RIETI Discussion Paper Series 18-E-007

Empirical Evidence for Collective Motion of Prices with Macroeconomic Indicators in Japan

KICHIKAWA Yuichi
Niigata University

IYETOMI Hiroshi
Niigata University

AOYAMA Hideaki
RIETI

YOSHIKAWA Hiroshi
RIETI

Research Institute of Economy, Trade & Industry, IAA

The Research Institute of Economy, Trade and Industry
http://www.rieti.go.jp/en/

RIETI Discussion Paper Series 18-E-007

February 2018

Empirical Evidence for Collective Motion of Prices with Macroeconomic Indicators in Japan*

KICHIKAWA Yuichi [1], IYETOMI Hiroshi [1], AOYAMA Hideaki [2,4], and YOSHIKAWA Hiroshi [3,4]

[1]Faculty of Science, Niigata University

[2]Graduate School of Science, Kyoto University

[3]Faculty of Economics, Rissho University

[4]Research Institute of Economy, Trade and Industry (RIETI)

Abstract

We apply a complex Hilbert principal component analysis (CHPCA) to a set of Japanese economic data collected over the last 32 years, comprising individual price indices of middle classification level (imported goods, producer goods, consumption goods and services), indices of business conditions (leading, coincident, lagging), yen-dollar exchange rate, monetary stock, and monetary base. The CHPCA gives new insight into the dynamical linkages of price movements with business cycles and financial conditions. A statistical test identifies two significant eigenmodes with the largest and second largest eigenvalues. The lead-lag relations among domestic prices in the two modes are quite similar, indicating the individual prices behave in a collective way. However, the collective motion of prices is driven differently, namely, by the exchange rate at the upper stream side in the first mode and domestic demand at the lower stream side in the second mode. In contrast, the monetary variables play no important role in the two modes.

Keywords: Price changes, Business cycles, Exchange rate, Hilbert transformation, Principal component analysis, Dynamical linkage

JEL classification: E31, D12, C40

*This study has been conducted as a part of the project "Large-scale Simulation and Analysis of Economic Network for Macro Prudential Policy" undertaken at the Research Institute of Economy, Trade and Industry (RIETI). This research was also supported by MEXT as Exploratory Challenges on Post-K computer (Studies of Multi-level Spatiotemporal Simulation of Socioeconomic Phenomena) and by JSPS KAKENHI Grant Number 25400393. We would like to thank the seminar participants at RIETI for their helpful comments and suggestions.

I. Introduction

The economy should be regarded as a system of closely interrelated components. This is a fundamental idea of complexity science which was established in early 1980's. In economics, however, such a view is traced back to Alfred Marshall more than a century ago, who appealed "The Mecca of the economist lies in economic biology rather than in economic dynamics". A living body is a typical example of complex systems. Mutual interactions of microscopic elements give rise to completely new phenomena such as collective motion of the components at a macroscopic scale; for instance, life is an outcome of coherent behavior of molecules. Microeconomics and macroeconomics had been standing in parallel as two independent disciplines. These days, however, macroeconomics is absorbed into microeconomics. This is because macroeconomics is lacking concrete empirical evidences which approve its necessity.

Very recently, an empirical analysis of individual prices of goods and services for Japan has been carried out (Yoshikawa et al., 2015). The frequency of individual price changes and synchronization are not constant but instead are time-varying, while the existing literature (see Klenow and Malin (2010) and references therein) routinely assumes otherwise. Moreover, they change in clusters, not simultaneously in the economy as a whole. According to the current standard theory, for instance, changes in money, supposedly the most important macro disturbance, would affect all prices more or less uniformly (Klenow and Malin, 2010). We thus recognize a significant gap between observed facts and theory. Furthermore, examination of the autocorrelations of individual prices reveals the importance of interdependence of individual prices with lead-lag relations; prices do not move independently each other.

The Phillips curve is the earliest empirical indication of a close relationship between aggregated price dynamics (as measured by inflation rate) and economic conditions (as measured by unemployment rate). It is an urgent issue for the Japanese economy how to get rid of the long-standing deflation. The Bank of Japan (BOJ) is drastically increasing the supply of money with inflation targetting. However, no one has a conclusive answer to the question, which comes first inflation/deflation or economic growth/recession? The BOJ expects inflation is ahead of economic growth. We will answer the question by an empirical analysis, which confirms that individual prices move in a coherent fashion with definite lead/lag relations and elucidates how business cycles are linked dynamically to the collective motion of prices.

In recent years, some of physicists have paid attention to socioeconomic phenomena. Rapid development of computers and advancement of information processing technology made it possible to obtain a variety of economic data in large quantities. The principal component analysis (PCA) and the random matrix theory (RMT) was successfully combined to detect correlations hidden in multivariate time series data (Laloux et al., 1999; Plerou et al., 2002; Utsugi et al., 2004; Kwapien and Drozdz, 2012). The RMT serves as a theoretically sound criterion to determine if eigenmodes of the correlation matrix are statistically significant; this is the critical issue that the PCA always encounters. However, the PCA assisted by the RMT is not so capable of extracting correlation structures with lead/lag relations, because it totally depends on correlations at equal time. Correlations between time series data are not always present in a simultaneous manner.

In order to explore dynamic correlations in climate data, the complex Hilbert principal component analysis (CHPCA) was developed by meteorologists (Horel, 1984; Barnett, 1983; Stein et al., 2011). The CHPCA is based on complexification of real data using the Hilbert transformation. Lead/lag relations in original data are manifested in a form of instantaneous phases of the complex time series thus constructed. Recently, the RMT has been extended so that it works as a null hypothesis for the CHPCA. If time series data have appreciable autocorrelations, however, the RMT criterion tends to predict more significant modes than it should do. This is because autocorrelations deceive us by giving rise to spurious cross-correlations for time series of finite length, especially in the case that their length is comparable with the number of species of data. To overcome such limitation of the RMT, the rotational random shuffling (RRS) method (Iyetomi et al., 2011a) was devised. This is a numerical method which destroys cross-correlations with autocorrelations preserved in time series data. Recently, the CHPCA assisted by the RMT or the RRS has been applied to various multivariate data such as stock market data (Arai et al., 2013) and world-wide financial data of markets and currencies (Vodenska et al., 2016),

The aforementioned study by Yoshikawa et al. (2015) also took advantage of the state-of-the-art methodology to analyze lead/lag dynamics of individual prices and to find out what are the major macroeconomic variables leading to systemic changes in aggregate prices. The analysis was based on a large set of micro prices at the most detailed level: prices for 75 imported goods, 420 producer goods, and 335 consumer goods and services. The data may be too disaggregated to detect possible collective behavior of prices from a sea of noises. In this study we thereby focus on price indices of middle classification level: prices for 10 imported goods, 23 producer goods, and 47 consumer goods and services. And we combine the set of prices with indices of business conditions (leading, coincident, lagging), yen-dollar exchange rate, M2 and Monetary Base. Applying the CHPCA to such integrated data enables us to elucidate dynamical linkage of comovement of prices with macroeconomic variables.

In the next section, details of the data set used here is given with preprocessing procedure. In Sec. 3 the CHPCA is reviewed for self-containment of the paper, and Sec. 4 is devoted to a brief description of the RMT and the RRS for detection of statistically meaningful eigenmodes. In Sec. 5 the CHPCA results are presented with an interpretation of correlation structures of the significant eigenmodes in terms of a simple collective-motion model. In Sec. 6 the paper is concluded. Some mathematical and statistical details are left for the appendices.

II. Data Set

We have collected the Japanese monthly data of the following categorized individual prices and macroeconomic variables for the period, January 1985 through December 2016:

- Consumer Price Index (CPI)[1] with 47 prices
- Producer Price Index (PPI)[2] with 23 prices
- Import Price Index (IPI)[3] with 10 prices
- US Dollar to Japanese Yen Exchange Rate (USD/JPY)[3]
- Index of Business Condition[4] with 3 indicators (Leading, Coincident, Lagging)
- Money Stock (M2)[5]
- Monetary Base

The totally 86 time series with length of 384 months as shown in Tables 1 and 2 were combined into a multivariate data set. Assuming the prices and the economic variables basically obey geometric brownian motion, we took logarithmic difference of their time series:

$$r_\mu(t) = \log_{10}\left[\frac{p_\mu(t+1)}{p_\mu(t)}\right] \ , \tag{1}$$

where $p_\mu(t)$ ($\mu = 1, \cdots, 86$) are the original time series data. Since the CPI data show jumps when sales tax was imposed (3% in April, 1989) and its rate was raised (from 3% to 5% in April, 1997 and from 5% to 8% in April, 2014) , we removed the sales tax effects simply by taking average of the values just before and after the sales tax shocks. Stationarity of the preprocessed data was then verified by the Phillips-Perron unit root test; all of the time series data are stationary at the 5% significance level. Also, the augmented Dickey-Fuller test has verified that most of the data are stationary except the price #14 (CPI repairs & maintenance) and the price #41 (CPI personal care services).[6]

III. Complex Hilbert Principal Component Analysis

Let us suppose that we have N different time series $x_\mu(t)$ ($\mu = 1, \cdots, N$; $t = 1, \cdots, T$) of length T, which have been standardized with zero mean and unit variance in advance. We first obtain complex time series $\xi_\mu(t)$ out of $x_\mu(t)$ through the relation,

$$\xi_\mu(t) = x_\mu(t) + i y_\mu(t) \ , \tag{2}$$

[1] 2015 base Middle classification, Statistics Bureau of Japan.

[2] 2015 base Middle classification, excluding consumption tax, Bank of Japan.

[3] Tokyo market, monthly average, Bank of Japan.

[4] Composite Index 2015 base, outlier processed, Cabinet Office, Government of Japan

[5] Time series created by connecting the current M2 statistics and the past M2 + CD statistics, Bank of Japan.

[6] The results of the p-value are 0.36 and 0.39 for #14 and #41, respectively.

where the imaginary part $y_\mu(t)$ is Hilbert transform of $x_\mu(t)$ defined by

$$y_\mu(t) = -\frac{1}{\pi}\int_{-\infty}^{\infty} \frac{x_\mu(u)}{t-u}\mathrm{d}u \ . \tag{3}$$

The integration over u in Eq. (3) should be interpreted as Cauchy's principal integration. In the actual calculations, we used a discretized version of the Hilbert transformation (Barnett, 1983), which is expressible in terms of the discretized Fourier transform of x_t:

$$X(k) = \sum_{t=0}^{T-1} x(t) e^{-i\frac{2\pi k t}{T}} . \tag{4}$$

The discrete Hilbert transform of x_t is given by

$$y(t) = \sum_{t=0}^{T-1} X(k) e^{-i\frac{\pi}{2}} e^{i\frac{2\pi k t}{T}} \mathrm{sgn}(k - \frac{T}{2}) \ , \tag{5}$$

with

$$\mathrm{sgn}(k - \frac{T}{2}) = \begin{cases} 1 & (k > T/2) \\ 0 & (k = T/2) \\ -1 & (k < T/2) \end{cases} . \tag{6}$$

We thus see that the Hilbert transformation has the effect of shifting the phase of $x(t)$ at every frequency by $\pi/2$ comparing with the inverse Fourier transformation for $x(t)$:

$$x(t) = \frac{1}{T}\sum_{k=0}^{T-1} X(k) e^{i\frac{2\pi k t}{T}} \ . \tag{7}$$

We then construct the complex correlation matrix $\tilde{\boldsymbol{C}}$ from the complex time series $\{\xi_\mu(t)\}$:

$$\tilde{\boldsymbol{C}} = \frac{1}{T}\boldsymbol{\Xi}\boldsymbol{\Xi}^\dagger, \tag{8}$$

where $\boldsymbol{\Xi}$ denotes $N \times T$ data matrix whose component is $\xi_\mu(t)$ and $\boldsymbol{\Xi}^\dagger$ is Hermite conjugate of $\boldsymbol{\Xi}$.

The complex principal component analysis (CHPCA) computationally amounts to the eigenvalue problem for $\tilde{\boldsymbol{C}}$. Since $\tilde{\boldsymbol{C}}$ is a Hermitian matrix, its eigenvalues are real and furthermore positive definite because of the dyadic form (8). On the other hand, the components of the eigenvectors are complex. The absolute values and the phases of the eigenvector components provide us with information on strength of correlations and lead-lag relationships embedded in multivariate time series. The correlation matrix $\tilde{\boldsymbol{C}}$ is expressible in terms of its eigenvalues and eigenvectors as

$$\tilde{\boldsymbol{C}} = \sum_{\ell=1}^{N} \lambda_\ell \boldsymbol{\alpha}_\ell \boldsymbol{\alpha}_\ell^\dagger \ , \tag{9}$$

where λ_ℓ and $\boldsymbol{\alpha}_\ell$ are the ℓ-th eigenvalue and its associated eigenvector, respectively, and we align the eigenvalues in descending order, that is, $\lambda_1 > \lambda_2 > \cdots > \lambda_N$.

Since the eigenvectors $\boldsymbol{\alpha}_\ell$'s form an orthonormal complete basis set, we can rewrite $\boldsymbol{\xi}(t)$ represented in the standard basis set $\{\boldsymbol{e}_\mu\}$ as

$$\boldsymbol{\xi}(t) = \sum_{\mu=1}^{N} \xi_\mu(t)\boldsymbol{e}_\mu = \sum_{\ell=1}^{N} a_\ell(t)\boldsymbol{\alpha}_\ell \ , \tag{10}$$

where

$$a_\ell(t) = \boldsymbol{\alpha}_\ell^\dagger \cdot \boldsymbol{\xi}(t) \ . \tag{11}$$

We refer to the coefficient $a_\ell(t)$ as *mode signal* of the ℓ-th eigenmode. The mode signals represent temporal behavior of the eigenmodes and their strength is measured by

$$I_\ell(t) = |a_\ell(t)|^2 \ . \tag{12}$$

In addition, we define relative mode intensity $\tilde{I}_\ell(t)$ by

$$\tilde{I}_\ell(t) = \frac{|a_\ell(t)|^2}{\sum_{\ell=1}^{N} |a_\ell(t)|^2} \ , \tag{13}$$

which calculates the fractional contribution of each eigenmode to the overall strength of price fluctuations at every instant of time. Also we note the following identity due to mutual orthogonality of $\boldsymbol{\alpha}_\ell$'s:

$$\boldsymbol{\xi}(t)^\dagger \cdot \boldsymbol{\xi}(t) = \sum_{\mu=1}^{N} |\xi_\mu(t)|^2 = \sum_{\ell=1}^{N} |a_\ell(t)|^2. \tag{14}$$

It is a crucial issue for the CHPCA as well as the PCA how to identify eigenmodes which are statistically significant. The random matrix theory (RMT) serves as a sound null hypothesis for such a statical significance test. However, autocorrelations involved in multivariate data reduce the usefulness of the RMT in removing statistical noise from them. The rotational random shuffling (RRS) method provides us with a null hypothesis alternative to the RMT in such a case (Iyetomi et al., 2011a,b). We impose the periodic boundary condition on each time series to make a "ring" in the time direction and randomly shuffle the data in a rotational way. The randomization destroys only cross-correlations preserving autocorrelations. The RRS serves as a robuster null hypothesis than the RMT. However, we have to numerically solve the eigenvalue problem of the complex correlation matrix for randomized data in the RRS.

IV. Results and Discussion

A. *Significance test of principal components*

We computed eigenvalues of the complex correlation matrix $\tilde{\boldsymbol{C}}$ constructed from the price data. Figure 1 is a parallel (rank-by-rank) comparison of the actual eigenvalues and engenvalues of data after RRS processing (sampled 1000 times). In the eigenvalues after RRS processing, it indicates the average value and standard deviation σ. Here, the average value $+3\sigma$ of the eigenvalues after the RRS processing was used as a criterion of significant eigenvalues. In this criterion, the top seven eigenvalues are significant. This result shows that the eigenvectors associated with those eigenvalues are regarded as manifestation of statistically meaningful correlations among individual prices.

Then, we calculated power spectrum of the mode signals associated with $\boldsymbol{\alpha}_\ell$ ($\ell = 1, \cdots, 7$) as shown in Figure 2. We observe the mode signals of higher order with $\ell \geq 3$ have large peaks corresponding to seasonal variations.

Table 3 spells out similarity between the two sets of eigenvectors. The one is a set of the significant eigenvectors $\boldsymbol{\alpha}_\ell$ ($\ell = 1, \cdots, 7$) obtained for the original data and the other, that of the significant eigenvectors $\boldsymbol{\beta}_m$ ($m = 1, \cdots, 6$) for seasonally adjusted data. We prepared the latter by taking year-to-year change of the original time series; this is the most primitive way for seasonal adjustment. The similarity is measured by calculating the inner product of $\boldsymbol{\alpha}_\ell$ and $\boldsymbol{\beta}_m$ for all pairs. Its computational details are given in Appendix A. The similarity measure η between a pair of complex vectors as defined by (A.2) is a natural extension of the cosine similarity between real vectors. Since the space of eigenvectors in this study is of very high dimension (86 dimensions), the p-value corresponding to $\eta = 0.9$ takes an extremely small value, 4.94×10^{-62}. On the other hand, the similarity corresponding to $p = 0.05$ is 0.186. An analytic formula for the p-value of the similarity is also given in Appendix A.

We thus see that the first two eigenvectors in both sets are in excellent agreement with each other. The remainder in the set $\{\boldsymbol{\alpha}_\ell\}$ has no notable counterparts in the set $\{\boldsymbol{\beta}_m\}$. This is because from the third to the sixth mode signals mainly describe seasonal components of fluctuations in the original data, not involved in the seasonally adjusted data. We thereby focus only on the first and second eigenmodes for the original data. The cumulative contribution ratio of the eigenvalues of those modes is 18.8 percent. We thus see that about 20% of the total price fluctuations can be explained by the collective motion of prices free from seasonal variations.

B. *Interpretation of the first and second eigenmodes*

In Figure 3, the complex components of the first and second eigenvectors are represented in terms of their absolute values and phases. The absolute value of each component in a significant eigenvector measures to what extent the corresponding price contributes to the eigenmode. The phase difference between a pair of components in a significant eigenvector represents lead-lag relationship between their corresponding prices in the eigenmode.

Prices whose components have large magnitude in the eigenvectors play an important role in their correlation structures. However, limited length of the price data would allow a component of completely random time series in the eigenvectors to have finite magnitude. To determine whether prices have statistically relevant components or not in the eigenvectors, we reiterated the CHPCA for the price data to which an auxiliary random time series was added as the 87th component[7]. We then determined the 5% significance level for each eigenmode as regards magnitude of the eigenvector components by collecting 10,000 samples with different random time series. Although the basic structures of the two eigenvectors are robust against addition of such a random time series, all of the components are not statistically meaningful. Here we dismiss components having magnitude below the 5% significance level.

Both in the two significant eigenvectors, most prices are distributed just on a half of complex plane. Such confinement of the phases of prices demonstrates their coherent behavior. In general, the lead-lag relations between prices in phase is not straightforwardly translated to lead-lag relations between them in real time. This is because the CHPCA entirely depends on correlation coefficients averaged over frequency, although information on correlations between time series and their quadrature-phase companions is retained. In our case, however, we recall the business cycle indicators, incorporated into the present analysis, have rather clear lead-lag relations in real time. It is officially said that the leading index is several months ahead of the coincident index, which in turn leads the lagging index by several months to six months. Collecting these results, we may estimate that phase difference of about 180 degrees which the components in the eigenvectors roughly span corresponds to 2 years time difference.

In the first eigenmode, obviously, changes of the exchange rate (#81) induce changes of domestic prices belonging to the PPI and CPI categories through import prices. The business cycle indicators, the leading (#82), the coincident (#83), and the lagging (#84) indices, accompany the exchange rate; the leading index is slightly ahead of the exchange rate. Also prices of raw materials and energy sources such as scrap & waste (#70), nonferrous metals (#57), petroleum & coal (#53) and other fuel & light (#17) synchronize with the exchange rate. And then the remaining PPI prices react to the financial shocks with some degrees of delay and then the CPI prices follow. We note the shocks gradually attenuate in the course of their propagation from upstream to downstream across domestic prices.

In the second eigenmode, on the other hand, domestic demand monitored by the business cycle indicators is assigned as a driving force for domestic prices. Propagation of shocks across prices does not have such damping behavior as observed in the first eigenmode. This is understandable because domestic demand is responsible for the overall economic rise or downturn including changes of prices on downstream side. The exchange rate and import prices except for price of petroleum, coal & natural gas (#75) also have components of large magnitude in the second eigenvector, but their role is not transparent at all. In fact, the curious relationship of the exchange rate with the dynamics of domestic prices is nothing more than a mathematical consequence as shown in Appendix B.

[7] We refer to this significance test as the auxiliary random variable method.

Finally, we note that neither of the two monetary variables, money stock (#85) and monetary base (#86), is an important player in the two eigenmodes. The magnitude of the components of both variables in the two eigenvectors is below the 5% significance level laid down by the auxiliary random variable method.

C. *A comovement of domestic prices*

Closer look at the first and the second eigenvectors suggests that the lead-lag relationship among domestic prices in the two eigenmodes is quite similar to each other. Figure 5 directly compares phases of the significant domestic prices in the first eigenvector with the corresponding phases in the second eigenvector. The prices are well aligned on the correlation plot.[8] It means that there exists robust internal dynamics of domestic prices irrespective of their driving forces, that is, the exchange rate accompanied by import prices in the first eigenmode and domestic demand in the second eigenmode. This empirical fact allows us to claim that domestic prices are interconnected by their mutual interactions to form a chain-like dynamic structure with definite lead-lag relations.

Let us further concentrate on dynamics of prices in the PPI and CPI categories by eliminating import prices and macroeconomic variables from our sights. When the CHPCA is applied to the reduced data set in which only the domestic prices are retained, only the largest eigenvalue exceeds the upper limit of the largest eigenvalue predicted by the RRS. Figure 6 shows the results for the eigenvalues and the complex components of the eigenvector associated with the largest eigenvalue. The eigenvector once again demonstrates collective behavior of domestic prices.

In fact, this collective behavior of domestic prices is quite similar to that revealed by the first and second eigenvectors of the CHPCA for the full data set. The similarity η, Eq. (A.2), of the first eigenvector of the reduced data set to its two counterparts in the first and second modes is calculated as 0.962 and 0.863, respectively. This is the reason why only a single principal component is identified as being significant without import prices and macroeconomic variables. As has been already remarked, comovement of domestic prices is driven in a different way in the two dominant eigenmodes obtained for the full data set; its driving factor is the exchange rate, accompanied by import prices, in the first mode and domestic demand in the second mode. However, shocks propagate across domestic prices sequentially aligned from upstream to downstream in a universal way, irrespective of the origin of shocks. This result further ascertains the existence of mutual interactions among domestic prices leading to such a universal comovement structure of them as schematically depicted in Figure 7.

[8] We can also confirm the strong resemblance between the lead-lag relations of domestic prices in the two eigenmodes by computing the generalized cosine similarity η, Eq. (A.2), between the corresponding complex vectors. The result is 0.73, which is highly significant in reference to the similarity between two random complex vectors. The associated p-value takes an extremely small value, 1.5×10^{-23}.

D. *Relationship with business cycles*

As indicated by the Phillips curve, in fact, the two significant eigenmodes establish a strong connection between collective dynamics of individual prices and business cycles. Furthermore, the business condition indices go ahead of the comovement of individual prices.

The mode signals, Eq. (11), enable us to see to what extent business cycles are dynamically linked with comovement of prices in the first and second eigenmodes. The contribution of the ℓ-th eigenmode to the μ-th component of the multivariate data set is given by

$$\xi_{\ell,\mu}(t) = a_\ell(t)\boldsymbol{\alpha}_{\ell,\mu} \ . \tag{15}$$

Summation of the contributions over all the eigenmodes restores the original complex time series $\xi_\mu(t)$:

$$\xi_\mu(t) = \sum_{\ell=1}^{N} \xi_{\ell,\mu}(t) \ . \tag{16}$$

The observed data is the real part of the corresponding complex time series and may be approximated by the contributions of the first and second eigenmodes:

$$x_\mu(t) = \Re\left[\xi_\mu(t)\right] \simeq \Re\left[\xi_{1,\mu}(t)\right] + \Re\left[\xi_{2,\mu}(t)\right] \ . \tag{17}$$

The original coincident index ($\mu = 83$) is compared with the corresponding contributions of the first and second eigenmodes and their superposition in Fig. 8, where we show the results obtained by successively accumulating their standardized logarithmic difference to make the comparison more lucid. The economic fluctuations are decomposable into two components in conjunction with the comovement of prices. Although the two eigenmodes are responsible only for about 20% of intensity of the total fluctuations as has been already remarked, they can reproduce quite well the business cycles as a whole. We thus see that comovement of individual prices are coupled to the long-term behavior of the economy to a large extent.

It is also true that there are economic peaks and troughs which are associated with neither the first nor the second eigenmode. During the Lost Decade from 1991 through 2002, there are two such peaks observed in the coincident index. The peak early in 1997 arises from the economic boom due to the government's large-scale economic pump-priming measures and the one late in 2000, from the IT bubble in Japan. This result confirms that those booming economies were far from recovery of the real economy. Also, the economic downturn caused by the Great East Japan Earthquake in March 2011 activated neither of the two eigenmodes, indicating that the disaster caused no critical damage to the whole economy in Japan. On the other hand, at the time of the global financial crisis triggered by the collapse of Lehman Brothers in September 2008, both eigenmodes were strongly excited by the economic shock. The mode signals clearly differentiate the nature of impact on the Japanese real economy of the world financial crisis from that of the great earthquake.

The second-mode component of business cycles should be thereby focused on more seriously because it gives information on the condition of the whole economy, i.e, on whether current economic growth is driven by an increase in domestic demand or not. From Fig. 8, we can learn that the economic upturn toward the crash of the bubble economy late in 1990 was basically led by the first-mode component. Also the similar situation is observable immediately after launch of Abe's second Cabinet, in December 2012. In corporation with the cabinet, the Bank of Japan has set the inflation target of 2% with large-scale monetary easing to promote recovery of the Japanese economy from the long-standing depression. It is clear that so-called Abenomics was initially successful. However, the doctrine is not so influential on the economy in the sense that it does not excite the comovement of prices in the second mode; even the second-mode component began to decline early in 2014. This should be continued to be monitored.

V. Summary

This study aimed to empirically elucidate collective behavior of individual prices and its dynamical linkage with macroeconomic variables representing the business and financial conditions in Japan. For the purpose, we applied the CHPCA to the composite monthly data set constructed from the individual price indices constituting IIP, PPI, and CPI, the leading, coincident, and lagging indices of business conditions, the yen-dollar exchange rate, the money stock (M2), and the monetary base, spanning the period from January 1985 to December 2016. The statistical test of the principal components with the RRS as a null hypothesis combined with the spectral analysis identified two principal components as being statistically meaningful. The lead-lag relations among domestic prices in the two modes are quite similar, indicating the individual prices behave in a collective way. However, the collective motion of prices is driven differently, that is, by the exchange rate at the upper stream side in the first mode and domestic demand at the lower stream side in the second mode. In contrast, the monetary variables play no important role in the two modes.

The empirical evidence for comovement of individual prices and its lead-lag relationship with business cycles reaffirms the importance of macroeconomics. We also expect that our findings here provide a sound basis for evidence-based policymaking.

Acknowledgements

This study has been conducted as a part of the Project "Large-scale Simulation and Analysis of Economic Network for Macro Prudential Policy" undertaken at Research Institute of Economy, Trade and Industry (RIETI). This research was also supported by MEXT as Exploratory Challenges on Post-K computer (Studies of Multi-level Spatiotemporal Simulation of Socioeconomic Phenomena) and by JSPS KAKENHI Grant Number 25400393. We would like to thank the seminar participants at RIETI for their helpful comments and suggestions.

Appendix A. Similarity Between Complex Vectors

One can measure similarity between two complex vectors $\boldsymbol{v}, \boldsymbol{w}$ with arbitrary phase factors by

$$\gamma(\boldsymbol{v}, \boldsymbol{w}) = \min_{\varphi} \left[\left\| \frac{\boldsymbol{v}}{\|\boldsymbol{v}\|} - e^{i\varphi} \frac{\boldsymbol{w}}{\|\boldsymbol{w}\|} \right\|^2 \right] = 2(1-\eta) \ , \tag{A.1}$$

with

$$\eta = \frac{|\boldsymbol{u} \cdot \boldsymbol{v}^*|}{\|\boldsymbol{u}\| \|\boldsymbol{v}\|} \ , \tag{A.2}$$

where $\|\boldsymbol{u}\|$ stands for the norm of $\boldsymbol{u}$. The distance $\gamma(\boldsymbol{v}, \boldsymbol{w})$ satisfies inequality $0 \leq \gamma \leq 2$ $(1 \leq \eta \leq 0)$ and takes $\gamma = 0$ $(\eta = 1)$ only when $\boldsymbol{v}$ and $\boldsymbol{w}$ coincide except for the degree of freedom of phase factor φ. The similarity η between complex vectors is a natural extension of the cosine similarity between real vectors.

The similarity η Assuming $\boldsymbol{u} = (1, 0, \cdots, 0)$ and $\boldsymbol{v} = (x_1 + iy_1, x_2 + iy_2, \cdots, x_n + iy_n)$ without loss of generality, one can calculate η as

$$\eta = \frac{r}{\sqrt{r^2 + R^2}} \ , \tag{A.3}$$

where

$$r =: \sqrt{x_1^2 + y_1^2} \ , \tag{A.4}$$

$$R =: \sqrt{x_2^2 + y_2^2 + \cdots + x_n^2 + y_n^2} \ , \tag{A.5}$$

If each component of $\boldsymbol{v}$ obey the standardized normal distribution, the probability density function of (r, R) is given by

$$f(r, R) \propto r R^{2n-3} e^{-\frac{(r^2+R^2)}{2}} \ . \tag{A.6}$$

Then, the relation (A.3) derives the probability density function of η from Eq. (A.6):

$$f(\eta) = 2(n-1)\eta(1-\eta^2)^{n-2}. \tag{A.7}$$

Also, one can calculate the p-value corresponding to η from Eq. (A.7) as

$$p = \int_{\eta}^{1} f(\eta') \mathrm{d}\eta' = (1-\eta^2)^{n-1} \ . \tag{A.8}$$

Figure 9 confirms Eq. (A.7) through comparison with the probability density function of η between random complex vectors generated by numerical simulations.

Appendix B. Mathematical Structure of Eigenvectors of Two-variable Model

In Section IV.B, we observe that the exchange rate and import prices, though they have large absolute values, lie behind domestic prices in the second mode (Figure 3(b)). In this appendix, we show that it is nothing but a mathematical necessity in two-variable model. To understand the correlation structures observed in the first and second eigenmodes, we introduce a simple two-variable model. For this purpose, we first replace the group motion of domestic prices by a single collective coordinate, that is, the mode signal of the first eigenmode of the CHPCA applied to the reduced data set in which only domestic pries are retained. Also we replace the dollar-yen exchange rate and import prices by another collective coordinate. Adopting the two collective coordinates reduces the economic system under study to a two-variable model.

In this two-variable model, the complex correlation matrix $\tilde{\boldsymbol{C}}$ has such a reduced form as

$$\tilde{\boldsymbol{C}} = \begin{pmatrix} \sigma_1 & \sigma_{12} \\ \sigma_{12}^* & \sigma_2 \end{pmatrix} , \tag{B.9}$$

where σ_1 and σ_2 are the variances of the collective coordinates for domestic prices and the exchange rate accompanied by import prices, respectively, and σ_{12} is a complex correlation coefficient between the two coordinates.

If σ_1 and σ_2 take an identical value σ, the two eigenvalues $\lambda_\pm$ are calculated as

$$\lambda_\pm = \sigma \pm |\sigma_{12}| , \tag{B.10}$$

with their eigenvectors $\boldsymbol{V}_\pm$ given by

$$\boldsymbol{V}_+ = \begin{pmatrix} 1 \\ \exp(-i\theta) \end{pmatrix} , \ \boldsymbol{V}_- = \begin{pmatrix} 1 \\ -\exp(-i\theta) \end{pmatrix} , \tag{B.11}$$

where θ is the phase angle of σ_{12}. We see that the relationship between the comovement of domestic prices and the exchange rate in $\boldsymbol{V}_+$ is reversed in $\boldsymbol{V}_-$. When $0 < \theta < \pi/2$, for example, the exchange rate leads the comovement of domestic prices with phase difference θ in $\boldsymbol{V}_+$, while the exchange rate follows the comovement of domestic prices with phase difference $\pi - \theta$.

In the actual data, we obtain $\sigma_1 = 7.81$ and $\sigma_2 = 7.42$. The former is the largest eigenvalue of the submatrix of $\tilde{\boldsymbol{C}}$ for domestic prices, and the latter, that for the exchange rate and import prices. Thus, the condition $\sigma_1 = \sigma_2$ is approximately satisfied. The model with (B.11) of $\boldsymbol{V}_+$ and $\boldsymbol{V}_-$ therefore well explains the correlation structures in the two dominant eigenmodes. The exchange rate drives the comovement of domestic prices in the first eigenmode to fix the phase difference θ between the two collective coordinates. On the other hand, in the second eigenmode, the lead-lag relationship of the exchange rate with the comovement of domestic prices is automatically determined by $\pi - \theta$. This is basically what we observe in Figure 3(b). It is simply a mathematical necessity in a two-variable model.

Given this mathematical fact, it is not the end of the story. Because replacement of the exchange rate by a completely random time series would result in the same mathematical relation between $\boldsymbol{V}_+$ and $\boldsymbol{V}_-$ as long as the condition $\sigma_1 \simeq \sigma_2$ is satisfied. The random time series is fixed to the comovement of domestic prices at any phase angle. The remaining issue to be addressed is thereby whether the fixed phase difference θ between the two collective coordinates in the first eigenmode is statistically significant or not. To test statistical significance of the phase angle between the comovement of domestic prices and the exchange rate, we reiterated the CHPCA calculation for the data set in which the exchange rate and import prices are substituted by a random time series with the variance kept the same; the new results serve as a null model. The strength of coupling between the two collective coordinates is represented by the magnitude of σ_{12} and hence by difference of the two dominant eigenvalues as shown in Eq. (B.10). Figure 10 demonstrates distribution of $\lambda_1 - \lambda_2$ in the null model. On the other hand, the actual result for $\lambda_1 - \lambda_2$ is 2.401 and its p-value is given as 0.006 according to the null hypothesis. This comparison allows us to infer that the fixed phase angle between the comovement of domestic prices and the exchange rate is statistically meaningful.

In conclusion, the correlation structures in the two dominant eigenmodes are fully understandable with a two-variable model. And also we confirm that the exchange rate is certainly a driving factor for the first eigenmode.

References

Arai, Yuta, Takeo Yoshikawa, and Hiroshi Iyetomi, "Complex principal component analysis of dynamic correlations in financial markets," *Frontiers in Artificial Intelligence and Applications*, 2013, *255*, 111 – 119.

Barnett, TP, "Interaction of the monsoon and Pacific trade wind system at interannual time scales Part I: the equatorial zone," *Monthly Weather Review*, 1983, *111* (4), 756–773.

Bils, Mark and Peter J Klenow, "Some evidence on the importance of sticky prices," Technical Report, National Bureau of Economic Research 2002.

Horel, J D, "Complex Principal Component Analysis: Theory and Examples," *Journal of Climate and Applied Meteorology*, December 1984, *23* (12), 1660–1673.

Iyetomi, Hiroshi, Yasuhiro Nakayama, Hideaki Aoyama, Yoshi Fujiwara, Yuichi Ikeda, and Wataru Souma, "Fluctuation-dissipation theory of input-output interindustrial relations," *Physical Review E*, 2011, *83* (1), 016103.

– , – , Hiroshi Yoshikawa, Hideaki Aoyama, Yoshi Fujiwara, Yuichi Ikeda, and Wataru Souma, "What causes business cycles? Analysis of the Japanese industrial production data," *Journal of the Japanese and International Economies*, 2011, *25* (3), 246–272.

Klenow, Peter J and Benjamin A Malin, "Microeconomic evidence on price-setting," Technical Report, National Bureau of Economic Research 2010.

Kwapien, Jaroslaw and Stanislaw Drozdz, "Physical approach to complex systems," *Physics Reports*, 2012, *515* (3 4), 115 – 226. Physical approach to complex systems.

Laloux, Laurent, Pierre Cizeau, Jean-Philippe Bouchaud, and Marc Potters, "Noise Dressing of Financial Correlation Matrices," *Phys. Rev. Lett.*, Aug 1999, *83*, 1467–1470.

Mankiw, N. Gregory, "Small Menu Costs and Large Business Cycles: A Macroeconomic Model of Monopoly," *The Quarterly Journal of Economics*, 1985, *100* (2), 529–537.

Plerou, Vasiliki, Parameswaran Gopikrishnan, Bernd Rosenow, Luís A. Nunes Amaral, Thomas Guhr, and H. Eugene Stanley, "Random matrix approach to cross correlations in financial data," *Phys. Rev. E*, Jun 2002, *65*, 066126.

Stein, Karl, Axel Timmermann, and Niklas Schneider, "Phase Synchronization of the El Niño-Southern Oscillation with the Annual Cycle," *Phys. Rev. Lett.*, Sep 2011, *107*, 128501.

Utsugi, Akihiko, Kazusumi Ino, and Masaki Oshikawa, "Random matrix theory analysis of cross correlations in financial markets," *Phys. Rev. E*, Aug 2004, *70*, 026110.

Vodenska, I., H. Aoyama, Y. Fujiwara, H. Iyetomi, and Y. Arai, "Interdependencies and Causalities in Coupled Financial Networks," *PLoS ONE*, 2016, *11* (3), e0150994.

Yoshikawa, Hiroshi, Hideaki Aoyama, Yoshi Fujiwara, and Hiroshi Iyetomi, "Deflation/Inflation Dynamics: Analysis based on micro prices," *Available at SSRN:http://ssrn.com/abstract=2565599 or http://dx.doi.org/10.2139/ssrn.2565599*, 2015.

Table 1. List of items for CPI.

ID	CPI
1	Cereals
2	Fish & seafood
3	Meats
4	Dairy products & eggs
5	Vegetables & seaweeds
6	Fruits
7	Oils, fats & seasonings
8	Cakes & candies
9	Cooked food
10	Beverages
11	Alcoholic beverages
12	Meals outside the home
13	Rent
14	Repairs & maintenance
15	Electricity
16	Gas
17	Other fuel & light
18	Water & sewerage charges
19	Household durable goods
20	Interior furnishings
22	Bedding
22	Domestic utensils
23	Domestic non-durable goods
24	Domestic services
25	Clothes
26	Shirts, sweaters & underwear
27	Footwear
28	Other clothing
29	Services related to clothing
30	Medicines & health fortification
31	Medical supplies & appliances
32	Medical services
33	Public transportation
34	Private transportation
35	Communication
36	School fees
37	School textbooks & reference books for study
38	Tutorial fees
39	Recreational durable goods
40	Recreational goods
41	Books & other reading materials
42	Recreational services
43	Personal care services
44	Toilet articles
45	Personal effects
46	Tobacco
47	Other miscellaneous

Table 2. List of items for PPI and IPI with macroeconomic variables.

ID	PPI
48	Food, beverages, tobacco & feedstuffs
49	Textile products
50	Lumber & wood products
51	Pulp, paper & related products
52	Chemicals & related products
53	Petroleum & coal products
54	Plastic products
55	Ceramic, stone & clay products
56	Iron & steel
57	Nonferrous metals
58	Metal products
59	General purpose machinery
60	Production machinery
61	Business oriented machinery
62	Electronic components & devices
63	Electrical machinery & equipment
64	Information & communications equipment
65	Transportation equipment
66	Other manufacturing industry products
67	Agriculture, forestry & fishery products
68	Minerals
69	Electric power, gas & water
70	Scrap & waste
ID	IPI
71	Foodstuffs & feedstuffs
72	Textiles
73	Metals & related products
74	Wood, lumber & related products
75	Petroleum, coal & natural gas
76	Chemicals & related products
77	General purpose, production & business oriented machinery
78	Electric & electronic products
79	Transportation equipment
80	Other primary products & manufactured goods
ID	Macroeconomic variable
81	US Dollar to Japanese Yen Exchange Rate
82	Index of Business Condition Leading Index
83	Index of Business Condition Coincident Index
84	Index of Business Condition Lagging Index
85	M2(seasonally adjusted)
86	Monetary Base (seasonally adjusted)

Table 3. Similarity $\eta = |\boldsymbol{\alpha}_\ell \cdot \boldsymbol{\beta}_m^*|$ between the statistically significant eigenvectors $\boldsymbol{\alpha}_\ell$ ($\ell = 1, \cdots, 7$) obtained for the original data and those $\boldsymbol{\beta}_m$ ($m = 1, \cdots, 6$) for the seasonally adjusted data.

$\ell \backslash m$	1	2	3	4	5	6
1	**0.956**	0.115	0.039	0.009	0.011	0.053
2	0.063	**0.847**	0.190	0.122	0.056	0.098
3	0.149	0.312	0.156	0.138	0.065	0.168
4	0.12	0.247	0.171	0.204	0.146	0.215
5	0.025	0.084	0.422	0.198	0.444	0.039
6	0.029	0.034	0.526	0.06	0.151	0.121
7	0.022	0.026	0.047	0.522	0.199	0.511

Table 4. Similarity $\eta = |\boldsymbol{\alpha}'_\ell \cdot \boldsymbol{\gamma}_m^*|$ between the domestic price portion $\boldsymbol{\alpha}'_\ell$ ($\ell = 1, \cdots, 7$) of the statistically significant eigenvectors for the original data and the eigenvector $\boldsymbol{\gamma}_m$ ($m = 1, \cdots, 6$) for the domestic price data.

$\ell \backslash m$	1	2	3	4	5	6
1	**0.939**	0.052	0.056	0.097	0.153	0.054
2	**0.826**	0.099	0.108	0.163	0.268	0.097
3	0.014	**0.997**	0.045	0.014	0.022	0.013
4	0.053	0.062	**0.992**	0.023	0.038	0.023
5	0.016	0.023	0.047	**0.974**	0.167	0.067
6	0.030	0.029	0.044	0.074	**0.911**	0.245
7	0.020	0.009	0.008	0.062	0.165	**0.949**

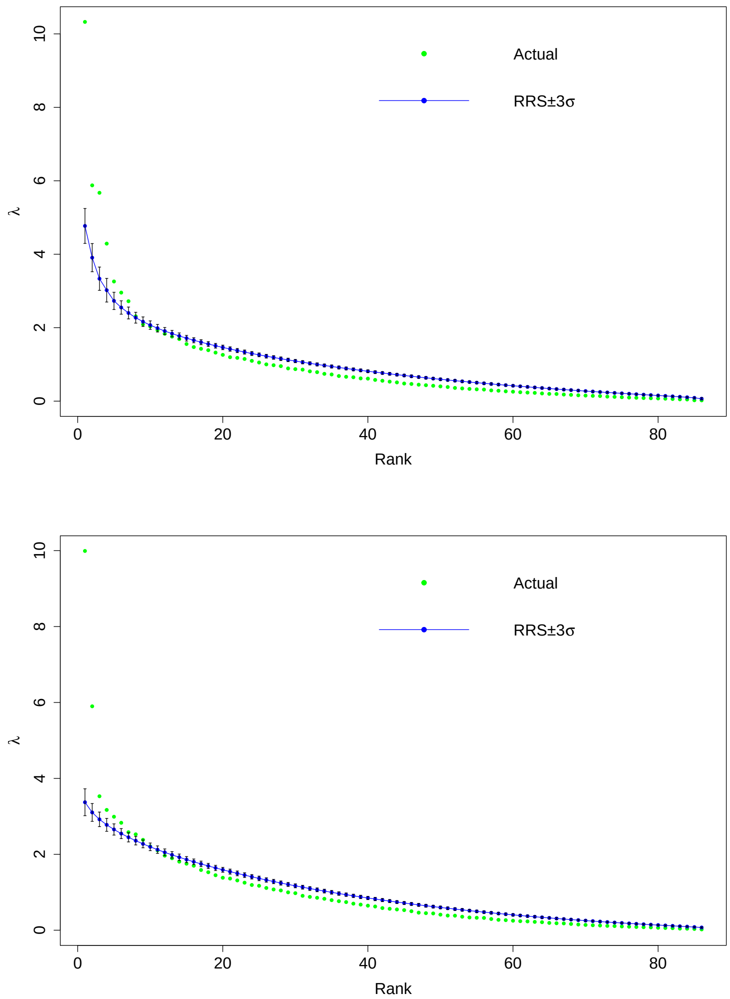

Figure 1. Parallel comparison of the eigenvalues with the RRS preprocessing (sampled 1000 times). Top and bottom panels show results based on the original data and the seasonally adjusted data, respectively.

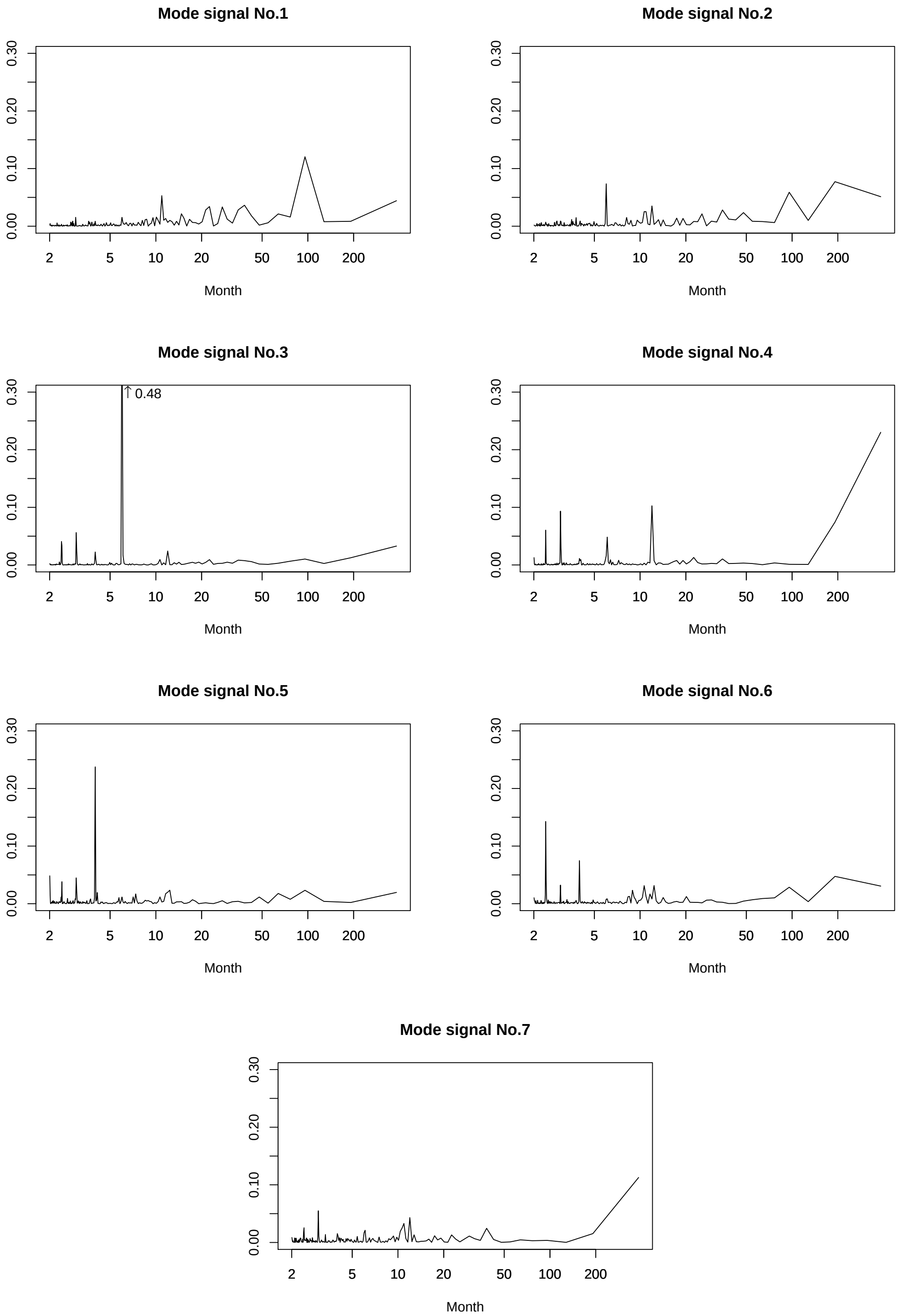

Figure 2. Power spectral density of the mode signal $a_\ell(t)$ $(\ell = 1, \cdots, 7)$ of the significant eigenmodes obtained for the original data.

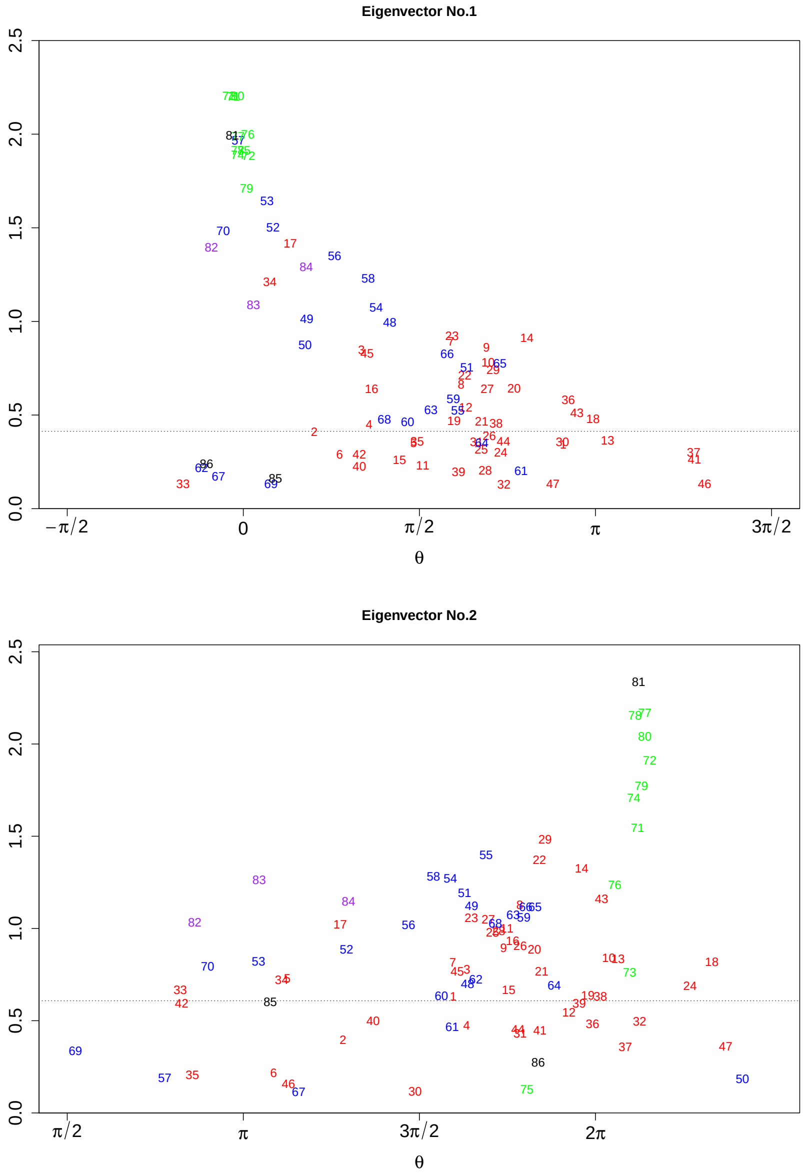

Figure 3. Lead-lag relationship between components, individual prices and macroeconomic variables, in the first and second eigenmodes. The horizontal and vertical axes show phase and absolute value of each component in the eigenmodes, respectively. The dotted line is a criterion to identify significant components with 5% significance level. Time direction is from left to right, that is, left components are ahead of right ones in time.

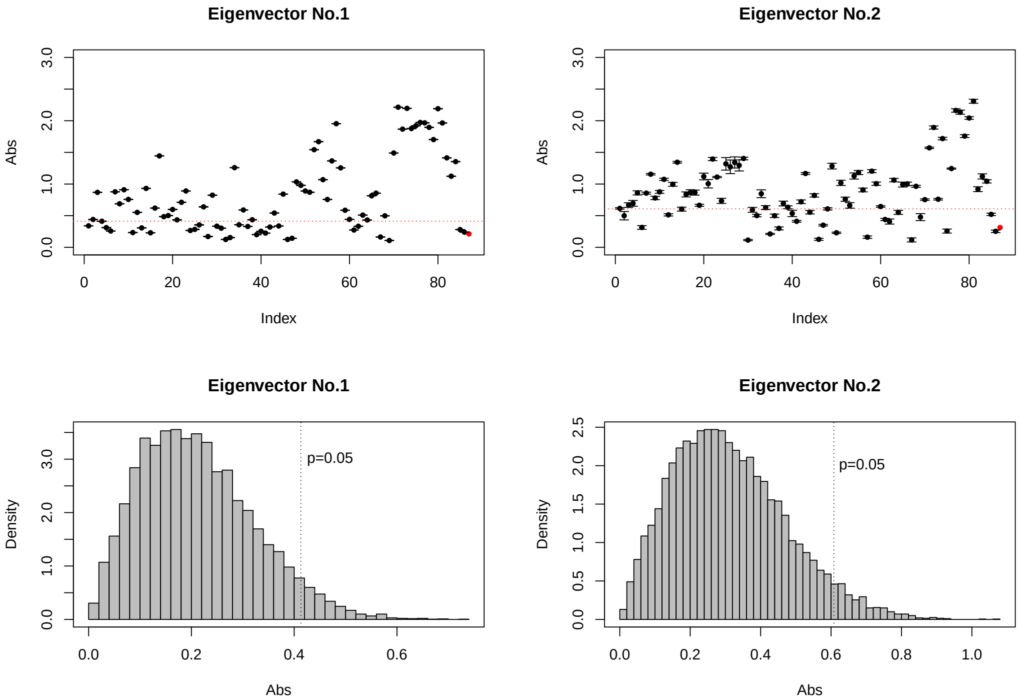

Figure 4. Results of the CHPCA applied to the original data with a random time series added. The upper panels show absolute value of each component in the two dominant eigenvectors, where its mean and $\pm 2\sigma$ error bar are plotted. The horizontal dotted line indicates the significance level of 0.05 determined from statistical variation of the extra random variable, the 87th component. The lower panels are the density distributions of absolute value of the extra random component in the two eigenmodes, which determine the significance level for each mode.

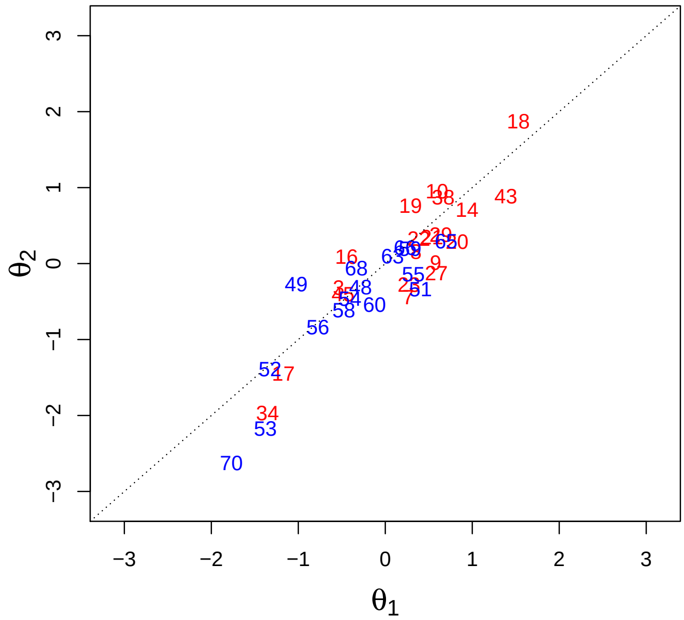

Figure 5. Comparison of the lead-lag relationship between the significant domestic prices in the first and the second eigenvectors. The phase θ_2 of each component in the second eigenvector is plotted against the phase θ_1 of the corresponding component in the first eigenvector.

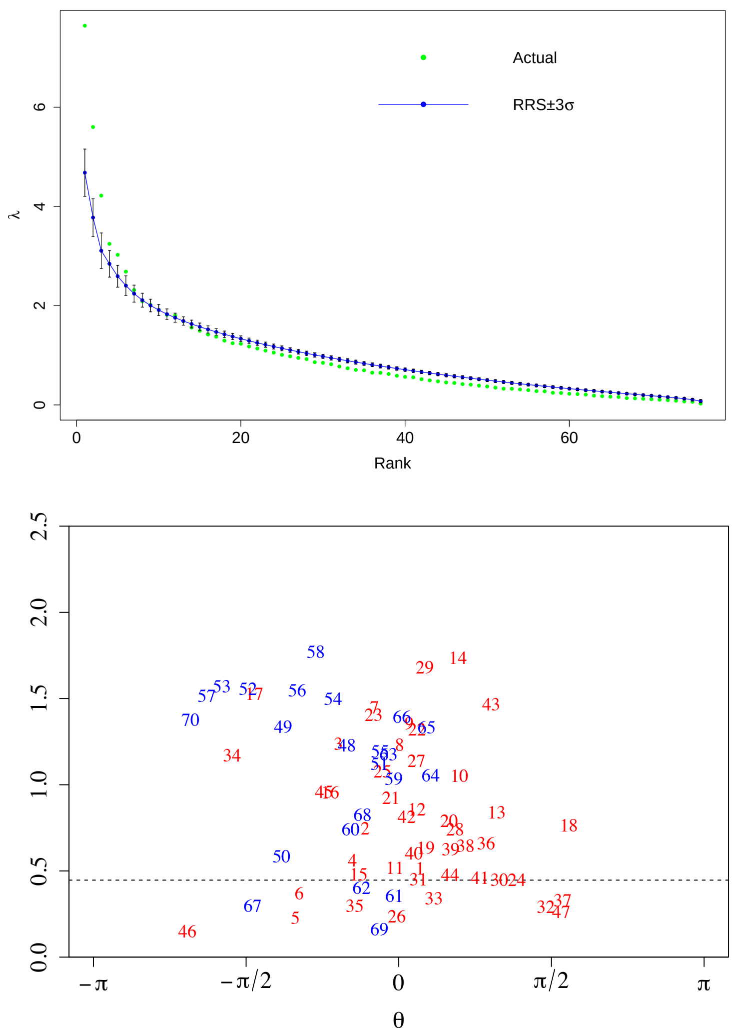

Figure 6. Results of the CHPCA applied to the reduced data without IPI's and the macroeconomic variables. The upper panel shows the eigenvalue distribution and the lower panel, the components in the first eigenvector (their magnitude is plotted against their phase).

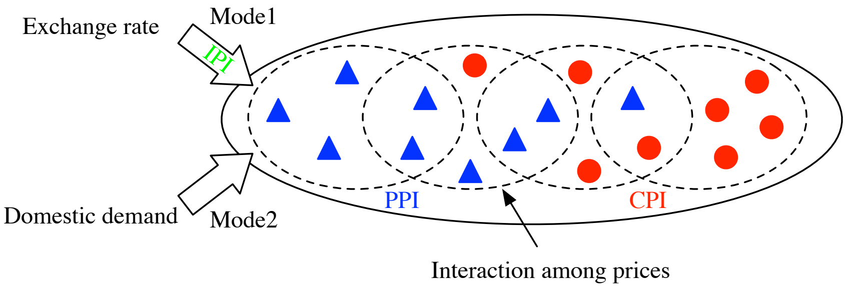

Figure 7. Schematic diagram of comovement of domestic prices, PPI's and CPI's, originating from their mutual interactions with its driving factors, the exchange rate through IPI's in the first mode and domestic demand in the second mode.

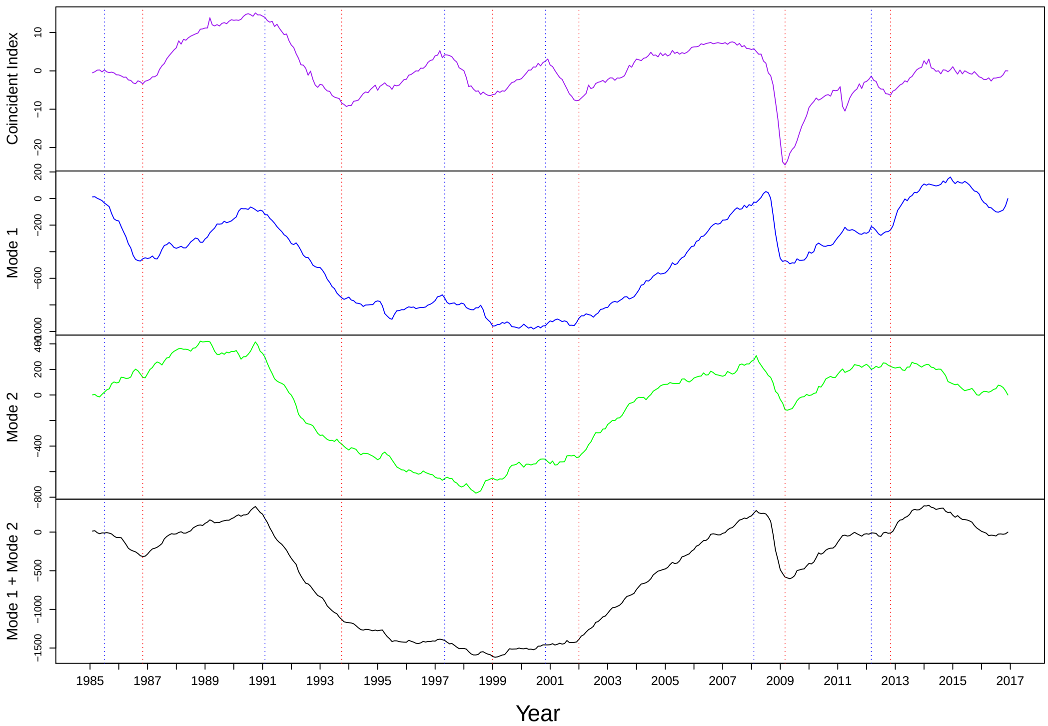

Figure 8. Coincident index reconstructed from its standardized logarithmic difference compared with the corresponding contributions in the first and second eigenmodes and their sum. The red and blue dotted vertical lines represent the economic troughs and peaks, respectively, drawn with the dates determined by the Cabinet Office.

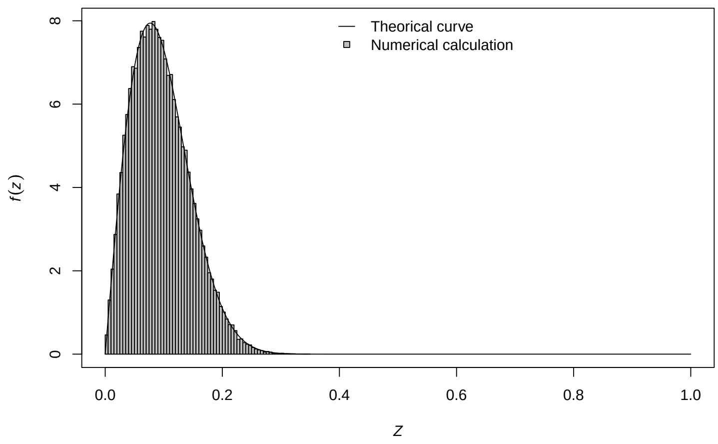

Figure 9. Numerical result of the probability density distribution of the similarity η between 86-dimensional random complex vectors for 10,000 samples, compared with the theoretical result, Eq. (A.8).

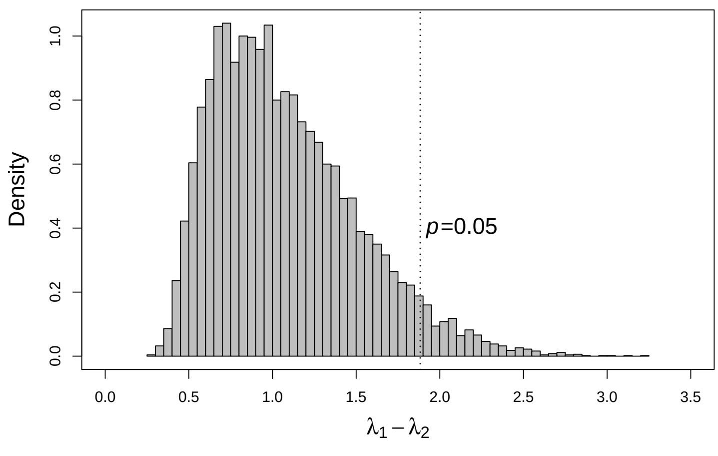

Figure 10. Distribution of the eigenvalue separation, $\lambda_1 - \lambda_2$, in the CHPCA with a random time series in place of the exchange rate and import prices (sampled 10,000 times), where the variance of random time series is kept the same as the original one.