Use of Information Variables in Inflation Forecasting

Jong Hwa Kim* and Joong Shik Lee**

Inflation can be seen as a phenomenon caused by a mixture of various international and domestic economic factors including demand, supply, and cost conditions. Consequently, extensive use of information variables that reflect key economic trends and factors is critical for accurate inflation forecasting. Despite their importance, studies on information variables - especially on the concept and their practical application - are relatively rare.

This paper analyses the scope for enhancing prediction accuracy in inflation forecasting by employing out-of-sample predictability tests, when information taken by way of principal component analysis from 42 individual economic indicators of the real, external, price, and financial-market sectors is used appropriately. The empirical experiments show that the accuracy of the model incorporating information variables is substantially superior to other time-series models. These results support the notion that efforts to apply more sophisticated processes to a broad spectrum of information variables are required in order to increase the predictability of inflation forecasting.

JEL Classification Number: C32, E37

Key words : inflation forecasting, information variables, principal component analysis, out-of-sample predictability tests

* Head, Economic Research Team, Daejeon & Chungnam Branch, the Bank of Korea (E-mail: jhkim@bok.or.kr)

** Senior Economist, Monetary Policy Analysis Team, Monetary Policy Department, the Bank of Korea (E-mail: jslee@bok.or.kr)

We would like to thank two anonymous referees for their comments.

I. Introduction

Since the April 1998 revision of the Bank of Korea Act, monetary policy in Korea has operated basically within an inflation targeting framework. Inflation targeting is characterized by the securing of public confidence in monetary policy, building on the central bank's accurate forecasting of the future course of inflation. Because it takes from six to twenty-four months for monetary policy to have an effective impact on inflation, by way of various channels including interest rates, exchange rates and asset prices, achieving an inflation target requires that monetary policy be implemented pre-emptively on the basis of inflation forecasts.

The success of inflation targeting, therefore, depends largely on the ability to predict future inflation. Large disparities between predicted and actual inflation rates would undermine the credibility of a pre-emptive monetary policy. Likewise, the central bank must be able to communicate effectively with the general public about the risk factors associated with uncertainties in the domestic and international economies. The failure of the central bank to provide timely and potent economic forecasts may result in unnecessary widespread public confusion, decreasing the efficiency of the transmission mechanism of monetary policy.

Against this background, central banks have focused on increasing the accuracy of inflation prediction through improving econometric methods, such as establishing more sophisticated macroeconomic models and developing inflation indicators. Inflation, however, is caused by a mixture of many factors including domestic and international supply and demand, cost conditions, and expectations of private economic agents. Moreover, the interaction of these factors tends to be irregular across time series. It follows that a comprehensive model is necessary to increase the precision of inflation forecasting; also called for is a method by which a broad range of information variables can be monitored and considered in an expedient manner to gauge inflationary pressures.

Information variables for the central banks are defined as intermediate economic indices used to help predict the final objective of price stability in the process of monetary policy implementation. Korea, like other countries that have opted for inflation targeting, utilizes wide range of information variables for various time horizons - real-time, short-term, and long-term periods. In addition, other forms of information variables include stock indices and exchange rates, which reflect the public's expectations of inflation. The importance of information variables has been heightened as the effectiveness of intermediate targets, such as monetary and credit aggregates, in monetary targeting, has decreased. Whereas the intermediate target places emphasis on the behavioral causality to the final objective, information variables are considered to be meaningful in inflation targeting regime as long as they have a high cross-correlation with the final objective of the monetary policy.

An individual information variable, however, reflects not only inflationary pressures, but also the individual attributes of the index, like short-term irregularities, making it difficult to predict future inflation by examining the movement of any single information variable. Additionally, even data extracted through sophisticated econometric techniques or detailed analysis eliminating irregularities may vary or even conflict with data from other information variables. For this reason, information variables are best utilized in a synthetic approach, where idiosyncratic and diverse data are combined into a comprehensive mechanism to examine inflationary pressures. Recently, Stock and Watson (1999) utilized principal component analysis to derive key trends from the movements of various information variables; the appropriate usage of data derived from these studies yielded increased accuracy in inflation predictions, regardless of the timeframe over which the observations were conducted. While this study was conducted in the United States, repeated trials in major European nations, Canada, and Japan confirmed its conclusions.[1)]

This study seeks to examine the implications of applying the Stock and Watson method to Korean information variables: whether the inflation prediction increases in accuracy as a result of applying the method, and whether this result is consistent with monetary policy framework under inflation targeting, especially with regards to basing policy decisions on the application of information variables. Most studies about the Korean economy focus on narrowly specific areas, or center on very few information variables; there are few or no studies that consider in a general sense information variables and their usefulness in the current framework of monetary policy. As Korea is in its seventh year of inflation targeting, this study examines the role of information variables in the establishment of a pre-emptive monetary policy.

The paper is organized as follows: Section II introduces existing studies about information variables in inflation forecasting; Section III discusses the methodology of comparing inflation forecasting and the information variables

1) Refer to sub-section 2 of Section II.

considered in inflation forecasting in Korea. Section IV discusses the empirical results of the study, and Section V concludes the study and outlines topics for future research.

II. Overview of Related Studies

Inflation forecasting has depended largely on modeling the behavior of inflation and important macroeconomic variables in accordance with the Phillips curve; much research has been conducted in this area because of the importance of inflation forecasting in macroeconomics. Early research confined the variables used in the Phillips curve to those closely related to production activities, like the unemployment rate and industrial production. Recently, however, the perception that various demand and cost factors influence inflation has prompted the utilization of diverse information variables to forecast inflation.

This section examines studies that have highlighted the role of information variables in the inflation forecasting process. It groups existing studies into two categories: studies that discuss the different types of available information variables, and studies that combine the data yielded by information variables in a systematic manner. The defining characteristic of current research efforts is that they represent the current state of inflation forecasting used by the central banks more accurately than previous such efforts; this is especially true as the accuracy of methods utilizing information variables is measured in comparison to other forecasting methods, placing the focus on the effect of information variables in improving prediction accuracy rather than adequacy of individual behavioral equation. Moreover, in order to realistically reflect the information sets that the central bank faces at any given moment, the forecast horizon is updated recursively in increments of one or two years; the accuracy thereof being measured through an out-of-sample forecast.

1. Research on Types of Information Variables

Kozicki (2001) examined economic variables, including GDP, interest rates, monetary aggregates, exchange rates, and unemployment rates, of eleven different countries in order to find if there were variables that could, regardless of time period or nation, increase the accuracy of inflation prediction. The research, however, found that individual variables varied largely in accuracy from nation to nation, and that the variables' degree of accuracy depended largely on the time period over which they were used. The study concluded that inflation prediction necessitated the monitoring of a wide range of information variables, while analysis of their movements needed to be interpreted according to the state of the economy and financial markets.

Stock and Watson (2003) researched the effect of asset prices - measured by interest rates, interest rate spread, stock indices, and exchange rates - on the predictability of inflation and economic growth rates. In this process, they surveyed 93 related studies that have been conducted in the past 15 years. They concluded that there was no specific variable that increased the accuracy of inflation prediction in a stable manner, thereby supporting the idea that no individual information variable led to satisfactory results in terms of consistency. At the same time, in an empirical analysis, they concurred with Kozicki (2001) in stating that the relevance of individual variables varied largely according to target variable, time period, and nation. However, a combination forecast, which averaged the predictions of all considered variables with specific weights, was consistently more accurate than forecasts derived from individual variables, which in turn supported the argument that the central bank can increase prediction accuracy by the appropriate use of multiple information variables.

In Korea, Yoo and Noh (1999) focused on approximately 20 different information variables, including monetary aggregates, inflation expectations, demand, cost, and indices of aggregate economic activity, and compared the inflation forecasting accuracy of each of them. Their research was based on an out-of-sample forecast with a forecast horizon of one quarter. The findings of the study were that, with the exception of M2, exchange rate, and unit labor costs, all the variables considered were significant in enhancing inflation prediction. Kim (2002) examined the effect of asset prices such as interest rate spread, stock indices, and real estate prices on inflation prediction, and found that, on a forecast horizon of one to two years, interest spread and housing prices affected inflation significantly. These studies, however, have the limitation of relatively narrow ranges of information variables, or not reflecting the timeframe necessary for a pre-emptive monetary policy under inflation targeting.

2. Research on the Combination of Information Variables

Studies in the previous sub-section usually categorized information variables, and then compared the inflation prediction accuracies of different categories, or variables. The inflation forecasting process of central banks, however, is closer to a 'look-at-everything approach', which examines a combination of all information variables, as opposed to individual variables, in making judgments related to inflationary pressures. Because the combination of data involves subjective decision-making, dependent on previous experience, it is difficult to simulate the process through models or formal equations.

Recent technology has, however, enabled the accurate collection of information from a large set of data, clearing the way for a more systematic method by which to combine and analyze data derived from information variables. For example, Stock and Watson (1999) produced an economic index that combined 168 information variables, including real economic activity, monetary aggregates, interest rates, and inflation expectation. The study proved that the accuracy of inflation forecasting could be highly improved through the combination of information variables. Through principal component analysis, they found trends common to the 168 variables, and they extracted several factors that were responsible for these movements. They reported that the appropriate use of these factors improved greatly the inflation forecasting accuracy of the Phillips curve. That economic indices derived from combining the data from many information variables can improve inflation prediction is possible because inflation itself is caused by a mixture of various economic factors including supply, demand, and cost conditions. Strikingly, despite differences in target information variables or weights of variables, the method of processing the data in this study was very similar to the actual method by which the central bank gauges and analyzes inflationary pressures.

Atkeson and Ohanian (2001) noted that, if U.S. inflation was predicted from 1984 to 1999 with the Stock and Watson method, it has like coin flipping a lower accuracy than a random walk model. Fisher, et al. (2002), however, argued that this lack of accuracy is limited to a specific period in which inflation variability was minimal; the inclusion of the 1970s, or the substitution of the GDP-based private consumption expenditure deflator for the Consumer Price Index (CPI) as the price indicator both yielded data that demonstrated the superiority of the Stock and Watson methodology.

The Federal Reserve Bank of Chicago adopted the Stock and Watson methodology to combine the data from 85 economic indicators, and called the first principal component the Chicago Fed National Activity Index (CFNAI). As noted by Fisher (2000), the CFNAI was intended to be an index for forecasting inflation. The Federal Reserve Bank of Chicago (2001) analyzed the empirical relationship between CFNAI and the American inflation cycle. Lee (2004) presented a new economic index, based on the CFNAI but adjusted for the Korean economy. Although discussing the identification of turning points of the business cycle, he does not tackle the question of inflation forecasting.

That the appropriate use of information variables increased the accuracy of inflation forecasting has been verified in research that extended beyond the United States economy. Forni, et al. (2003) found that an economic index that combined 447 European macroeconomic variables, including financial variables, monetary aggregates, industrial production, prices, and expectations surveys, improved the accuracy of inflation forecasting. Gosselin and Tkacz (2001) presented concurring research findings for Canada, while Shintani (2005) and Kitamura and Koike (2003) applied the Stock and Watson methodology to the Japanese economy and found that the combined economic indices clearly improve inflation forecasting accuracy.

III. Applying Principal Component Analysis to Information Variables

This section lays out the basic set-up for an inflation forecasting equation based on the Phillips curve, and examines the method by which the movements of information variables are integrated. Also discussed in this process are the different types and characteristics of information variables that are applicable in Korea.

1. Set-up of an Inflation Forecasting Equation to Compare Prediction Accuracy

We establish a behavioral equation for inflation forecasting utilizing a generalized Phillips curve as follows.

$$\pi^h_{t+h} = \alpha + \beta(L)F_t + \phi(L)\pi_t + \varepsilon_t \qquad (1)$$

In the above equation, π^h_{t+h} is the conditional forecast of inflation (π_t) at time t for h period ahead, while F_t is a variable that captures co-movement among the information variables used to predict inflation.[2] Coefficients $\beta(L)$ and $\phi(L)$ are

2) The general form of the Phillips curve is a special case of equation (1) where F_t is confined to a single variable such as the unemployment rate or gross production.

polynomial lagged equations, and h is the inflation forecast horizon for the central bank's pre-emptive monetary policy. Specific values (e.g. 6, 12, 18, 24 months) are assigned to h.

While equation (1) is simple, it reflects realistically the process by which an inflation-targeting central bank engages in inflation forecasting. First, the equation reflects not the current inflation, but inflation after period h, which is indicative of the pre-emptive monetary policy that is the central bank's *modus operandi*. Also, in the comparison of forecasting accuracy, the use of recursive estimation allows for comparison in a specific time period. By finding the values for a specific time period, one can conduct an out-of-sample forecast, allowing the reflection in the inflation forecasting process of the actual information set in data owned by central banks. Inoue and Kilian (2002) noted that comparing the in-sample forecast for accuracy brought a spurious predictability bias due in part to the lack of adequacy of the data set. For this reason, a test based on an out-of-sample forecast is more objective and practical method of gauging the accuracy of the model than an in-sample forecast.

Whereas the movement of π_t can be observed in real life, F_t is an abstract concept which requires an adequate treatment of the information variables utilized by the central bank. Stock and Watson (1999), however, provided a systematic framework for extracting information from the multiple variables by using principal component analysis. If X_{it} represents the information variables available to the central bank at time t, the number of which is N, and the sample period is $t = 1, \cdots, T$, then F_t can be, by definition, represented as follows.

$$X_{it} = \lambda_i(L)F_t + e_{it} \tag{2}$$

In this equation, $\lambda_i(L)$ is the polynomial coefficients of F_t variable that represents common trends in X_{it}. If we assume that the components of F_t and e_{it} are independent of each other, F_t is the value that minimizes the following non-linear objective function.

$$\min V(F, \Lambda) = \frac{\sum_{i=1}^{N}\sum_{t=1}^{T}(X_{it} - \lambda_i' F_t)^2}{NT} \tag{3}$$

Notice that λ_i is the ith component in the diagonal matrix Λ composed of coefficients of $\lambda_i(L)$. Stock and Watson (2002) showed that equation (3) is equal to the maximization of $tr[\Lambda'(X'X)\Lambda]$[3] under the constraint of $\Lambda'\Lambda/N = I$. In this case, the estimated value of F as derived from principal component analysis is as follows.

$$\hat{F} = \frac{X\hat{\Lambda}}{N} \tag{4}$$

Meanwhile, Stock and Watson (2003) regarded F_t as a single information variable (X_{it}), and calculated a forecast based on equation (1) for all i. These values were then re-combined into a combination forecast. From an econometric perspective, this combination forecast can be more sophisticated if its weights are assigned so as to minimize the forecast error. Realistically, however, it is rare for a central bank to forecast inflation by a mechanical combination of different forecasts that are each based on only one variable. Although principal component analysis is also a simple statistical method, it is much closer to the real-world activity of a central bank in its method and process of combining information variables. For this reason, this study will utilize principal component analysis as the means by which to compare inflation prediction accuracy.

2. Coverage of Information Variables and Their Characteristics

This study used 42 information variables (X_{it}) for inflation forecasting, ranging widely in coverage including real, overseas, price and financial-market sectors. The specific information variables are identified in [Table 1] by category. Because most real sector variables have a relatively limited time series, the sample period was set from January 1985 to December 2004. While the number of information variables considered is less than those utilized in the studies for other countries, almost all the variables that could realistically be taken into consideration in the Korean situation have been included.

Utilizing principal component analysis[4] in integrating movements of variables in [Table 1] seems to be consistent with the look-at-everything approach taken by inflation-targeting central banks in inflation forecasting. It is more advanced in that it does not put any prior restrictions on the weights of variables, unlike Yoo and Noh (1998) where inflation leading indicators are estimated under the ad hoc assumption that included variables are weighed equally.

---

3) X represents $N \times T$ matrix composed of vector X_{it} as its columns.

4) A more sophisticated econometric method is to apply maximum likelihood estimation through state-space model and Kalman filter algorithm as in Stock and Watson (1999). However, the number of variables that can be accommodated in this method is typically much smaller than the number of indicators applicable with principal component analysis.

Specifically, matrix X is comprised of economic indicators included in [Table 1], and F, capturing co-movement among indicators, is derived by estimating equation (3) and (4). All indicators, except the real effective exchange rate and the OECD leading indicator, are seasonally adjusted, and log-differenced (or differenced) to ensure the series is stationary, if necessary.[5] Following existing studies, each stationary series is adjusted for outliers - defined to be an observation whose distance away from the median is greater than six times the interquartile range of the series. In such cases, the original observation is replaced with the upper or lower value of the outlier criterion range, depending on the direction of deviation. As a next step, each series is standardized to have a mean of zero and standard deviation of one. Generally, new components derived from principal component analysis are statistically processed variables that are orthogonal to each other, and it is difficult to identify their economic meaning. However, through the value of the eigenvector corresponding to the largest eigenvalue of the second-moment matrix, we can recognize the degree of contribution of each principal component. After performing principal component analysis for the whole sample period, we sort out five component indicators in the order of highest absolute value of eigenvector for the first, second, and third principal components in [Table 2][6]. The contribution of these three principal components to the variability is about 40 percent, far higher than the 7 percent resulting from the assumption that individual components are perfectly orthogonal.

5) To be more specific, the inventory circulation indicator and labor entrance/separation ratio are level; the unemployment rate and yields on corporate bonds are first differenced; and remaining indicators are first log-differenced.

6) The principal components of 4th and onwards do not provide clear interpretation for the eigenvectors of composed variables. Furthermore, variables in the components are rarely included in the inflation equations based on the Bayesian Information Criterion (BIC) in the empirical results of Section IV.

Table 1 List of Information Variables

Sector	Information variables
Production	① Industrial production, all industries; ② Industrial production, manufacturing; ③ Industrial production, capital goods; ④ Industrial production, consumer goods; ⑤ Capacity utilization ratio, manufacturing; ⑥ Inventory circulation indicator; ⑦ Electric power consumption, manufacturing
Consumption	① Shipment, nondurable consumer goods; ② Shipment, durable consumer goods; ③ Wholesale & retail sales; ④ Retail sales in non-specialized stores; ⑤ Household consumption expenditures, all cities; ⑥ Automobile sales
Investment	① Shipment, intermediate goods; ② Shipment, machinery equipment; ③ Shipment, capital goods; ④ Machinery orders received; ⑤ Machinery imports; ⑥ Building permits; ⑦ Domestic construction orders received
Labor	① Labor input; ② Labor entrance/separation ratio; ③ Employed, nonagricultural; ④ Regular employees; ⑤ Unemployment rate; ⑥ Average monthly wages, all industries; ⑦ Average monthly wages, manufacturing
Price	① Producer price, commodities; ② Producer price, manufacturing industry products; ③ Import price, raw materials & intermediate goods; ④ Import price, capital goods; ⑤ Import price, consumer goods;
Financial markets	① M1; ② M2; ③ M3; ④ Domestic credit (M3); ⑤ Yields on corporate bonds; ⑥ Korea composite stock price index (KOSPI)
External	① Exports; ② Imports; ③ Real effective exchange rate; ④ OECD leading indicator

Table 2 Estimation Results of Principal Component Analysis[1)]

Principal component	Major information variables of corresponding component
The first principal component	① Industrial production, all industries (0.275); ② Industrial production, manufacturing (0.272); ③ Industrial production, consumer goods (0.251); ④ Capacity utilization ratio, manufacturing (0.246); ⑤ Shipment, intermediate goods (0.241)
The second principal component	① M3 (0.401); ② Domestic credit (M3) (0.363); ③ M2 (0.336); ④ Average monthly wages, all industries (0.329); ⑤ Average monthly wages, manufacturing (0.296)
The third principal component	① Producer price, manufacturing industry products (0.343); ② Import price, raw materials & intermediate goods (0.331); ③ Producer price, commodities (0.281); ④ Labor entrance/separation ratio (0.268); ⑤ Inventory circulation indicator (0.254)

Note: 1) Figures in parenthesis are absolute values of eigenvector of the largest eigenvalue in the second-moment matrix.

When we examine the estimation results of principal component analysis, economic variables representing the business cycle like industrial production lead the first principal component. This fits naturally with Burns and Mitchell's (1946) classical definition of the business cycle, which is that of abstract and aggregate movement occurring at about the same time in many economic activities. The movement of the second principal component is dominated by monetary aggregates and wages, similar to demand factors in inflationary pressures. As Milton Friedman asserted, inflation, defined as the persistent increase of the overall price level, reflects monetary phenomena eventually, and monetary aggregates can be a major factor in the demand-pull inflation in this respect. Wages can be regarded as a cost factor in the sense that they are value added to labor input, but can also be interpreted as a demand factor, since wages tend to have a close linkage with the overall inflation rate under an economic structure where they are determined by labor contracts. For example, real wages are defined as the moving average of the increasing rates of the consumer price index as in Batini and Haldane (1999). On the other hand, the indices of producer price, import price, and labor entrance/separation ratio, which reflect the cost structure of firms, are estimated to have relatively higher values of eigenvectors in the third principal component. This seems to be closely related to factors causing cost-push inflation.

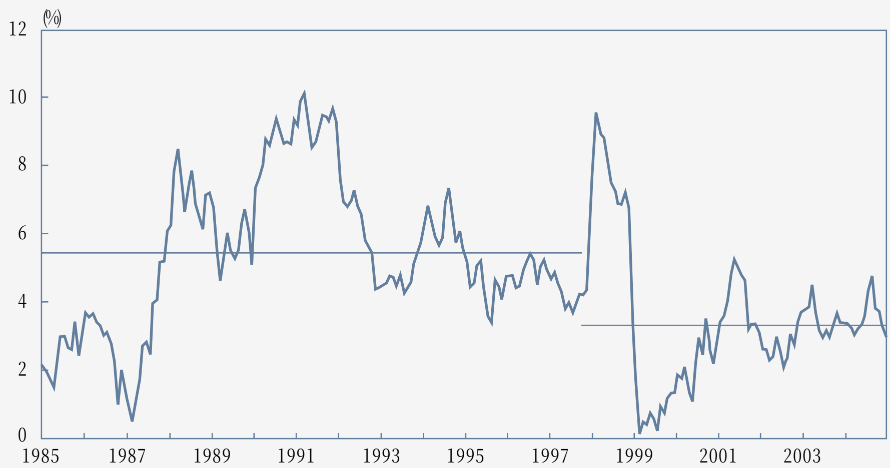
IV. Empirical Results

In this section, we will examine behavioral change in Korean inflation since the 1980s focusing on the consumer price index (CPI). We will compare various models′ inflation predictive power using the inflation equation set up in the previous section. The chief emphasis is placed on whether previous research from other countries is replicated in Korea – that is, whether an inflation forecasting model using information variables aggregating factors from 42 individual economic indicators from the real and financial-market sectors is superior to benchmark models.

1. Inflation Behavior in Korea

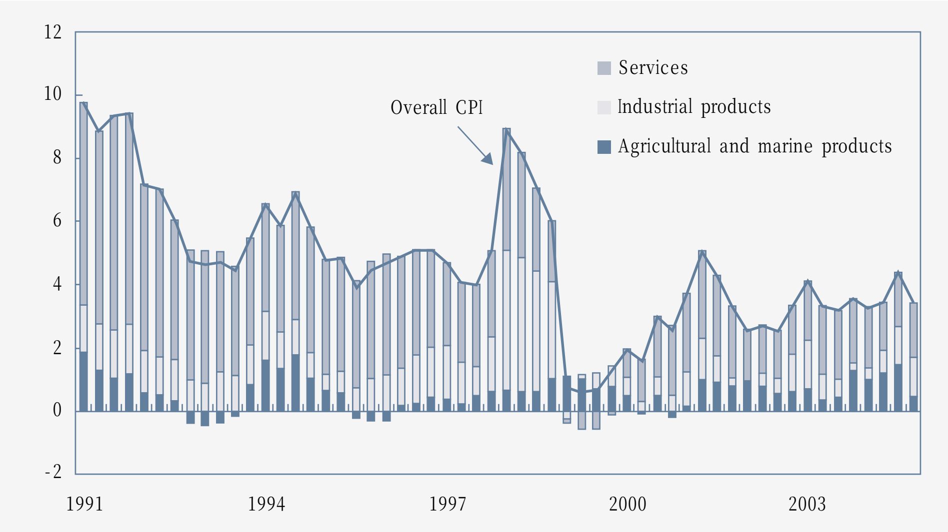
The Korean CPI's rate of increase was running in the double-digits during the late 1980s and early 1990s, mainly in response to growth-oriented economic policy and frequent bad harvests. On average, it remained around 5.5 percent a year until the financial crisis in the late 1990s. However, it fell to the 3.5 percent level after the financial crisis as a result of stabilization policies, which involved the introduction of inflation targeting, and intensified price competition through the expansion of market liberalization, reform of the distribution structure, and the proliferation of large-scale discount stores [Figure 1].

Figure 1 Trends of CPI Inflation in Korea
(Compared with the same month in the previous year)



Examination of the major characteristics of inflation by product groups demonstrates that the contribution of service prices to overall CPI inflation has increased in significance. [Figure 2] breaks down the overall rates of increase in the CPI into the contributions of agricultural, livestock, and marine products; industrial products; and services. We can see that movements of overall CPI inflation have been dominated by service CPIs (rents, public services and individual services) since the 1990s, when the importance of the service sector in the economy grew very rapidly. The CPI of industrial products, however, which had shown a relatively stable pattern of movements, acted to increase the volatility of overall CPI inflation because the scale of its contribution increased greatly for some time because of the massive deprecation of the Korean won following the financial crisis, although it subsequently narrowed again. On the other hand, the CPI of agricultural, livestock, and marine products demonstrated movements similar to CPI as a whole, contributing around 1 percentage point to inflation.

Figure 2 Trends of Product Groups' Contributions to Overall CPI Inflation in Korea
(Compared with the same quarter in the previous year)

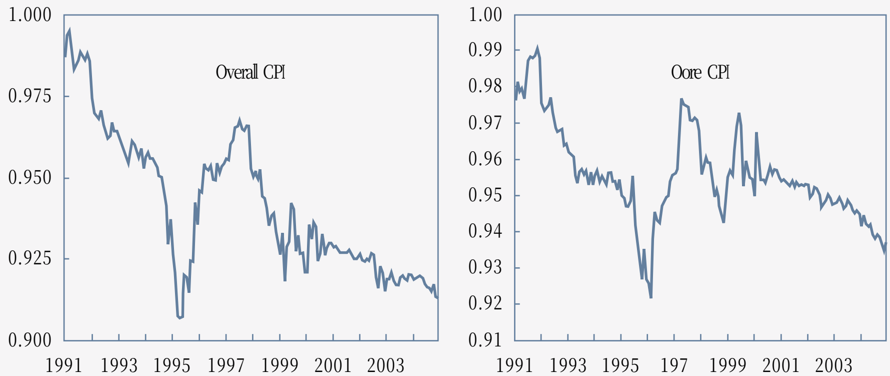


Levin and Piger (2003) reported that the persistence of inflation[7)] had recently dropped dramatically in major developed countries. Pagan (2003) addressed the difficulty of inflation forecasting as a result of the decrease in inflation persistence. Wilis (2003) attributed the decrease in inflation persistence to the increased flexibility of factors affecting production costs as a result of intensified global competition, advances in technology, and the enlarged weights of temporary and daily workers. We estimated the AR(3) model in Korea by a rolling regression, in which the monthly rates of increase in the overall CPI and core CPI[8)] are used as a dependent variable for a fixed 8-year window from the beginning of the 1980s, and autoregressive coefficients are summed to evaluate inflation persistence [Figure 3].

7) Inflation persistence means the extent of time required to return to baseline when inflation deviates from it due to the unexpected shocks.

8) There are various ways to estimate core CPI such as adjusting specific components or econometric method. In this paper, we use core CPI excluding agriculture and oil products (49 items in total, 11.7% percent in weights), compiled by the National Statistics Office and utilized as the official index in the medium term inflation targeting at the Bank of Korea.

Figure 3 Trends of Inflation Persistence[1] in Korea



Note: 1) Sum of autoregressive coefficients in AR(3) model for inflation estimated by a rolling regression for a fixed 8-year window.

We estimate that inflation persistence continued to demonstrate a general downward trend in Korea except during the mid-1990s, when it rose temporarily. The decrease is gradual, however, in comparison to developed countries, as the coefficients remained at a level higher than 0.9. This implies that the price structure in Korea is very rigid compared to industrialized countries in spite of efforts to stimulate the workings of the market mechanism. In other words, it can be concluded that the inflexible labor market structure resulting from immature market conditions caused firms' pricing power and the flexibility of the market mechanism to be lower than in advanced countries, contributing to the relatively high persistence of inflation in Korea.

2. An Outline of Empirical Analysis

We selected three models to compare the usefulness of information variables as is done in the papers by Atkeson and Ohanian (2001) and Kozicki (2001): (i) the AR model, (ii) the Factor model comprising information variables in the Phillips curve, and (iii) the Random Walk (RW) model, implying that the current conditional forecast for the future is equal to the present value. Based on equation (1), we can express these three models as follows.

$$\left.\begin{aligned}&\text{AR Model}: \hat{\pi}^h_{t+h} = \hat{\alpha} + \sum_{i=0}^{k} \hat{\phi}_i\, \pi_{t-i}\\&\text{Factor Model}: \hat{\pi}^h_{t+h} = \hat{\alpha} + \sum_{i=1}^{l} \hat{\beta}_i \hat{F}_i + \sum_{i=0}^{m} \hat{\phi}_i\, \pi_{t-i}\\&\text{RW Model}: \hat{\pi}^h_{t+h} = \pi_t\end{aligned}\right\} h = 6, 12, 18, 24 \text{ months} \qquad (5)$$

In the existing research, the AR model and RW model are widely used to evaluate the predictive power of the Factor model, because they have the following three characteristics. First, the AR model is based on pure time-series models of Box-Jenkins type, which are known to be very effective in short-term forecasting provided the time lag of the autoregressive and moving-average coefficients is appropriately set. Therefore, if the predictive power of the Factor model is better than that of an appropriately identified AR model, it could be concluded that information variables are useful in inflation forecasting. On the other hand, the RW model is the most naive type of forecasting, as it assumes that the current level of inflation equals the predicted level of inflation. If the predictive power of the Factor model is inferior to that of the RW model, it could be concluded that information variables are not useful in inflation forecasting. Meese and Rogoff (1983) showed that, for exchange rates that are highly volatile and affected by factors other than economic fundamentals, the predictive power of other models is not superior to that of the RW model. Furthermore, Atkeson and Ohanian (2001) asserted that, in the United States, the predictive power of both the AR model and of the Factor model is less than that of the RW model when measured over a specific period of time.

Since the main objective of empirical analysis lies in gauging the contribution of information variables in inflation forecasting, the ratio of predictive power between models could be a more relevant criterion of evaluation than the predictive power of individual models. For this reason, we employ the ratio of the mean squared forecast error (MSFE), widely used in out-of-sample forecasts, as the estimator of predictive power comparison. The AR model is used as the benchmark. If we define $\hat{\pi}^h_{b,t+h/t}$ and $\hat{\pi}^h_{a,t+h/t}$ as estimates of inflation forecasts of the benchmark model (in this case, the AR model) and the comparison model (the Factor or RW model) after the h-period, the comparison estimator could be written as follows.

$$\text{MSFE ratio} = \frac{\dfrac{1}{T_2 - T_1 - h + 1}\sum_{t=T_1}^{T_2-h} \left(\hat{\pi}^h_{t+h} - \hat{\pi}^h_{a,t+h/t}\right)^2}{\dfrac{1}{T_2 - T_1 - h + 1}\sum_{t=T_1}^{T_2-h} \left(\hat{\pi}^h_{t+h} - \hat{\pi}^h_{b,t+h/t}\right)^2} \qquad (6)$$

In the above equation, T_1 and T_{2-h} represent the starting and ending points of the out-of-sample period. Therefore, the predictive power of the various models is compared for the fixed period of $t = T_1+h, \cdots, T_2$. If the MSFE ratio is 1, it indicates that the AR model has the same predictive power as that of the model used for comparison. If the MSFE ratio is smaller (larger) than 1, the predictive power of the model used for comparison is superior (interior) to the benchmark. At every point, the optimum lag of the AR model (k in equation (5)), the optimum component (l in equation (6)), and the autoregressive coefficient (m in equation (6)) are determined by tests based on the Bayesian Information Criterion (BIC).

When we estimate the probability distribution to test the statistical significance of the MSFE ratio, we are not able to use traditional estimators like the simple standard deviation as the estimated coefficient, and the numbers of autoregressive coefficients and principal components change at every point of time in recursive regression. Therefore, we estimate the distribution of the MSFE ratio using the Bartlett kernel function as in West (1996), and the width of the kernel is set to be $h-1$ as in Stock and Watson (1999).

3. Results of Comparison of Inflation Prediction Power

The inflation equation is estimated for the period of January 1985 to December 2004 and the out-of-sample period forecast is utilized beginning in January 1992. For example, in a forecast horizon case where h is six, we make an inflation forecast for January 1992 based on the estimated coefficient for the period of January 1985 to July 1991. An out-of-sample forecast for February 1992 is made based on the estimated coefficient for the period of February 1985 to August 1991.[9] As we have examined in section IV.1, inflation in Korea experienced a structural change - decrease in level and in persistence - after the financial crisis. Therefore, we implement out-of-sample forecasts for the whole period and two sub-periods, with December 1997 as the mid-point.

[Table 3] shows the comparison results of out-of-sample inflation (overall CPI) forecasts for three models. For a forecast horizon within 12 months, the predictive power of the Factor model is clearly superior to that of the AR model, except for periods before 1997; for a forecast horizon over 12 months, however, the predictive powers of both models are about the same. It is noteworthy that

9) In this process, there exists a possibility that the importance of components may vary over time since *the weights aggregating information variables* are updated each time in the process of estimation for the Factor model, just as they are in the CFNAI estimated by principal component analysis.

after the financial crisis, the predictive power of the Factor model is superior regardless of the length of the forecast horizon, with the improvement being most significant for the 12-month horizon. Considering that Kim (2001) and Lee (2003) estimated that the optimal monetary policy horizon for Korea is around 1 year, these results imply the possibility of inflation targeting regime becoming firmly established.

Table 3 Comparison of Inflation Prediction Power for Overall CPI Inflation[1), 2)]

	Total period (92.1~04.12)	Before the financial crisis (92.1~97.12)	After the financial crisis (98.1~04.12)
<h = 6 months>			
Factor Model	0.949(0.02)	1.052(0.01)	0.882(0.04)
RW Model	10.505(1.22)	5.713(0.41)	13.620(2.17)
<h = 12 months>			
Factor Model	0.877(0.04)	1.050(0.01)	0.763(0.06)
RW Model	22.238(2.88)	13.860(1.36)	27.758(5.01)
<h = 18 months>			
Factor Model	1.012(0.01)	1.065(0.01)	0.977(0.01)
RW Model	23.222(2.95)	17.490(2.36)	27.001(4.88)
<h = 24 months>			
Factor Model	0.999(0.01)	1.059(0.01)	0.960(0.01)
RW Model	22.980(2.71)	18.820(2.97)	25.675(4.14)

Notes: 1) MSFE ratio relative to the AR Model.
2) Figures in parenthesis are standard errors according to West (1996).

On the other hand, in contrast to Atkeson and Ohanian (2001), the predictive power of inflation derived from the RW model is inferior to that from the AR model, regardless of the sample period or the forecast horizon. This seems to result from the fact that the volatility of Korean macroeconomic conditions including economic activity is much higher than that of the United States, and this phenomenon seems to be incorporated unchanged in the case of inflation as well. Therefore the predictive power of both the Factor model, which uses new variables derived from aggregating information variables, and the AR model, which derives inflation forecast from previous inflation behavior, are superior to that of the RW model, which assumes that the future forecast equals the current inflation rate.

[Table 4] compares the three models in predicting core CPI. The results were very similar to [Table 3], since core CPI has mirrored the movement of overall

CPI. Apart from some periods and forecast horizons, the ability of the Factor model to predict core CPI is superior to that of the AR model, and this trend becomes evident for the period after 1998. Since the values of the MSFE ratio in [Table 4] are less than those in [Table 3] for the same sample period and forecast horizon, the forecast accuracy of core CPI seems to be generally better than that of the overall CPI.

Table 4 Comparison of Inflation Prediction Power for Core CPI Inflation[1), 2)]

	Total period (92.1~04.12)	Before the financial crisis (92.1~97.12)	After the financial crisis (98.1~04.12)
<h = 6 months>			
Factor Model	0.923(0.01)	1.044(0.00)	0.839(0.02)
RW Model	16.089(0.74)	8.751(0.25)	21.185(1.31)
<h = 12 months>			
Factor Model	0.785(0.02)	0.950(0.02)	0.668(0.04)
RW Model	36.564(2.01)	19.583(0.76)	48.556(3.50)
<h = 18 months>			
Factor Model	0.990(0.01)	0.973(0.00)	1.003(0.01)
RW Model	39.085(2.08)	24.193(1.13)	49.668(3.47)
<h = 24 months>			
Factor Model	1.014(0.01)	1.041(0.01)	0.993(0.01)
RW Model	36.598(1.84)	27.666(1.60)	43.095(2.86)

Notes: 1) MSFE ratio relative to the AR Model.
2) Figures in parenthesis are standard errors according to West (1996).

V. Concluding Remarks

This paper examines the roles of information variables with regard to inflation forecasting. We selected 42 individual economic indicators from the real, external, price and financial-market sectors that are widely used in the process of inflation forecasting by monetary policy authorities, and formed a Factor model that uses information taken by way of principal component analysis. We then compared the usefulness of the Factor model with that of the AR model and the Random Walk (RW) model in terms of out-of-sample forecast.

Empirical results show that the Factor model is superior in predictability to the AR model regardless of inflation variables or sample periods, and that the forecast errors derived from the RW model seem to be excessive due to huge fluctuation in macro-economic variables in countries like Korea. Considering the usefulness of time-series models, such as the Box-Jenkins model, in short-term forecasting, the fact that the Factor model consistently outperforms the AR model in forecasting implies that information variables have a potential to play an important role in inflation forecasting. If we utilize information variables in a systematic manner, we could reduce inflation forecast errors significantly and contribute to enhancing the credibility of monetary policy.[10)]

These empirical results in Korea seem to be generally consistent with a monetary policy framework of inflation targeting. It is common knowledge that the Monetary Policy Committee of the Bank of Korea considers a wide range of economic variables including the real, financial and external sectors when they determine monetary policy every month. That the predictive power of the 12-month forecast was the most outstanding confirms that a pre-emptive monetary policy based on a highly precise inflation forecast should be implemented, considering that there exists a time lag of 6 to 24 months until monetary policy is transmitted to real sector variables such as inflation. Also, the fact that inflation forecasting utilizing core CPI inflation is better than that using headline CPI inflation augurs well for the Bank of Korea's adoption of core CPI inflation as its target inflation variable since 2000.

Although it can be said that the main contribution made by this paper lies in its testing of the usefulness of information variables in an econometric setting, more research should follow in order to consolidate the empirical findings. It cannot reasonably be said that the number of variables used in this study - forty two - is small, considering the restricted number of variables for which we have series available in a consistent manner. However, the number is fairly limited in comparison to earlier papers such as Stock and Watson (1999). More effort is needed to bring in the use of additional information variables and, in this regard, consideration should be given to the incorporation of real estate prices and household debt, both of which appear to have had an increasingly large impact on inflation in recent years. New variables need to be constructed and applied that relate closely to future inflation, such as inflation expectations, consumer expectation surveys, and other surveys widely employed by major central banks in advanced countries.

Finally, the development of a new system is called for to synthesize and analyze information variables systematically for inflation forecasting. In the short run, we need to capture the information by applying rather more

10) However, the systematic use of information variables is only a necessary condition for inflation forecasting. More efforts are called for in order to enhance the predictability of inflation models by taking general supply-demand conditions and price inflation into consideration.

sophisticated econometric methods, since information variables are intrinsically characterized by the coexistence of information on inflationary pressure and other destabilizing factors. Furthermore, it would be helpful to conduct research concerning the statistical threshold value for the assessment of inflationary pressure; this could be accomplished through an analysis of the empirical relationship between information derived from inflation variables and inflation itself, as modeled by the CFNAI of the Federal Reserve Bank of Chicago.[11)]

11) Refer to Federal Reserve Bank of Chicago (2001) for more more details on applying the CFNAI to assess inflationary pressures.

References

[In Korean]

Kim, Jong Wook, "Analysis of Changes in Asset Prices as Predictors of Inflation," *Monthly Bulletin*, Bank of Korea, pp. 25-47, July 2002.

Kim, Yang Woo, "Optimal Time Lag of Monetary Policy under Inflation Targeting," *Quarterly Economic Analysis* 7(3), Bank of Korea, pp.1-25, September 2001.

Lee, Joong Shik, "Identifying Business Cycle Turning Points with a New Index of Aggregate Economic Activity," *Quarterly Economic Analysis* 10(1), Bank of Korea, pp.107-134, March 2004.

Lee, Myong-Hwal, "Optimal Monetary Policy Rule under Inflation Targeting," *Financial Research* 17(1), Korea Finance Institute, pp. 1-23, June 2003.

Yoo, Chang Ho, and Eun Young Noh, "Studies on Predictors of Inflation," *Monthly Bulletin*, Bank of Korea, pp. 11-35, September 1998.

Yoo, Chang Ho, and Eun Young Noh, "Studies on the Relationship between Major Information Variables and Inflation," *Monthly Bulletin*, Bank of Korea, pp. 3-25, March 1999.

[In English]

Atkeson, A. and L. Ohanian, "Are Phillips Curves Useful for Forecasting Inflation?," *Quarterly Review* 25(1), FRB Minneapolis, pp. 2-11, 2001.

Batini, N. and A. Haldane, "Forward-looking Rules for Monetary Policy," in J. Taylor (ed.), *Monetary Policy Rules*, University of Chicago Press, pp.157-192, 1999.

Bernanke, B. and J. Boivin, "Monetary Policy in a Data-rich Environment," *Journal of Monetary Economics* 50, pp. 525-546, 2003.

Burns, A. F. and W.C. Mitchell, *Measuring Business Cycles*, National Bureau of Economic Research, New York, 1946.

Federal Reserve Bank of Chicago, *CFNAI Background Release*,(http://www.chicagofed.org/economicresearchanddata/national/pdffiles/CFNAIbga.pdf), 2001.

Fisher, Jonas, "Forecasting Inflation with a Lot of Data," *Chicago Fed Letter*, FRB Chicago, No. 151, 2000.

Fisher, J., C. Liu, and R. Zhou, "When Can We Forecast Inflation?," *Economic Perspectives* 26(1), FRB Chicago, pp. 32-44, 2002.

Forni M., M. Hallin, M. Lippi, and L. Reichlin, "Do Financial Variables Help

Forecasting Inflation and Real Activity in the Euro-area?," *Journal of Monetary Economics* 50, pp. 1243-1255, 2003.

Gosselin, M. and G. Tkacz, "Evaluating Factor Models: An Application to Forecasting Inflation in Canada," *Working Paper* 2001-18, Bank of Canada, November 2001.

Inoue, A. and L. Kilian, "In-sample or Out-of-sample Tests of Predictability: Which One Should We Use?," *Working Paper* 195, European Central Bank, November 2002.

Kitamura, T. and R. Koike, "The Effectiveness of Forecasting Methods Using Multiple Information Variables," *Monetary and Economic Studies*, Bank of Japan, pp. 105-143, February 2003.

Kozicki, S., "Why Do Central Banks Monitor So Many Inflation Indicators?," *Economic Review*, FRB Kansas City 86(3), pp. 5-42, 2001.

Levin, A. and J. Piger, "Is Inflation Persistence Intrinsic in Industrial Economies?," FRB St. Louis, *Working Paper* 2002-23, November 2003.

Meese, R. and K. Rogoff, "Empirical Exchange Rate Models of the Seventies: Do They Fit Out of Sample?," *Journal of International Economics* 14, pp. 3-24, 1983.

Pagan, A., "Report on Modelling and Forecasting at the Bank of England," *mimeo*, Bank of England, 2003.

Shintani, M., "Nonlinear Forecasting Analysis Using Diffusion Indexes: An Application to Japan," *Journal of Money, Credit, and Banking* 37(3), pp. 517-538, 2005.

Stock, J. and M. Watson, "New Indexes of Coincident and Leading Economic Indicators," *NBER Macroeconomics Annual*, pp. 351-394, 1989.

Stock, J. and M. Watson, "Forecasting Inflation," *Journal of Monetary Economics* 44, pp. 293-335, 1999.

Stock, J. and M. Watson, "Macroeconomic Forecasting Using Diffusion Indexes," *Journal of Business and Economic Statistics* 20(2), pp.147-162, 2002.

Stock, J. and M. Watson, "Forecasting Output and Inflation: The Role of Asset Prices," *Journal of Economic Literature* 41(3), pp.788-829, 2003.

West, K., "Asymptotic Inference about Predictive Ability," *Econometrica* 64, pp.1067-1084, 1996.

Wilis, J., "Implications of Structural Changes in the U.S. Economy for Pricing Behavior and Inflation Dynamics," *Economic Review*, FRB Kansas City 88(1), pp.5-27, 2003.