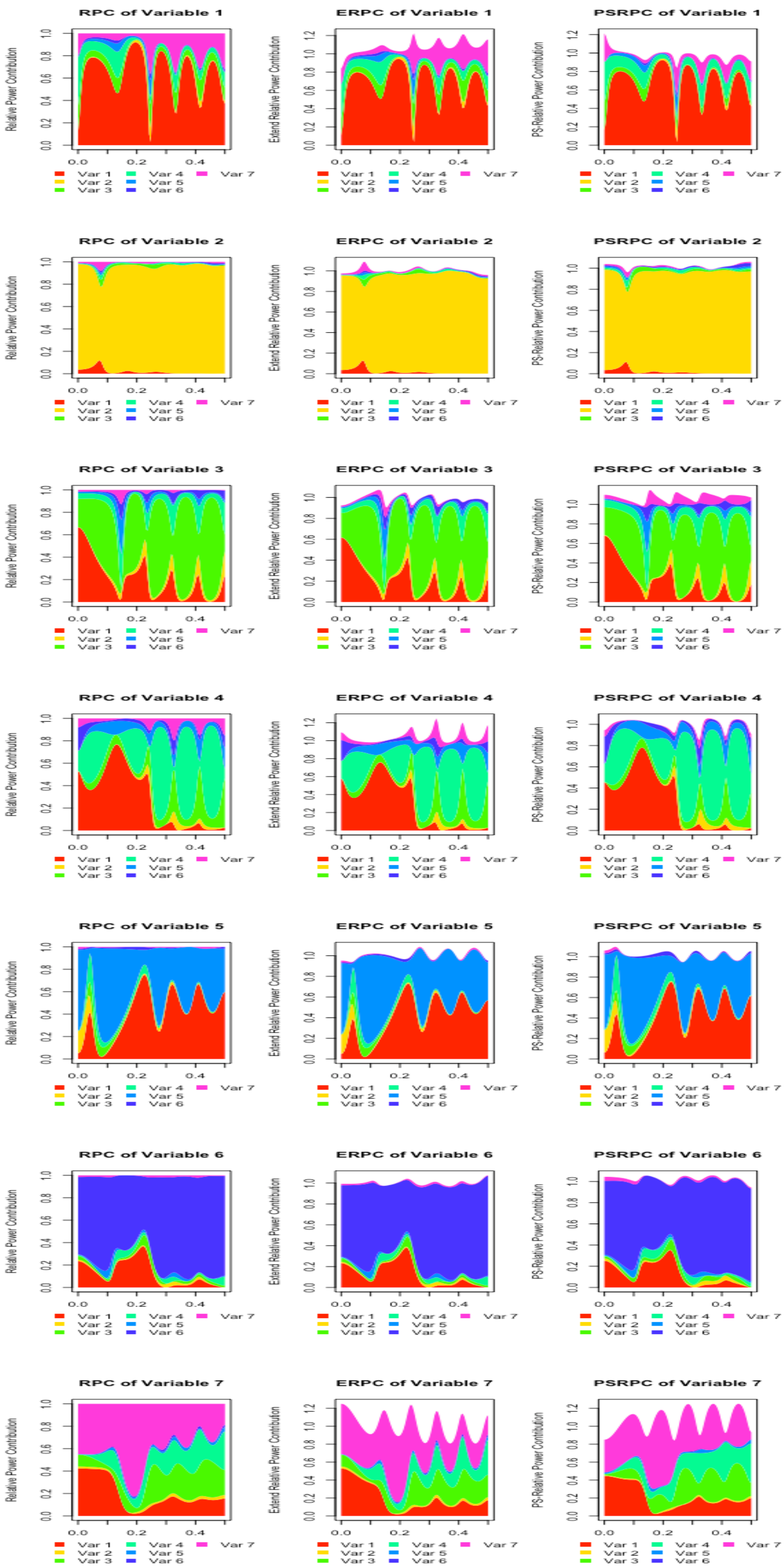

Figure 12: Hakusan data $H = 11$

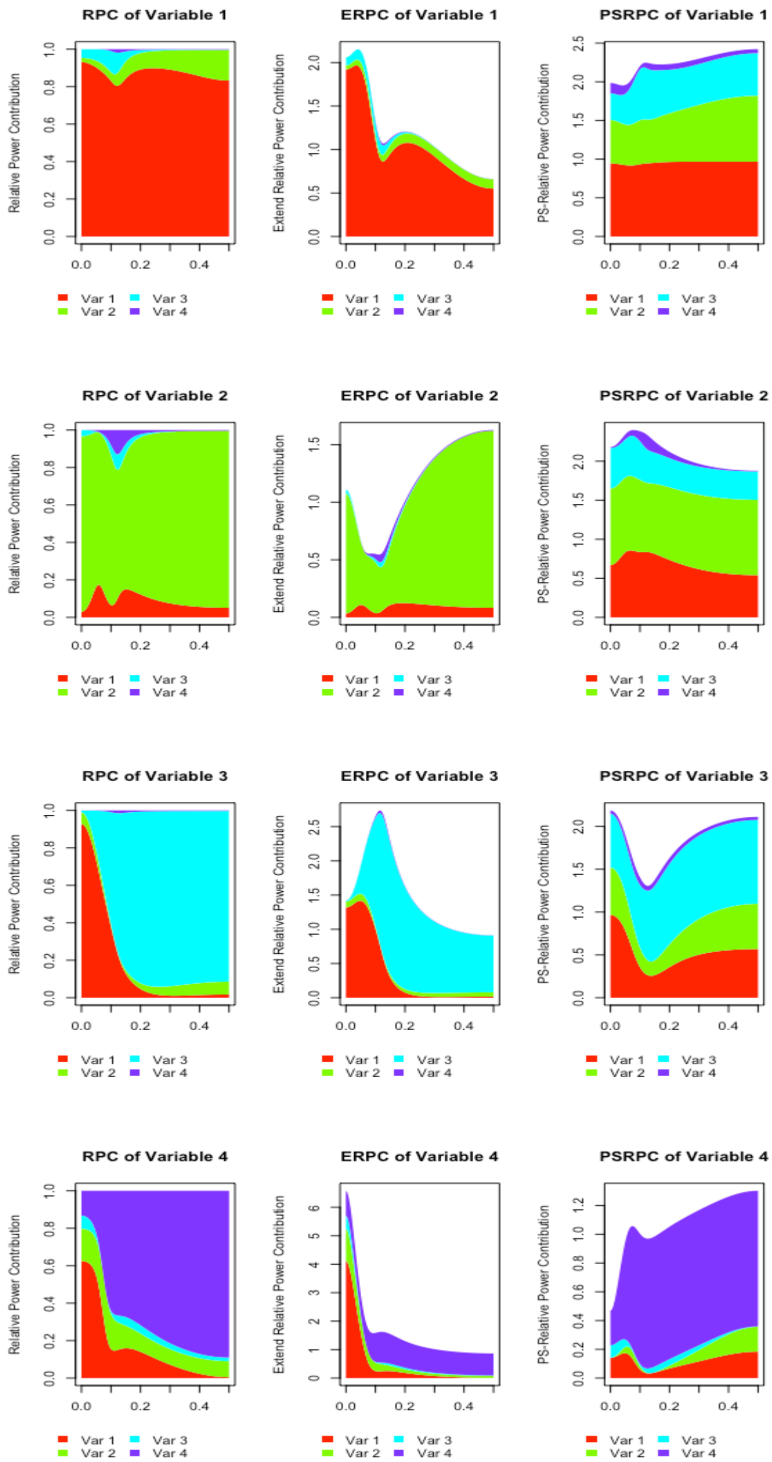

Figure 13: Macroeconomic data $H = \infty$

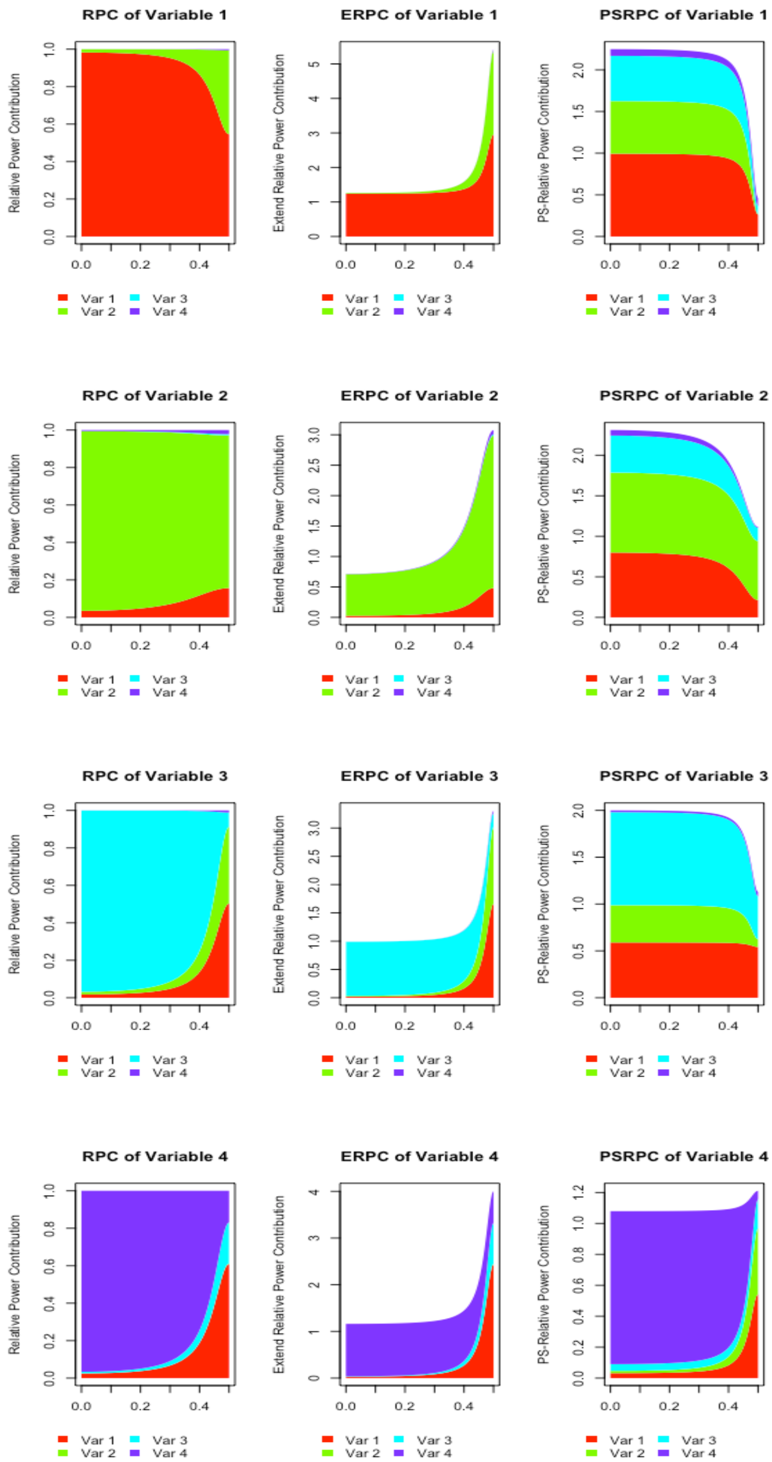

Figure 14: macroeconomic data $H = 2$

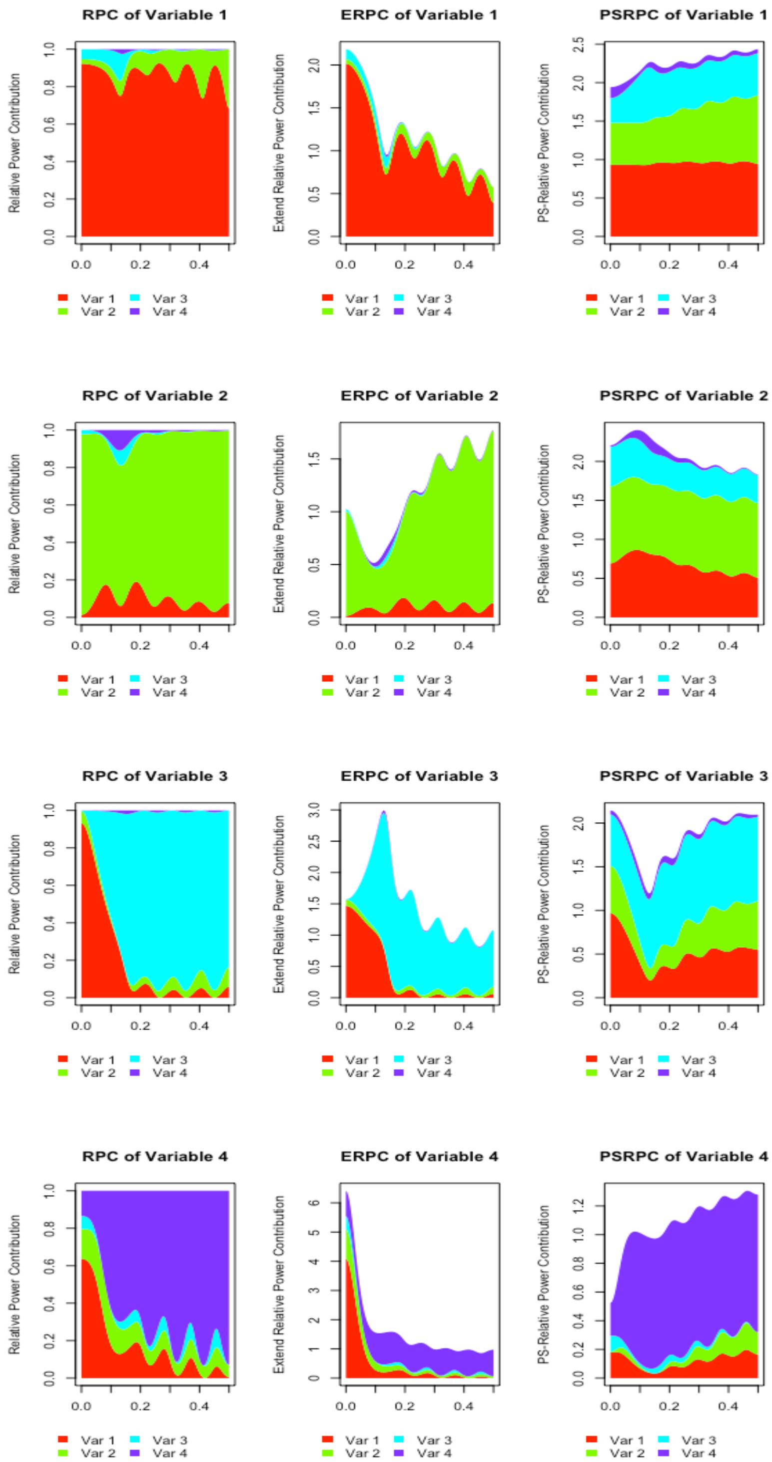

Figure 15: macroeconomic data $H = 11$