A Nonlinear Factor Analysis of Business Cycles with A Large Data Set: Evidence from Japan and the U.S. *

Mototsugu Shintani†

January 2002
This version: May 2002

Abstract

This paper first constructs a business cycle index of Japanese economic activity based on a dynamic factor model with a large data set, using a principal components method employed by Stock and Watson (1998) in their analysis of the U.S. diffusion index. As in the U.S., the factor-diffusion index in Japan is found to be useful in the context of out-of-sample forecasting. Secondly, the business cycle characteristics of the U.S. and Japan are further investigated by the two-step estimation of the dynamic factor structure. The evidence suggests the possibility of nonlinearity in both the U.S. and Japan while it excludes the class of nonlinearity that can generate endogenous fluctuation or chaos.

Keywords: Diffusion Index; Dynamic Factor Model; Nonlinearity; Prediction

JEL classification: F31; F41

---

*This research was conducted during the author's visit to the Economic and Social Research Institute (ESRI), Cabinet Office of the Japanese government. The author would like to thank Yanqin Fan, Koichi Hamada, Toshiaki Watanabe and Mark Watson for their helpful comments and discussions. The author also thanks Yasuyuki Komaki and officers of ESRI for their help in obtaining the data set.

†Department of Economics, Vanderbilt University, Nashville, TN 37235, USA. e-mail: `mototsugu.shintani@vanderbilt.edu`

1 Introduction

The components of the official Japanese business cycle index have recently been revised. This is the eighth major revision since the government first introduced the index in 1960. Such a revision, however, requires an expert judgment on the merits of a limited number of economic variables.

A dynamic factor approach based on a large number of time-series variables, on the other hand, is not subject to this variable selection problem since the common factors are systematically extracted using all available information from each series. Stock and Watson (1998) designed a new class of diffusion indexes based on the principal components estimator of the common factors, using a large data set. The index for the United States, constructed from 215 macroeconomic series, was then used in out-of-sample forecasts of several key U.S. economic variables. The diffusion index forecast outperformed conventional time-series forecasts.

The first goal of this paper is to construct a diffusion index of Japanese economic activity similar to Stock and Watson's, and to see if the diffusion index forecast works as well for Japan as it does for the U.S. We estimate dynamic factors from the principal components of 234 monthly Japanese macroeconomic variables. The out-of-sample forecasts are then conducted for six important variables, including industrial production and inflation. The diffusion index forecast for Japan leads to a considerable reduction in forecast mean square errors compared to those from benchmark autoregressive models. Therefore, this evidence from Japan, as do the results from the U.S., confirms the usefulness of the factor-diffusion index.

The empirical success of diffusion index forecasts suggests that the driving forces of business cycles responsible for covariation of the macroeconomic variables are well captured by a small set of estimated factors. In other words, the generating mechanism of the common factors offers important information about the characteristics and sources of business cycles. One advantage of the principal components method over other methods, such as the one based on exact MLE, is that it allows a very general class of the dynamic structure of the factors. In consequence, we can estimate the linear or nonlinear model of the factors, using two-step estimation method. This approach looks promising if we can justify replacing the unobservable latent factors with the estimated factors in the second step estimation.

The second goal of the paper is to show the validity of such a two-step estimation of the factor structure and to investigate the possibility of nonlinear factor structure, using the data from the U.S. and Japan. For both parametric and nonparametric estimators of the factor dynamics, conditions required for the two-step estimation are derived. When the number of the series (N) increases at a sufficiently fast rate compared to the time-series observations (T), the effect of the estimation error of the factor is shown to be negligible in the limit distribution of the second-step estimators. Based on this fact, the dynamics of the common factors are investigated using various methods, including kernel-based nonparametric autoregression and neural networks.

The results from the test for linearity suggest the possibility of nonlinear factor structures in the U.S. and Japan. Finally, a nonparametric test for chaos is applied to the factors to see if the nonlinearity in the data is complex enough to generate endogenous fluctuations in the economy. For most cases, the hypothesis of chaos based on common factors is significantly rejected for both the U.S. and Japan.

The remainder of the paper is organized as follows: Section 2 constructs a Japanese diffusion index based on a dynamic factor model and applies the index to forecast selected macroeconomic time-series. Section 3 provides a nonlinear analysis of business cycle based on the factor-diffusion index using both data from the U.S. and Japan. Some concluding remarks are made in Section 4.

2 Diffusion Index Forecasts in Japan

2.1 Background

In January 2002, the Cabinet Office of Japanese government (formerly the Economic Planning Agency) revised the components of official business cycle indexes — the diffusion index and the composite index. The most drastic change was made regarding the leading indicators, as five out of the former eleven series were replaced and one new series was added. For the coincident indicators, the raw materials consumption index was replaced with the producers goods shipment index. The total number of the lagged indicators was reduced from eight to seven. The revision of the components always involves extensive review of many candidate series.[1] The new series that replaces the former series is expected to have a stronger connection to the current business cycle. However, since the indexes are based on a small number of economic variables, there is a possibility of losing useful information in some of the candidate series that are not selected. In short, the official diffusion index measures the average direction of the selected series and the official composite index measures the average growth rate of the series. In other words, both indexes can be considered as a weighted average of the series with unit weight on a limited number of the series and zero weight on all other remaining variables that are available. Since such a weighting scheme requires expert judgment, it is more desirable to construct an index with the automatic weight selection procedure.[2]

In this section, we construct an index of Japanese economic activity based on the factor model with a large data set recently proposed by Stock and Watson (1998). We then apply this index to out-of-sample forecasting of selected individual time-series. This type of diffusion index forecast has been proved to be successful in

[1] The report by the Cabinet Office (1997) contains a list of 253 candidate series used in the seventh revision of the Japanese official business cycle index. Candidate variables employed for the eighth revision are not published but are similar to those used in the previous revision.

[2] Such a variable selection issue on the Japanese official diffusion index is also discussed in Kanoh (1990).

forecasting the macroeconomic time-series in the U.S. (Stock and Watson, 1998, 1999, 2001).[3] Furthermore, real-time version of the diffusion index in the U.S. economy based on such a model is currently available from Federal Reserve Bank of Chicago. However, any studies of its performance in Japan are not yet available. Since the index is constructed by applying the principal components method to a large number of variables, all the candidate series considered by the Cabinet Office can be included in the analysis of Japanese business cycles.

We first summarize the basic feature of the factor model which has been used in the literature. Let x_{it} be an i-th component of N-dimensional multiple time-series $X_t = (x_{1t}, \ldots, x_{Nt})'$ and $t = 1, ..., T$. A natural way to explain the comovement of x_{it}'s caused by a single factor, such as productivity shocks, demand shocks or monetary policy shocks, is to use a simple one-factor model

$$x_{it} = \lambda_i^0 f_t^0 + e_{it} \tag{1}$$

for $i = 1, ..., N$, where λ_i^0's are factor loadings with respect to i-th series, f_t^0 is a scalar common factor, and e_{it}'s are uncorrelated idiosyncratic shocks. While the factor f_t^0 is not directly observable, it is known that the f_t^0 can be consistently estimated by using the first principal component of the $N \times N$ covariance matrix $X'X$ where X is the $T \times N$ data matrix with t-th row X_t', or by using the first eigenvector of the $T \times T$ matrix XX'. There are several ways to generalize this simple model. First, multiple factors can be included. Second, instead of using the static factor model, a dynamic structure can be introduced by allowing (i) a dynamic data generating process for f_t^0, (ii) lags of f_t^0 in (1), and (iii) serial correlation in e_{it}'s. The factor model with such structure is called as a dynamic factor model and has become popular in macroeconomic analysis after influential works by Sargent and Sims (1977), Geweke (1977) and Stock and Watson (1989). Third, e_{it}'s can be correlated across series. Such a correlated case is sometimes referred to as the approximate factor model as opposed to the exact factor model that does not allow cross-sectional correlation. This third generalization seems to be important for the purpose of constructing a diffusion index from the large N series since variables in the same group (e.g., real output) are likely to be mutually correlated but less correlated with variables in the other groups.

The generalized models with small N may be estimated using maximum likelihood method combined with the Kalman filter technique. However, the theoretical results for the generalized dynamic factor models allowing for a large N have not been available until recently. Stock and Watson (1998) have shown the consistency of the principal components estimator as well as the first-order forecasting efficiency, using the estimated factors when both T and N tend to infinity. Forni and Reichlin (1998) and Forni, Hallin, Lippi and Reichlin (2000) have considered alternative estimation methods and have proven the validity of the methods with large N. Bai and Ng (2002) have considered the issue of selecting the number of factors, while Bai (2001)

[3] Marcellino, Stock and Watson (2000) applied a similar approach to European data.

has derived the asymptotic distribution of the principal components estimators of factors.

Before presenting the estimation results of a factor model with large N, we briefly review related works on the factor model-based business cycle index in Japan. The idea of using the principal components method to extract the common business cycle factors is not new in the analysis of Japanese business cycles. Kariya (1993) used the first principal component to define the common trend factor and the second principal component to extract the common cyclical factor in his MTV model. Applicability of Stock and Watson (1989) type dynamic factor models in Japan is considered by Mori, Satake and Ohkusa (1993). The performance of the model using more recent data is investigated in Fukuda and Onodera (2001) and Yoshizoe, et al. (2001). The Japan Center for Economic Research also reports the latest estimates under the name of "JCER Business Index." While these studies employ factor models to construct business cycle indexes, there are no studies applying the factor model with a large data set in Japan.

2.2 Estimation of a Factor-Diffusion Index

Our estimation of a factor utilizes a balanced panel of 234 monthly series from 1973:2 to 2000:12 (see Appendix B for the list of variables). It should be noted that a large number of the series overlap the candidate series considered by the Cabinet Office in the revision of the official business cycle index (see also variable list in Cabinet Office, 1997). Most variables are expressed in first differences of logs of seasonally adjusted series or seasonal growth rates of unadjusted series to obtain the I(0) stationarity. In addition, all the series are standardized to have sample mean zero and unit sample variance. A common factor estimator, $\widetilde{f}_t$, is defined by the first principal component of all the data with a normalization $T^{-1}\sum_{t=1}^{T}\widetilde{f}_t^2 = 1$.

The index of industrial production in mining and manufacturing (hereafter referred to as IP) is one of the coincident indicators used in the current diffusion index by the Cabinet Office. Figure 1 plots IP with recessionary episodes shown in the shaded area. Since the cyclical peaks and troughs of IP is known to be very close to the official business cycle dating, it is of interest to compare the factor-diffusion index with IP. Figures 2 and 3 show the cumulated factor $\sum_{i=1}^{t}\widetilde{f}_i$ and the factor $\widetilde{f}_t$, respectively, with recessionary episodes shown in the shaded area. It can be seen that the cyclical peaks and troughs of the IP in Figure 1 and that of the common factor in Figure 2 do not always coincide.

For the purpose of determining the turning point based on the common factor, Diebold and Rudebusch (1996) have proposed a dynamic factor model in which the factor follows Hamilton's (1989) Markov switching model. Kim and Nelson (1998, 1999) further developed a Markov chain Monte Carlo procedure to estimate such a combined model in a single step when N is small. The procedure of Kim and Nelson was applied to Japanese data by Watanabe (2001). However, when N is large, such a procedure is not applicable because of the computational problem. Instead, we follow Diebold's (2000) simple two-step procedure to estimate a Markov switching model combined with a large N dynamic factor model. Figure 4 shows the recession probabilities computed from fitting a Markov switching model with AR(2) dynamics to a common factor obtained from the first principal component. The extracted recession probability does not differ much from the Cabinet Office's official recessionary episodes shown by the shaded area. However, one interesting observation is that the probability of recession in 1985 is lower than that of the 1995 period, which is not included in the Cabinet Office business cycle chronology. In section 3, we apply a similar two-step procedure to investigate the possibility of nonlinear dynamic structure of the factor-diffusion index.

2.3 Forecasting Performance

One of the most important area of applying the estimated common factor, or factor-diffusion index, is to use such extracted information from many variables in forecasting macroeconomic time-series. This aspect is emphasized in Stock and Watson (1998, 1999, 2001) and Marcellino, Stock and Watson (2000). We follow their procedure and check the forecasting performance of the factor using a simulated out-of-sample methodology. We consider the following (h-period ahead) forecasting equation with a dynamic factor structure,

$$y_{t+h} = \beta(L)f_t^0 + \gamma(L)y_t + u_{t+h} \tag{2}$$

where y_{t+h} is the scalar time series variable of interest, $\beta(L)$ and $\gamma(L)$ are finite lag polynomials and u_{t+h} is the forecast error satisfying $E(u_{t+h}|\mathcal{F}_t) = 0$ where $\mathcal{F}_t = (f_t^0, y_t, X_t, f_{t-1}^0, y_{t-1}, X_{t-1}, \ldots)$. While the forecast with true factor f_t^0 is infeasible, OLS regression of (2) using estimated factor $\widetilde{f}_t$ is proved to provide first-order efficient forecast by Stock and Watson (1998).

For the forecasting variable y_t, five measures of aggregate activity from the coincident indicators used in current Cabinet Office diffusion index are considered. They are: the index of industrial production (IP); the index of producer's shipments (SHIP); the index of capacity utilization ratio (CAP); the index of sales in small and medium-sized enterprises (SALE); and the index of non-scheduled worked hours (HOUR). In addition, while it is not included in the official diffusion index, the forecasting inflation rate based on the consumer price index (CPI) is considered.[4]

All factors and forecasts are computed using a fully recursive, or simulated out-of-sample, methodology. At each date in the simulated out-of-sample period, the data is standardized and the factors and forecasting models are reestimated. The simulated

[4] Effects of the introduction of the consumption tax in April 1989 and the tax raise in April 1997 on the CPI have been adjusted using the X12-ARIMA program. We employ the I(1) specification of the price index for Japan rather than the I(2) specification which has been used for the U.S. by Stock and Watson (1999).

out-of-sample forecast periods are 1991:1 to 2000:12. The forecasting performance of the models to be examined are as follows. The first model (DI-AR-LAG) includes multiple factors, lags of the factors and lags of y_t with a number of factors and both lag orders selected by BIC (maximum numbers for each are 12, 3 and 6, respectively). The second model (DI-AR) excludes lags of the factors from the DI-AR-LAG model. The third model (DI) includes only current factors. In addition to the basic method of estimating the factors described in the previous subsections, we also employed the principal components method applied to a stacked data matrix for the purpose of allowing lagged f_t^0 in (1). The performance of the forecasts based on the stacked factor model with lags of y_t (DIS-AR) and without lags of y_t (DIS) is also investigated.

The forecasts from all five models above ($y_{t+h|t}$) are compared with the forecast from the benchmark univariate AR model ($y_{t+h|t}^{AR}$) where the lag order is selected using BIC. For each model, forecasting performance is evaluated based on the relative mean squared error (MSE),

$$\frac{MSE(y_{t+h|t})}{MSE(y_{t+h|t}^{AR})}$$

and the estimated coefficient on the candidate forecast from the forecast combining regression

$$y_{t+h} = \alpha y_{t+h|t} + (1-\alpha) y_{t+h|t}^{AR} + \eta_{t+h}. \quad (3)$$

Tables 1 and 2 show the results on the 1-year (12-month) ahead forecast and 6-month ahead forecast, respectively. The forecasts based on factor-diffusion index models perform very well relative to the AR benchmark. In many cases, a 20 to 30 percent reduction in MSE is obtained and the hypothesis of $\alpha = 1$ in the forecast combining regression is not rejected. This outcome is very encouraging giving the fact that major public and private research institutes failed to provide a satisfactory forecast of business cycles and prolonged recessions during the 1990s (see Fukuda and Onodera, 2001).

As a second application of the factor-diffusion index, we consider the possibility of using extracted information to aid in improving preliminary estimate of Japanese GDP, which has been criticized recently for its poor performance. During the 1990s, the nominal and real growth rates of Japanese GDP dropped rapidly and became close to zero in the late 1990s. Unlike the high economic growth period with a ten percent growth rate, relatively small revision may result in flipping the sign of growth rate during the zero growth period. In effect, the accuracy of the preliminary GDP estimates has become very important issue and the Cabinet Office has been criticized for the magnitude of revision in GDP. The Cabinet Office periodically revises the official GDP. The first preliminary estimate (Y1) of GDP for a given quarter is released approximately seventy days after the end of the quarter. The second preliminary estimate (Y2) is released approximately 100 days after the end of the quarter. This revision is followed by further revisions, the annual revision (Y3) at the end of the next year and the second annual revision (Y4) at the end of the following year. In addition to these four estimates, we consider the latest available estimate at the time of this study and refer it as the final (Y5). Table 3 shows the descriptive statistics of GDP revisions in Japan.[5] For both nominal and real GDP, the decreasing standard deviations of the revisions imply inaccuracy of the early estimates.

Mankiw and Shapiro (1986), in their study of GNP revisions in the U.S., argued that the early estimates could be considered as forecasts of revised estimates. If we interpret the preliminary estimates Y1 and Y2, as forecasts of Y5, we may be able to improve the accuracy of the forecast by using additional information available at the time of releasing Y1 or Y2. Along this line of approach, we consider the possibility of using the information contained in the common factor, or the factor-diffusion index, for the purpose of improving the GDP estimates and further understanding of current economic condition.

Since GDP is quarterly data, we consider three different models to obtain a forecast of Y5 using diffusion index. The first model (DI-M) only uses the monthly diffusion index available at the timing of release. The second model (DI-Q) uses a quarterly sum of monthly diffusion indexes, and the third model (DI-Q2) uses three monthly diffusion indexes separately, in each forecasting regression. Since the timing of release of Y5 is not fixed, we consider in-sample forecast only. Table 4 presents the results in terms of the coefficient on factor-diffusion index forecast in the forecasting combining regression similar to (3). For all cases, the coefficients are close to 0.5 and are significantly different from zero, implying the possibilities of usefulness information in the factor-diffusion index on forecasting GDP revisions. While further analysis is required, we expect the forecast combination method (of preliminary estimates and the factor estimates) to improve the accuracy of GDP estimates.

3 Nonlinear Analysis of Business Cycles

3.1 Two-Step Estimation of Factor Structure

In the previous section, the factor-diffusion index forecast for Japan leads to a considerable reduction in forecast mean square errors compared to those from benchmark autoregressive models. Therefore, the evidence from Japan, along with the results from the U.S., confirms the usefulness of the factor-diffusion index. The empirical success of diffusion index forecasts both in the U.S. and Japan suggests that the driving forces of business cycles responsible for covariation of the macroeconomic variables are well captured by a small set of estimated factors. In other words, the generating mechanism of the common factors offers important information about the characteristics and sources of business cycles. One advantage of the principal components method is that it allows a very general class of the dynamic structure of the factors, because of the nonparametric nature of the approach.

[5]The data is kindly provided by Yasuyuki Komaki.

In this section, we focus on the dynamic structure of the factor rather than using the factor for the forecasting purpose. We estimate the factor dynamics by means of a two-step estimation. In the first step, we estimate the factor using the principal components method that is employed in the previous section. In the second step, we then use an estimated factor to estimate the various models. For example, the persistence of business cycles can be studied by the impulse responses or half-lives obtained from the second-step estimation of the traditional linear models using estimated factors. Suppose the factor dynamic structure to be a linear AR(1) model,

$$f_t^0 = \phi f_{t-1}^0 + \varepsilon_t \tag{4}$$

where $E\left[\varepsilon_t | f_{t-1}^0\right] = 0$ and $E\left[\varepsilon_t^2 | f_{t-1}^0\right] = \sigma^2$. If f_t^0 is available, the model can be estimated by OLS,

$$\widehat{\phi} = \left(\sum_{t=1}^{T} \left(f_{t-1}^0\right)^2\right)^{-1} \sum_{t=1}^{T} f_{t-1}^0 f_t^0.$$

However, since f_t^0 is not observable, we replace f_t^0 by $\widetilde{f}_t$ and the feasible estimator is

$$\widetilde{\phi} = \left(\sum_{t=1}^{T} \left(\widetilde{f}_{t-1}\right)^2\right)^{-1} \sum_{t=1}^{T} \widetilde{f}_{t-1} \widetilde{f}_t.$$

Alternatively, a nonlinear process can be considered and estimated to incorporate more complex business cycle characteristics. A nonlinear model can be parametrically specified or nonparametrically estimated depending on the purpose of the analysis. Suppose we take the latter approach and consider general nonlinear AR(1) model without specifying the functional form,

$$f_t^0 = m\left(f_{t-1}^0\right) + \varepsilon_t \tag{5}$$

where $m(f)$ is a function. If f_t^0 is observable, we can estimate the function $m(f)$ using a Nadaraya-Watson type nonparametric regression estimator,

$$\widehat{m}(f) = \sum_{t=1}^{T} f_t^0 K\left(\frac{f_{t-1}^0 - f}{h}\right) \Big/ \sum_{t=1}^{T} K\left(\frac{f_{t-1}^0 - f}{h}\right)$$

where K is a kernel function and h is bandwidth satisfying $h \to 0$ and $Th \to \infty$. However, the feasible two-step estimator is

$$\widetilde{m}(f) = \sum_{t=1}^{T} \widetilde{f}_t K\left(\frac{\widetilde{f}_{t-1} - f}{h}\right) \Big/ \sum_{t=1}^{T} K\left(\frac{\widetilde{f}_{t-1} - f}{h}\right).$$

The remaining issue is to see if we can justify replacing the unobservable latent factors with the estimated factors in the second-step estimation. Below, we investigate the validity of such a two-step estimation methods. We first employ the following assumptions.

Assumption F (factors): (i) $E\left(f_t^0\right) = 0$, $E\left((f_t^0)^2\right) = \Sigma_F = 1$, $E\left((f_t^0)^4\right) < \infty$, and (ii)

$$\sqrt{T}\left(F^{0\prime}F^0/T - \Sigma_F\right) = O_p\left(1\right).$$

where $F^0 = [f_1^0, \cdots, f_T^0]'$.

Assumption FL (factor loadings): (i) $\left|\lambda_i^0\right| \leq \overline{\lambda} < \infty$, and (ii)

$$\sqrt{N}\left(\Lambda^{0\prime}\Lambda^0/N - \Sigma_\Lambda\right) = O_p\left(1\right).$$

where $\Lambda^0 = \left[\lambda_1^0, \cdots, \lambda_N^0\right]'$.

Assumption E (errors): For some positive constant $\kappa < \infty$, such that for all N and T,

(i) $E\left(e_{it}\right) = 0$, $E\left(e_{it}{}^8\right) \leq \kappa$,
(ii) $\sum_{s=1}^T \left|\gamma_N(s,t)\right| \leq M$ for all t, where $\gamma_N(s,t) = E\left(e_s' e_t/N\right) = E(N^{-1}\sum_{i=1}^N e_{is}e_{it})$,
(iii) $E(e_{it}e_{jt}) = \tau_{ij,t}$ where $|\tau_{ij,t}| \leq |\tau_{ij,t}|$ for all t and $\sum_{i=1}^N |\tau_{ij}| \leq \kappa$ for all j,
(iv) $E(e_{is}e_{jt}) = \tau_{ij,st}$ where $\sum_{i=1}^N \sum_{j=1}^N \sum_{s=1}^T \sum_{t=1}^T |\tau_{ij,st}| \leq \kappa$,
(v) For all (s,t), $E\left|N^{-1/2}\sum_{i=1}^N \left[e_{is}e_{it} - E(e_{is}e_{it})\right]\right|^4 \leq \kappa$,
(vi) $E\left((TN)^{-1}\sum_{i=1}^N \left(\sum_{t=1}^T f_t^0 e_{it}\right)^2\right) \leq \kappa$, and
(vii) $E\left|(TN)^{-1/2}\sum_{t=1}^T \sum_{k=1}^N f_t^0 \lambda_k^0 e_{kt}\right|^2 \leq \kappa$.

These assumptions are standard and employed by Stock and Watson (1998), Bai (2001) and Bai and Ng (2002). The main results of this subsection are summarized in the following Propositions.

Proposition 1. *Let x_{it} and f_t^0 be generated from (1) and (4), respectively, and suppose that assumptions F, FL and E are satisfied. Then, $\widetilde{\phi} - \phi = O_p(T^{-1/2})$ if $N/T \to \kappa$, and $\sqrt{T}\left(\widetilde{\phi} - \phi\right) \xrightarrow{d} N(0, 1-\phi^2)$ if $N/T \to \infty$.*

Proposition 2. *Let x_{it} and f_t^0 be generated from (1) and (5), respectively, and suppose that assumptions F, FL and E are satisfied. Then $\widehat{m}\left(f\right) - \widetilde{m}\left(f\right) = o_p(1)$ for all f if $N/T \to \infty$ and $h = \kappa T^{-\alpha}$ where $0 < \alpha < 1/4$.*

Proofs are given in Appendix A. The results above provide the conditions for the two-step estimation to be valid. When the number of the series (N) increases at a sufficiently fast rate compared to the time-series observations (T), the effect of the estimation error of the factor becomes negligible in the limit distribution of the second-step estimators. While we prove only for the case of autoregressive models of order one, we can generalize the results for the case of higher order models. Furthermore, these asymptotic results can be also used for the hypothesis testing regarding the specification of the dynamic factor structure.

3.2 Testing for Linearity

In the applications of factor models in macroeconomics, the linear model is often employed for the dynamic structure of the factor. In this subsection, we investigate the plausibility of linear assumption in the business cycles of the U.S. and Japan using the estimated factor-diffusion index. A number of tests for linearity (or neglected nonlinearity) have been proposed in the specification testing literature. We employ three different tests for the null hypothesis of $m(f) = \phi f$ for some ϕ, which have power against a wide range of nonlinear alternatives. All the tests can be viewed as a type of conditional moment test with a moment restriction being $E\left[\varepsilon_t | f_{t-1}^0\right] = 0$ where $\varepsilon_t = f_t^0 - \phi f_{t-1}^0$.

The first test is the kernel-based consistent specification test proposed by Zheng (1996) and extended to linear time-series case by Fan and Li (1997). It utilizes the kernel estimation of the moment condition of the form $E\left[\varepsilon_t E\left(\varepsilon_t | f_{t-1}^0\right) g\left(f_{t-1}^0\right)\right] = 0$ where $g(f)$ is the probability density function. With appropriate standardization, the test statistic follows asymptotically normal under the null hypothesis of linear specification.

The second test we employ is the neural network test proposed by White (1989). The test utilizes the moment condition $E[\varepsilon_t \Psi_t] = 0$ where $\Psi_t = \left(\psi\left(\gamma_1 f_{t-1}^0\right), \ldots, \psi\left(\gamma_q f_{t-1}^0\right)\right)'$ is a $q \times 1$ vector of activation functions ψ. The test statistics can be simply computed using T times R^2 of the regression of the residuals on f_{t-1}^0 and activation functions Ψ_t with the coefficients γ_j's being randomly drawn. In this paper, we follow a suggestion by Lee, White and Granger (1993) and include only the second and third principal components of Ψ_t with $q = 10$ hidden units in the auxiliary regression to avoid collinearity of f_{t-1}^0 and Ψ_t. The test statistic follows χ^2 distribution with two degrees of freedom under the null hypothesis of linearity.

One drawback of the White neural network test is the unidentifiability of γ_j's under the null hypothesis. Instead of using random γ_j's, Teräsvirta, Lin and Granger (1993) replaced the activation functions by their Volterra expansion up to the third order under the null. The third test we employ is this LM type neural network test based on the significance of coefficients on cubic terms in the auxiliary regression.

Table 5 shows the results of all three tests of linearity, with autoregressive orders ranging from one to four. Based on the kernel-based tests, the linear hypothesis of factor-diffusion index (DI1) is significantly rejected for half the time for both the U.S. and Japan. On the other hand, the White neural network test and LM type test both reject the linearity hypothesis for all the cases at the conventional significance level. In addition to DI1, the test is also conducted for the second and third principal components, DI2 and DI3. In many cases, the linearity is rejected for these additional indexes, while the evidence of nonlinearity in DI2 is not as strong as DI1 and DI3.

3.3 Stability Analysis Based on the Lyapunov Exponent

Understanding the structure of business cycles has been one of the most important objectives in of macroeconomics. Traditionally, we treat expansions and contractions of the economy as a result of exogenous random shocks explained by the change in policy, change in demand, technological change and other supply shocks. An alternative view is to consider the endogenous aggregate fluctuation via a chaotic system, or a simple nonlinear deterministic system that can have stochastic-like unpredictable behavior. It should be noted that nonlinearity is allowed in the first view but not necessary as opposed to the second view. For both views, the source of the fluctuation can be summarized in the small number of common components, and thus the dynamic factor model is expected to provide useful information in distinguishing chaos from exogenous fluctuation in the economy.

To provide empirical evidence regarding the two competing views of the business cycle, we compute a stability measure called Lyapunov exponent based on the factor-diffusion index. For the nonlinear AR model of order one, the Lyapunov exponent of the system is defined as

$$\lambda \equiv \lim_{M \to \infty} M^{-1} \sum_{t=1}^{M} \ln \left| Dm(f_{t-1}^{0}) \right|$$

where $Dm(f)$ is the first derivative of the function $m(f)$. A chaotic system has a positive Lyapunov exponent while a exogenous system with a unique and globally stable steady state has a negative Lyapunov exponent. We estimate the nonlinear AR model using both kernel-based method (local quadratic smoother) and neural network method and employ the sample analogue estimator of λ based on the estimated function $\widetilde{m}\,(f)$. Hypothesis testing regarding the stability of the system can be conducted using the standard error of the Lyapunov exponent (see Shintani and Linton, 2001, for this procedure in detail).

We point out that using the common factors rather than individual series in this type of test is advantageous for the following reason. If the true system consists of N equations, we require $2N+1$ lags in nonlinear AR model based on a single series for this method to be valid. However, in general, nonparametric estimation of such a high dimensional model involves computational difficulties. The common factor approach, by construction, achieves dimensional reduction, akin to the single index model in the microeconometrics literature.

Table 6 shows the Lyapunov exponent estimates of the factor-diffusion index based on the nonlinear AR model of order from one to four. Both full sample estimates

($M = T$) and block estimates ($M < T$) are presented. For all cases, the Lyapunov exponents of the factor-diffusion index (DI1) are significantly negative, implying the evidence against chaotic explanation of business cycle. The additional indexes, DI2 and DI3, based on the second and third principal components, respectively, also provide the evidence against chaos while DI2 has a larger exponent in comparison with the other two indexes. It is interesting to note that very similar results are obtained for both U.S. and Japan. The negative Lyapunov exponents for both the U.S. and Japan can be considered as an empirical justification of the impulse response analysis that is commonly used among macroeconomists since it requires the assumption of the exogenous shocks and a stable steady state in the system.

4 Conclusion

This paper first constructed a business cycle index of Japanese economic activity based on a dynamic factor model with a large data set, using a principal components method employed by Stock and Watson (1998) in their analysis of the U.S. diffusion index. As in the U.S., the factor-diffusion index in Japan is found to be useful in the context of out-of-sample forecasting. Secondly, the business cycle characteristics of the U.S. and Japan are further investigated by the two-step estimation of the dynamic factor structure. The evidence suggests the possibility of nonlinearity in both the U.S. and Japan while it excludes the class of nonlinearity that can generate endogenous fluctuation or chaos.

There are several remaining issues for further analysis. First, it should be noted that we only considered linear forecasting in the first half of the paper, while we found some evidence of nonlinearity in the factor structure in the latter half. Therefore, there may be some gain from employing nonlinear model in forecasting. Secondly, by construction, the principal components method extracts factors by the linear transformation of the data. The index based on the nonlinear transformation of the data may be considered as an alternative to the index employed in the paper.

Appendix A: Proofs

Proof of Proposition.1.

The principal components estimator $\widetilde{F} = \left[\widetilde{f}_1, \cdots, \widetilde{f}_T\right]'$ is the first eigenvector of the $T \times T$ matrix XX' with normalization $T^{-1}\sum_{t=1}^{T}\widetilde{f}_t^2 = 1.$, where

$$X = \begin{bmatrix} X_1' \\ \vdots \\ X_T' \end{bmatrix} = \begin{bmatrix} x_{11} & \cdots & x_{N1} \\ \vdots & \ddots & \vdots \\ x_{1T} & \cdots & x_{NT} \end{bmatrix}.$$

Following Bai (2001) and Bai and Ng (2002), our results are based on the result with its linear transformation

$$\widehat{F} = \left(\frac{1}{TN}XX'\right)\widetilde{F} = \widetilde{F}v_{NT} \tag{A.1}$$

or

$$\widehat{f}_t = v_{NT}\widetilde{f}_t$$

where v_{NT} is the largest eigenvalue of XX'/TN. From Bai's (2001) Lemma A.2,

$$\widehat{f}_t - H_{NT}f_t^0 = O_p(T^{-1/2}C_{NT}^{-1}) + O_p(N^{-1/2}C_{NT}^{-1}) + O_p(N^{-1/2})$$

where $H_{NT} = (\widetilde{F}'F^0/T)(\Lambda^{0\prime}\Lambda^0/N)$ and $C_{NT} = \min(\sqrt{T}, \sqrt{N})$. This implies $\widehat{f}_t - H_{NT}f_t^0 = O_p(T^{-1/2})$ if $N/T \to \kappa$ and $\widehat{f}_t - H_{NT}f_t^0 = o_p(T^{-1/2})$ if $N/T \to \infty$. Let $\widetilde{H}_{NT} = H_{NT}/v_{NT}$. Then since $v \equiv \operatorname*{plim}_{T,N\to\infty} v_{NT} = \Sigma_\Lambda\Sigma_F = \Sigma_\Lambda$ from Bai's (2001) Lemma A.3, and $H \equiv \operatorname*{plim}_{T,N\to\infty} H_{NT} = \Sigma_\Lambda\Sigma_F^{1/2} = \Sigma_\Lambda$, we have $\widetilde{H} \equiv \operatorname*{plim}_{T,N\to\infty} \widetilde{H}_{NT} = \Sigma_\Lambda\Sigma_F^{1/2}/\Sigma_\Lambda\Sigma_F = \Sigma_F^{-1/2} = 1.$

$$\begin{aligned}
\sqrt{T}\left(\widetilde{\phi} - \phi\right) &= \sqrt{T}\left(\sum_{t=1}^{T}\left(\widetilde{f}_{t-1}\right)^2\right)^{-1}\sum_{t=1}^{T}\widetilde{f}_{t-1}\left(\widetilde{f}_t - \phi\widetilde{f}_{t-1}\right) \\
&= \sqrt{T}\left(\sum_{t=1}^{T}\left(\widetilde{f}_t\right)^2 - \left(\widetilde{f}_T\right)^2\right)^{-1}\sum_{t=1}^{T}\widetilde{f}_{t-1}\left(\widetilde{f}_t - \phi\widetilde{f}_{t-1}\right) \\
&= T^{-1/2}\sum_{t=1}^{T}\widetilde{f}_{t-1}\left(\widetilde{f}_t - \phi\widetilde{f}_{t-1}\right) + o_p(1) \\
&= T^{-1/2}\sum_{t=1}^{T}\widetilde{f}_{t-1}\left\{\left(\widetilde{f}_t - \widetilde{H}_{NT}f_t^0\right) - \phi\left(\widetilde{f}_{t-1} - \widetilde{H}_{NT}f_{t-1}^0\right)\right\} \\
&\quad + T^{-1/2}\widetilde{H}_{NT}\sum_{t=1}^{T}\widetilde{f}_{t-1}\varepsilon_t + o_p(1) \\
&= T^{-1/2}\sum_{t=1}^{T}\widetilde{f}_{t-1}\left(\widetilde{f}_t - \widetilde{H}_{NT}f_t^0\right) - T^{-1/2}\phi\sum_{t=1}^{T}\widetilde{f}_{t-1}\left(\widetilde{f}_{t-1} - \widetilde{H}_{NT}f_{t-1}^0\right) \\
&\quad + T^{-1/2}\widetilde{H}_{NT}\sum_{t=1}^{T}\left(\widetilde{f}_{t-1} - \widetilde{H}_{NT}f_{t-1}^0\right)\varepsilon_t + T^{-1/2}\widetilde{H}_{NT}^2\sum_{t=1}^{T}f_{t-1}^0\varepsilon_t + o_p(1) \\
&= a_T + b_T + c_T + \sqrt{T}\left(\widehat{\phi} - \phi\right) + o_p(1), \text{ say.}
\end{aligned}$$

For a_T, we have

$$
\begin{aligned}
a_T &= \frac{1}{\sqrt{T}}\sum_{t=1}^{T}\widetilde{f}_{t-1}\left(\widetilde{f}_t-\widetilde{H}_{NT}f_t^0\right)\\
&= \frac{1}{\sqrt{T}}v_{NT}^{-1}\sum_{t=1}^{T}\widetilde{f}_{t-1}\left(\widehat{f}_t-H_{NT}f_t^0\right)\\
&\le v_{NT}^{-1}\left(\sum_{t=1}^{T}\left(\widetilde{f}_{t-1}\right)^2\right)^{1/2}\left(\frac{1}{T}\sum_{t=1}^{T}\left(\widehat{f}_t-H_{NT}f_t^0\right)^2\right)^{1/2}\\
&= \begin{cases} O_p(T^{1/2})\times O_p(T^{-1/2})=O_p(1) & \text{if } N/T\to\kappa,\\ O_p(T^{1/2})\times o_p(T^{-1/2})=o_p(1) & \text{if } N/T\to\infty.\end{cases}
\end{aligned}
$$

Similarly, for b_T,

$$
b_T=\begin{cases} O_p(1) & \text{if } N/T\to\kappa,\\ o_p(1) & \text{if } N/T\to\infty.\end{cases}
$$

For c_T,

$$
\begin{aligned}
c_T &= \frac{1}{\sqrt{T}}\widetilde{H}_{NT}\sum_{t=1}^{T}\left(\widetilde{f}_{t-1}-\widetilde{H}_{NT}f_{t-1}^0\right)\varepsilon_t\\
&= \frac{1}{\sqrt{T}}\widetilde{H}_{NT}v_{NT}^{-1}\sum_{t=1}^{T}\left(\widehat{f}_{t-1}-H_{NT}f_{t-1}^0\right)\varepsilon_t\\
&\le \sqrt{T}\widetilde{H}_{NT}v_{NT}^{-1}\left(\frac{1}{T}\sum_{t=1}^{T}\varepsilon_t^2\right)^{1/2}\left(\frac{1}{T}\sum_{t=1}^{T}\left(\widehat{f}_{t-1}-H_{NT}f_{t-1}^0\right)^2\right)^{1/2}\\
&= \begin{cases} \sqrt{T}\times O_p(1)\times O_p(T^{-1/2})=O_p(1) & \text{if } N/T\to\kappa,\\ \sqrt{T}\times O_p(1)\times o_p(T^{-1/2})=o_p(1) & \text{if } N/T\to\infty.\end{cases}
\end{aligned}
$$

Therefore, $\sqrt{T}\left(\widetilde{\phi}-\phi\right)-\sqrt{T}\left(\widehat{\phi}-\phi\right)$ is $O_p(1)$ if $N/T\to\kappa$, and $o_p(1)$ if $N/T\to\infty$.
■

The proof of Proposition 1 uses the following lemma.

Lemma A.1. *Suppose assumptions F, FL and E are satisfied.. Then,*

$$
\widetilde{H}_{NT}-\widetilde{H}=O_p(C_{NT}^{-1})
$$

where $C_{NT}=\min(\sqrt{T},\sqrt{N})$.

Proof of Lemma A.1.

$$
\begin{aligned}
H_{NT}-H &= (\widetilde{F}'F^0/T)(\Lambda^{0\prime}\Lambda^0/N)-\Sigma_\Lambda\Sigma_F^{1/2}\\
&= (\widetilde{F}'F^0/T)\left[(\Lambda^{0\prime}\Lambda^0/N)-\Sigma_\Lambda\right]\\
&\quad+\left[(\widetilde{F}'F^0/T)-\Sigma_F^{1/2}\right]\Sigma_\Lambda.
\end{aligned}
$$

The first term is $O_p(N^{-1/2})$ from the assumption FL. To obtain the rate of $(\widetilde{F}'F^0/T) - \Sigma_F^{1/2}$, we first derive the rate of $v_{NT} - v$.

Multiplying (A.1) by $T^{-1}\widetilde{F}'$ yields

$$v_{NT} = \frac{1}{T}\widetilde{F}'\left(\frac{1}{TN}XX'\right)\widetilde{F}$$

and expanding XX' with $X = F^0\Lambda^{0\prime} + e$ where

$$e = \begin{bmatrix} e_1' \\ \vdots \\ e_T' \end{bmatrix} = \begin{bmatrix} e_{11} & \cdots & e_{N1} \\ \vdots & \ddots & \vdots \\ e_{1T} & \cdots & e_{NT} \end{bmatrix}$$

yields

$$\begin{aligned}
& \frac{1}{T}\widetilde{F}'\left(\frac{1}{TN}XX'\right)\widetilde{F} \\
= & \frac{1}{T}\widetilde{F}'\left(\frac{1}{TN}\left(F^0\Lambda^{0\prime} + e\right)\left(\Lambda^0 F^{0\prime} + e'\right)\right)\widetilde{F} \\
= & \frac{1}{T}\widetilde{F}'\left(\frac{1}{TN}\left(F^0\Lambda^{0\prime}\Lambda^0F^{0\prime} + F^0\Lambda^{0\prime}e' + e\Lambda^0F^{0\prime} + ee'\right)\right)\widetilde{F} \\
= & \left(\frac{\widetilde{F}'F^0}{T}\right)\left(\frac{\Lambda^{0\prime}\Lambda^0}{N}\right)\left(\frac{F^{0\prime}\widetilde{F}}{T}\right) + \frac{2}{T}\widetilde{F}'F^0\Lambda^{0\prime}e'\widetilde{F}/(TN) + \frac{1}{TN}\widetilde{F}'ee'\widetilde{F}/T \\
= & \left(\frac{\widetilde{F}'F^0}{T}\right)\left(\frac{\Lambda^{0\prime}\Lambda^0}{N}\right)\left(\frac{F^{0\prime}\widetilde{F}}{T}\right) + 2T^{-2}\sum_{t=1}^{T}\sum_{s=1}^{T}\widetilde{f}_s\xi_{st}\widetilde{f}_t \\
& + T^{-2}\sum_{t=1}^{T}\sum_{s=1}^{T}\widetilde{f}_s\gamma_N(s,t)\widetilde{f}_t + T^{-2}\sum_{t=1}^{T}\sum_{s=1}^{T}\widetilde{f}_s\zeta_{st}\widetilde{f}_t
\end{aligned}$$

where $\xi_{st} = f_t^0\Lambda^{0\prime}e_s/N$ and $\zeta_{st} = e_s'e_t/N - \gamma_N(s,t)$. For the second term,

$$\begin{aligned}
T^{-2}\sum_{t=1}^{T}\sum_{s=1}^{T}\widetilde{f}_s\xi_{st}\widetilde{f}_t &= T^{-1}\sum_{t=1}^{T}\left(T^{-1}\sum_{s=1}^{T}\widetilde{f}_s\xi_{st}\right)\widetilde{f}_t \\
&\leq \left(T^{-1}\sum_{t=1}^{T}\widetilde{f}_t^2\right)^{1/2}\left(T^{-1}\sum_{t=1}^{T}\left(T^{-1}\sum_{s=1}^{T}\widetilde{f}_s\xi_{st}\right)^2\right)^{1/2} \\
&= O_p\left(\frac{1}{\sqrt{N}C_{NT}}\right)
\end{aligned}$$

since $T^{-1}\sum_{t=1}^{T}\widetilde{f}_t^2 = 1$ and $T^{-1}\sum_{s=1}^{T}\widetilde{f}_s\xi_{st} = O_p\left(\frac{1}{\sqrt{N}C_{NT}}\right)$ from Lemma A.2(d) in Bai (2001). Similarly, for the third and fourth terms,

$$\begin{aligned}
T^{-2}\sum_{t=1}^{T}\sum_{s=1}^{T}\widetilde{f}_s\gamma_N(s,t)\widetilde{f}_t &= T^{-1}\sum_{t=1}^{T}\left(T^{-1}\sum_{s=1}^{T}\widetilde{f}_s\gamma_N(s,t)\right)\widetilde{f}_t = O_p\left(\frac{1}{\sqrt{T}C_{NT}}\right) \\
T^{-2}\sum_{t=1}^{T}\sum_{s=1}^{T}\widetilde{f}_s\zeta_{st}\widetilde{f}_t &= T^{-1}\sum_{t=1}^{T}\left(T^{-1}\sum_{s=1}^{T}\widetilde{f}_s\zeta_{st}\right)\widetilde{f}_t = O_p\left(\frac{1}{\sqrt{N}C_{NT}}\right)
\end{aligned}$$

since $T^{-1}\sum_{s=1}^{T}\widetilde{f}_s\gamma_N(s,t) = O_p\left(\frac{1}{\sqrt{T}C_{NT}}\right)$ and $T^{-1}\sum_{s=1}^{T}\widetilde{f}_s\zeta_{st} = O_p\left(\frac{1}{\sqrt{N}C_{NT}}\right)$ from Lemma A.2(a) and (b) in Bai (2001). Therefore,

$$v_{NT} = \left(\frac{\widetilde{F}'F^0}{T}\right)\left(\frac{\Lambda^{0\prime}\Lambda^0}{N}\right)\left(\frac{F^{0\prime}\widetilde{F}}{T}\right) + O_p\left(\frac{1}{C_{NT}^2}\right)$$

Next, multiplying (A.1) by $T^{-1}F^{0\prime}$ yields

$$\frac{1}{T}F^{0\prime}\left(\frac{1}{TN}XX'\right)\widetilde{F} = \left(\frac{F^{0\prime}\widetilde{F}}{T}\right)v_{NT}$$

and expanding XX' with $X = F^0\Lambda^{0\prime} + e$ yields

$$\begin{aligned}
&\frac{1}{T}F^{0\prime}\left(\frac{1}{TN}XX'\right)\widetilde{F} \\
= &\frac{1}{T}F^{0\prime}\left(\frac{1}{TN}\left(F^0\Lambda^{0\prime}+e\right)\left(\Lambda^0F^{0\prime}+e'\right)\right)\widetilde{F} \\
= &\frac{1}{T}F^{0\prime}\left(\frac{1}{TN}\left(F^0\Lambda^{0\prime}\Lambda^0F^{0\prime}+F^0\Lambda^{0\prime}e'+e\Lambda^0F^{0\prime}+ee'\right)\right)\widetilde{F} \\
= &\left(\frac{F^{0\prime}F^0}{T}\right)\left(\frac{\Lambda^{0\prime}\Lambda^0}{N}\right)\left(\frac{F^{0\prime}\widetilde{F}}{T}\right) \\
&+\left(\frac{F^{0\prime}F^0}{T}\right)\Lambda^{0\prime}e'\widetilde{F}/(TN) + \frac{1}{TN}F^{0\prime}e\Lambda^0\left(\frac{F^{0\prime}\widetilde{F}}{T}\right) + \frac{1}{TN}F^{0\prime}ee'\widetilde{F}/T
\end{aligned}$$

and

$$\begin{aligned}
v_{NT} = &\left(\frac{F^{0\prime}F^0}{T}\right)\left(\frac{\Lambda^{0\prime}\Lambda^0}{N}\right) + \frac{1}{TN}F^{0\prime}e\Lambda^0 \\
&+\left(\frac{F^{0\prime}\widetilde{F}}{T}\right)^{-1}\left(\frac{F^{0\prime}F^0}{T}\right)\Lambda^{0\prime}e'\widetilde{F}/(TN) + \left(\frac{F^{0\prime}\widetilde{F}}{T}\right)^{-1}\frac{1}{TN}F^{0\prime}ee'\widetilde{F}/T.
\end{aligned}$$

For the first term, from the assumptions F and FL, we have

$$\left(\frac{F^{0\prime}F^0}{T}\right)\left(\frac{\Lambda^{0\prime}\Lambda^0}{N}\right) - v = O_p\left(\frac{1}{C_{NT}}\right).$$

For the second term,

$$\frac{1}{TN}F^{0\prime}e\Lambda^0 = \frac{1}{TN}\sum_{t=1}^{T}\sum_{k=1}^{N}f_t^0\lambda_k^0e_{kt} = O_p\left(\frac{1}{\sqrt{NT}}\right)$$

from assumption E. For the third term, since

$$\begin{aligned}
\frac{1}{TN}\widetilde{F}'e\Lambda^0 - \widetilde{H}_{NT}\left(\frac{1}{TN}F^{0\prime}e\Lambda^0\right) &= \frac{1}{TN}\left(\widetilde{F}-\widetilde{H}_{NT}F^0\right)'e\Lambda^0 \\
&= \frac{1}{TN}\sum_{t=1}^{T}\sum_{k=1}^{N}\left(\widetilde{f}_t-\widetilde{H}_{NT}f_t^0\right)\lambda_k^0e_{kt} \\
&\le \frac{1}{\sqrt{N}}\left(\frac{1}{T}\sum_{t=1}^{T}\left(\widetilde{f}_t-\widetilde{H}_{NT}f_t^0\right)^2\right)^{1/2}\left(\frac{1}{T}\sum_{t=1}^{T}\sum_{k=1}^{N}\left(\frac{1}{\sqrt{N}}\lambda_k^0e_{kt}\right)^2\right)^{1/2} \\
&= O_p\left(\frac{1}{\sqrt{NT}C_{NT}}\right) + O_p\left(\frac{1}{NC_{NT}}\right) + O_p\left(\frac{1}{N}\right),
\end{aligned}$$

from $\widetilde{f}_t - \widetilde{H}_{NT} f_t^0 = O_p(T^{-1/2}C_{NT}^{-1}) + O_p(N^{-1/2}C_{NT}^{-1}) + O_p(N^{-1/2})$ and the assumption E, we have

$$\frac{1}{TN}\widetilde{F}' e \Lambda^0 = O_p\left(\frac{1}{\sqrt{NT}C_{NT}}\right) + O_p\left(\frac{1}{N}\right).$$

For the fourth term,

$$\begin{aligned}
\frac{1}{TN}F^{0\prime}ee'\widetilde{F}/T &= T^{-2}\sum_{t=1}^{T}\sum_{s=1}^{T}\widetilde{f}_s\gamma_N(s,t)f_t^0 + T^{-2}\sum_{t=1}^{T}\sum_{s=1}^{T}\widetilde{f}_s\zeta_{st}f_t^0 \\
&= T^{-1}\sum_{t=1}^{T}\left(T^{-1}\sum_{s=1}^{T}\widetilde{f}_s\gamma_N(s,t)\right)f_t^0 + T^{-1}\sum_{t=1}^{T}\left(T^{-1}\sum_{s=1}^{T}\widetilde{f}_s\zeta_{st}\right)f_t^0 \\
&= O_p\left(\frac{1}{\sqrt{T}C_{NT}}\right) + O_p\left(\frac{1}{\sqrt{N}C_{NT}}\right) = O_p\left(\frac{1}{C_{NT}^2}\right).
\end{aligned}$$

Combining the results with $\left(F^{0\prime}\widetilde{F}/T\right)^{-1} = O_p(1)$ and $F^{0\prime}F^0/T = O_p(1)$ yields

$$v_{NT} - v = O_p(C_{NT}^{-1}).$$

This implies

$$\left(\frac{\widetilde{F}'F^0}{T}\right)\left(\frac{\Lambda^{0\prime}\Lambda^0}{N}\right)\left(\frac{F^{0\prime}\widetilde{F}}{T}\right) - v = O_p(C_{NT}^{-1})$$

and

$$(\widetilde{F}'F^0/T) - \Sigma_F^{1/2} = O_p(C_{NT}^{-1}).$$

Therefore,

$$\begin{aligned}
H_{NT} - H &= (\widetilde{F}'F^0/T)\left[(\Lambda^{0\prime}\Lambda^0/N) - \Sigma_\Lambda\right] \\
&\quad + \left[(\widetilde{F}'F^0/T) - \Sigma_F^{1/2}\right]\Sigma_\Lambda. \\
&= O_p(N^{-1/2}) + O_p(C_{NT}^{-1}) = O_p(C_{NT}^{-1}).
\end{aligned}$$

Combining $H_{NT} - H = O_p(C_{NT}^{-1})$ and $v_{NT} - v = O_p(C_{NT}^{-1})$ yields

$$\widetilde{H}_{NT} - \widetilde{H} = H_{NT}/v_{NT} - H/v = O_p(C_{NT}^{-1}).$$

■

Proof of Proposition.2.

$$\begin{aligned}
&\widetilde{m}(f) - \widehat{m}(f) \\
= &\left[\frac{1}{Th}\sum_{t=1}^{T}\widetilde{f}_t K\left(\frac{\widetilde{f}_{t-1} - f}{h}\right)\right] \big/ \widetilde{g}(f) - \left[\frac{1}{Th}\sum_{t=1}^{T}f_t^0 K\left(\frac{f_{t-1}^0 - f}{h}\right)\right] \big/ \widehat{g}(f) \\
= &\left\{\frac{1}{Th}\sum_{t=1}^{T}\widetilde{f}_t K\left(\frac{\widetilde{f}_{t-1} - f}{h}\right) - \frac{1}{Th}\sum_{t=1}^{T}f_t^0 K\left(\frac{f_{t-1}^0 - f}{h}\right)\right\} \big/ \widetilde{g}(f) \\
&+ \left(\frac{1}{\widetilde{g}(f)} - \frac{1}{\widehat{g}(f)}\right)\left[\frac{1}{Th}\sum_{t=1}^{T}f_t^0 K\left(\frac{f_{t-1}^0 - f}{h}\right)\right]
\end{aligned}$$

$$
\begin{aligned}
&= \left\{\frac{1}{Th}\sum_{t=1}^{T}\widetilde{f}_t K\left(\frac{f_{t-1}^0 - f}{h}\right) - \frac{1}{Th}\sum_{t=1}^{T} f_t^0 K\left(\frac{f_{t-1}^0 - f}{h}\right)\right\} \diagup \widetilde{g}(f) \\
&\quad + \left\{\frac{1}{Th}\sum_{t=1}^{T}\widetilde{f}_t K\left(\frac{\widetilde{f}_{t-1} - f}{h}\right) - \frac{1}{Th}\sum_{t=1}^{T}\widetilde{f}_t K\left(\frac{f_{t-1}^0 - f}{h}\right)\right\} \diagup \widetilde{g}(f) \\
&\quad + \left(\frac{1}{\widetilde{g}(f)} - \frac{1}{\widehat{g}(f)}\right)\left[\frac{1}{Th}\sum_{t=1}^{T} f_t^0 K\left(\frac{f_{t-1}^0 - f}{h}\right)\right] \\
&= a_T^* + b_T^* + c_T^*, \text{ say,}
\end{aligned}
$$

where

$$
\widehat{g}(f) = \frac{1}{Th}\sum_{t=1}^{T} K\left(\frac{f_{t-1}^0 - f}{h}\right) \quad \text{and} \quad \widetilde{g}(f) = \frac{1}{Th}\sum_{t=1}^{T} K\left(\frac{\widetilde{f}_{t-1} - f}{h}\right).
$$

For the denominator of a_T^*, we have

$$
\begin{aligned}
&\frac{1}{Th}\sum_{t=1}^{T}\left(\widetilde{f}_t - f_t^0\right) K\left(\frac{f_{t-1}^0 - f}{h}\right) \\
&= \frac{1}{Th}\sum_{t=1}^{T}\left(\widetilde{f}_t - \widetilde{H}_{NT} f_t^0\right) K\left(\frac{f_{t-1}^0 - f}{h}\right) + \frac{1}{Th}\left(\widetilde{H}_{NT} - \widetilde{H}\right)\sum_{t=1}^{T} f_t^0 K\left(\frac{f_{t-1}^0 - f}{h}\right).
\end{aligned}
$$

For the first term,

$$
\begin{aligned}
&\frac{1}{Th}\sum_{t=1}^{T}\left(\widetilde{f}_t - \widetilde{H}_{NT} f_t^0\right) K\left(\frac{f_{t-1}^0 - f}{h}\right) \\
&\leq T^{1/2}\left(\frac{1}{T}\sum_{t=1}^{T}\left(\widetilde{f}_t - \widetilde{H}_{NT} f_t^0\right)^2\right)^{1/2}\left(\frac{1}{Th}\sum_{t=1}^{T} K^2\left(\frac{f_{t-1}^0 - f}{h}\right)\right)^{1/2} \\
&= T^{1/2} \times o_p(T^{-1/2}) \times O_p(1) = o_p(1).
\end{aligned}
$$

For the second term,

$$
\begin{aligned}
&\left(\widetilde{H}_{NT} - \widetilde{H}\right)\frac{1}{Th}\sum_{t=1}^{T} f_t^0 K\left(\frac{f_{t-1}^0 - f}{h}\right) \\
&= O_p(T^{-1/2}) \times O_p(1) = o_p(1).
\end{aligned}
$$

For the denominator of b_T^*, by mean value theorem, we have

$$
\begin{aligned}
&\frac{1}{Th}\sum_{t=1}^{T}\widetilde{f}_t K\left(\frac{\widetilde{f}_{t-1} - f_{t-1}^0 + f_{t-1}^0 - f}{h}\right) - \frac{1}{Th}\sum_{t=1}^{T}\widetilde{f}_t K\left(\frac{f_{t-1}^0 - f}{h}\right) \\
&= \frac{1}{Th}\sum_{t=1}^{T}\widetilde{f}_t K'\left(\frac{\left(f_{t-1}^0 - f\right)^*}{h}\right)\left(\frac{\widetilde{f}_{t-1} - f_{t-1}^0}{h}\right) \\
&= \frac{1}{Th^2}\sum_{t=1}^{T}\widetilde{f}_t\left(\widetilde{f}_{t-1} - \widetilde{H}_{NT} f_{t-1}^0\right) K'\left(\frac{\left(f_{t-1}^0 - f\right)^*}{h}\right) \\
&\quad + \frac{1}{Th^2}\left(\widetilde{H}_{NT} - \widetilde{H}\right)\sum_{t=1}^{T}\widetilde{f}_t f_{t-1}^0 K'\left(\frac{\left(f_{t-1}^0 - f\right)^*}{h}\right)
\end{aligned}
$$

where $\left(f_{t-1}^0 - f\right)^*$ lies between $\widetilde{f}_{t-1} - f_{t-1}^0$ and $f_{t-1}^0 - f$.

For the first term,

$$\begin{aligned}
& \frac{1}{Th^2}\sum_{t=1}^{T}\widetilde{f}_t\left(\widetilde{f}_{t-1} - \widetilde{H}_{NT}f_{t-1}^0\right)K'\left(\frac{\left(f_{t-1}^0 - f\right)^*}{h}\right) \\
\leq\ & \frac{1}{h^2}\sup\left|K'(x)\right|\left(\frac{1}{T}\sum_{t=1}^{T}\left(\widetilde{f}_t\right)^2\right)^{1/2}\left(\frac{1}{T}\sum_{t=1}^{T}\left(\widetilde{f}_{t-1} - \widetilde{H}_{NT}f_{t-1}^0\right)^2\right)^{1/2} \\
=\ & \frac{1}{h^2}\times O_p(1)\times o_p(T^{-1/2}) = o_p(T^{2\alpha-1/2}) = o_p(1).
\end{aligned}$$

For the second term,

$$\begin{aligned}
& \frac{1}{Th^2}\left(\widetilde{H}_{NT} - \widetilde{H}\right)\sum_{t=1}^{T}\widetilde{f}_t f_{t-1}^0 K'\left(\frac{\left(f_{t-1}^0 - f\right)^*}{h}\right) \\
\leq\ & \frac{1}{h^2}\left(\widetilde{H}_{NT} - \widetilde{H}\right)\sup\left|K'(x)\right|\left(\frac{1}{T}\sum_{t=1}^{T}\widetilde{f}_t f_{t-1}^0\right) \\
=\ & \frac{1}{h^2}\times O_p(T^{-1/2})\times O_p(1) = O_p(T^{2\alpha-1/2}) = o_p(1).
\end{aligned}$$

Similarly, for c_T^*, by mean value theorem,

$$\begin{aligned}
\widetilde{g}(f) - \widehat{g}(f) &= \frac{1}{Th}\sum_{t=1}^{T}K\left(\frac{\widetilde{f}_{t-1} - f_{t-1}^0 + f_{t-1}^0 - f}{h}\right) - \frac{1}{Th}\sum_{t=1}^{T}K\left(\frac{f_{t-1}^0 - f}{h}\right) \\
&= \frac{1}{Th}\sum_{t=1}^{T}K'\left(\frac{\left(f_{t-1}^0 - f\right)^*}{h}\right)\left(\frac{\widetilde{f}_{t-1} - f_{t-1}^0}{h}\right) \\
&= \frac{1}{Th^2}\sum_{t=1}^{T}\left(\widetilde{f}_{t-1} - \widetilde{H}_{NT}f_{t-1}^0\right)K'\left(\frac{\left(f_{t-1}^0 - f\right)^*}{h}\right) \\
&\quad + \frac{1}{Th^2}\left(\widetilde{H}_{NT} - \widetilde{H}\right)\sum_{t=1}^{T}f_{t-1}^0 K'\left(\frac{\left(f_{t-1}^0 - f\right)^*}{h}\right) \\
&\leq \frac{1}{h^2}\sup\left|K'(x)\right|\left(\frac{1}{T}\sum_{t=1}^{T}\left(\widetilde{f}_{t-1} - \widetilde{H}_{NT}f_{t-1}^0\right)\right) \\
&\quad + \frac{1}{h^2}\left(\widetilde{H}_{NT} - \widetilde{H}\right)\sup\left|K'(x)\right|\left(\frac{1}{T}\sum_{t=1}^{T}f_{t-1}^0\right) \\
&= o_p(T^{2\alpha-1/2}) + O_p(T^{2\alpha-1/2}) = o_p(1).
\end{aligned}$$

Combining three results yield

$$\widetilde{m}(f) - \widehat{m}(f) = o_p(1).$$

■

Appendix B: Data

This appendix lists the series used to construct the diffusion index based on the factor model described in the main text.

Series Number	
	Real Output
1 (IP)	Index of Industrial Production (Mining and manufacturing)
2	Index of Industrial Production (Manufacturing)
3	Index of Industrial Production (Mining)
4	Index of Industrial Production (Iron and steel)
5	Index of Industrial Production (Non-ferrous metals)
6	Index of Industrial Production (Fabricated metals)
7	Index of Industrial Production (General machinery)
8	Index of Industrial Production (Electrical machinery)
9	Index of Industrial Production (Transport equipment)
10	Index of Industrial Production (Precision instruments)
11	Index of Industrial Production (Ceramics, clay and stone products)
12	Index of Industrial Production (Chemicals)
13	Index of Industrial Production (Petroleum and coal products)
14	Index of Industrial Production (Plastic products)
15	Index of Industrial Production (Pulp, paper and paper products)
16	Index of Industrial Production (Textiles)
17	Index of Industrial Production (Foods and tobacco)
18	Index of Industrial Production (Other manufacturing)
19	Index of Industrial Production (Final demand goods)
20	Index of Industrial Production (Producer goods)
21	Index of Industrial Production (Producer goods for mining and manufacturing)
22	Index of Industrial Production (Producer goods for others)
23	Index of Producer's Shipments (Final demand goods)
24 (SHIP)	Index of Producer's Shipments (Producer goods)
25	Index of Producer's Shipments (Producer goods for mining and manufacturing)
26	Index of Producer's Shipments (Producer goods for others)
27	Index of Raw Materials Consumption (Manufacturing)
28	Large Consumption of Electric Energy (Total)
29 (CAP)	Index of Capacity Utilization Ratio (Manufacturing)
30	Index of Capacity Utilization Ratio (Iron and steel)
31	Index of Capacity Utilization Ratio (Non-ferrous metals)
32	Index of Capacity Utilization Ratio (Fabricated metals)
33	Index of Capacity Utilization Ratio (General machinery)
34	Index of Capacity Utilization Ratio (Electrical machinery)
35	Index of Capacity Utilization Ratio (Transport equipment)
36	Index of Capacity Utilization Ratio (Precision instruments)
37	Index of Capacity Utilization Ratio (Ceramics, clay and stone products)
38	Index of Capacity Utilization Ratio (Chemicals)
39	Index of Capacity Utilization Ratio (Petroleum and coal products)
40	Index of Capacity Utilization Ratio (Textiles)
41	Index of Capacity Utilization Ratio (Rubber products)
42	Index of Capacity Utilization Ratio (Machinery)
43(SALE)	Index of Sales in Small and Medium Sized Enterprises (Manufacturing)
44	Index of Tertiary Industry Activity (Total)
45	Index of Tertiary Industry Activity (Electricity, gas, heat and water supply)
46	Index of Tertiary Industry Activity (Transport and Communication)
47	Index of Tertiary Industry Activity (Transport)
48	Index of Tertiary Industry Activity (Wholesale, retail trade, eating and drinking places)
49	Index of Tertiary Industry Activity (Eating and drinking places)
50	Index of Tertiary Industry Activity (Finance and insurance)
51	Index of Tertiary Industry Activity (Real estate)
52	Index of Tertiary Industry Activity (Services)
53	Index of Tertiary Industry Activity (Personal services)
54	Index of Tertiary Industry Activity (Business services)

Inventories

55 Index of Producer's Inventory Ratio of Finished Goods (Mining and manufacturing)
56 Index of Producer's Inventory Ratio of Finished Goods (Final demand goods)
57 Index of Producer's Inventory Ratio of Finished Goods (Investment goods)
58 Index of Producer's Inventory Ratio of Finished Goods (Capital goods)
59 Index of Producer's Inventory Ratio of Finished Goods (Construction goods)
60 Index of Producer's Inventory Ratio of Finished Goods (Consumer goods)
61 Index of Producer's Inventory Ratio of Finished Goods (Durable consumer goods)
62 Index of Producer's Inventory Ratio of Finished Goods (Non-durable consumer goods)
63 Index of Producer's Inventory Ratio of Finished Goods (Producer goods)
64 Index of Producer's Inventory Ratio of Finished Goods (Producer goods for mining and manufacturing)
65 Index of Producer's Inventory Ratio of Finished Goods (Producer goods for others)
66 Index of Raw Materials Inventory Ratio (Manufacturing)
67 Index of Producer's Inventory of Finished Goods (Mining and manufacturing)
68 Index of Producer's Inventory of Finished Goods (Final demand goods)
69 Index of Producer's Inventory of Finished Goods (Investment goods)
70 Index of Producer's Inventory of Finished Goods (Capital goods)
71 Index of Producer's Inventory of Finished Goods (Construction goods)
72 Index of Producer's Inventory of Finished Goods (Consumer goods)
73 Index of Producer's Inventory of Finished Goods (Durable consumer goods)
74 Index of Producer's Inventory of Finished Goods (Non-durable consumer goods)
75 Index of Producer's Inventory of Finished Goods (Producer goods)
76 Index of Producer's Inventory of Finished Goods (Producer goods for mining and manufacturing)
77 Index of Producer's Inventory of Finished Goods (Producer goods for others)
78 Index of Inventory (Final demand goods)

Investments

79 Index of Producer's Shipments (Investment goods excluding transport equipments)
80 Index of Producer's Shipments (Producer goods)
81 Index of Industrial Production (Investment goods)
82 Index of Industrial Production (Capital goods)
83 Index of Industrial Production (Construction goods)
84 Index of Production Capacity (Manufacturing)
85 Machinery Orders (Total excluding ships)
86 Machinery Orders (Private sector excluding volatile orders)
87 Machinery Orders (Manufacturing)
88 Machinery Orders (Non-manufacturing excluding volatile orders)
89 Machinery Orders (Government)
90 Order Received for Construction (Grand Total)
91 Order Received for Construction (Private)
92 Order Received for Construction (Manufacturing)
93 Order Received for Construction (Non-manufacturing)
94 Order Received for Construction (Public)
95 Total Floor Area of Building Construction Started (Grand Total)
96 Total Floor Area of Building Construction Started (Mining, Manufacturing and Commercial Use)
97 Total Floor Area of Building Construction Started (Mining)
98 Total Number of New Housing Construction Started (Total)
99 Total Number of New Housing Construction Started (Owned)
100 Total Number of New Housing Construction Started (Rented)
101 Total Number of New Housing Construction Started (Built for sale)
102 Total Number of New Housing Construction Started (Government housing loan corporation)
103 Total Floor Area of New Housing Construction Started (Total)
104 Total Floor Area of New Housing Construction Started (Owned)
105 Total Floor Area of New Housing Construction Started (Rented)
106 Total Floor Area of New Housing Construction Started (Built for sale)

Employment

107 Index of Non-scheduled Worked Hours (All industries - 30 or more persons)
108(HOUR) Index of Non-scheduled Worked Hours (Manufacturing)
109 Index of Total Worked Hours (All industries - 30 or more persons)
110 Index of Total Worked Hours (Manufacturing)
111 Ratio of Non-scheduled to Total Worked Hours (All industries - 30 or more persons)
112 Ratio of Non-scheduled to Total Worked Hours (Manufacturing)
113 New Job Offers
114 Effective Job Offers

115 New Job Offer Rate
116 Effective Job Offer Rate
117 New Job Offers (Part-time)
118 Effective Job Offers (Part-time)
119 New Job Offer Rate (Part-time)
120 Effective Job Offer Rate (Part-time)
121 Index of Regular Workers Employment (All industries - 30 or more persons)
122 Index of Regular Workers Employment (All industries excluding services)
123 Index of Regular Workers Employment (Mining)
124 Index of Regular Workers Employment (Construction)
125 Index of Regular Workers Employment (Manufacturing)
126 Index of Regular Workers Employment (Electricity, gas, heat supply)
127 Index of Regular Workers Employment (Transport and communication)
128 Index of Regular Workers Employment (Wholesale and retail trade)
129 Index of Regular Workers Employment (Finance and insurance)
130 Index of Regular Workers Employment (Real estate)
131 Index of Regular Workers Employment (Services)
132 Number of Unemployment
133 Unemployment Rate
134 Number of Beneficiaries of Unemployment Insurance (Initial claimants)
135 Number of Beneficiaries of Unemployment Insurance (Total)
136 Number of Persons with Unemployment Insurance
137 Real Wage Index (Contractual cash earnings in all industries - 30 or more persons)

Consumption

138 Sales at Department Stores (Total)
139 Sales at Department Stores (Per square meter floor space)
140 Index of Sales (Total)
141 Index of Sales (Wholesale)
142 Index of Sales (General Merchandize Retail)
143 Number of New Passenger Car Registrations and Reports (Total)
144 Number of New Passenger Car Registrations and Reports (excluding cars under 550cc)
145 Household Consumption Expenditure (Workers)
146 Household Consumption Expenditure (Food)
147 Household Disposable Income (Workers)
148 Index of Industrial Production (Consumer goods)
149 Index of Industrial Production (Durable consumer goods)
150 Index of Industrial Production (Non-durable consumer goods)
151 Index of Producer's Shipments (Consumer goods)
152 Index of Producer's Shipments (Durable consumer goods)
153 Index of Producer's Shipments (Non-durable consumer goods)

Firms

154 Index of Investment Climate (Manufacturing)
155 Corporation Tax Revenue
156 Suspension of Business Transaction with Bank

Money, stock price and interest rate

157 Money Supply (M2+CD, average outstanding)
158 Money Supply (M1, average outstanding))
159 Monetary Base (Average outstanding)
160 Bank Notes Issued (Average outstanding)
161 Bank Clearings (Number)
162 Bank Clearings (Value)
163 Nikkei Stock Average 225 Selected Stocks (Average of month)
164 Nikkei Stock Average 500 Selected Stocks
165 Stock Price Index (TOPIX)
166 Stock Price Average (Tokyo stock market first section)
167 Stock Price Index (Fisheries, agriculture and forestry)
168 Stock Price Index (Mining)
169 Stock Price Index (Construction)
170 Stock Price Index (Foods)
171 Stock Price Index (Textiles)
172 Stock Price Index (Pulp and paper)
173 Stock Price Index (Oil and coal products)

174 Stock Price Index (Rubber products)
175 Stock Price Index (Glass and ceramics product)
176 Stock Price Index (Iron and steel)
177 Stock Price Index (Non-ferro metals)
178 Stock Price Index (Metal products)
179 Stock Price Index (Machinery)
180 Stock Price Index (Electrical machinery)
181 Stock Price Index (Transportation equipment)
182 Stock Price Index (Precision instrument)
183 Stock Price Index (Other products)
184 Stock Price Index (Electric and gas)
185 Stock Price Index (Land transportation)
186 Stock Price Index (Marine transportation)
187 Stock Price Index (Air transportation)
188 Stock Price Index (Warehouse and transport related)
189 Stock Price Index (Communication)
190 Stock Price Index (Real estate)
191 Stock Price Index (Service)
192 Sales Volume (Daily Average, Tokyo stock market first section)
193 Sales Value (Daily Average, Tokyo stock market first section)
194 Official Discount Rates
195 Short-term Prime Lending Rates
196 Long-term Prime Lending Rates
197 Average Contracted Interest Rate on Loans and Discounts (Domestically licensed bank)
198 Yields of Bond Traded with Repurchase Agreement (3 months, month average)
199 Call Rates (Collateralized Overnight, month average)
200 Bill Rates (2 months, month average)
201 Yields of Short-term Government Securities (13 weeks)
202 Yields of Interest Bearing Bank Debentures (5 years)
203 Yields of Interest Bearing Government Bonds (10 years)
204 Yields of Government Guaranteed Bonds (10 years)
205 Yields of Local Government Bonds (10 years)
206 Yields to Maturity of Listed Government Bond (Longest term until redemption day)

Price indexes

207 Nikkei Commodity Price Index (17items)
208 Nikkei Commodity Price Index (42items)
209 Wholesale Price Index (All commodities)
210 Wholesale Price Index (Manufacturing industry products)
211 Wholesale Price Index (Raw materials)
212 Wholesale Price Index (Intermediate materials)
213 Wholesale Price Index (Final goods)
214 Wholesale Price Index (Capital goods)
215 Wholesale Price Index (Consumer goods)
216 Wholesale Price Index (Durable consumer goods)
217 Wholesale Price Index (Nondurable consumer goods)
218 Consumer Price Index (General)
219 (CPI) Consumer Price Index (General excluding fresh food)
220 Consumer Price Index (General excluding fresh food and imputed rent)
221 Consumer Price Index (Food)
222 Consumer Price Index (Housing)
223 Consumer Price Index (Fuel light and water charges)
224 Consumer Price Index (Furniture and household utensils)
225 Consumer Price Index (Clothes and footwear)
226 Consumer Price Index (Medical care)
227 Consumer Price Index (Transportation and communication)
228 Consumer Price Index (Reading and recreation)
229 Consumer Price Index (Miscellaneous)

Trade

230 Terms of Trade Index (All commodities)
231 Quantum Index of Exports (Total)
232 Quantum Index of Imports (Total)
233 Customs Clearance (Value of exports, grand total)
234 Foreign Exchange Rate (Yen per US dollar, Spot)

References

Bai, J. (2001). "Inference on factor models of large dimensions." manuscript presented at NBER Summer Institute 2001.

Bai, J. and S. Ng (2002). "Determining the number of factors in approximate factor models." Econometrica 70(1): 191-221.

Cabinet Office (1997). "30 keiretsu ni yoru keiki doukou shisuu (Business cycle index based on 30 series)." Economic Social Research Institute (formerly Economic Planning Agency), in Japanese.

Diebold, F.X. (2000). ""Big Data" dynamic factor models for macroeconomic measurement and forecasting." manuscript presented at World Congress of the Econometric Society 2000.

Diebold, F.X. and G. D. Rudebusch (1996). "Measuring business cycles: A modern perspectives." Review of Economics and Statistics 78, 67-77.

Fan, Y. and Q. Li (1997). "A consistent nonparametric test for linearity of AR(p) models." Economics Letters 55: 53-59.

Forni, M., M. Hallin, M. Lippi and L. Reichlin (2000). "The generalized dynamic-factor model: identification and estimation." Review of Economics and Statistics 82(4): 540-554.

Forni, M. and L. Reichlin (1998). "Let's get real: a factor analytical approach to disaggregated business cycle dynamics." Review of Economic Studies 65: 453-473.

Fukuda, S. and T. Onodera (2001). "A new composite index of coincident economic indicators in Japan: how can we improve the forecast performance?" International Journal of Forecasting, 17: 483-498.

Geweke, J. (1977). "The dynamic factor analysis of economic time-series models." in D. J. Aigner and A.S. Goldberger (eds.), Latent Variable in Socioeconomic Models, Amsterdam, North-Holland, 365-387.

Hamilton, J. (1989). "A new approach to the economic analysis of nonstationary time series and the business cycle." Econometrica 57: 357-384.

Kanoh, S. (1990). "The statistical reconsideration of the EPA diffusion index." Journal of the Japanese and International Economies 4: 139-156.

Kariya, T. (1993). Quantitative Methods of Portfolio Analysis, Dordrecht, Kluwer Academic Publishers.

Kim, C.-J. and C. R. Nelson (1998). "Business cycle turning points, a new coincident index, and tests of duration dependence based on a dynamic factor model with regime-switching." Review of Economics and Statistics 80, 188-201.

Kim, C.-J. and C. R. Nelson (1999). State Space Models with Regime Switching, (Cambridge, MIT Press).

Lee, T.-H., H. White and C. W. J. Granger (1993). "Testing for neglected nonlinearity in time series models: A comparison of neural network methods and alternative tests." Journal of Econometrics 56, 269-290.

Mankiw, N.G. and M.D. Shapiro (1986). "News or noise: An analysis of GNP revisions." Survey of Current Business, 20-25.

Mori, K., Satake, M. and Ohkusa, Y. (1993). "Stock-Watson taipu no keiki shisuu: Nihon keizai he no ouyou (Stock-Watson type business cycle index: Application to Japanese economy," Doushisha University Keizigaku Rousyu, 45, 28-50, in Japanese.

Sargent, T. J. and C. Sims (1977). "Business cycle modeling without pretending to have too much a priori theory." In C. Sims (ed.), New Methods of business Cycle Research (Minneapolis: Federal Reserve Bank of Minneapolis).

Shintani, M. and O. Linton (2001). "Is there chaos in the world economy? A nonparametric test using consistent standard errors." International Economic Review, forthcoming.

Stock, J. H. and M. W. Watson (1989). "New indexes of coincident and leading economic indicators." In O. Blanchard and S. Fischer (eds.), NBER Macroeconomics Annual (Cambridge, Mass. MIT Press).

Stock, J. H. and M. W. Watson (1998). "Diffusion indexes." NBER Working Paper no.6702.

Stock, J. H. and M. W. Watson (1999). "Forecasting inflation." Journal of Monetary Economics 44: 293-335.

Stock, J. H. and M. W. Watson (2001). "Macroeconomic forecasting using diffusion indexes." manuscript, Harvard University and Princeton University.

Teräsvirta, T., C-F.J.Lin and C.W.J.Granger (1993). "Power of the neural network linearity test." Journal of Time Series Analysis 14, 209-220.

Watanabe, T. (2001). "Measuring business cycle turning points in Japan with a dynamic Markov switching factor model." manuscript, Tokyo Metropolitan University.

White, H., (1989) “An additional hidden unit test for neglected nonlinearity in multilayer feedforward networks.” Proceedings of the international joint conference on neural networks, Washington, DC, IEEE Press, New York, NY, II, 451-455.

Yoshizoe, Y, S. Ohira, E. Shioji, M. Katsuura, H. Motoyama, K. Takase, T. Onishi, A. Sawada, S. Aoki, T. Kitaoka, R. Serizawa, H. Maejima (2001). “Keizai doko shihyo no saikento (Re-examination of indices of economic activities).” Keizai Bunseki (Economic Analysis), Viewpoints of policy analysis series No. 19, Economic and Social Research Institute, Cabinet Office of Japan, in Japanese.

Zheng, J. X. (1996). “A consistent test of functional form via nonparametric estimation techniques.” Journal of Econometrics 75, 263-289.

Table 1
Simulated out-of-sample forecasting results: 12-month horizon

Series	Benchmark	Factor Model						Stacked Factor Model			
	(1)AR	(2)DI-AR-LAG		(3)DI-AR		(4)DI		(5)DIS-AR		(6)DIS	
	RMSE	Rel.MSE	$\widehat{\alpha}$	Rel.MSE	$\widehat{\alpha}$	Rel.MSE	$\widehat{\alpha}$	Rel.MSE	$\widehat{\alpha}$	Rel.MSE	$\widehat{\alpha}$
IP	0.065	0.74	1.11	0.76	1.16	0.76	1.16	0.74	1.20	0.67	1.30
		(0.22)	(0.38)	(0.21)	(0.45)	(0.21)	(0.45)	(0.21)	(0.40)	(0.24)	(0.38)
SHIP	0.060	0.71	1.13	0.81	1.00	0.81	1.00	0.81	0.98	0.77	1.08
		(0.20)	(0.32)	(0.16)	(0.37)	(0.16)	(0.37)	(0.18)	(0.35)	(0.18)	(0.34)
CAP	0.061	0.72	1.22	0.81	1.12	0.81	1.12	0.70	1.24	0.70	1.24
		(0.18)	(0.35)	(0.15)	(0.41)	(0.15)	(0.41)	(0.17)	(0.30)	(0.17)	(0.30)
SALE	0.067	0.79	1.06	0.79	1.07	0.79	1.07	0.82	1.06	0.81	1.10
		(0.18)	(0.38)	(0.18)	(0.38)	(0.18)	(0.38)	(0.16)	(0.40)	(0.16)	(0.41)
HOUR	0.146	0.79	0.92	0.88	0.88	1.02	0.46	0.83	0.97	0.89	0.73
		(0.24)	(0.46)	(0.18)	(0.50)	(0.17)	(0.40)	(0.20)	(0.47)	(0.20)	(0.41)
CPI	0.012	0.80	0.71	0.81	0.69	0.89	0.54	0.79	0.64	0.64	0.68
		(0.20)	(0.16)	(0.20)	(0.16)	(0.27)	(0.10)	(0.27)	(0.18)	(0.32)	(0.12)

Notes: RMSE is the root MSE of a benchmark univariate autoregressive forecast with lag length selected by BIC. Rel.MSE is the ratio of the MSE of the forecast to the MSE of the benchmark model. $\widehat{\alpha}$ is the forecast combining coefficient estimate. Numbers in parentheses are HAC standard errors. Forecasting series are: index of industrial production (IP); index of producer's shipments (SHIP); index of capacity utilization ratio (CAP); index of sales in small and medium-sized enterprises (SALE); index of non-scheduled worked hours (HOUR); and consumer price index (CPI).

Table 2
Simulated out-of-sample forecasting results: 6-month horizon

Series	Benchmark	Factor Model						Stacked Factor Model			
	(1)AR	(2)DI-AR-LAG		(3)DI-AR		(4)DI		(5)DIS-AR		(6)DIS	
	RMSE	Rel.MSE	$\widehat{\alpha}$	Rel.MSE	$\widehat{\alpha}$	Rel.MSE	$\widehat{\alpha}$	Rel.MSE	$\widehat{\alpha}$	Rel.MSE	$\widehat{\alpha}$
IP	0.036	0.72	1.11	0.82	1.18	0.79	1.23	0.75	1.22	0.76	1.19
		(0.17)	(0.28)	(0.12)	(0.33)	(0.13)	(0.32)	(0.13)	(0.26)	(0.15)	(0.31)
SHIP	0.036	0.69	1.07	0.83	0.98	0.83	0.98	0.79	1.07	0.80	1.02
		(0.18)	(0.23)	(0.13)	(0.31)	(0.13)	(0.31)	(0.15)	(0.33)	(0.15)	(0.31)
CAP	0.037	0.63	1.27	0.74	1.19	0.74	1.19	0.68	1.12	0.67	1.13
		(0.14)	(0.19)	(0.12)	(0.24)	(0.12)	(0.24)	(0.14)	(0.19)	(0.14)	(0.19)
SALE	0.033	0.89	0.74	0.93	0.66	0.92	0.67	0.86	0.94	0.92	0.76
		(0.10)	(0.22)	(0.11)	(0.23)	(0.10)	(0.23)	(0.12)	(0.31)	(0.08)	(0.26)
HOUR	0.068	0.90	0.68	0.88	0.81	1.03	0.45	0.87	0.75	0.91	0.62
		(0.16)	(0.28)	(0.13)	(0.33)	(0.14)	(0.22)	(0.16)	(0.30)	(0.16)	(0.23)
CPI	0.006	0.86	0.66	0.86	0.66	1.24	0.41	0.78	0.71	0.84	0.58
		(0.16)	(0.14)	(0.16)	(0.14)	(0.36)	(0.10)	(0.22)	(0.18)	(0.22)	(0.10)

Notes: See notes for Table 1.

Table 3
DGP Revisions

		Y1	Y2	Y3	Y4	Y5
		(A) Nominal GDP				
Growth rates	mean	0.94	1.00	0.97	0.99	1.04
	s.d.	1.18	1.17	1.04	0.92	1.01
	min	-1.90	-2.10	-1.60	-1.50	-1.60
	max	3.60	3.50	2.90	2.70	3.10
Total revisions	mean	0.10	0.04	0.07	0.05	–
	s.d.	0.71	0.70	0.55	0.42	–
	min	-1.90	-1.90	-1.40	-1.20	–
	max	2.00	1.90	1.10	1.50	–
		(B) Real GDP				
Growth rates	mean	0.68	0.73	0.71	0.75	0.69
	s.d.	1.07	1.06	0.95	0.82	0.85
	min	-2.90	-2.80	-2.50	-2.00	-2.00
	max	3.00	2.90	2.40	2.70	2.60
Total revisions	mean	0.01	-0.04	-0.02	-0.06	–
	s.d.	0.73	0.71	0.58	0.40	–
	min	-2.10	-2.30	-1.60	-1.30	–
	max	1.80	1.60	1.10	1.00	–

Notes: Percent changes from the same quarter of the previous year. Total revisions are differences from Y5 (e.g. Y5-Y1). Sample period: 1980:I-1999:I

Table 4
Forecasting GDP using diffusion index

	(1) DI-M	(2) DI-Q	(3) DI-Q2
(A) Nominal GDP growth rates (Y5)			
Y1 and DI	0.42	0.48	0.44
	(0.06)	(0.05)	(0.06)
Y2 and DI	0.43	0.45	0.44
	(0.06)	(0.06)	(0.06)
(B) Real GDP growth rates (Y5)			
Y1 and DI	0.42	0.51	0.46
	(0.07)	(0.08)	(0.07)
Y2 and DI	0.43	0.47	0.44
	(0.07)	(0.07)	(0.08)

Notes: Forecast combining coefficient estimates on DI-based forecasts in the in-sample forecasts of Y5 are reported. Numbers in parentheses are standard errors. Sample period: 1980:I-1999:I

Table 5
Tests for linearity

	US				Japan			
lags	1	2	3	4	1	2	3	4
				(A) Kernel-based test				
DI1	0.249	0.573	0.008	<0.001	0.017	<0.001	0.179	0.315
DI2	0.922	0.906	0.745	0.666	0.777	0.056	0.069	0.007
DI3	<0.001	0.067	0.473	0.216	0.888	0.067	0.343	0.373
				(B) Neural network test				
DI1	0.004	0.036	0.009	0.001	<0.001	<0.001	0.008	0.016
DI2	0.211	0.004	0.090	<0.001	0.078	0.087	0.001	<0.001
DI3	<0.001	<0.001	0.041	0.001	0.039	0.056	0.022	0.007
				(C) LM type neural network test				
DI1	0.004	0.002	<0.001	<0.001	<0.001	<0.001	0.024	0.021
DI2	0.073	0.002	<0.001	<0.001	0.039	0.022	<0.001	<0.001
DI3	<0.001	0.001	0.005	0.003	0.020	0.006	0.004	0.003

Notes: Numbers are p-values of the tests for the null hypothesis of linearity. See Fan and Li (1997) for the kernel-based test, White (1989) for the neural network test and Teräsvirta, Lin and Granger (1993) for the LM type neural network test.

Table 6
Lyapunov exponent estimates

	US				Japan			
lags	1	2	3	4	1	2	3	4
	(A) Kernel estimation							
	(1) Full sample							
DI1	-0.532	-0.311	-0.220	-0.138	-1.843	-0.845	-0.326	-0.203
	(0.020)	(0.026)	(0.014)	(0.009)	(0.062)	(0.036)	(0.013)	(0.010)
DI2	-0.211	-0.093	-0.069	-0.089	-0.521	-0.152	-0.060	-0.039
	(0.011)	(0.008)	(0.004)	(0.007)	(0.030)	(0.009)	(0.004)	(0.004)
DI3	-1.680	-0.724	-0.312	-0.250	-1.114	-0.937	-0.588	-0.371
	(0.069)	(0.028)	(0.010)	(0.007)	(0.024)	(0.027)	(0.020)	(0.018)
	(2) Block							
DI1	-0.552	-0.312	-0.225	-0.126	-1.854	-0.803	-0.310	-0.193
	(0.043)	(0.055)	(0.031)	(0.022)	(0.133)	(0.078)	(0.031)	(0.030)
DI2	-0.204	-0.085	-0.064	-0.078	-0.487	-0.131	-0.052	-0.033
	(0.015)	(0.007)	(0.007)	(0.012)	(0.023)	(0.009)	(0.005)	(0.006)
DI3	-1.645	-0.645	-0.302	-0.237	-1.141	-0.940	-0.558	-0.353
	(0.142)	(0.054)	(0.023)	(0.020)	(0.047)	(0.062)	(0.046)	(0.040)
	(B) Neural network estimation							
	(1) Full sample							
DI1	-0.559	-0.284	-0.181	-0.119	-3.724	-1.454	-0.328	-0.101
	(0.026)	(0.021)	(0.028)	(0.023)	(0.078)	(0.058)	(0.014)	(0.009)
DI2	-0.191	-0.085	-0.066	-0.076	-0.487	-0.151	-0.089	-0.063
	(0.017)	(0.007)	(0.005)	(0.008)	(0.028)	(0.021)	(0.015)	(0.008)
DI3	-1.947	-0.679	-0.327	-0.263	-1.170	-0.902	-0.648	-0.736
	(0.050)	(0.032)	(0.021)	(0.020)	(0.039)	(0.029)	(0.044)	(0.050)
	(2) Block							
DI1	-0.587	-0.294	-0.202	-0.121	-3.719	-1.382	-0.321	-0.089
	(0.055)	(0.046)	(0.054)	(0.046)	(0.110)	(0.122)	(0.029)	(0.009)
DI2	-0.202	-0.088	-0.066	-0.073	-0.470	-0.123	-0.044	-0.041
	(0.029)	(0.013)	(0.010)	(0.014)	(0.018)	(0.017)	(0.013)	0.016
DI3	-1.876	-0.627	-0.285	-0.228	-1.183	-0.906	-0.681	-0.674
	(0.121)	(0.059)	(0.043)	(0.049)	(0.077)	(0.066)	(0.087)	(0.128)

Notes: Numbers in parentheses are standard errors.

Figure 1: Index of Industrial Production (IIP)

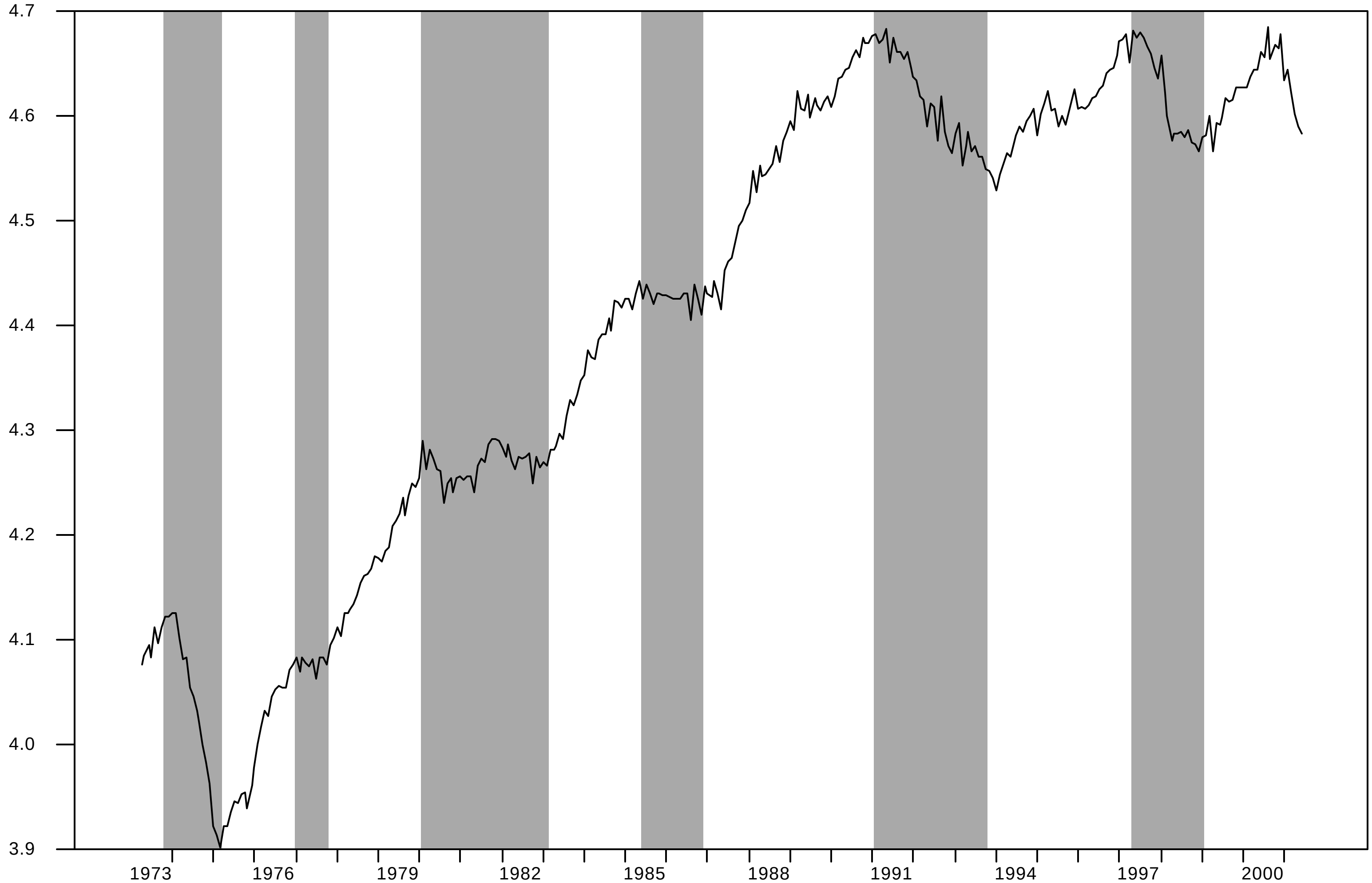

Figure 2: Factor Model DI (Cumulated Sum)

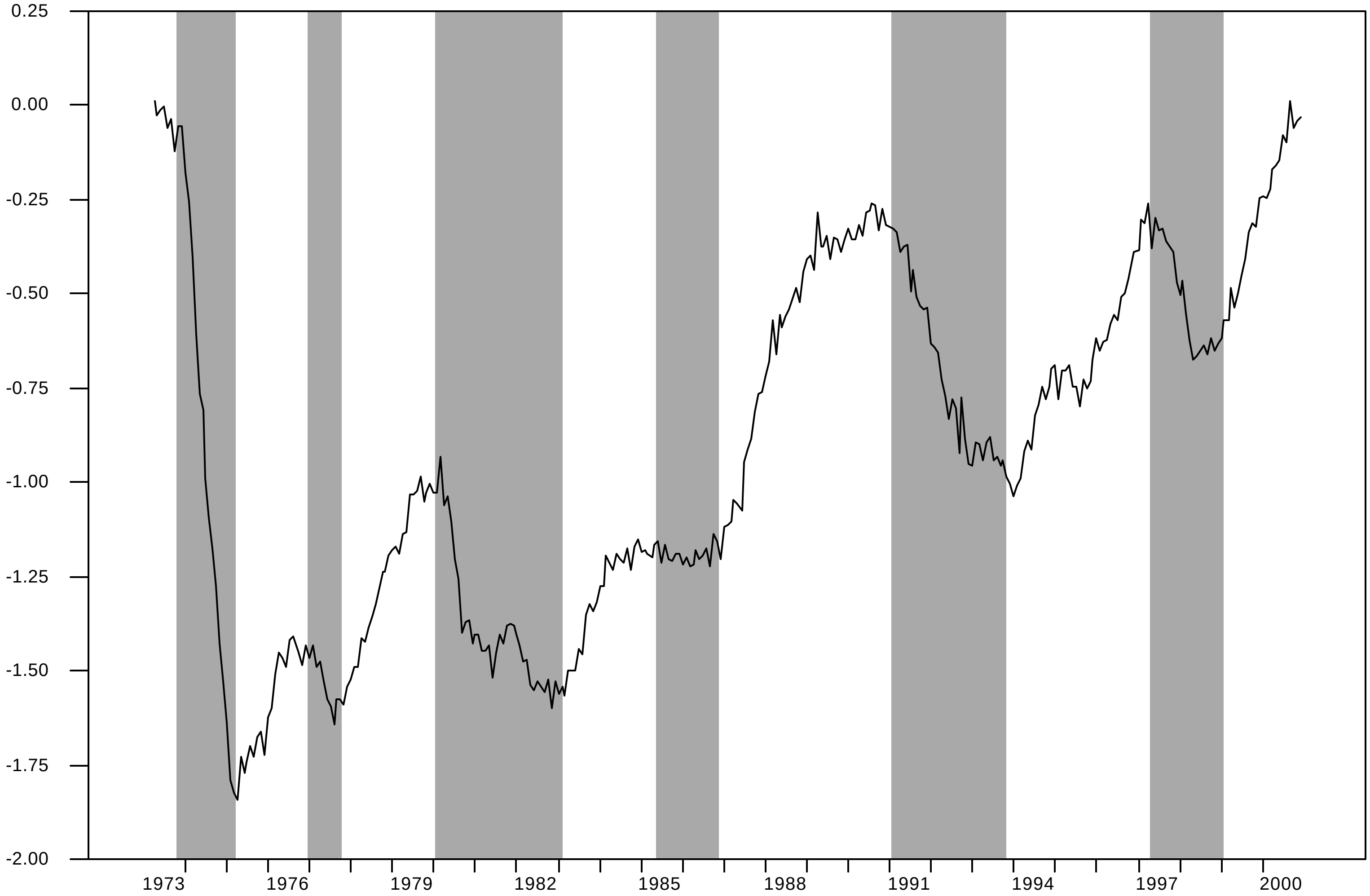

Figure 3: Factor Model DI

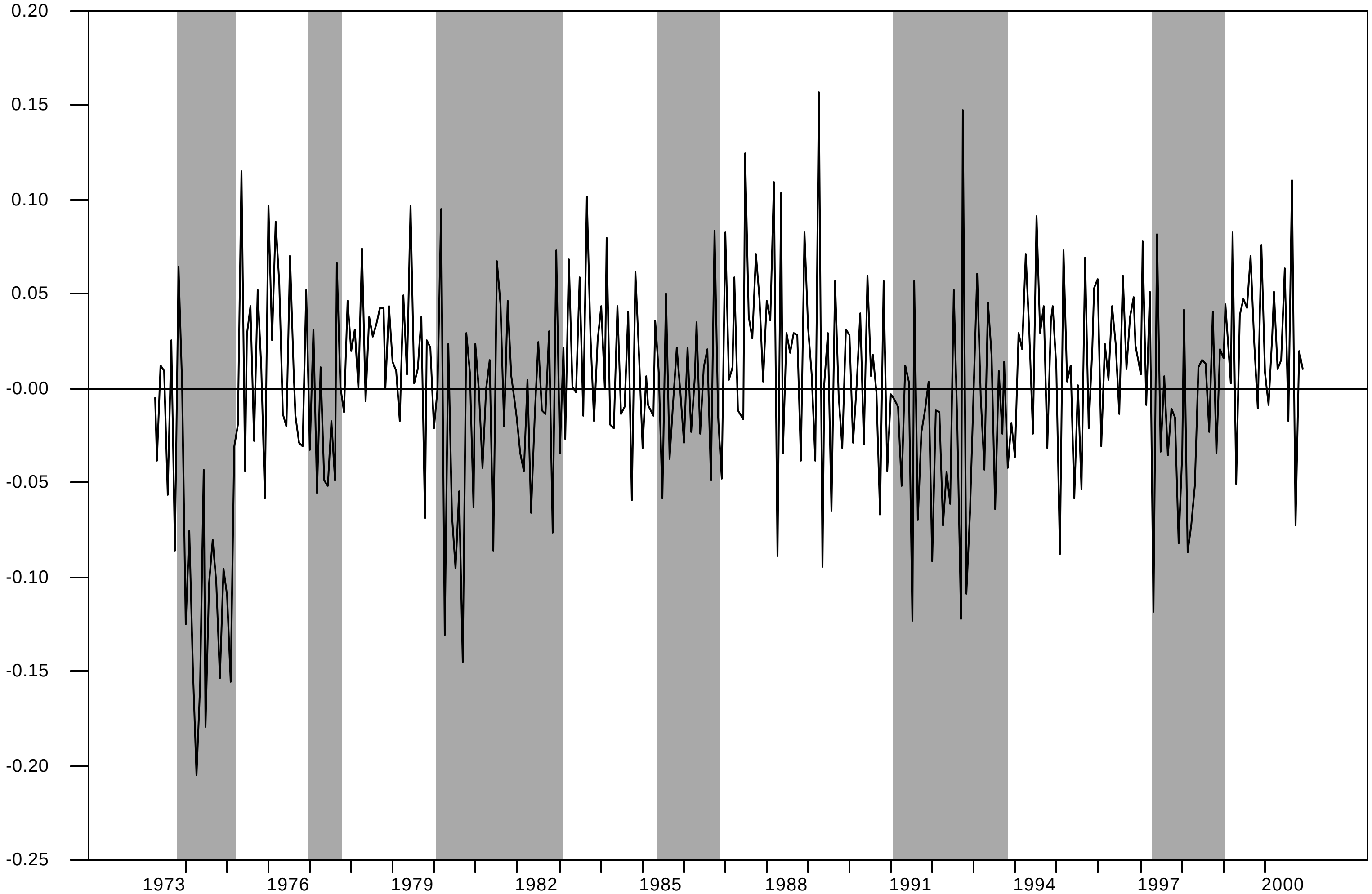

Figure 4: Smoothed Probability of Recession