Auburn University
Department of Economics
Working Paper Series

Improving Forecast Accuracy of Financial Vulnerability: PLS Factor Model Approach

Hyeongwoo Kim† and Kyunghwan Ko *

†Auburn University; *Bank of Korea

AUWP 2019-03

Improving Forecast Accuracy of Financial Vulnerability: PLS Factor Model Approach*

Hyeongwoo Kim† Kyunghwan Ko‡
Auburn University Bank of Korea

April 2019

Abstract

We present a factor augmented forecasting model for assessing the financial vulnerability in Korea. Dynamic factor models often extract latent common factors from a large panel of time series data via the method of the principal components (PC). Instead, we employ the partial least squares (PLS) method that estimates *target specific* common factors, utilizing covariances between predictors and the target variable. Applying PLS to 198 monthly frequency macroeconomic time series variables and the Bank of Korea's Financial Stress Index (KFSTI), our PLS factor augmented forecasting models consistently outperformed the random walk benchmark model in out-of-sample prediction exercises in all forecast horizons we considered. Our models also outperformed the autoregressive benchmark model in short-term forecast horizons. We expect our models would provide useful early warning signs of the emergence of systemic risks in Korea's financial markets.

Keywords: Partial Least Squares; Principal Component Analysis; Financial Stress Index; Out-of-Sample Forecast; RRMSPE

JEL Classifications: C38; C53; C55; E44; E47; G01; G17

---

*Financial support for this research was provided by the Bank of Korea. The views expressed herein are those of the authors and do not necessariliy reflect the views of the Bank of Korea.

†Patrick E. Molony Professor, Department of Economics, Auburn University, 138 Miller Hall, Auburn, AL 36849. Tel: +1-334-844-2928. Fax: +1-334-844-4615. Email: gmmkim@gmail.com.

‡International Finance Research Team, International Department, The Bank of Korea, 39, Namdaemun-ro, Jung-gu, Seoul 04531, Tel: +82-2-759-5966. Email: kyunghwan.ko@bok.or.kr.

1 Introduction

Financial crises often come to a surprise realization with no systemic warnings. Furthermore, as Reinhart and Rogoff (2014) point out, harmful spillover effects on other sectors of the economy are likely to be severe because recessions followed by financial crises are often longer and deeper than other economic downturns. To avoid financial crises, Reinhart and Rogoff (2009) suggest to use an early-warning system (EWS) that alerts policy makers and financial market participants to incoming danger signs.

To design an EWS, it is crucially important to obtain a proper measure of the financial vulnerability that quantifies the potential risk in financial markets. One may consider the conventional Exchange Market Pressure (EMP) index proposed by Girton and Roper (1977). Instead, this paper employs an alternative measure known as financial stress index (FSTI) that is rapidly gaining popularity since the recent financial crisis.

The EMP index is computed using a small number of monetary variables such as exchange rate depreciations and changes in international reserves. On the other hand, FSTI is constructed utilizing a broad range of key financial market variables. In the US, 12 financial stress indices have currently become available (Oet, Eiben, Bianco, Gramlich, and Ong (2011)) since the recent financial crisis. The Bank of Korea also developed FSTI (KFSTI) in 2007 and started to report it on a yearly basis in their Financial Stability Report.

In this paper, we employ the monthly frequency KFSTI data as a proxy variable for financial market risk in Korea, and propose an out-of-sample forecasting procedure that extracts potentially useful predictive contents for KFSTI from a large panel of monthly frequency macroeconomic data.[1]

Conventional approaches to predict financial crises include the following. Frankel and Saravelos (2012) and Sachs, Tornell, and Velasco (1996) used linear regression approaches to test the statistical significance of various economic variables on the occurrence of historical crisis episodes. Others employed discrete choice models including parametric probit or logit models (Frankel and Rose (1996); Eichengreen, Rose, and Wyplosz (1995); Cipollini and Kapetanios (2009)) and nonparametric

[1] High frequency KFSTI data are for internal use only. We appreciate the Bank of Korea for giving permission to use the monthly frequency data.

signals approach (Kaminsky, Lizondo, and Reinhart (1998); Edison (2003); EI-Shagi, Knedlik, and von Schweinitz (2013); Christensen and Li (2014)).

Our forecasting procedure is different from these earlier studies in the sense that we extract potentially useful predictive contents for a new measure of the financial vulnerability such as the KFSTI from a broad range of macroeconomic time series data. Our proposed method is suitable in a data-rich environment, and may be considered as an alternative to dynamic factor models that are widely employed in the recent macroeconomic forecasting literature.

Since the influential work of Stock and Watson (2002), factor models often utilize principal components (PC) analysis to extract latent common factors from a large panel of predictor variables. Estimated factors, then, can be used to formulate forecasts of a target variable employing linear regressions of the target on estimated common factors. It should be noted that the PC method constructs common factors based solely on predictor variables.[2] Boivin and Ng (2006), however, pointed out that the performance of the PC method may be poor in forecasting the target variable if predictive contents are in a certain factor that may be dominated by other factors.

To overcome this issue, we employ the partial least squares (PLS) method that is proposed by Wold (1982). The method constructs *target specific* common factors from linear, orthogonal combinations of predictor variables taking the *covariance* between the target variable and predictor variables into account. Even though Kelly and Pruitt (2015) demonstrate that PC and PLS generate asymptotically similar factors when the data has a strong factor structure, Groen and Kapetanios (2016) show that PLS models outperform PC-based models in forecasting the target variable in the presence of a weak factor structure.

In this paper, we estimate multiple common factors using PLS from a large panel of 198 monthly frequency macroeconomic data in Korea and the KFSTI from October 2000 to June 2016. We apply PLS to the first *differenced* macroeconomic data and the KFSTI to avoid issues that are associated with nonstationarity in the data.[3] Then, we augment two types of benchmark models, the nonstationary

[2]Cipollini and Kapetanios (2009) employed the dynamic factor model via the PC method for their out-of-sample forecasting exercises for financial crisis episodes.

[3]Bai and Ng (2004) propose a similar method for their panel unit root test procedure that uses PC to estimate latent factors.

random walk (RW) and the stationary autoregressive (AR) models, with estimated PLS factors to out-of-sample forecast the KFSTI foreign exchange market index (KFSTI-FX) and the KFSTI stock market index (KFSTI-Stock).

We evaluate the out-of-sample forecast accuracy of our PLS-based models relative to these benchmark models using the ratio of the root mean squared prediction errors ($RRMSPE$) and the Diebold-Mariano-West (DMW) test statistics. We employed both the recursive (expanding window) method and the fixed-size rolling window method. Based on the $RRMSPE$ and the DMW statistics, our models consistently outperform the benchmark RW models in out-of-sample predictability in all forecast horizons we consider for up to one year. On the other hand, our models outperform the AR benchmark model only in short-term forecast horizons.

Financial market stability is viewed as an important objective of many central banks. To the best of our knowledge, the present paper is the first to predict the emergence of systemic risks in financial markets in Korea using PLS-based dynamic factor models.[4] We expect our models help provide useful early warning indicators of financial distress that may become prevalent in Korea's financial markets, resulting in harmful spillovers to other sectors of the economy.

The rest of the paper is organized as follows. Section 2 explains how we extract latent common factors and formulate out-of-sample forecasts using PLS factor-augmented forecasting models. We also describe our out-of-sample forecast strategies and model evaluation methods. In Section 3, we provide data descriptions and report our major empirical findings. Section 4 concludes.

2 The Econometric Method

2.1 The Method of the Principal Components

Consider a panel of N macroeconomic time series predictor variables, $\mathbf{x} = [\mathbf{x}_1, \mathbf{x}_2, ..., \mathbf{x}_N]$, where $\mathbf{x}_i = [x_{i,1}, x_{i,2}, ..., x_{i,T}]'$, $i = 1, ..., N$. Dynamic factor models that are based on the principal component (PC) method (e.g., Stock and Watson (2002)) assume

[4] Kim, Shi, and Kim (2016) implemented similar forecasting exercises using factor estimates from the PC method, which utilizes 198 predictor variables but not the target variable.

the following factor structure for $\mathbf{x}$. Abstracting from deterministic terms,

$$x_{i,t} = \boldsymbol{\lambda}_i^{'}\mathbf{f}_t + \varepsilon_{i,t}, \tag{1}$$

where $\mathbf{f}_t = [f_{1,t}, f_{2,t}, \cdots, f_{R,t}]^{'}$ is an $R \times 1$ vector of *latent* common factors at time t and $\boldsymbol{\lambda}_i = [\lambda_{i,1}, \lambda_{i,2}, \cdots, \lambda_{i,R}]^{'}$ denotes an $R \times 1$ vector of time-invariant associated factor loading coefficients. $\varepsilon_{i,t}$ is the idiosyncratic error term.

As shown by Nelson and Plosser (1982), most macroeconomic time series variables are better approximated by a nonstationary stochastic process. Further, Bai and Ng (2004) pointed out that the PC estimator for $\mathbf{f}_t$ from (1) may be inconsistent when $\varepsilon_{i,t}$ is an integrated process. As Bai and Ng (2004) suggested, one may estimate $\mathbf{f}_t$ and $\boldsymbol{\lambda}_i$ via the PC method for the first-differenced data. For this, rewrite (1) as follows.

$$\Delta x_{i,t} = \boldsymbol{\lambda}_i^{'}\Delta\mathbf{f}_t + \Delta\varepsilon_{i,t} \tag{2}$$

for $t = 2, \cdots, T$. After normalizing $\Delta\mathbf{x} = [\Delta\mathbf{x}_1, \Delta\mathbf{x}_2, ..., \Delta\mathbf{x}_N]$, we apply PC to $\Delta\mathbf{x}\Delta\mathbf{x}^{'}$ to obtain the factor estimates $\Delta\hat{\mathbf{f}}_t$ along with their associated factor loading coefficients $\hat{\boldsymbol{\lambda}}_i$.[5] Estimates for the idiosyncratic components are naturally given by the residuals $\Delta\hat{\varepsilon}_{i,t} = \Delta x_{i,t} - \hat{\boldsymbol{\lambda}}_i^{'}\Delta\hat{\mathbf{f}}_t$. Level variables are recovered as follows,

$$\hat{\varepsilon}_{i,t} = \sum_{s=2}^{t}\Delta\hat{\varepsilon}_{i,s}, \;\; \hat{\mathbf{f}}_t = \sum_{s=2}^{t}\Delta\hat{\mathbf{f}}_s \tag{3}$$

2.2 The Partial Least Squares Method

Partial least squares (PLS) models for a scalar target variable y_t are motivated by the following linear regression model. Abstracting from deterministic terms,

$$y_t = \Delta\mathbf{x}_t^{'}\boldsymbol{\beta} + u_t, \tag{4}$$

where $\Delta\mathbf{x}_t = [\Delta x_{1,t}, \Delta x_{2,t}, ..., \Delta x_{N,t}]'$ is an $N \times 1$ vector of predictor variables at time $t = 1, ..., T$, $\boldsymbol{\beta}$ is an $N \times 1$ vector of associated coefficients, and u_t is an error term. Note that we use the first-differenced predictor variables, assuming that $\mathbf{x}_t$ is a vector of integrated processes.

[5] We first normalize the data prior to estimations, because the method of the principal components is not scale invariant.

PLS models are useful especially when N is large. Instead of running a regression for (4), one may employ a data dimensionality reduction method via the following regression with an $R\times 1$ vector of components $\Delta\mathbf{c}_t = [\Delta c_{1,t}, \Delta c_{2,t}, ..., \Delta c_{R,t}]^{'}$, $R < N$ as follows,

$$
\begin{aligned}
y_t &= \Delta\mathbf{x}_t^{'}\mathbf{w}\boldsymbol{\theta} + u_t \\
&= \Delta\mathbf{c}_t^{'}\boldsymbol{\theta} + u_t
\end{aligned}
\tag{5}
$$

That is,

$$
\Delta\mathbf{c}_t = \mathbf{w}^{'}\Delta\mathbf{x}_t, \tag{6}
$$

and $\mathbf{w} = [\mathbf{w}_1, \mathbf{w}_2, ..., \mathbf{w}_R]$ is an $N\times R$ matrix of each column $\mathbf{w}_r = [w_{1,r}, w_{2,r}, ..., w_{N,r}]^{'}$, $r = 1, ..., R$, is an $N\times 1$ vector of weights on predictor variables for the r^{th} component or factor. $\boldsymbol{\theta}$ is an $R \times 1$ vector of PLS regression coefficients.

PLS regression minimizes the sum of squared residuals from the equation (5) for $\boldsymbol{\theta}$ instead of $\boldsymbol{\beta}$ in (4). It should be noted, however, that we do not directly utilize $\boldsymbol{\theta}$ in the present paper. In what follows, we employ a two-step forecasting method so that our models are comparable with the PC-based forecasting models. That is, we estimate $\Delta\mathbf{c}_t$ via the PLS method, then augment our benchmark forecasting model with PLS factor estimates for $\Delta\mathbf{c}_t$.

There are many available PLS algorithms (Andersson (2009)) that work well. Among others, one may use the algorithm proposed by Helland (1990) to forecast the j-period ahead target variable y_{t+j}, $j = 1, 2, .., k$. One may obtain these factors recursively as follows. First, $\Delta c_{1,j,t}$ is determined by the following linear combinations of the predictor variables in $\Delta\mathbf{x}_t$.

$$
\Delta\hat{c}_{1,j,t} = \sum_{i=1}^{N} w_{i,j,1}\Delta x_{i,t}, \tag{7}
$$

where the loading (weight) $w_{i,j,1}$ is given by $Cov(y_{t+j}, \Delta x_{i,t})$.

Next, regress y_{t+j} and $\Delta x_{i,t}$ on $\Delta\hat{c}_{1,j,t}$ to get residuals, $\tilde{y}_{t+j}$ and $\Delta\tilde{x}_{i,t}$, respectively. The second factor estimate $\Delta\hat{c}_{2,j,t}$ is then obtained similarly as in (7) with $w_{i,j,2} = Cov(\tilde{y}_{t+j}, \Delta\tilde{x}_{i,t})$. We repeat until the R^{th} factor $\Delta\hat{c}_{R,j,t}$ is obtained.

2.3 The PLS Factor Forecast Models

Our first PLS factor forecast model, the PLS-RW model, is motivated by a *nonstationary* random walk process augmented by $\Delta\hat{\mathbf{c}}_t$. Abstracting from deterministic terms,

$$y_{t+j}^{PLS_{RW}} = y_t + \boldsymbol{\gamma}_j^{'}\boldsymbol{\Delta}\hat{\mathbf{c}}_t + e_{t+j},\ j = 1, 2, .., k, \tag{8}$$

that is, when $\boldsymbol{\gamma}_j = \mathbf{0}$, y_t obeys the random walk (RW) process.[6]

Since the coefficient on y_t is fixed, we cannot use the unrestricted least squares estimator for (8). We resolve this problem by regressing $y_{t+j} - y_t$ on $\Delta\hat{\mathbf{c}}_t$ first to obtain the consistent estimate $\hat{\boldsymbol{\gamma}}_j$.[7] Adding y_t back to the fitted value, we obtain the following j-period ahead forecast for y_{t+j},

$$\widehat{y}_{t+j|t}^{PLS_{RW}} = y_t + \hat{\boldsymbol{\gamma}}_j^{'}\Delta\hat{\mathbf{c}}_t \tag{9}$$

The natural benchmark (BM) model of the PLS-RW model (8) is the following RW model.

$$y_{t+1}^{BM_{RW}} = y_t + \eta_{t+1}, \tag{10}$$

where e_{t+j} in (8) is a partial sum of the white noise process η_t, that is, $e_{t+j} = \sum_{s=1}^{j} \eta_{t+s}$. It should be noted that our PLS-RW model (8) nests this RW benchmark model (10) when $\boldsymbol{\gamma}_j = \mathbf{0}$. The j-period ahead forecast from this benchmark RW model is,

$$\widehat{y}_{t+j|t}^{BM_{RW}} = y_t \tag{11}$$

Our second PLS factor forecast model, the PLS-AR model, is motivated by a *stationary* AR(1)-type stochastic process augmented by PLS factor estimates $\boldsymbol{\Delta}\hat{\mathbf{c}}_t$. Abstracting from deterministic terms,

$$y_{t+j}^{PLS_{AR}} = \alpha_j y_t + \boldsymbol{\beta}_j^{'}\boldsymbol{\Delta}\hat{\mathbf{c}}_t + u_{t+j},\ j = 1, 2, .., k, \tag{12}$$

[6] Note that this specification is inconsistent with our earlier specification described in (4) that requires stationarity of the target variable y_t. Practically speaking, the random walk type models often perform well in forecasting persistent variables such as the KFSTI. Furthermore, it is often difficult to distinguish highly persistent or near unit root variables from stationary variables (observational equivalence). With these in mind, we employ two mutually exclusive stochastic processes described in (8) and (12). We thank the referee who pointed out this issue.

[7] That is, we assume that $y_{t+j} - y_t$ is stationary.

where α_j is less than one in absolute value for stationarity.

We again employ a *direct* forecasting approach by regressing the j-period ahead target variable (y_{t+j}) directly on the current period target variable (y_t) and the estimated factors ($\Delta\hat{\mathbf{c}}_t$). Note that (12) is an AR(1) process for $j = 1$ extended by covariates $\Delta\hat{\mathbf{c}}_t$. Applying the ordinary least squares (LS) estimator for (12), we obtain the following j-period ahead forecast for the target variable,

$$\widehat{y}_{t+j|t}^{PLS_{AR}} = \hat{\alpha}_j y_t + \hat{\boldsymbol{\beta}}_j^{'} \Delta\hat{\mathbf{c}}_t, \tag{13}$$

where $\hat{\alpha}_j$ and $\hat{\boldsymbol{\beta}}_j$ are the least squares coefficient estimates.

Naturally, the benchmark model for the PLS-AR (12) is the following stationary AR(1)-type or simply the AR model,

$$y_{t+j}^{BM_{AR}} = \alpha_j y_t + u_{t+j},\ j = 1, 2, .., k, \tag{14}$$

which relates y_{t+j} directly with the current value y_t. The j-period ahead forecast from this model is,

$$\hat{y}_{t+j}^{BM_{AR}} = \hat{\alpha}_j y_t, \tag{15}$$

where $\hat{\alpha}_j$ is obtained by regressing y_{t+j} directly on y_t as in (14).[8] Note that the PLS-AR model (12) nests the stationary benchmark model (14) when $\boldsymbol{\Delta}\hat{\mathbf{c}}_t$ does not contain any useful predictive contents for y_{t+j}, that is, $\boldsymbol{\beta}_j = 0$.

2.4 Out-of-Sample Forecast Strategies

We first implement out-of-sample forecast exercises employing a recursive (expanding window) scheme. After estimating PLS factors $\{\Delta\hat{\mathbf{c}}_t\}_{t=1}^{T_0}$ using the initial $T_0 < T$ observations, $\{y_t, \Delta x_{i,t}\}_{t=1}^{T_0}$, $i = 1, 2, ..., N$, we obtain the $j-$period ahead out-of-sample forecast for the target variable, y_{T_0+j} by (9) or (13). Then, we expand the data by adding one more observation, $\{y_t, \Delta x_{i,t}\}_{t=1}^{T_0+1}$, $i = 1, 2, ..., N$, and re-estimate $\{\Delta\hat{\mathbf{c}}_t\}_{t=1}^{T_0+1}$ which is used to formulate the next forecast, y_{T_0+j+1}. We repeat this until we forecast the last observation, y_T. We implement forecasting exercises under this expanding window scheme for up to 12-month forecast horizons, $j = 1, 2, ..., 12$.

[8] One may employ a *recursive* approach with an AR(1) model, $y_{t+1} = \alpha y_t + \varepsilon_{t+1}$. Given the estimate of the persistence parameter, one may formulate the j-period ahead forecast by $\hat{\alpha}^j y_t$.

We also employ a fixed-size rolling window method, which performs better than the recursive method in the presence of structural breaks. After we obtain the first forecast y_{T_0+j} using the initial $T_0 < T$ observations, $\{y_t, \Delta x_{i,t}\}_{t=1}^{T_0}$, $i = 1, 2, ..., N$, we add one observation but drop one earliest observation for the next round forecasting. That is, we re-estimate $\{\Delta \hat{\mathbf{c}}_t\}_{t=2}^{T_0+1}$ from $\{y_t, \Delta x_{i,t}\}_{t=2}^{T_0+1}$, $i = 1, 2, ..., N$, maintaining the same number of observations (T_0) to obtain the second round forecast, y_{T_0+j+1}. Again, we repeat until we forecast the last observation, y_T.

For model evaluations regarding the out-of-sample prediction accuracy, we use the ratio of the root mean square prediction error ($RRMSPE$) defined as follows,

$$RRMSPE(j) = \frac{\sqrt{\frac{1}{T-T_0-j}\sum_{t=T_0+j}^{T}\left(\varepsilon_{t+j|t}^{BM_m}\right)^2}}{\sqrt{\frac{1}{T-T_0-j}\sum_{t=T_0+j}^{T}\left(\varepsilon_{t+j|t}^{PLS_m}\right)^2}},\ m = AR, RW, \tag{16}$$

where

$$\varepsilon_{t+j|t}^{BM_m} = y_{t+j} - \widehat{y}_{t+j|t}^{BM_m},\ \varepsilon_{t+j|t}^{PLS_m} = y_{t+j} - \widehat{y}_{t+j|t}^{PLS_m} \tag{17}$$

Note that our PLS models outperform the benchmark models when $RRMSPE$ is greater than 1.[9]

We supplement our analyses by employing the Diebold-Mariano-West (DMW) test. For this, we define the following loss differential function,

$$d_t = (\varepsilon_{t+j|t}^{BM_m})^2 - (\varepsilon_{t+j|t}^{PLS_m})^2,\ m = AR, RW, \tag{18}$$

where the squared loss function can be replaced with the absolute value loss function.

The DMW statistic is defined as follows to test the null of equal predictive accuracy, that is, $H_0 : Ed_t = 0$,

$$DMW(j) = \frac{\bar{d}}{\sqrt{\widehat{Avar}(\bar{d})}}, \tag{19}$$

where $\bar{d}$ is the sample average, $\bar{d} = \frac{1}{T-T_0-j}\sum_{t=T_0+j}^{T} d_t$. In the presence of serial

[9] We also employed a similar approach with the ratio of the root absolute mean square prediction error (RAMSPE). That is, the loss function is defined with the absolute value instead of the squared value. Results are qualitatively similar and available upon requests from authors.

correlations, $\widehat{Avar}(\bar{d})$ denotes the long-run variance of $\bar{d}$,

$$\widehat{Avar}(\bar{d}) = \frac{1}{T - T_0} \sum_{i=-q}^{q} k(i, q)\hat{\mathbf{\Gamma}}_i, \tag{20}$$

where $k(\cdot)$ is a kernel function with the bandwidth parameter q, and $\hat{\Gamma}_i$ is the i^{th} autocovariance function estimate.

3 Empirical Findings

3.1 Data Descriptions

We employ the financial stress index (KFSTI) data to quantify the financial vulnerability in Korea. The Bank of Korea introduced the index in 2007 and report KFSTI on a yearly basis in their Financial Stability Report. We obtained monthly frequency data, which in principle are for internal use only.[10] The data is available from May 1995, but our sample period covers from October 2000 until August 2016 to obtain a large panel of predictor variables.

We use the following two KFSTI sub-indices, one for the foreign exchange market (KFSTI-FX) and the other one for the stock market (KFSTI-Stock). We do not report forecasting exercise results for the two other KFSTI sub-indices for the bond market and for the financial industry, since our model performed relatively poorly for these two indices. Such limited performances of our factor models might be due to the fact that our common factors are extracted only from macroeconomic variables even though the financial industries and bond markets are often influenced by non-economic political factors.

Figure 1 provides graphs of the KFSTI-FX and the KFSTI-Stock. We note that both indices exhibit a sharp spike during the recent financial crisis that began in 2008. KFSTI-Stock exhibits more frequent turbulent periods in comparison with dynamics of the KFSTI-FX.

Figure 1 around here

[10] We obtained permission from the Bank of Korea to use the data for this research.

We obtained 198 predictor variables from the Bank of Korea. Observations are monthly frequency and span from October 2000 to August 2016. All variables other than those in percent (e.g., interest rates and unemployment rates) are log-transformed prior to estimations. We categorized these 198 time series data into 13 groups as summarized in Table 1.

Group #1 includes 14 domestic and world nominal interest rates. Groups #2 through #4 are an array of prices and monetary aggregate variables, while group #5 consist of bilateral nominal exchange rates. That is, groups #1 through #5 represent nominal sector variables in Korea. On the other hand, groups #6 through #11 entail various kinds of real activity variables such as production, inventory, and labor market variables. The last two groups represent business condition indices and stock market indices in Korea, respectively.

Table 1 around here

3.2 Evaluations of the Model

This subsection discusses the in-sample fit and the out-of-sample prediction performance of our PLS factor models relative to those of the benchmark and PC factor models.

3.2.1 In-Sample Fit Analysis

Figure 2 reports estimated *level* PC factors, $\hat{\mathbf{f}}_t = \sum_{s=2}^{t} \Delta\hat{\mathbf{f}}_s$, for up to 6 factors, along with their associated factor loading coefficient estimates ($\hat{\boldsymbol{\lambda}}$). In Figures 3 and 4, we report *level* PLS factors $\hat{\mathbf{c}}_t = \sum_{s=2}^{t} \Delta\hat{\mathbf{c}}_s$ for the KFSTI-FX and the KFSTI-Stock, respectively, and their weight matrix estimates ($\hat{\mathbf{w}}$). Note that we report two sets of PLS factors whereas only one set of PC factors is presented. This is because the PLS method utilizes the covariance between the predictor variables and the target variable, whereas the PC method does not consider the target variable when it extracts the common factors.

We noticed that PC factors are very different from PLS factors for each KF-STI index. Further, we note that $\hat{\boldsymbol{\lambda}}$ estimates are very different from $\hat{\mathbf{w}}$, meaning that PLS and PC factor estimates are obtained from utilizing different combinations of the predictor variables $\mathbf{x}$. Since we are mainly interested in out-of-sample predictability performances of the PLS method relative other models, we do not attempt to trace the sources of these factors. However, distinct factor estimates from the PLS and the PC methods imply that the performance of these methods would differ in out-of-sample forecasting exercises we report in what follows.

Figures 2, 3, and 4 around here

We also report R^2 values in Figure 5, obtained from LS regressions of the target variable y_t on estimated factors, $\Delta\hat{\mathbf{c}}_t$ and $\Delta\hat{\mathbf{f}}_t$, for up to 12 factors. Not surprisingly, PLS factors provide much better in-sample fit performance than PC factors, because $\Delta\hat{\mathbf{c}}_t$ is estimated using the covariance between the target and the predictor variables. For example, R^2 from $\Delta\hat{c}_1$ is over 0.3, whereas that from $\Delta\hat{f}_1$ is slightly over 0.02 for the KFSTI-FX. In the case of the KFSTI-Stock, R^2 from $\Delta\hat{c}_1$ is about 0.2, while $\Delta\hat{f}_1$ virtually has no explanatory power.

Note that $\Delta\hat{f}_{10}$ and $\Delta\hat{f}_2$ have the highest R^2 for the KFSTI-FX and for the KFSTI-Stock, respectively, whereas contributions of PLS factors are the highest for the first factor estimate $\Delta\hat{c}_1$. That is, marginal R^2 decreases when we regress the target variable to the next PLS factors. This is because we extract *orthogonal* PLS factors sequentially, utilizing the *remaining* covariances of the target and the predictor variables. Since the PC method uses only the predictor variables without considering the target variable, marginal R^2 values do not necessarily decrease. Cumulative R^2 value with up to 12 PLS factors is about 0.8 for both indices, whereas that with PC factors is less than 0.3 and 0.2 for the foreign exchange index and the stock index, respectively. In a nutshell, the PLS method yields superior in-sample fit performance in comparison with the PC method.

Figure 5 around here

3.2.2 Out-of-Sample Forecasting Performance

In Tables 2 and 3, we report *RRMSPE*'s and the DMW statistics of the PLS-RW forecasting model (9) relative to the performance of the RW benchmark model (11) for the KFSTI-FX and the KFSTI-Stock, respectively. We implement out-of-sample forecast exercises using up to 12 (k) factor estimates obtained from PLS for $\{y_{t+j}, \Delta x_{i,t}\}$ for up to 12-month forecast horizons (h). We used $p_{50\%}$ for the sample split point, that is, initial 50% observations were used to formulate the first out-of-sample forecast in implementing forecasting exercises via the recursive (expanding window) scheme as well as the fixed-size rolling window scheme.

Most *RRMSPE* values are strictly greater than 1, and the DMW test rejects the null of equal predictability favoring our factor models. That is, our PLS-RW model consistently outperforms the RW benchmark model in all forecast horizons and in both the recursive and the rolling window method. It should be noted that we use critical values from McCracken (2007) instead of the asymptotic critical values from the standard normal distribution, because the PLS-RW model nests the RW benchmark model.[11]

Tables 2 and 3 around here

Tables 4 and 5 report the forecasting performance of the PLS-AR model (13) relative to the AR benchmark model (15). Results sharply contrast with earlier results reported in Tables 2 and 3. The PLS-AR model outperforms the AR model only in the short-term forecast horizons. More specifically, the PLS-AR model outperforms the AR model in 1-month ahead out-of-sample forecast for the KFSTI-FX under the recursive forecasting scheme, while the AR model performs better in most other cases. The PLS-AR model performs relatively better for the KFSTI-Stock, as *RRMSPE* values are greater than 1 at least in one-month ahead forecast for the index under the both schemes.

Even though the performance of the PLS-AR model relative to the AR benchmark is not overwhelmingly good, it should be noted that the PLS-AR model can still provide useful early warning indicators of incoming danger to Korea's financial market. Financial crises often occur abruptly and unexpectedly. Given such

[11] Asymptotic critical values are not valid when one model nests the other model.

tendency, it is good to have an instrument that generates warning signs before the systemic risks materialize in the financial market.

Tables 4 and 5 around here

We repeat the same exercises using combinations of $\Delta\hat{\mathbf{c}}_t$ and $\Delta\hat{\mathbf{f}}_t$ and report the results in Tables 6 through 9. That is, we extended the benchmark forecasting models using equal numbers of factors obtained from the PLS and the PC methods. For example, $k = 4$ means that $\Delta\hat{c}_1$, $\Delta\hat{c}_2$, $\Delta\hat{f}_1$, and $\Delta\hat{f}_2$ are used as condensed predictor variables. Results are qualitatively similar to previous performances reported in Tables 2 through 5. That is, marginal contributions of using PC factors ($\Delta\hat{\mathbf{f}}_t$) in addition to PLS factors ($\Delta\hat{\mathbf{c}}_t$) are mostly negligibly small.

Tables 6 through 9 around here

3.2.3 Comparisons with the PC Models

This sub-section compares the out-of-sample prediction performances of the PLS models relative to those of the PC models using the *RRMSPE* criteria, the *RMSPE* from the PLS model divided by the *RMSPE* from the corresponding PC model. That is, *RRMSPE* greater than 1 implies a better performance of the PLS model.

As can be seen in Figure 6 for the KFSTI-FX, the PLS-RW model outperforms the PC-RW model in all forecast horizons we consider. It is interesting to see that the PLS-RW model's relative performance becomes better as we employ more factor estimates or when forecast horizons become longer. On the other hand, we observed qualitatively similar performance of the PLS-AR model and the PC-AR model in predicting the KFSTI-FX, even though the PLS-AR model tend to perform better in short-term forecast horizons with many factor estimates.

Figure 6 around here

The PLS-RW model again demonstrates substantially better performance than the PC-RW model in predicting the KFSTI-Stock in all forecast horizons under both the recursive and the fixed-size rolling window schemes. Interestingly, the PC-AR model overall outperforms the PLS-AR model for the KFSTI-Stock under the recursive scheme, while the latter outperforms the former under the fixed-size rolling window scheme. This seems to explain slight improvements in forecasting performance, see Tables 5 and 9, under the recursive scheme when we combine PLS and PC factors together.

Figure 7 around here

Lastly, we compare the performances of the PLS-AR model and the PLS-RW model using the *RRMSPE* criteria. *RRMSPE* greater than 1 implies that the PLS-AR model outperforms the PLS-RW model. Results are reported in Figure 8. It should be noted that both PLS models perform similarly well in short-term forecast horizons unless very small numbers of factors are employed. However, as the forecast horizon increases, the PLS-AR model tend to outperform the PLS-RW model. Note that the PLS-RW is based on the RW model, which is a "no change" prediction model. If the KFSTI obeys a mean reverting stochastic process, RW type models would not perform well in long-term forecast horizons. To check this possibility, we employed the conventional ADF test, which rejected the null of nonstationarity at the 5% significance level for both indices, confirming the conjecture described earlier.[12]

Figure 8 around here

4 Concluding Remarks

This paper proposes a factor-augmented forecasting model for the systemic risks in Korea's financial markets using the partial least squares (PLS) method as an alternative to the method of the principal components (PC). Unlike PC factor models that

[12] Results are available upon requests.

estimate common factors solely from predictor variables, the PLS approach generates the *target specific common factors* utilizing covariances between the predictors and the target variable.

Taking the Bank of Korea's Financial Stress Index (KFSTI) as a proxy variable of the financial vulnerability in Korea, we applied PLS to a large panel of 198 monthly frequency macroeconomic variables and the KFSTI from October 2000 to June 2016. Obtaining PLS common factors, we augmented the two benchmark models, the random walk (RW) model and the stationary autoregressive (AR) type model, with estimated PLS factors to out-of-sample forecast the KFSTI for the foreign exchange market and the stock market. We then implemented an array of out-of-sample prediction exercises using the recursive (expanding window) and the fixed-size rolling window schemes for 1-month to 1-year forecast horizons.

We evaluate our proposed PLS factor-augmented forecasting models via the ratio of the root mean squared prediction error and the Diebold-Mariano-West statistics. Our PLS-RW models consistently outperform the nonstationary random walk benchmark model. On the other hand, the PLS-AR forecasting models perform better than the AR models only for short-term forecast horizons. That is, unlike the PLS-RW model, the performance of the PLS-AR model is not overwhelmingly better than its benchmark. However, it should be noted that the PLS-AR model, and of course the PLS-RW model, can still provide potentially useful early warning signs of financial distress before the systemic risks materialize in Korea's financial market within a month. Combining all together, the PLS factor models perform much better than the PC factor models especially when the models are combined with the nonstationary random walk benchmark model.

References

ANDERSSON, M. (2009): "A Comparison of Nine PLS1 Algorithms," *Journal of Chemometrics*, 23, 518–529.

BAI, J., AND S. NG (2004): "A PANIC Attack on Unit Roots and Cointegration," *Econometrica*, 72(4), 1127–1177.

BOIVIN, J., AND S. NG (2006): "Are more data always better for factor analysis?," *Journal of Econometrics*, 132(1), 169–194.

CHRISTENSEN, I., AND F. LI (2014): "Predicting Financial Stress Events: A Signal Extraction Approach," Working Paper 2014-37. Bank of Canada.

CIPOLLINI, A., AND G. KAPETANIOS (2009): "Forecasting Financial Crises and Contagion in Asia using Dynamic Factor Analysis," *Journal of Empirical Finance*, 16(2), 188–200.

EDISON, H. J. (2003): "Do indicators of financial crises work? An evaluation of an early warning system.," *International Journal of Finance and Economics*, 8 (1), 11–53.

EI-SHAGI, M., T. KNEDLIK, AND G. VON SCHWEINITZ (2013): "Predicting Financial Crises: The (Statistical) Significance of the Signals Approach," *Journal of International Money and Finance*, 35, 75–103.

EICHENGREEN, B., A. K. ROSE, AND C. WYPLOSZ (1995): "Exchange Market Mayhem: The Antecedents and Aftermath of Speculative Attacks," *Economic Policy*, 10 (21), 249–312.

FRANKEL, J., AND G. SARAVELOS (2012): "Can leading indicators assess country vulnerability? Evidence from the 2008–09 global financial crisis," *Journal of International Economics*, 87, 216 –231.

FRANKEL, J. A., AND A. K. ROSE (1996): "Currency crashes in emerging markets: an empirical treatment," *Journal of International Economics*, 41 (3/4), 351–366.

Girton, and Roper (1977): "A Monetary Model of Exchange Market Pressure Applied to the Postwar Canadian Experience," *American Economic Review*, 67, 537–548.

Groen, J., and G. Kapetanios (2016): "Revisiting useful approaches to data-rich macroeconomic forecasting," *Computational Statistics and Data Analysis*, 100, 221–239.

Helland, I. S. (1990): "Partial Least Squares Regression and Statistical Models," *Scandinavian Journal of Statistics*, 17, 97–114.

Kaminsky, G., S. Lizondo, and C. Reinhart (1998): "Leading indicators of currency crises," IMF Working Paper No. 45 International Monetary Fund.

Kelly, B., and S. Pruitt (2015): "The three-pass regression filter: A new approach to forecasting using many predictors," *Journal of Econometrics*, 186(2), 294–316.

Kim, H., W. Shi, and H. H. Kim (2016): "Forecasting Financial Stress Indices in Korea: A Factor Model Approach," Auburn Economics Working Paper Series auwp2016-10, Department of Economics, Auburn University.

McCracken, M. W. (2007): "Asymptotics for out of sample tests of Granger causality," *Journal of Econometrics*, 140, 719–752.

Nelson, C. R., and C. I. Plosser (1982): "Trends and random walks in macroeconmic time series : Some evidence and implications," *Journal of Monetary Economics*, 10(2), 139–162.

Oet, M. V., R. Eiben, T. Bianco, D. Gramlich, and S. J. Ong (2011): "The Financial Stress Index: Identification of System Risk Conditions," Working Paper No. 1130, Federal Reserve Bank of Cleveland.

Reinhart, C. M., and K. S. Rogoff (2009): *This Time Is Different: Eight Centuries of Financial Folly*, vol. 1 of *Economics Books*. Princeton University Press.

——— (2014): "Recovery from Financial Crises: Evidence from 100 Episodes," *American Economic Review Papers and Proceedings*, 104(5), 50–55.

Sachs, J., A. Tornell, and A. Velasco (1996): "Financial Crises in Emerging Markets: The Lessons from 1995," *Brookings Papers on Economic Activity*, 27(1), 147–199.

Stock, J. H., and M. W. Watson (2002): "Macroeconomic Forecasting using Diffusion Indexes," *Journal of Business and Economic Statistics*, 20(2), 147–162.

Wold, H. (1982): *Soft modelling: the basic design and some extensions*, vol. 1 of *Systems under indirect observation, Part II*. North-Holland, Amsterdam.

Table 1. Macroeconomic Data Descriptions

Group ID	Data ID	Data Descriptions
#1	1-14	Domestic and World Interest Rates
#2	15-35	Exports/Imports Prices
#3	36-54	Producer/Consumer/Housing Prices
#4	55-71	Monetary Aggregates
#5	72-83	Bilateral Exchange Rates
#6	84-110	Manufacturers'/Construction New Orders
#7	111-117	Manufacturers' Inventory Indices
#8	118-135	Housing Inventories
#9	136-157	Sales and Capacity Utilizations
#10	158-171	Unemployment/Employment/Labor Force Participation
#11	172-180	Industrial Production Indices
#12	181-186	Business Condition Indices
#13	187-198	Stock Indices

Table 2. PLS-RW vs. RW: Foreign Exchange Market

$$\widehat{y}^{PLS_{RW}}_{t+j|t} = y_t + \hat{\gamma}_j^{'} \Delta \mathbf{c}_t \text{ vs. } \widehat{y}^{BM_{RW}}_{t+j|t} = y_t$$

		Recursive		Rolling Window				Recursive		Rolling Window	
k	h	*RRMSPE*	*DMW*	*RRMSPE*	*DMW*	k	h	*RRMSPE*	*DMW*	*RRMSPE*	*DMW*
1	1	0.924	-1.447	**1.028**	**1.327**	6	1	0.993	-0.090	**1.057**	**1.306**
	2	**1.006**	0.130	**1.052**	**1.854**		2	**1.040**	**0.416**	**1.090**	**2.014**
	4	**1.030**	0.498	**1.047**	**1.607**		4	**1.172**	**1.548**	**1.145**	**3.485**
	6	**1.157**	**1.166**	**1.056**	**1.340**		6	**1.336**	**2.188**	**1.201**	**3.386**
	9	**1.291**	**3.760**	**1.050**	**1.278**		9	**1.328**	**3.126**	**1.151**	**2.753**
	12	**1.377**	**2.142**	**1.046**	**0.959**		12	**1.544**	**2.012**	**1.155**	**2.657**
2	1	0.960	-0.558	**1.019**	**0.806**	8	1	0.985	-0.171	**1.084**	**2.556**
	2	0.990	-0.143	0.979	-0.485		2	**1.049**	**0.462**	**1.088**	**1.667**
	4	**1.086**	**1.093**	**1.035**	**1.032**		4	**1.238**	**1.999**	**1.196**	**3.536**
	6	**1.159**	**1.360**	**1.056**	**1.551**		6	**1.295**	**1.956**	**1.169**	**3.081**
	9	**1.215**	**3.008**	**1.052**	**1.303**		9	**1.356**	**3.978**	**1.217**	**3.783**
	12	**1.360**	**2.276**	**1.060**	**1.193**		12	**1.470**	**2.016**	**1.189**	**2.949**
4	1	0.964	-0.456	**1.033**	**0.640**	10	1	0.992	-0.092	**1.038**	**0.972**
	2	**1.031**	0.382	**1.059**	**1.363**		2	**1.077**	**0.708**	**1.055**	**0.929**
	4	**1.111**	**1.184**	**1.128**	**3.594**		4	**1.290**	**2.241**	**1.135**	**1.803**
	6	**1.281**	**2.139**	**1.213**	**3.176**		6	**1.330**	**2.149**	**1.072**	**1.259**
	9	**1.337**	**4.173**	**1.171**	**3.116**		9	**1.356**	**3.626**	**1.182**	**3.059**
	12	**1.550**	**2.281**	**1.132**	**2.395**		12	**1.572**	**2.391**	**1.213**	**2.990**

Note: *RRMSPE* denotes the ratio of the root mean squared prediction errors, which is the mean squared prediction error (*RMSPE*) from the benchmark model divided by the *RMSPE* from the competing Partial Least Squares factor model. DMW is the Diebold-Mariano-West statistics. We repeat estimations and forecasting starting from the first 50% observations until we (out-of-sample) forecast the last observation of the KFSTI. DMW statistics in bold denote the rejection of the null hypothesis of equal predictability at the 5% significance level in favor of our factor models. The critical values are from McCracken (2007) to avoid size distortion because the benchmark model is nested by our factor model.

Table 3. PLS-RW vs. RW: Stock Market

$$\widehat{y}^{PLS_{RW}}_{t+j|t} = y_t + \hat{\gamma}_j^{'}\Delta\mathbf{c}_t \text{ vs. } \widehat{y}^{BM_{RW}}_{t+j|t} = y_t$$

		Recursive		**Rolling Window**				**Recursive**		**Rolling Window**	
k	h	*RRMSPE*	*DMW*	*RRMSPE*	*DMW*	k	h	*RRMSPE*	*DMW*	*RRMSPE*	*DMW*
1	1	0.992	-1.241	0.993	-3.170	6	1	**1.045**	**1.227**	**1.127**	**3.155**
	2	0.992	-0.980	**1.009**	**0.933**		2	**1.078**	**1.242**	**1.271**	**3.553**
	4	0.992	-0.544	**1.010**	**0.730**		4	**1.111**	**1.763**	**1.334**	**3.175**
	6	**1.016**	**0.888**	**1.003**	0.190		6	**1.107**	**2.038**	**1.333**	**3.277**
	9	**1.024**	0.703	**1.007**	0.244		9	**1.114**	**1.245**	**1.341**	**2.697**
	12	**1.017**	0.610	**1.010**	0.381		12	**1.107**	**1.836**	**1.338**	**3.240**
2	1	**1.020**	**1.362**	**1.058**	**2.802**	8	1	**1.052**	**1.249**	**1.137**	**2.891**
	2	**1.019**	**0.765**	**1.089**	**2.423**		2	**1.064**	**0.932**	**1.282**	**3.194**
	4	**1.015**	0.571	**1.128**	**2.775**		4	**1.104**	**1.550**	**1.317**	**3.004**
	6	**1.047**	**1.554**	**1.119**	**2.666**		6	**1.121**	**2.048**	**1.337**	**3.262**
	9	**1.032**	**0.702**	**1.137**	**2.051**		9	**1.114**	**1.184**	**1.331**	**2.630**
	12	**1.011**	0.411	**1.091**	**2.004**		12	**1.102**	**1.703**	**1.377**	**3.179**
4	1	**1.022**	**0.631**	**1.132**	**4.068**	10	1	**1.097**	**1.534**	**1.147**	**3.248**
	2	**1.056**	**0.878**	**1.253**	**3.608**		2	**1.060**	**0.812**	**1.280**	**3.076**
	4	**1.065**	**1.079**	**1.314**	**3.460**		4	**1.125**	**1.762**	**1.321**	**2.787**
	6	**1.099**	**1.839**	**1.304**	**3.391**		6	**1.126**	**2.075**	**1.384**	**3.098**
	9	**1.126**	**1.340**	**1.419**	**3.242**		9	**1.134**	**1.312**	**1.358**	**2.591**
	12	**1.132**	**2.207**	**1.294**	**2.982**		12	**1.147**	**2.330**	**1.482**	**3.476**

Note: *RRMSPE* denotes the ratio of the root mean squared prediction errors, which is the mean squared prediction error (*RMSPE*) from the benchmark model divided by the *RMSPE* from the competing Partial Least Squares factor model. DMW is the Diebold-Mariano-West statistics. We repeat estimations and forecasting starting from the first 50% observations until we (out-of-sample) forecast the last observation of the KFSTI. DMW statistics in bold denote the rejection of the null hypothesis of equal predictability at the 5% significance level in favor of our factor models. The critical values are from McCracken (2007) to avoid size distortion because the benchmark model is nested by our factor model.

Table 4. PLS-AR vs. AR: Foreign Exchange Market

$$\widehat{y}^{PLS_{AR}}_{t+j|t} = \hat{\alpha}_j y_t + \hat{\beta}^{\prime}_j \Delta \mathbf{c}_t \text{ vs. } \widehat{y}^{BM_{AR}}_{t+j|t} = \hat{\alpha}_j y_t$$

		Recursive		**Rolling Window**				**Recursive**		**Rolling Window**	
k	*h*	*RRMSPE*	*DMW*	*RRMSPE*	*DMW*	*k*	*h*	*RRMSPE*	*DMW*	*RRMSPE*	*DMW*
1	1	**1.048**	**1.721**	0.997	-0.320	6	1	**1.022**	**0.602**	0.953	-3.801
	2	**1.000**	0.002	0.973	-1.839		2	0.989	-0.371	0.944	-2.565
	4	**1.001**	0.262	0.995	-1.035		4	0.995	-0.135	0.968	-1.872
	6	0.984	-0.527	0.993	-1.414		6	0.896	-2.981	0.984	-1.023
	9	0.951	-2.547	0.998	-0.657		9	0.905	-2.576	0.997	-0.259
	12	0.953	-1.333	0.996	-0.717		12	0.979	-0.502	0.969	-1.354
2	1	**1.054**	**1.658**	0.985	-0.970	8	1	**1.029**	**0.710**	0.954	-3.566
	2	0.999	-0.044	0.955	-2.196		2	0.999	-0.043	0.938	-2.825
	4	**1.010**	0.532	0.981	-1.647		4	**1.031**	**0.900**	0.972	-1.517
	6	0.983	-0.428	0.991	-0.992		6	0.868	-2.883	0.954	-1.949
	9	0.960	-2.237	**1.005**	**0.582**		9	0.907	-2.643	0.981	-0.910
	12	0.963	-1.396	0.993	-0.457		12	0.963	-0.619	0.945	-1.714
4	1	**1.019**	**0.703**	0.978	-2.320	10	1	**1.024**	**0.512**	0.933	-3.466
	2	**1.001**	0.039	0.962	-2.214		2	0.987	-0.315	0.923	-2.945
	4	**1.001**	0.048	0.960	-3.191		4	**1.066**	**1.579**	0.934	-1.226
	6	0.964	-0.878	0.988	-0.980		6	0.877	-2.568	0.897	-2.589
	9	0.921	-2.291	**1.025**	**0.976**		9	0.837	-2.935	0.976	-0.689
	12	0.941	-1.961	0.974	-1.363		12	0.993	-0.160	0.926	-2.081

Note: *RRMSPE* denotes the ratio of the root mean squared prediction errors, which is the mean squared prediction error (*RMSPE*) from the benchmark model divided by the *RMSPE* from the competing Partial Least Squares factor model. DMW is the Diebold-Mariano-West statistics. We repeat estimations and forecasting starting from the first 50% observations until we (out-of-sample) forecast the last observation of the KFSTI. DMW statistics in bold denote the rejection of the null hypothesis of equal predictability at the 5% significance level in favor of our factor models. The critical values are from McCracken (2007) to avoid size distortion because the benchmark model is nested by our factor model.

Table 5. PLS-AR vs. AR: Stock Market

$$\widehat{y}_{t+j|t}^{PLS_{AR}} = \hat{\alpha}_j y_t + \hat{\beta}_j' \Delta \mathbf{c}_t \text{ vs. } \widehat{y}_{t+j|t}^{BM_{AR}} = \hat{\alpha}_j y_t$$

		Recursive		**Rolling Window**				**Recursive**		**Rolling Window**	
k	*h*	*RRMSPE*	*DMW*	*RRMSPE*	*DMW*	*k*	*h*	*RRMSPE*	*DMW*	*RRMSPE*	*DMW*
1	1	**1.018**	**1.717**	**1.032**	**3.786**	6	1	**1.035**	**1.703**	**1.032**	**1.197**
	2	**1.001**	0.171	**1.023**	**3.315**		2	**1.011**	**0.699**	**1.030**	**1.162**
	4	0.991	-1.763	**1.037**	**3.116**		4	**1.012**	**0.727**	0.976	-0.670
	6	0.994	-3.829	**1.032**	**2.585**		6	0.991	-0.777	0.945	-1.276
	9	0.992	-3.448	**1.011**	**1.357**		9	**1.002**	0.152	0.894	-1.614
	12	0.992	-1.170	**1.008**	0.373		12	0.989	-0.511	0.910	-1.359
2	1	**1.021**	**1.789**	**1.035**	**2.765**	8	1	**1.053**	**2.049**	**1.013**	**0.477**
	2	**1.001**	0.139	**1.019**	**1.644**		2	**1.006**	**0.344**	**1.002**	**0.078**
	4	0.993	-1.312	**1.039**	**2.821**		4	**1.016**	**0.991**	0.925	-1.793
	6	0.990	-3.169	**1.019**	**0.871**		6	0.990	-0.792	0.930	-1.309
	9	0.986	-2.030	0.979	-0.826		9	**1.004**	**0.263**	0.838	-2.262
	12	0.992	-0.470	0.942	-1.311		12	**1.001**	0.030	0.884	-1.636
4	1	**1.013**	**0.856**	**1.043**	**2.642**	10	1	**1.075**	**1.934**	**1.018**	**0.731**
	2	**1.003**	0.154	**1.051**	**2.939**		2	0.997	-0.128	0.996	-0.127
	4	0.995	-0.500	**1.052**	**2.503**		4	**1.021**	**1.167**	0.906	-1.905
	6	0.986	-2.355	0.997	-0.120		6	0.983	-1.189	0.886	-2.004
	9	0.989	-0.764	0.983	-0.375		9	**1.020**	**0.988**	0.814	-2.186
	12	0.979	-0.972	0.887	-1.908		12	**1.009**	**0.339**	0.862	-1.972

Note: *RRMSPE* denotes the ratio of the root mean squared prediction errors, which is the mean squared prediction error (*RMSPE*) from the benchmark model divided by the *RMSPE* from the competing Partial Least Squares factor model. DMW is the Diebold-Mariano-West statistics. We repeat estimations and forecasting starting from the first 50% observations until we (out-of-sample) forecast the last observation of the KFSTI. DMW statistics in bold denote the rejection of the null hypothesis of equal predictability at the 5% significance level in favor of our factor models. The critical values are from McCracken (2007) to avoid size distortion because the benchmark model is nested by our factor model.

Table 6. PLS-PCA-RW vs. RW: Foreign Exchange Market

$$\widehat{y}_{t+j|t}^{PLS/PC_{RW}} = y_t + \hat{\varphi}_j^{\prime} \Delta \mathbf{z}_t \text{ vs. } \widehat{y}_{t+j|t}^{BM_{RW}} = y_t$$

		Recursive		**Rolling Window**				**Recursive**		**Rolling Window**	
k	*h*	*RRMSPE*	*DMW*	*RRMSPE*	*DMW*	*k*	*h*	*RRMSPE*	*DMW*	*RRMSPE*	*DMW*
2	1	0.938	-1.136	**1.029**	**1.696**	8	1	**1.009**	0.103	**1.038**	**0.802**
	2	0.995	-0.086	**1.013**	0.415		2	**1.076**	**0.750**	**1.080**	**1.688**
	4	**1.057**	**0.846**	**1.045**	**1.466**		4	**1.226**	**1.727**	**1.154**	**2.993**
	6	**1.145**	**1.104**	**1.050**	**1.214**		6	**1.327**	**2.222**	**1.180**	**3.005**
	9	**1.171**	**2.460**	**1.027**	**0.655**		9	**1.363**	**3.871**	**1.146**	**2.566**
	12	**1.359**	**2.049**	**1.047**	**0.965**		12	**1.493**	**1.999**	**1.135**	**2.056**
4	1	0.962	-0.518	**1.037**	**1.417**	10	1	0.954	-0.508	**1.064**	**2.185**
	2	**1.016**	0.216	**1.029**	**0.709**		2	**1.050**	**0.461**	**1.061**	**1.410**
	4	**1.086**	**0.966**	**1.102**	**2.752**		4	**1.217**	**1.937**	**1.129**	**2.640**
	6	**1.222**	**1.609**	**1.154**	**2.995**		6	**1.285**	**1.860**	**1.137**	**2.506**
	9	**1.330**	**4.297**	**1.133**	**2.907**		9	**1.340**	**3.518**	**1.167**	**2.959**
	12	**1.536**	**2.336**	**1.095**	**1.643**		12	**1.443**	**1.893**	**1.144**	**2.185**
6	1	0.959	-0.579	**1.014**	**0.252**	12	1	0.944	-0.586	**1.033**	**0.847**
	2	**1.025**	**0.330**	**1.067**	**1.599**		2	**1.045**	**0.411**	**1.028**	**0.419**
	4	**1.101**	**1.005**	**1.124**	**3.425**		4	**1.165**	**1.377**	**1.078**	**0.809**
	6	**1.277**	**2.008**	**1.209**	**2.751**		6	**1.163**	**0.910**	**1.092**	**1.485**
	9	**1.324**	**3.590**	**1.132**	**2.450**		9	**1.331**	**3.149**	**1.160**	**2.825**
	12	**1.590**	**2.271**	**1.109**	**2.097**		12	**1.576**	**2.037**	**1.131**	**2.755**

Note: *RRMSPE* denotes the ratio of the root mean squared prediction errors, which is the mean squared prediction error (*RMSPE*) from the benchmark model divided by the *RMSPE* from the competing Partial Least Squares factor model. DMW is the Diebold-Mariano-West statistics. We repeat estimations and forecasting starting from the first 50% observations until we (out-of-sample) forecast the last observation of the KFSTI. DMW statistics in bold denote the rejection of the null hypothesis of equal predictability at the 5% significance level in favor of our factor models. The critical values are from McCracken (2007) to avoid size distortion because the benchmark model is nested by our factor model.

Table 7. PLS-PC-RW vs. RW: Stock Market

$$\widehat{y}_{t+j|t}^{PLS/PC_{RW}} = y_t + \hat{\varphi}_j' \Delta \mathbf{z}_t \text{ vs. } \widehat{y}_{t+j|t}^{BM_{RW}} = y_t$$

		Recursive		**Rolling Window**				**Recursive**		**Rolling Window**	
k	*h*	*RRMSPE*	*DMW*	*RRMSPE*	*DMW*	*k*	*h*	*RRMSPE*	*DMW*	*RRMSPE*	*DMW*
2	1	**1.001**	0.167	**1.006**	0.645	8	1	**1.047**	**1.226**	**1.121**	**2.881**
	2	0.996	-0.238	**1.016**	**0.965**		2	**1.079**	**1.228**	**1.243**	**3.171**
	4	0.993	-0.402	**1.018**	**1.001**		4	**1.121**	**1.960**	**1.289**	**2.829**
	6	**1.014**	**0.757**	0.999	-0.066		6	**1.102**	**2.064**	**1.289**	**2.964**
	9	**1.016**	0.488	0.984	-0.545		9	**1.108**	**1.242**	**1.270**	**2.340**
	12	**1.006**	0.254	0.970	-1.180		12	**1.103**	**2.005**	**1.304**	**2.956**
4	1	**1.042**	**1.001**	**1.067**	**3.071**	10	1	**1.053**	**1.306**	**1.147**	**2.863**
	2	**1.050**	**0.882**	**1.149**	**3.312**		2	**1.061**	**0.924**	**1.273**	**2.976**
	4	**1.044**	**0.841**	**1.164**	**3.281**		4	**1.116**	**1.754**	**1.296**	**2.878**
	6	**1.092**	**1.852**	**1.178**	**3.173**		6	**1.116**	**2.222**	**1.258**	**2.626**
	9	**1.128**	**1.395**	**1.213**	**2.813**		9	**1.127**	**1.399**	**1.248**	**2.139**
	12	**1.105**	**2.099**	**1.162**	**2.576**		12	**1.113**	**2.072**	**1.297**	**2.652**
6	1	**1.042**	**1.164**	**1.112**	**2.865**	12	1	**1.086**	**1.397**	**1.165**	**3.144**
	2	**1.058**	**0.898**	**1.234**	**3.150**		2	**1.058**	**0.769**	**1.277**	**2.916**
	4	**1.061**	**1.055**	**1.263**	**2.745**		4	**1.123**	**1.692**	**1.320**	**2.808**
	6	**1.086**	**1.689**	**1.288**	**3.070**		6	**1.131**	**1.966**	**1.359**	**3.040**
	9	**1.120**	**1.233**	**1.285**	**2.411**		9	**1.149**	**1.399**	**1.315**	**2.407**
	12	**1.117**	**2.046**	**1.261**	**2.672**		12	**1.137**	**2.123**	**1.399**	**3.064**

Note: *RRMSPE* denotes the ratio of the root mean squared prediction errors, which is the mean squared prediction error (*RMSPE*) from the benchmark model divided by the *RMSPE* from the competing Partial Least Squares factor model. DMW is the Diebold-Mariano-West statistics. We repeat estimations and forecasting starting from the first 50% observations until we (out-of-sample) forecast the last observation of the KFSTI. DMW statistics in bold denote the rejection of the null hypothesis of equal predictability at the 5% significance level in favor of our factor models. The asymptotic critical values from the standard normal distribution are used.

Table 8. PLS-PCA-AR vs. AR: Foreign Exchange Market

$$\widehat{y}_{t+j|t}^{PLS/PC_{AR}} = \hat{\alpha}_j y_t + \hat{\omega}_j' \Delta \mathbf{z}_t \text{ vs. } \widehat{y}_{t+j|t}^{BM_{AR}} = \hat{\alpha}_j y_t$$

		Recursive		**Rolling Window**				**Recursive**		**Rolling Window**	
k	h	*RRMSPE*	*DMW*	*RRMSPE*	*DMW*	k	h	*RRMSPE*	*DMW*	*RRMSPE*	*DMW*
2	1	**1.033**	**1.244**	0.997	-0.268	8	1	**1.030**	**0.752**	0.939	-3.054
	2	**1.002**	0.063	0.970	-1.838		2	0.998	-0.071	0.932	-2.881
	4	0.999	-0.134	0.996	-0.784		4	**1.020**	**0.434**	0.949	-1.320
	6	0.984	-0.367	0.993	-0.874		6	0.877	-2.386	0.931	-2.562
	9	0.947	-2.168	**1.020**	**0.711**		9	0.899	-2.191	0.967	-1.092
	12	0.951	-1.033	0.991	-1.046		12	0.944	-0.964	0.947	-1.880
4	1	**1.033**	**0.977**	0.973	-2.458	10	1	**1.005**	**0.098**	0.937	-3.363
	2	0.997	-0.117	0.962	-2.206		2	0.998	-0.056	0.915	-3.006
	4	0.988	-0.352	0.988	-1.091		4	0.978	-0.558	0.925	-2.026
	6	0.954	-1.054	0.990	-0.882		6	0.853	-1.860	0.934	-2.107
	9	0.926	-1.991	**1.026**	**1.369**		9	0.877	-1.757	0.952	-1.384
	12	0.980	-0.516	0.983	-1.049		12	**1.007**	**0.063**	0.903	-2.821
6	1	**1.027**	**0.712**	0.966	-2.322	12	1	0.983	-0.299	0.928	-3.176
	2	0.998	-0.061	0.944	-2.557		2	0.964	-0.806	0.894	-2.714
	4	**1.012**	**0.329**	0.971	-1.856		4	0.970	-0.434	0.885	-1.253
	6	0.922	-2.153	0.991	-0.549		6	0.806	-1.785	0.905	-2.078
	9	0.913	-1.836	0.988	-0.659		9	0.865	-2.357	0.963	-1.101
	12	0.963	-1.138	0.965	-1.314		12	0.988	-0.111	0.894	-2.771

Note: *RRMSPE* denotes the ratio of the root mean squared prediction errors, which is the mean squared prediction error (*RMSPE*) from the benchmark model divided by the *RMSPE* from the competing Partial Least Squares factor model. DMW is the Diebold-Mariano-West statistics. We repeat estimations and forecasting starting from the first 50% observations until we (out-of-sample) forecast the last observation of the KFSTI. DMW statistics in bold denote the rejection of the null hypothesis of equal predictability at the 5% significance level in favor of our factor models. The critical values are from McCracken (2007) to avoid size distortion because the benchmark model is nested by our factor model.

Table 9. PLS-PCA-AR vs. AR: Stock Market

$$\widehat{y}_{t+j|t}^{PLS/PC_{AR}} = \hat{\alpha}_j y_t + \hat{\omega}_j' \Delta \mathbf{z}_t \text{ vs. } \widehat{y}_{t+j|t}^{BM_{AR}} = \hat{\alpha}_j y_t$$

		Recursive		**Rolling Window**				**Recursive**		**Rolling Window**	
k	*h*	*RRMSPE*	*DMW*	*RRMSPE*	*DMW*	*k*	*h*	*RRMSPE*	*DMW*	*RRMSPE*	*DMW*
2	1	**1.021**	**1.510**	**1.039**	**2.691**	8	1	**1.044**	**1.733**	**1.020**	**0.768**
	2	**1.001**	0.135	**1.022**	**1.897**		2	**1.011**	**0.609**	**1.002**	**0.100**
	4	0.990	-1.818	**1.038**	**2.722**		4	**1.021**	**1.265**	0.928	-1.829
	6	0.993	-2.431	**1.018**	**0.954**		6	0.984	-1.202	0.922	-1.516
	9	0.997	-0.349	0.973	-0.886		9	**1.019**	**0.942**	0.820	-2.984
	12	0.999	-0.091	0.923	-1.604		12	**1.018**	**0.633**	0.874	-2.109
4	1	**1.033**	**1.344**	**1.028**	**1.901**	10	1	**1.053**	**1.936**	**1.026**	**0.890**
	2	**1.015**	**0.782**	**1.027**	**2.017**		2	**1.004**	**0.234**	0.998	**-0.049**
	4	0.999	-0.108	**1.015**	**0.902**		4	**1.021**	**1.307**	0.917	-1.880
	6	**1.002**	0.242	**1.002**	0.090		6	0.990	-0.634	0.887	-1.763
	9	**1.011**	**0.649**	0.957	-1.076		9	**1.022**	**1.009**	0.805	-2.280
	12	0.991	-0.471	0.916	-1.696		12	**1.027**	0.914	0.848	-1.750
6	1	**1.050**	**2.068**	**1.016**	**0.702**	12	1	**1.069**	**1.729**	**1.034**	**1.189**
	2	**1.010**	**0.485**	**1.019**	**0.775**		2	0.999	-0.062	0.993	**-0.183**
	4	**1.001**	0.056	0.953	-1.359		4	**1.020**	**1.069**	0.906	-1.853
	6	0.993	-0.784	0.918	-2.330		6	0.999	-0.081	0.887	-1.884
	9	**1.015**	**1.191**	0.846	-2.666		9	**1.047**	**1.846**	0.809	-2.298
	12	0.996	-0.191	0.869	-2.200		12	**1.024**	**0.832**	0.848	-1.903

Note: *RRMSPE* denotes the ratio of the root mean squared prediction errors, which is the mean squared prediction error (*RMSPE*) from the benchmark model divided by the *RMSPE* from the competing Partial Least Squares factor model. DMW is the Diebold-Mariano-West statistics. We repeat estimations and forecasting starting from the first 50% observations until we (out-of-sample) forecast the last observation of the KFSTI. DMW statistics in bold denote the rejection of the null hypothesis of equal predictability at the 5% significance level in favor of our factor models. The critical values are from McCracken (2007) to avoid size distortion because the benchmark model is nested by our factor model.

Figure 1. Korean Financial Stress Index

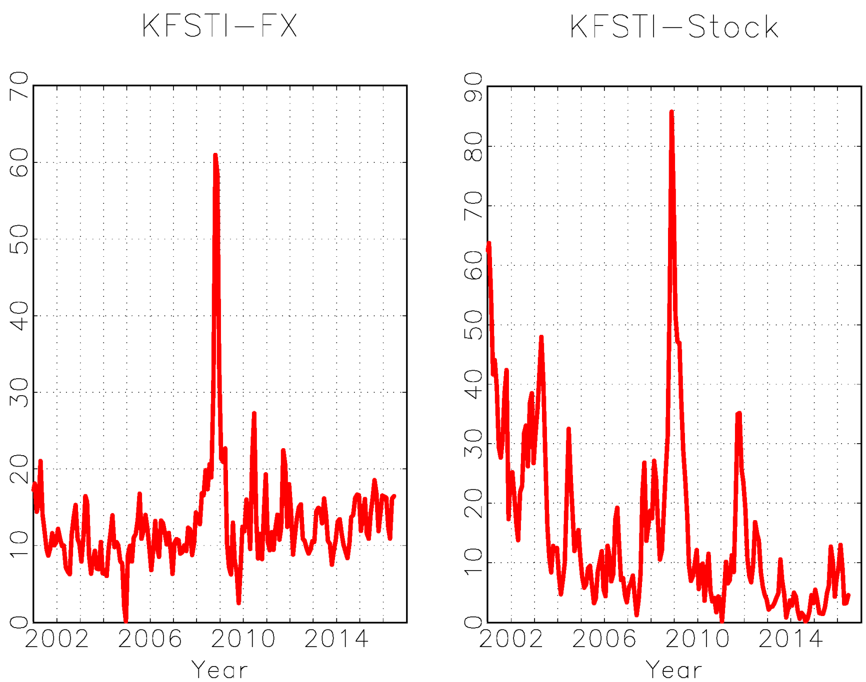
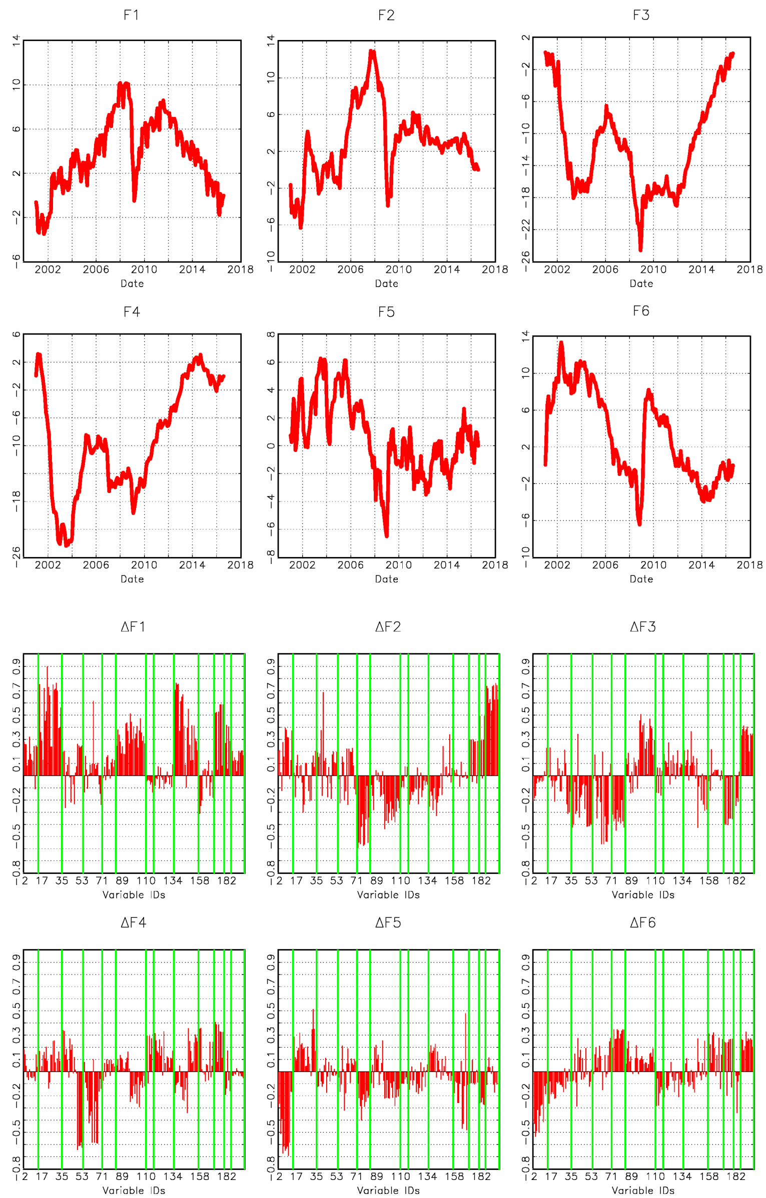
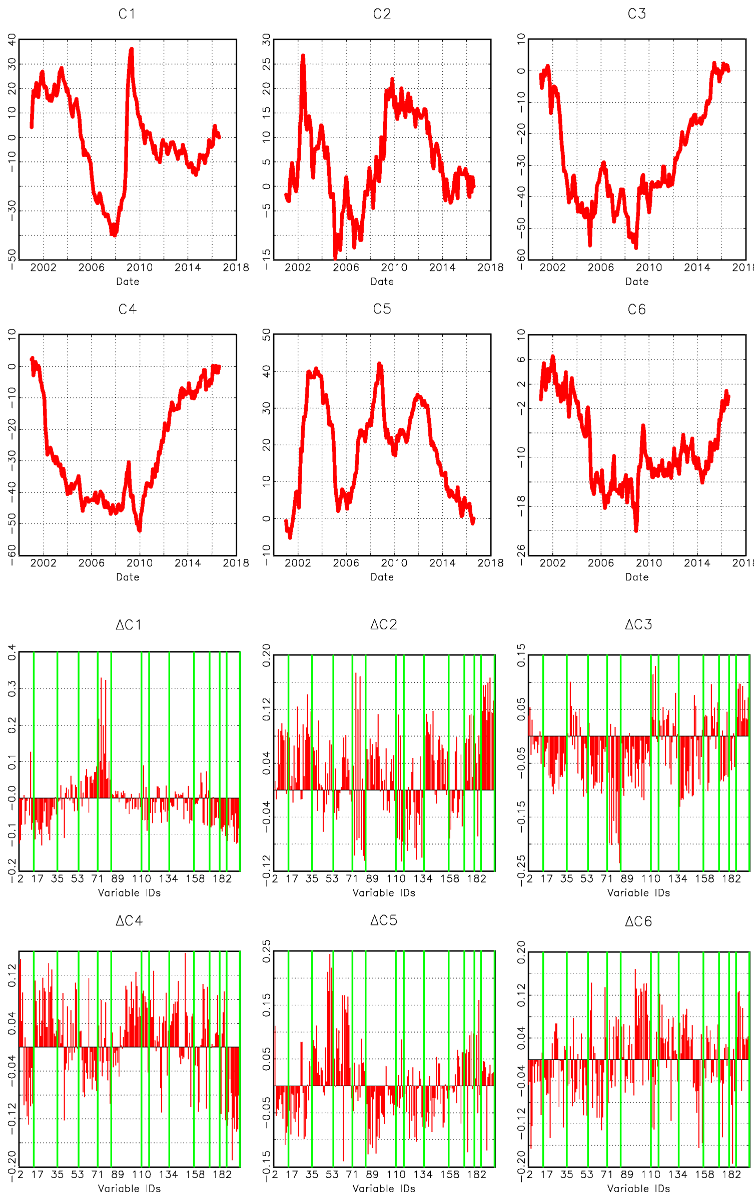


Figure 2. Principal Component Analysis



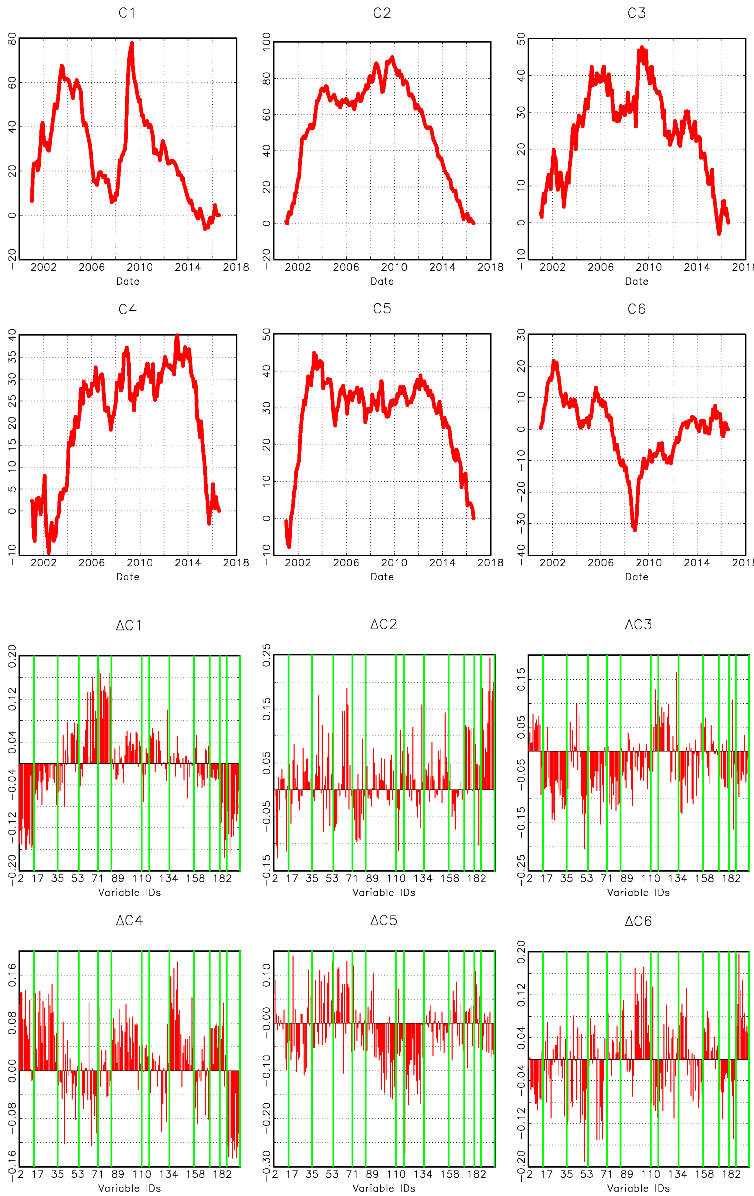
Note: Estimated *level* factors via the method of the principal component are reported in the top panel. Graphs in the bottom panel are factor loading coefficients estimates.

Figure 3. Partial Least Squares Estimation: Foreign Exchange Market



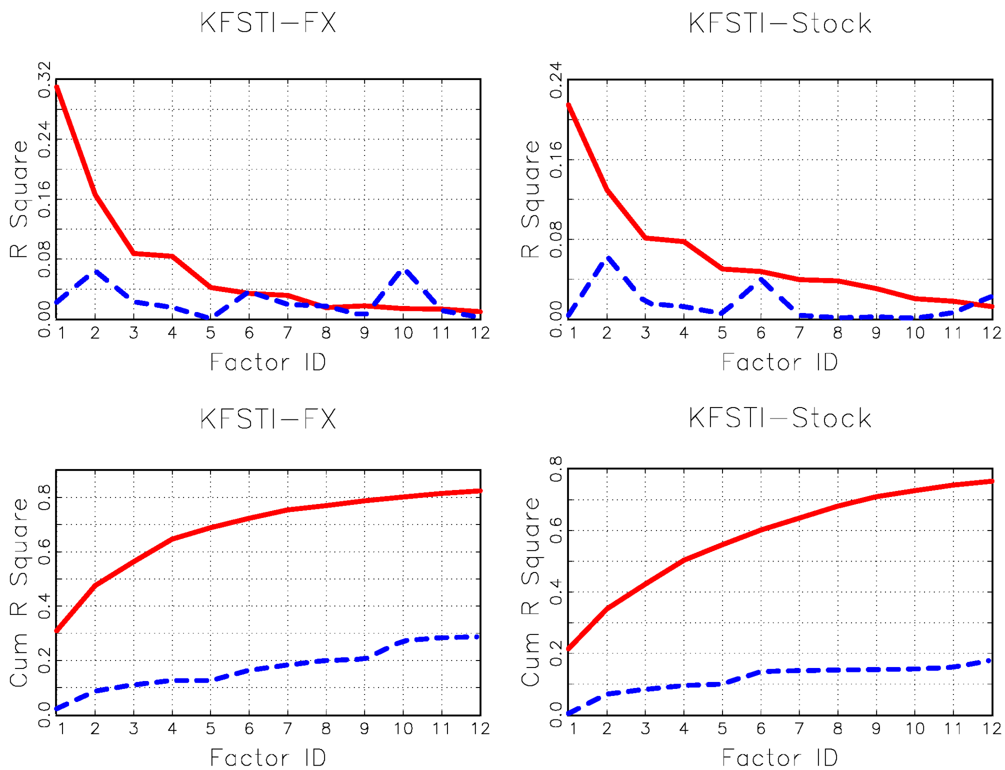
Note: Estimated *level* factors via the partial least squares method are reported in the top panel. Graphs in the bottom panel are weighting matrix estimates.

Figure 4. Partial Least Squares Estimation: Stock Market



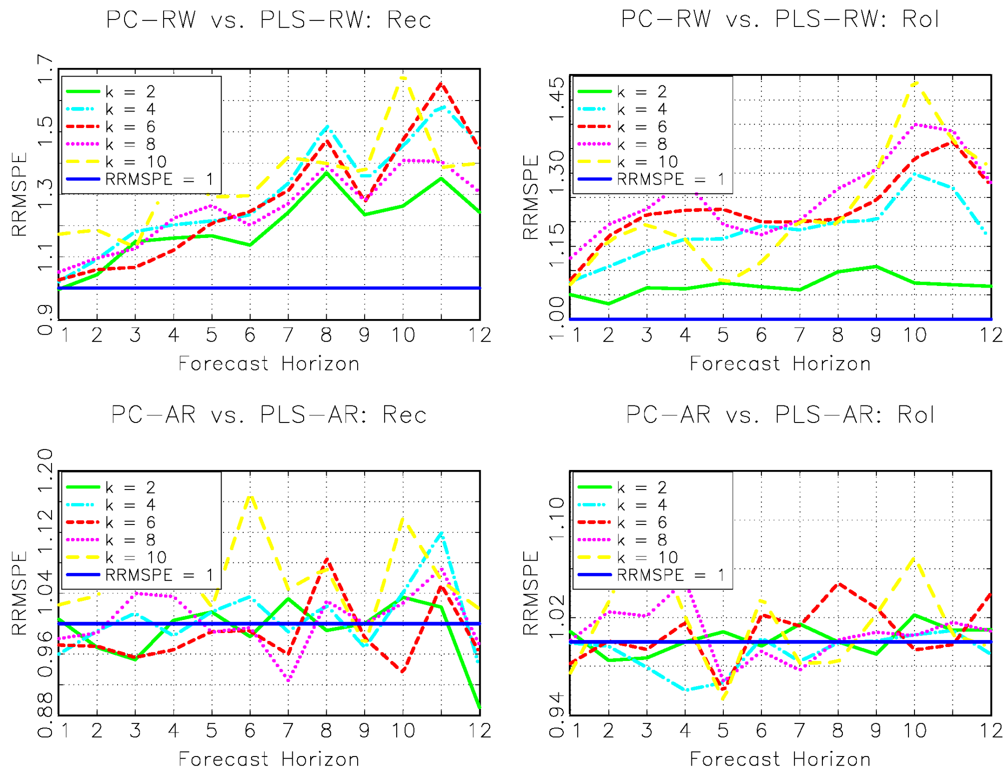
Note: Estimated *level* factors via the partial least squares method are reported in the top panel. Graphs in the bottom panel are weighting matrix estimates.

Figure 5. In-Sample Fit Analysis: R Squares



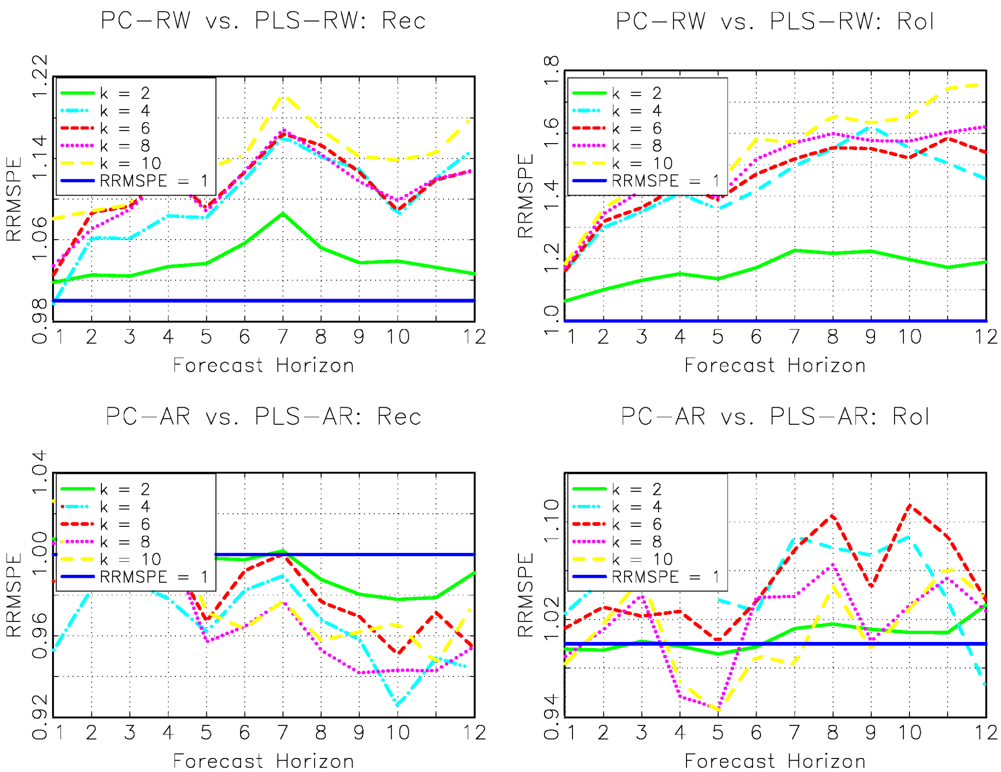
Note: We report R^2 and cumulative R^2 values in the top and lower panel, respectively.

Figure 6. Cross-Comparisons: Foreign Exchange Market



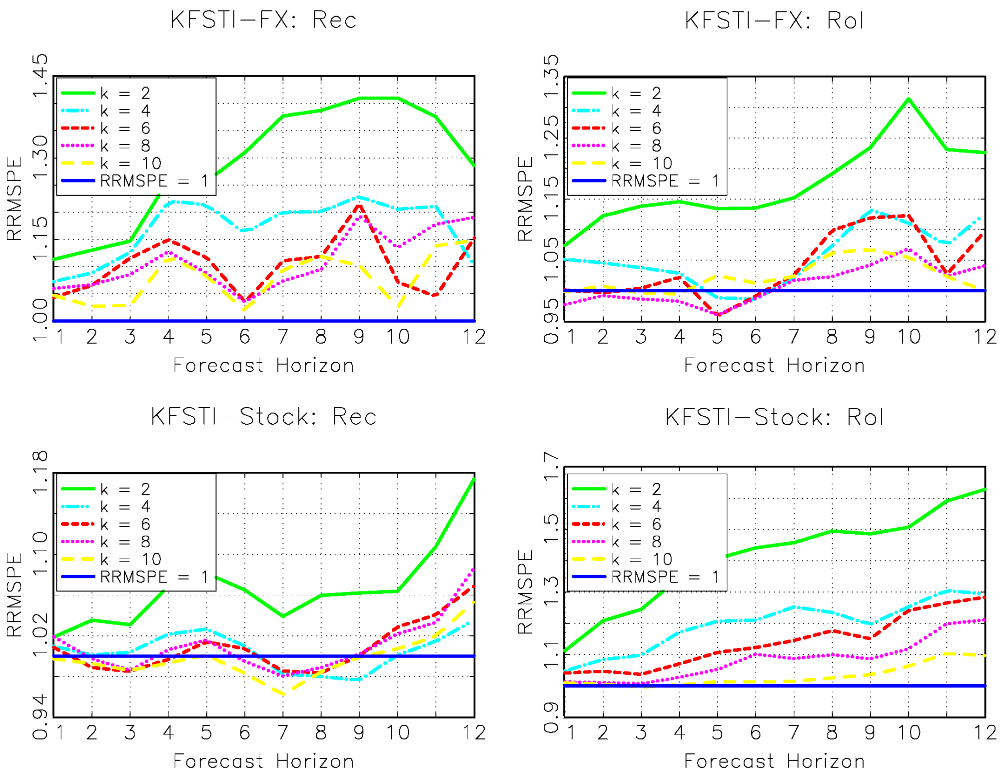
Note: We report the *RRMSPE* defined as the *RMSPE* of the PC method divided the *RMSPE* of the PLS. That is, the PLS method outperforms the PC method when *RRMSPE* is greater than one.

Figure 7. Cross-Comparisons: Stock Market



Note: We report the *RRMSPE* defined as the *RMSPE* of the PC method divided the *RMSPE* of the PLS. That is, the PLS method outperforms the PC method when *RRMSPE* is greater than one.

Figure 8. Cross-Comparisons: PLS-RW vs. PLS-AR



Note: We report the *RRMSPE* defined as the *RMSPE* of the PLS-RW model divided the *RMSPE* of the PLS-AR model. That is, the PLS-AR model outperforms the PLS-RW model when *RRMSPE* is greater than one.