Auburn University
Department of Economics
Working Paper Series

What Charge-Off Rates Are Predictable by Macroeconomic Latent Factors?

Hyeongwoo Kim and Jisoo Son

Auburn University

AUWP 2024-01

What Charge-Off Rates Are Predictable by Macroeconomic Latent Factors?

Hyeongwoo Kim* and Jisoo Son†
Auburn University

February 2024

Abstract

Charge-offs signal critical information regarding the risk level of loan portfolios in the banking system, and they indicate the potential for systemic risk towards deep recessions. Utilizing consolidated financial statements, we have compiled the net charge-off rate (COR) data from the 10 largest U.S. bank holding companies (BHCs) for disaggregated loans, including business loans, real estate loans, and consumer loans, as well as the average top 10 COR for each loan category. We propose factor-augmented forecasting models for CORs that incorporate latent common factor estimates, including targeted factors, via an array of data dimensionality reduction methods for a large panel of macroeconomic predictors. Our models have demonstrated superior performance compared with benchmark forecasting models especially well for business loan and real estate loan CORs, while predicting consumer loan CORs remains challenging especially at short horizons. Notably, real activity factors improve the out-of-sample predictability over the benchmarks for business loan CORs even when financial sector factors are excluded.

Keywords: Net Charge-Off Rate; Top 10 Bank Holding Companies; Disaggregated Loan CORs; Principal Component Analysis; Partial Least Squares; Out-of-Sample Forecast

JEL Classification: C38; C53; C55; G01; G17

*Department of Economics, Auburn University, 138 Miller Hall, Auburn, AL 36849. Tel: +1-334-844-2928. Email: gmmkim@gmail.com.

†Department of Economics, Auburn University, 101 Miller Hall, Auburn, AL 36849. Tel: +1-334-844-2903. Email: jzs0221@auburn.edu.

1 Introduction

This paper presents a novel framework for forecasting the net charge-off rate (COR) of the top 10 largest bank holding companies (BHCs) in the United States, specifically those with a balanced loan structure. Our approach leverages various data dimensionality reduction techniques to estimate latent common factors, including targeted ones, using a large panel of macroeconomic predictors in the U.S. We have compiled a comprehensive dataset that includes individual CORs for disaggregated loans, such as business loans, real estate loans, and consumer loans, from the top 10 BHCs, as well as the average COR for each loan category.

Our research stands in contrast to previous studies that primarily focused on aggregated CORs for all banks. By studying disaggregated CORs, we are able to avoid the potential inaccuracies that may arise due to the exit of small and medium-sized banks that are facing liquidity problems, subsequently replaced by other healthy banks.

the exit of small and intermediate banks exit the banking system. Furthermore, we reveal notable disparities in the predictive content of macroeconomic variables for CORs across specific loan segments, which underscores the critical importance of loan disaggregation when assessing the relationship between macroeconomic indicators and CORs.

Net charge-offs refer to the dollar amount of loans removed from the books (gross charge-offs) that are charged against loss reserves, adjusted for any subsequent recoveries. The net charge-off rate (COR) of a bank is calculated by dividing net charge-offs by its outstanding loans. The COR signals crucial information about the quality and risk level of a bank's loan portfolio, which can generate harmful ripple effects on other banks and other sectors of the economy.[1]

Lessons learned from the recent subprime mortgage market crisis and the subsequent Great Recession have highlighted the importance of well-functioning financial markets in promoting sustainable economic prosperity. Financial crises often catch us by surprise and have spillover effects on real activity sectors. As highlighted by Reinhart and Rogoff (2009), financial market meltdowns can result in prolonged and more painful recessions. As can be seen in Figure 1, the top 10 COR tends to increase rapidly prior to the onset of recessions. The delinquency rate of the top 100 banks also exhibits similar countercyclical dynamics. It should be noted that the recessions in the early 1990s and late 2000s were characterized by a rapid surge in COR and delinquency rate, leading to longer durations of these recessions. Note also that the COR fluctuates between 0.20% to 2.31%, while the delinquency rate ranges between 1.46% to 7.94%, which reflects the fact that delinquent loans are still considered as part of the active receivables that are potentially recoverable. In contrast, the COR exhibits less volatility reflecting actual realized loan losses, implying a perceived stronger connection to actual loan losses in comparison with delinquency rates.[2]

[1] We do not assert that COR is the sole best target for the purpose of mitigating financial instability. For example, the loan delinquency rate, shown in Figure 1, can be a viable candidate as a crucial variable for this early warning signal device. We employ the COR as the main target due to its data availability in FR Y-9C Reports (Schedule HI-B) and perceived stronger connections to actual loan losses compared to delinquency rates.

[2] Delinquency rates measure the percentage of loans that are overdue in terms of payments. In general, delinquency is categorized by the number of days a payment is overdue, e.g., 30 days, 60 days. Higher delinquency rates suggest potential future problems with loan repayments, although delinquent loans are still active receivables that can still

Figure 1 around here

The aforementioned observations imply the potential benefits of good forecasting models for CORs not only for bankers but also for policy makers. CORs can serve as valuable Early Warning Signals (EWS) for economic downturns, providing timely information on potential vulnerabilities in financial markets. The current literature offers a wide range of research works focused on predicting financial market stability.

For instance, linear regression frameworks have been employed by Eichengreen, Rose, and Wyplosz (1995), Sachs, Tornell, and Velasco (1996), and Frankel and Saravelos (2012) to investigate which economic variables can help predict the occurrence of crises. Parametric discrete choice models have been used by Frankel and Rose (1996) and Cipollini and Kapetanios (2009). Also, nonparametric signal detection approaches have been explored by Kaminsky, Lizondo, and Reinhart (1998), Brüggemann and Linne (1999), Berg and Pattillo (1999), Bussiere and Mulder (1999), Edison (2003), Berg, Borensztein, and Pattillo (2005), EI-Shagi, Knedlik, and von Schweinitz (2013), and Christensen and Li (2014), among others.

In the pursuit of selecting an appropriate measure to quantify potential risks in financial markets, the choice holds significant importance. Since the seminal work of Girton and Roper (1977), many researchers have used the Exchange Market Pressure (EMP) index, which is designed to detect the turbulence within money and foreign exchange markets. A comprehensive review of the EMP index can be found in Tanner (2002).

Alternatively, another measure that is gaining rapid popularity is the financial stress index (FSI). Unlike the EMP index, FSIs are typically constructed using a wide range of financial market variables. See Kliesen, Owyang, and Vermann (2012) for a survey on this topic. Recent studies have focused on examining the out-of-sample predictability of FSIs as a proxy for financial market vulnerability. Notable works exploring this area include those by Christensen and Li (2014), Kim, Shi, and Kim (2020), Kim and Ko (2020), and Kim and Shi (2021), among others.

In this paper, we propose factor-augmented forecasting models for an alternative measure of impending financial distress, that is, the net charge-off rate (COR), which provides insights into the loan portfolio quality of banks. To extract latent common factors, we employ data dimensionality reduction methods on a large panel of nonstationary macroeconomic predictors. Specifically, we utilize the principal components (PC) method and the partial least squares (PLS) method (Wold, 1982).

Building on the work of Stock and Watson (2002), there has been a growing body of literature that uses PC approach to perform predictions of key macroeconomic variables. For example, Engel, Mark, and West (2015), Greenaway-McGrevy, Mark, Sul, and Wu (2018), Kim and Park (2020), and Behera, Kim, and Kim (2023) demonstrate that factor-based models outperform the random walk model in out-of-sample forecasting exercises for exchange rates. Furthermore, West and Wong (2014), Chen, Jackson, Kim, and Resiandini (2014), and Chiaie, Ferrara, and Giannone

be paid off. On the other hand, charge-offs occur after a loan has been delinquent for an extended period, and the bank has determined that the debt is unlikely to be recovered.

(2022) highlight the usefulness of latent common factors in both in-sample fitting and out-of-sample prediction of commodity prices.

Notwithstanding its popularity in the current literature, the principal components (PC) method has certain limitations. As pointed out by Boivin and Ng (2006), the performance of the PC approach may be constrained if the useful predictive content for the target is contained within specific factors that are overshadowed by other factors. This is because PC extracts common factors solely from predictor variables, without considering the specific relationship with the target variable.

On the other hand, the partial least squares (PLS) method leverages the covariance structure between the target and predictor variables to generate customized target-specific factors. See Kelly and Pruitt (2015) and Groen and Kapetanios (2016) for comparisons between the PC and PLS approaches. In what follows, we demonstrate that the models with PLS factors indeed outperform both PC factor models, as well as benchmark models.

For our analysis, we constructed the net charge-off rate (COR) data for the top 10 largest U.S. bank holding companies (BHCs) utilizing consolidated financial statements (FR Y-9C: Schedules HI-B and HC-C). The data covers the period from 1986:III to 2021:I. We also conducted forecasting exercises for real estate loan CORs, which have a shorter sample period ranging from 1991:I to 2021:I. This is because real estate loans constitute a significant portion of the business activities of the top 10 BHCs as will be demonstrated in subsequent analysis.

To extract latent common factors, we compiled a large panel of 237 quarterly frequency predictors from the FRED-QD database. These predictors encompass a wide range of variables related to both real economic activity and the financial sector. We utilized this panel for the same sample period as the net COR data.

To evaluate and compare the out-of-sample predictability of our models, we employed the relative root mean square prediction error ($RRMSPE$) statistics.[3] We compared the performance of our models against benchmark models, including stationary autoregressive models and random walk models.

Our major findings can be summarized into the following key aspects. Firstly, our factor-augmented forecasting models consistently outperform the benchmark models, particularly when employing PLS factors. Secondly, our models demonstrate higher predictive accuracy for CORs of business loans and real estate loans, while consumer loan CORs are more challenging to predict. This suggests that business and real estate loan CORs are more closely linked to macroeconomic predictors that capture business cycle factors. Consumer loan CORs exhibit more persistent dynamics, making them more difficult to predict. Thirdly, we find that real activity factors play a crucial role in predicting business loan CORs, often dominating the performance of all other factor models. This aligns with the findings of Boivin and Ng (2006) who demonstrate that additional data may not necessarily improve predictions when noisy predictors are present. We also point out that our results complement the work of Liu, Moon, and Schorfheide (2023), who propose a panel

[3] For similar research conducted using the aggregate net charge-off rate in the U.S. banking sector, see Barth, Joo, Kim, Lee, Maglic, and Shen (2020).

Tobit model with heteroskedasticity to generate forecasts for bank-level loan charge-off rates in small banks. While their study focuses on a large cross-section (large N) of short time series (small T) of censored observations, our research sheds light on the forecasting of net charge-off rates for the top 10 largest BHCs in the U.S., providing valuable insights into the overall banking system.

The remainder of the paper is organized as follows. Section 2 provides a detailed description of our factor-augmented forecasting models, as well as the out-of-sample forecast schemes employed in our study. We also outline the evaluation methods used to assess the performance of our models. In Section 3, we offer data descriptions and provide an initial overview of the data. Some in-sample analysis of our models is also presented. Section 4 reports the results of our out-of-sample forecasts. We examine the performance of our models using all factors as well as subsets of the predictors, which enables us to assess the influence of different sets of factors sourced from various sectors of the economy on the accuracy of our forecasts. Section 5 provides concluding remarks and discuss the implications of our research.

2 The Forecasting Model with Latent Factors

This section presents our factor-augmented forecasting models for the charge-off rate (COR) of U.S. bank holding companies (BHCs). We consider two benchmark models: the nonstationary random walk (RW) model and a stationary autoregressive (AR) model. In what follows, we show that these benchmark models are augmented with latent common factors that are estimated via an array of data dimensionality reduction methods for a large panel of macroeconomic time series data, including the Principal Component (PC) and the Partial Least Squares (PLS) methods for nonstationary predictors.

2.1 Data Dimensionality Reduction Methods to Estimate Latent Factors

2.1.1 Principal Component Approach

Since the seminal work of Stock and Watson (2002), PC has been popularly employed in the current macroeconomic and international finance literature. To employ this approach, consider a large panel of N macroeconomic $T \times 1$ time series predictors/variables, $\mathbf{x} = [\mathbf{x}_1, \mathbf{x}_2, ..., \mathbf{x}_N]$, where $\mathbf{x}_i = [x_{i,1}, x_{i,2}, ..., x_{i,T}]'$, $i = 1, ..., N$. Abstracting from deterministic terms, we assume the following factor structure for each predictor $\mathbf{x}_i$,

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

where $\mathbf{f}_t = \left[f_{1,t}^{PC}, f_{2,t}^{PC}, \cdots, f_{R,t}^{PC} \right]'$ is an $R \times 1$ vector of *latent* time-varying *common* factors at time t. $\boldsymbol{\lambda}_i = [\lambda_{i,1}, \lambda_{i,2}, \cdots, \lambda_{i,R}]'$ denotes an $R \times 1$ vector of time-invariant but idiosyncratic factor loading coefficients for $\mathbf{x}_i$. That is, $\boldsymbol{\lambda}_i' \mathbf{f}_t^{PC}$ describes the underlying data generating process from the common source in the economy, while $\varepsilon_{i,t}$ is the idiosyncratic error term only for i^{th} predictor $x_{i,t}$.

It should be noted that estimating the latent common factors via PC may be spurious if $\varepsilon_{i,t}$ is nonstationary. Since most macroeconomic time series variables are better approximated by an integrated $I(1)$ stochastic process, see Nelson and Plosser (1982), we apply the PC method for the first-differenced data as follows to estimate the factors consistently.

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

for $t = 2, \cdots, T$. See Bai and Ng (2004) for more detailed explanation on this approach. Estimates for the idiosyncratic components are naturally given by the residuals $\Delta\hat{\varepsilon}_{i,t} = \Delta\tilde{x}_{i,t} - \hat{\boldsymbol{\lambda}}_i' \Delta\hat{\mathbf{f}}_t^{PC}$, where hatted variables indicate estimates. Level factors and level error terms are recovered via cumulative summation,

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

Note that our approach yields consistent factor estimates even when $\mathbf{x}$ includes stationary variables because differencing $I(0)$ variables result in $I(-1)$, which is still stationary.[4]

2.1.2 Target-Specific Factor Estimations via Partial Least Squares

Unlike PC, the PLS approach estimates target-specific factors that are customized for the variable of interest.[5] Let $co_{i,j,t}$ denote the net charge-off rate (COR) for loan type j of a bank holding company i at time t. Abstracting from deterministic terms, consider the following linear regression model.

$$co_{i,j,t} = \Delta\mathbf{x}_t' \boldsymbol{\beta} + e_{i,j,t}, \tag{4}$$

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 $t = 1, ..., T$, while $\boldsymbol{\beta}$ is an $N \times 1$ vector of associated coefficients. $e_{i,j,t}$ is an error term. Note that we employ the first-differenced predictor variables, considering nonstationarity of $\mathbf{x}_t$ as explained in the previous section for PC.

PLS is particularly useful for sparse regression models with many predictors. Rewrite (4) as follows,

$$\begin{aligned} co_{i,j,t} &= \Delta\mathbf{x}_t' \mathbf{w}\boldsymbol{\theta} + e_{i,j,t} \\ &= \Delta\mathbf{f}_{i,j,t}^{PLS'} \boldsymbol{\theta} + e_{i,j,t}, \end{aligned} \tag{5}$$

where $\Delta\mathbf{f}_{i,j,t}^{pls} = \left[\Delta f_{1,i,j,t}^{PLS}, \Delta f_{2,i,j,t}^{PLS}, ..., \Delta f_{R,i,j,t}^{PLS}\right]'$, $R < N$ is an $R \times 1$ vector of PLS factors for COR of a bank i for j type loan. Note that the PLS factor is a linear combination of *all* predictor

[4] Alternatively, one may continue to difference the variables until the null of nonstationarity hypothesis is rejected via a unit root test, e.g., augmented Dickey-Fuller test. Although this approach is statistically more rigorous, it may not be practically useful because unit root tests often provides contradicting statistical inferences in small samples when the test specification changes. See Cheung and Lai (1995).

[5] Kelly and Pruitt (2015) and Behera, Kim, and Kim (2023) estimated target specific latent common factors by combining least absolute shrinkage and selection operator (LASSO) with PLS and PC. Bai and Ng (2008) introduced an approach to apply the method of principal components to targeted predictors.

variables,

$$\Delta \mathbf{f}_{i,j,t}^{PLS} = \mathbf{w}_{i,j}^{'} \Delta \mathbf{x}_t, \tag{6}$$

where $\mathbf{w}_{i,j}$= $[\mathbf{w}_{1,i,j}, \mathbf{w}_{2,i,j}, ..., \mathbf{w}_{R,i,j}]$ is an $N \times R$ weighting matrix. That is, $\mathbf{w}_{r,i,j} = [w_{1,i,j,r}, .., w_{N,i,j,r}]^{'}$, $r = 1, ..., R$, is an $N \times 1$ vector of weights on predictor variables for the r^{th} PLS factor, $\Delta f_{r,i,j,t}^{PLS}$. $\boldsymbol{\theta}$ is an $R \times 1$ vector of PLS regression coefficients. Note that PLS regression minimizes the sum of squared residuals from the equation (5) for $\boldsymbol{\theta}$ instead of $\boldsymbol{\beta}$ in (4), resulting in target specific factor estimates for $co_{i,j,t}$. It should be also noted, however, that we augment the benchmark forecasting model with estimated PLS factors $\Delta \hat{\mathbf{f}}_{i,j,t}^{PLS}$ only to make our models to be comparable with the PC factors. That is, we do not utilize $\boldsymbol{\theta}$ for our out-of-sample forecasting exercises in the present paper.

We estimate PLS factors following the sequential procedure proposed by Helland (1990) as follows.[6] First, $\Delta \hat{f}_{1i,j,t}^{PLS}$ is pinned down by the following linear combinations of the predictors in $\Delta \mathbf{x}_t$.

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

where the loading (weight) $w_{s,1}$ is given by $Cov(co_{i,j,t}, \Delta x_{s,t})$. Next, we regress $co_{i,j,t}$ and $\Delta x_{s,t}$ on $\Delta \hat{f}_{1,i,j,t}^{PLS}$ then get the residuals to remove the explained component by the first factor $\Delta \hat{f}_{1,i,j,t}^{PLS}$. The second factor estimate $\Delta \hat{f}_{2,i,j,t}^{PLS}$ is then obtained similarly as in (7) with $w_{s,2} = Cov(\widetilde{co}_{i,j,t}, \Delta \tilde{x}_{s,t})$. We repeat until the R^{th} factor $\Delta \hat{f}_{R,i,j,t}^{PLS}$ is obtained.

2.2 Factor Augmented Forecasting Models

2.2.1 Factor Augmented Nonstationary Model

We augment two benchmark forecasting models, nonstationary random walk (RW) model and stationary autoregressive (AR) model by adding latent factor estimates to improve the out-of-sample predictability of the model.[7] For simplicity, we denote $\boldsymbol{\Delta} \hat{\mathbf{f}}_t$ a vector of latent factors obtained either by PC or PLS.

Our nonstationary RW benchmark model for COR (co_t) is,

$$co_{t+1}^{BM_{RW}} = co_t + \eta_{t+1}, \tag{8}$$

where η_{t+1} is a white noise process, which implies $co_{t+j}^{BM_{RW}} = co_t + \sum_{s=1}^{j} \eta_{t+s}$. Therefore, the j-period ahead forecast is the following.

$$\widehat{co}_{t+j|t}^{BM_{RW}} = co_t \tag{9}$$

[6] See Andersson (2009) for a brief survey on available PLS estimation algorithms.

[7] We report the DF-GLS test Elliott, Rothenberg, and Stock (1996) results in Table A2 in the Appendix. We implemented the test with an intercept and the optimal number of lags was chosen via the Bayesian Information Criteria, which resulted in choosing 1 lag for 22 out of 30 CORs. The test rejects the null hypothesis of nonstationarity for 22 (13) out of 30 CORs at the 10% (5%) significance level, yielding somewhat mixed results. This outcome may reflect Observational Equivalence, making it challenging to distinguish highly persistent time series variables from nonstationary variables. These findings motivate us to employ both AR and RW based prediction models.

Augmenting the RW model by adding $\Delta\hat{\mathbf{f}}_t$ to (8), we obtain the following. Abstracting from deterministic terms again,

$$co_{t+j}^{F_{RW}} = co_t + \boldsymbol{\gamma}_j^{\prime}\boldsymbol{\Delta}\hat{\mathbf{f}}_t + \sum_{s=1}^{j}\eta_{t+s},\ j = 1, 2, .., k, \tag{10}$$

Note that (10) nests the RW model (8) when $\boldsymbol{\gamma}_j = \mathbf{0}$.[8]

Note that we cannot use the unrestricted LS for (10) because the coefficient on co_t is restricted to be one. To resolve this problem, we first regress the long-differenced target variable $co_{t+j} - co_t$ on $\Delta\hat{\mathbf{f}}_t$ to obtain the consistent estimate $\hat{\boldsymbol{\gamma}}_j$, assuming that $co_{t+j} - co_t$ is stationary. Adding co_t back to the fitted value yields the following.

$$\widehat{co}_{t+j|t}^{F_{RW}} = co_t + \hat{\boldsymbol{\gamma}}_j^{\prime}\Delta\hat{\mathbf{f}}_t \tag{11}$$

2.2.2 Factor Augmented Stationary Forecasting Model

Our second benchmark model is motivated by the following *stationary* AR(1)-type stochastic process,[9,10]

$$co_{t+j}^{BM_{AR}} = \alpha_j co_t + u_{t+j},\ j = 1, 2, .., k, \tag{12}$$

where $|\alpha_j| < 1$ for stationarity. (12) implies the following j-period ahead forecast.

$$\widehat{co}_{t+j|t}^{BM_{AR}} = \hat{\alpha}_j co_t, \tag{13}$$

where $\hat{\alpha}_j$ is the LS estimate of α_j.

Similarly as in (10), our second factor-augmented forecasting model is,

$$co_{t+j}^{F_{AR}} = \alpha_j co_t + \boldsymbol{\beta}_j^{\prime}\boldsymbol{\Delta}\hat{\mathbf{f}}_t + u_{t+j},\ j = 1, 2, .., k \tag{14}$$

Therefore, we obtain the following j-period ahead forecast for the target variable,

$$\widehat{co}_{t+j|t}^{F_{AR}} = \hat{\alpha}_j co_t + \hat{\boldsymbol{\beta}}_j^{\prime}\Delta\hat{\mathbf{f}}_t, \tag{15}$$

[8] Note that this specification is inconsistent with our earlier specification described in (4) that requires stationarity of the target variable co_t. Practically speaking, however, the random walk type models often perform well in forecasting persistent variables. Furthermore, it is often difficult to distinguish highly persistent or near unit root variables from stationary variables (observational equivalence), leading us to the two mutually exclusive stochastic processes described in (10) and (14).

[9] We employ a direct forecasting model by regressing co_{t+j} *directly* on the current value co_t. Alternatively, one may employ a *recursive* forecasting approach with an AR(1) model, $co_{t+1} = \alpha co_t + \varepsilon_{t+1}$, which implies $\alpha_j = \alpha^j$ under this approach.

[10] It's worth noting that while AR(p) could serve as a benchmark, the key insights remain consistent, as we primarily assess the improvement in predictability over the AR(p) model. Similarly, although ARIMA or ARMA models could be used, the process would become considerably more complex. In essence, our main objective is to investigate how augmenting the benchmark model with macroeconomic factors enhances its out-of-sample forecastability. Furthermore, it's important to recall that the optimal number of lags determined via the Bayesian Information Criteria was 1 lag for 22 out of 30 CORs. This suggests that AR(1) appears to be a reasonably effective representation.

where $\hat{\alpha}_j$ and $\hat{\boldsymbol{\beta}}_j$ are the least squares coefficient estimates. Note that (14) nests the stationary benchmark model (12) when $\boldsymbol{\Delta}\hat{\mathbf{f}}_t$ does not contain any useful predictive contents for co_{t+j}, that is, $\boldsymbol{\beta}_j = 0$.

2.3 Evaluation Methods

We evaluate the out-of-sample predictability of our factor-augmented forecasting models using a recursive (expanding) window scheme as follows.[11]

We begin with estimating the first set of factors $\left\{\Delta\hat{\mathbf{f}}_t\right\}_{t=1}^{T_0}$ using either PC or PLS for the initial $T_0 < T$ observations, $\{co_t, \Delta x_{i,t}\}_{t=1}^{T_0}$, $i = 1, 2, ..., N$. Then, we formulate the first forecast $\widehat{co}_{T_0+j|T_0}$ as explained in the previous section. Then, one observation is added for the second round forecasting. That is, we re-estimate $\left\{\Delta\hat{\mathbf{f}}_t\right\}_{t=1}^{T_0+1}$ from $\{co_t, \Delta x_{i,t}\}_{t=1}^{T_0+1}$, $i = 1, 2, ..., N$, formulating the second round forecast, $co_{T_0+1+j|T_0+1}$. We repeat until we forecast the last observation, co_T.

To evaluate the out-of-sample prediction accuracy of our factor-augmented models, we use the ratio of the root mean square prediction error ($RRMSPE$) defined as follows,

$$RRMSPE(j) = \frac{\sqrt{\frac{1}{T-T_0-j}\sum_{t=T_0+j}^{T}\left(\varepsilon_{t+j|t}^{F}\right)^2}}{\sqrt{\frac{1}{T-T_0-j}\sum_{t=T_0+j}^{T}\left(\varepsilon_{t+j|t}^{BM}\right)^2}}, \tag{16}$$

where

$$\varepsilon_{t+j|t}^{BM} = co_{t+j} - \widehat{co}_{t+j|t}^{BM},\ \varepsilon_{t+j|t}^{F} = co_{t+j} - \widehat{co}_{t+j|t}^{F} \tag{17}$$

Note that our factor models outperform the benchmark model when $RRMSPE$ is less than 1.[12]

3 The Empirics

3.1 Data Descriptions and a Preliminary Examination of the Data

3.1.1 Net Charge-Off Rates of the Top 10 Bank Holding Companies

We constructed the net charge-off rate (COR) on disaggregated loans as well as total loans of the top 10 bank holding companies (BHCs) in the U.S., following the guidelines given in the FR Y-9C reports that are obtained from the Federal Reserve Bank of Chicago. The amount of gross charge-offs and recoveries are obtained from Schedule HI-B, while we acquired the amount of outstanding loans from Schedule HC-C. Observations are quarterly and span from 1986:III to 2021:I. We removed seasonality in the data using X-13ARIMA-SEATS prior to estimation.

[11] Alternatively, fixed-size rolling window schemes may be used which may perform better if the underlying data generating process changes. We do not employ this scheme as the results turn out to be less robust due to small number of observations.

[12] Alternatively, one may employ the ratio of the root mean absolute prediction error ($RRMAPE$). That is, the loss function is defined with the absolute value instead of the squared value. $RRMAPE$ tends to perform more reliably in the presence of outliers. Results are overall qualitatively similar.

The National Information Center (NIC) provides the relevant information on BHCs and other institutions, both domestic and foreign financial entities, that are operating in the U.S. under the supervision of the Federal Reserve system. We selected the top 10 BHCs based on the book or market value of total assets as of September 30, 2021 among the top 25 largest BHCs with a balanced loan structure with sufficient data availability.[13,14] See Table 1 for information about these top 10 BHCs used in this paper.

We excluded certain large BHCs such as Goldman Sachs and Morgan Stanley due to their transition to bank holding companies in 2008 during the financial crisis, resulting in significantly constrained sample periods for these BHCs.[15] Capital One and TD Bank Group were also excluded from our analysis, because they became bank holding companies in 2004 and 2008, respectively.[16] That is, the sample periods for these excluded BHCs are notably shorter than the one utilized in this paper, spanning from 1986:III to 2021:I.

Some other large bank holding companies were excluded as commercial banking does not represent their primary business focus. For example, the outstanding total loans of Charles Schwab and Bank of New York Mellon constituted 16.08% and 12.89% of their assets, respectively, representing significantly smaller proportions compared to Bank of America and Wells Fargo, which reported figures of 31.08% and 45.84%, respectively.

Table 1 around here

Table 1 also reports the average shares of business loans (BL), consumer loans (CL), and real estate loans (RL) out of the total outstanding loans of each BHC. For example, JPM's average shares of the business, consumer, and real estate loans are 26.3%, 20.1%, and 32.8%, respectively. Overall, business and real estate loans constitute a major portion of the top 10 BHCs' loan business areas. The sample period of real estate loans is shorter, ranging from 1991:I to 2021:I.

The rest of the total loans belongs to other categories such as credit card loans and other consumer loans. Their sample periods are also from 1991:I to 2021:I, and the quality of the data was clearly inferior to others.[17] Therefore, we implement our forecasting exercises mainly for CORs of all, business, and consumer loans for the full sample period. We also complement our exercises by implementing the same assessment for the real estate loan CORs notwithstanding their short sample period, but because the real estate loan business takes up the largest share for all 10 BHCs.

As can be seen in Figure 2, we also note that the shares of these loans are far from being stable over time. Shares of the real estate loans overall exhibit an upward trend until the beginning of

[13] The values of assets are measured by book value for the fixed assets and by the market value of the securities.

[14] We selected the top 10 BHCs to emphasize the predictive contents derived from disaggregated loan and COR data. In the expansive loan markets, the business loans for the top 5 BHCs alone surpassed $810 trillion in 2021, while the consumer loans amounted to over $525 trillion.

[15] Goldman Sachs and Morgan Stanley reported business loan figures of $50.3 trillion and $42.8 trillion, respectively.

[16] Charles Schwab's loan portfolio was much smaller, with business and consumer loans amounting to $1.6 trillion and $5.8 trillion, respectively.

[17] We observed frequent N.A. observations in these type COR data than those of the other major loan categories.

the sub-prime mortgage market crisis near 2005-6, followed by a negative trend as real estate loan activities declined since then. The shares of business loans often demonstrate a mirror image of the real estate loan shares, implying that BHCs may adjust their business loan activities considering the profitability of other type loan business. Consumer loan shares are overall the smallest in most BHCs.

Figure 2 around here

In addition to the individual BHC-level COR data, we created the top 10 average COR ($co_{t10,j,t}$) by utilizing the total loan amount of the top 10 BHCs and their associated total net charge-offs as follows.

$$co_{t10,j,t} = \frac{\sum_{i=1}^{10} co_{i,j,t}}{\sum_{i=1}^{10} loan_{i,j,t}}, \tag{18}$$

where $cor_{i,j,t}$ denotes the amount of net charge-offs on loan type j of a top 10 BHC i at time t while $loan_{i,j,t}$ is its associated amount of outstanding loans. We also employ the average CORs of top 100 and all U.S. banks, which are obtained from the FRED.

Figure 3 reports dynamics of the CORs of the top 10 BHCs as well as the top 10 average COR (thick solid lines) in the first column. As we mentioned earlier, CORs tend to rise rapidly before the onset of recessions such as the Great Recession in 2008-9. In the second column, we report figures of individual top 10 BHCs' COR deviations from the top 10 average COR. The top 10 average CORs seem to be reasonable approximation of overall dynamics of individual CORs. The business loan CORs seem to show more homogeneous dynamics while consumer loan CORs exhibit greater variability across BHCs.

Figure 3 around here

Table 2 present summary statistics of CORs of the top 10 individual BHCs as well as the three measures of aggregate CORs of top 10, top 100, and all banks. The mean (average) tends to be greater than the median value especially for business and all loans CORs, resulting in overall positive skewness. For consumer loan CORs, medians were roughly close to mean values. All three type loan CORs exhibit highly leptokurtic distributions, namely, fat-tail distributions that are likely to occur in financial market data. The Jarque-Bera statistics (Jarque and Bera, 1980, 1987) rejects the null hypothesis of normal distribution for all cases.[18] The consumer loan COR tend to show higher standard deviations as seen in Figure 3.

Table 2 around here

[18] We employ the critical values from Deb and Sefton (1996) to avoid size distortion problems in using the asymptotic critical values.

3.1.2 Cross-Section Properties of Net Charge-Off Rates

This subsection investigates the cross-section properties of CORs in the banking sector via the pair-wise cross-correlation analysis of CORs in each type loans. For this, we first remove serial correlation in $co_{i,t}$ using the following augmented Dickey-Fuller regression.[19]

$$co_{i,t} = c + \alpha co_{i,t} + \sum_{s=1}^{p} \beta_j \Delta co_{i,t} + \varepsilon_{i,t} \tag{19}$$

We then calculate the pair-wise correlation coefficients $\hat{\rho}_{i,j}$, $i, j = 1, ...N$ using the residuals $\hat{\varepsilon}_{i,t}$ and $\hat{\varepsilon}_{j,t}$ from (19) for top 10 individual BHCs and three aggregate measures, that is, average CORs of the top 10, top 100, and all banks. Also, we present the following cross-section dependence (CD) test statistic proposed by Pesaran (2021).

$$CD = \left(\frac{2T}{N(N-1)}\right)^{1/2} \left(\sum_{i=1}^{N-1} \sum_{j=i+1}^{N} \hat{\rho}_{i,j}\right) \to^d \mathcal{N}(0,1), \tag{20}$$

where T denotes the number of observations.

We report two heat maps in Figure 4 for the business loan CORs (upper panel) and the consumer loan CORs (lower panel). Excluding $\hat{\rho}_{i,j}$ of the aggregate measures, the cross-correlations of business loan CORs range from 0.010 (JPM and BAC) to 0.784 (BAC and KEY), whereas from -0.120 (PNC and BMO) to 0.651 (JPM and BAC) for consumer loan CORs. The correlations range from -0.165 (USB and PMO) and 0.538 (TFC and KEY) for all loan CORs.[20]

We note much lighter color in the upper-left area of the business loan COR heat map. In fact, the correlations with JPM, $\hat{\rho}_{JPM,j}$, tend to be low, similarly as those with WFC. As can be seen in Table 3, their average correlations are 0.193 and 0.204 for JPM and WFC, respectively, which are lower than those of other top 10 BHCs. The average correlation of all top 10 BHCs is 0.366 (0.330 including aggregate CORs). It should be noted, however, that $\hat{\rho}_{i,j}$ is overall higher for business loan CORs in comparison with consumer loan CORs. Average $\hat{\rho}_{i,j}$ of consumer loan CORs is 0.287 (0.255 including aggregate CORs), which is substantially lower than that of business loan CORs. Average correlations are lower for consumer loan CORs for 8 out of 10 BHCs with exceptions of JPM and CFG.

The cross-section dependence (CD) test statistics supports the presence of common drivers in CORs, rejecting the null hypothesis of cross-section independence at the 1% significance level for all three type loan CORs. We notice that the p-value of the business loan CORs is lower than that of consumer loan CORs, which implies a stronger cross-section dependence in the business loan CORs.

Figure 4 around here

[19] We use the general-to-specific rule with a maximum two lags to select the optimal number of lags.

[20] The heatmap of all loan CORs is not reported to save space. It is available upon request.

Table 3 around here

3.1.3 Large Panel of Macroeconomic Data

We employ 237 quarterly frequency macroeconomic time series variables from the FRED-QD database, matching the sample period with that of the COR data. We log-transformed all quantity variables prior to estimations, while those in percent such as interest rates and unemployment rates were divided by 100.

We categorized these macroeconomic variables into 14 groups. Groups #1 through #6 include 118 real activity predictors, while groups #7 to #14 are nominal/financial sector variables. In addition to extracting latent factors from all predictors, we also estimate real activity factors and financial factors separately to track the sources of the predictability, if any, for CORs. See Table A1 in the Appendix for more detailed information.

In what follows, we report greater predictive contents of macroeconomic latent factors for the business loans in comparison with consumer loans, which implies that our factor-based forecasting models would work better for the CORs of business loans than those of consumer loans. We also obtained substantial predictive contents of the macro factors for the top 10 real estate loan CORs.

3.2 Factor Model In-Sample Analysis

This section provides some useful in-sample properties of the factor estimates that are obtained from the average CORs of the top 10 BHCs and the large panel of macroeconomic predictors. In Figure 5, we first present estimated level factors, that is, $\hat{f}_{i,t} = \sum_{s=2}^{t} \Delta \hat{f}_{i,s}$, $i = 1, 2$, which are visually more tractable. PC factors are reported in the top left panel, whereas PLS factors appear in other three panels, because PLS yields customized factors to fit each target COR data.

As can be seen in Figure 5, the estimated level factors exhibit strong co-movement with each other. This implies that PLS level factors for each type CORs are likely to be correlated with business cycle dynamics, because PC factors are estimated utilizing differenced macro/finance predictors, generating business cycle factors instead of trend components. Also, this implies that both PC factors and PLS factors are likely to share predictive contents for the CORs.[21] We note, however, that PC factors overall demonstrate closer dynamics with PLS factors for all loan CORs and business loan CORs, while PLS factors for consumer loan CORs exhibit more pronounced dynamics in comparison with these other factors. In what follows, we report our factor-augmented forecasting models perform better for all loan CORs and business loan CORs than for consumer loan CORs.[22]

[21] PC factors are derived from the business cycle components of macroeconomic data, as consistent factor estimates are obtained through differenced time series that remove trend components. In cases where the target variable for the PLS factors is closely associated with the underlying business cycle components, both PLS factors and PC factors demonstrate similar dynamics, a pattern we observe in our specific context.

[22] We also observe that PLS factors often outperform PC factors in terms of predictive accuracy. For example, RW-based models demonstrate excellent predictability when extended by PLS factors (PLSRW), whereas the gains in predictability are limited when augmented with PC factors (PCRW), while both AR-based models (PCAR and PLSAR) showcase superior prediction performance.

Figure 5 around here

Figure 6 reports the R^2 statistics and the cumulative R^2 statistics of PC and PLS factors for up to 12 factors. By construction, PLS factors provide a better in-sample fit than PC factors because PLS utilizes the covariance structure between the target (top 10 average CORs) and the predictor variables, while PC factors are extracted only from the variance-covariance structure of macro/finance predictor variables. Putting it differently, the PLS method yields superior in-sample performance relative to the PC method by construction.

Note that, unlike PC factors, the cumulative R^2 statistics (second column) of PLS factors exhibit positive slopes at a decreasing rate. This is because our PLS algorithm sequentially estimates orthogonalized common factors using residuals of the target and the predictors, as explained earlier in Section 2. On the other hand, the PC method utilizes predictors only without considering the target variable, thus additional R^2 values do not necessarily decrease. For example, $\hat{f}^{PC}_{4,t}$ seems to have the highest in-sample explanatory power for all three CORs.

Figure 6 around here

Following Ludvigson and Ng (2009), we investigate the source of the estimated common factors via the marginal R^2 analysis. That is, we regress each predictor onto the common factor to measure how much of the variation in each predictor can be explained by the common factor. Results are reported in Figure 7.

The first PC common factor, $\Delta\hat{f}^{PC}_{1,t}$, seems to be heavily correlated with real activity predictors (groups #1 through #6) such as NIPA (#1, ID 1-22), industrial production (#2, ID 23-38), and labor market condition (#3, ID 39-87) macroeconomic variables. $\Delta\hat{f}^{PC}_{2,t}$ is likely to be coming mainly from price predictors (#7, ID 119-166), while $\Delta\hat{f}^{PC}_{3,t}$ explains substantial variations of financial market predictors such as exchange rates (#10, ID 202-206), stock markets (#11, ID 207-213), and household balance sheets variables (#13, ID 216-224). On the other hand, $\Delta\hat{f}^{PC}_{4,t}$ exhibits overall balanced marginal R^2 statistics distribution for both the real activity and the nominal/financial sector variable groups.

The marginal R^2 statistics of the PLS factors exhibit similar distributions, especially between $\Delta\hat{f}^{PLS}_{i,A,t}$ (all loans CORs) and $\Delta\hat{f}^{PLS}_{i,B,t}$ (business loans CORs). The marginal R^2 statistics of $\Delta\hat{f}^{PLS}_{1,A,t}$ and $\Delta\hat{f}^{PLS}_{1,B,t}$ are distributed overall evenly except the price predictors (#7), while $\Delta\hat{f}^{PLS}_{1,C,t}$ (consumer loans CORs) explains the variations of the most predictors including group #7 variables. Overall, the third and fourth PLS common factors, $\Delta\hat{f}^{PLS}_{i,j,t}, i = 3, 4$ and $j = A, B, C$, seem to explain the variations of the nominal/finance variables (#7 through #14) more, while the first and the second PLS factors are more closely correlated with real activity variables (#1 through #6).

Figure 7 around here

4 Out-of-Sample Prediction Performance

We implement an array of out-of-sample (OOS) forecast exercises for the CORs of the top 10 individual BHCs as well as the two aggregate CORs. Employing a recursive scheme, we evaluate the OOS predictability of our factor-augmented forecasting models in comparison with the two benchmark models, utilizing PC and PLS for 237 quarterly frequency time series predictors. Motivated by the work of Boivin and Ng (2006), we also assess the predictability of our models when factors are extracted from subsets of the panel data such as real activity groups (#1 through #6) and nominal/financial sector groups (#7 through #14).

4.1 Out-of-Sample Predictability of the Total Macro Factors

We report the $RRMSPE$ statistics (16) for an array of factor-augmented forecasting models in comparison with the random walk (RW) benchmark model. The $RRMSPE$ statistics with the stationary autoregressive (AR) model is also presented. Recall that competing models perform better than the benchmark RW model when the $RRMSPE$ is less than one.

We begin with the OOS forecasts for all loan CORs utilizing up to 10 latent factors.[23] Figure 8 compares the 1-quarter ahead out-of-sample prediction performance of the two factor-augmented stationary AR model forecasts, $\widehat{co}_{t+1|t}^{PLSAR}$ and $\widehat{co}_{t+1|t}^{PCAR}$, the two factor-augmented nonstationary RW model forecasts, $\widehat{co}_{t+1|t}^{PLSRW}$ and $\widehat{co}_{t+1|t}^{PCRW}$, and the AR benchmark model forecast, $\widehat{co}_{t+1|t}^{AR}$. Results overall imply that our factor-augmented forecasting models yield substantial improvement in short-term predictability over the both benchmark models. Detailed analysis is as follows.

We observe that $\widehat{co}_{t+1|t}^{AR}$ outperforms the benchmark $\widehat{co}_{t+1|t}^{RW}$ ($RRMSPE < 1$) for five BHCs (JPM, WFC, USB, PNC, BMO) but not for the rest of BHCs (BAC, TFC, FITB, CFG, KEY). We note that $\widehat{co}_{t+1|t}^{AR}$ performs worse than $\widehat{co}_{t+1|t}^{RW}$ for the two aggregate CORs, the top 10 average COR and the average COR of all banks.

In most cases, $\widehat{co}_{t+1|t}^{PLSAR}$ and $\widehat{co}_{t+1|t}^{PCAR}$ exhibit superior performance over the benchmark models. $\widehat{co}_{t+1|t}^{PLSRW}$ also outperforms $\widehat{co}_{t+1|t}^{RW}$ when sufficiently large number (around 4 or more) of factors are used, while $\widehat{co}_{t+1|t}^{PCRW}$ does not perform very well no matter how many factors are employed. In a nutshell, the PLS factors $\Delta\hat{f}_{i,A,t}^{PLS}$ seem to play an important role in enhancing the predictability consistently even with a single factor $\Delta\hat{f}_{1,A,t}^{PLS}$.

Figure 8 around here

Figure 9 provides the $RRMSPE$ statistics for the 2-quarter ahead OOS prediction models. $\widehat{co}_{t+2|t}^{AR}$ outperforms the benchmark $\widehat{co}_{t+2|t}^{RW}$ for five BHCs (JPM, WFC, USB, PNC, BMO) again

[23] We estimated the optimal number of factors by the Information Criteria proposed by Bai and Ng (2002), which resulted in 12 to 20 factors. It should be noted that such selection procedures are based on in-sample and large sample analysis. However, the chosen optimal number of factors does not inherently guarantee improved out-of-sample predictability in practical terms. Our exercises demonstrate solid evidence of additional predictability with 10 latent factors.

but not for the rest of BHCs. $\widehat{co}^{AR}_{t+2|t}$ also performs worse than $\widehat{co}^{RW}_{t+2|t}$ again for the two aggregate CORs, but the $RRMSPE$ statistics are closer to one when the forecast horizon rises from 1 to 2. In fact, the performance of $\widehat{co}^{AR}_{t+2|t}$ improved in most cases.

$\widehat{co}^{PLSAR}_{t+2|t}$ and $\widehat{co}^{PCAR}_{t+2|t}$ continue to outperform the benchmark model $\widehat{co}^{RW}_{t+2|t}$, and so does $\widehat{co}^{PLSRW}_{t+1|t}$ when sufficiently large number of factors are used. For the aggregate CORs, our factor-augmented forecasting models again demonstrate superior predictability over the benchmark models.

Figure 9 around here

Figures 10 and 11 report the $RRMSPE$ statistics for the 4-quarter (1-year) and 8-quarter (2-year) ahead OOS prediction models. It should be noted that the predictability of the stationary benchmark model, $\widehat{co}^{AR}_{t+4|t}$ and $\widehat{co}^{AR}_{t+8|t}$, continues to improve the predictability at longer-horizons, reflecting that the deviations of CORs tend to quickly revert back to their equilibrium paths.

Our factor-augmented models outperform the benchmark RW model. However, additional information gains by adding factors seem to diminish as we can see that $\widehat{co}^{PLSAR}_{t+8|t}$ and $\widehat{co}^{PCAR}_{t+8|t}$ perform similarly well as $\widehat{co}^{AR}_{t+8|t}$. See Table A2 in the Appendix for more detailed results for the aggregate CORs of the top 10 banks and all U.S. banks.

Figures 10 around here

Figures 11 around here

We now turn to the performance of our forecasting models for disaggregated COR data, that is, business loan CORs and consumer loan CORs as well as real estate loan CORs. Figure 12 reports the $RRMSPE$ statistics for the business loan aggregate CORs of the top 10 banks and those of all U.S. banks for the 1, 2, 4, and 8-quarter ahead forecasts.[24] $\widehat{co}^{PLSAR}_{t+j|t}$, $\widehat{co}^{PCAR}_{t+j|t}$, and $\widehat{co}^{PLSRW}_{t+j|t}$ again outperform the nonstationary RW model in most cases. These factor models overall outperform the stationary AR model at short horizons ($H = 1, 2$), whereas additional gains over $\widehat{co}^{AR}_{t+j|t}$ appear to diminish as the forecast horizon gets longer. See Table A3 in the Appendix for more detailed results.

Figure 12 around here

As can be seen in Figure 13, our forecasting models demonstrate mixed performance for consumer loan CORs in comparison with the performance for business loan CORs. $\widehat{co}^{PLSAR}_{t+j|t}$, $\widehat{co}^{PCAR}_{t+j|t}$, and $\widehat{co}^{PLSRW}_{t+j|t}$ overall outperform both benchmark models, $\widehat{co}^{AR}_{t+j|t}$ and $\widehat{co}^{RW}_{t+j|t}$, for consumer loan

[24] All results are available upon request.

CORs of all U.S. BHCs, but less satisfactorily for the top 10 average COR for consumer loans. See Table A4 in the Appendix for more detailed results.

One interesting finding is that $\widehat{co}^{AR}_{t+j|t}$ performs better than $\widehat{co}^{RW}_{t+j|t}$ only in 8-period ahead forecasts, which is in stark contrast with previous results for business loan CORs. This reminds our in-sample findings we reported earlier. Consumer loan CORs tend to exhibit greater degree of idiosyncratic dynamics (Figure 3) as well as substantially greater standard deviations (Table 2). Although most level factors tend to demonstrate a (near) unit root process, the level factors from consumer loan CORs, $\hat{f}^{PLS}_{i,C,t}$, show even more persistent dynamics (Figure 3), which may be related with the Martingale property of consumption. Putting it differently, consumption smoothing by optimizing agents may imply a martingale process of consumption that is difficult to forecast.

Figure 13 around here

Figure 14 reports the performance of our factor forecasting models for real estate loan CORs. Although direct comparisons with previous results are difficult due to the different sample period (1991:I-2021:I), we implement forecasting exercises for real estate loan COR, because real estate loans comprise one of the major business components of large U.S. BHCs (see Table 1 and Figure 2). We obtained the following interesting findings.

For the top 10 average COR for real estate loans, $\widehat{co}^{PLSAR}_{t+j|t}$and $\widehat{co}^{PCAR}_{t+j|t}$ outperform both benchmark models substantially especially at shorter horizons and when the number of factors is small. Both nonstationary factor models, $\widehat{co}^{PLSRW}_{t+j|t}$ and $\widehat{co}^{PCRW}_{t+j|t}$, perform overall poorly. On the other hand, we were unable to find superior performance of our factor models for all bank CORs. Furthermore, the out-of-sample forecasting performance tend to become worse when the number of factors increases.

In a nutshell, more factors do not necessarily yield useful predictive contents for real estate loan CORs, which implies that useful information for predicting real estate loan CORs may reside in first few macroeconomic factors, whereas other factors tend to provide noise in our forecasting exercises. Better performance of our factor forecasting models for the top 10 average COR in comparison with all bank COR may reflect the latter is calculated with the banks that survive the crisis. That is, small banks that exit the banking industry may not be used for the all bank CORs.

Figure 14 around here

4.2 Real Activity vs. Nominal/Financial Factors

As shown by Boivin and Ng (2006), more variables are not necessarily better for the purpose of forecasting if some predictors do not possess useful predictive contents. Including such variables can increase noise in formulating predictions. In a similar vein, Behera and Kim (2019) demonstrate that factors extracted from real activity variables, excluding financial sector variables, tend to yield greater predictive contents for U.S. real effective exchange rate at longer horizons.[25]

Figure 15 presents the $RRMSPE$ statistics of our PLS factor-augmented OOS forecasting models for the all loan COR of the top 10 BHCs, using total factors, real activity factors (groups #1 to #6, data ID 1-118), and financial/nominal factors (groups #7 to #14, data ID 119-237).[26] Results imply that the total factor model ($\widehat{co}_{t+j|t}^{PLSAR}$) and financial factor model ($\widehat{co}_{t+j|t}^{PLSAR-F}$) perform similarly well, outperforming both benchmark models. The real factor model ($\widehat{co}_{t+j|t}^{PLSAR-R}$) also overall outperforms both benchmark models but worse than other factor models. See Tables A5 and A6 in the Appendix for more detailed results.

Figure 15 around here

Figure 16 present our forecasting exercises with these subset factors for the business loan COR of the top 10 BHCs. Results are in stark contrast with those for the all loan COR. We note that $\widehat{co}_{t+j|t}^{PLSAR-R}$ overall outperform not only the benchmark models, $\widehat{co}_{t+j|t}^{AR}$ and $\widehat{co}_{t+j|t}^{RW}$, but also other factor-augmented models $\widehat{co}_{t+j|t}^{PLSAR}$ and $\widehat{co}_{t+j|t}^{PLSAR-F}$. The PLS real factor model and the total factor model both outperform other models substantially at the 1-quarter forecast horizon, implying that real activity predictors contain more important predictable contents for the business loan COR. $\widehat{co}_{t+j|t}^{PLSAR-R}$ strongly dominate other models at the 2-quarter and the 4-quarter forecast horizons. It continues to outperform others at 8-quarter horizon but marginally. These findings imply that business loan CORs are heavily influenced by macroeconomic real activity, whereas financial factors play a limited role in predicting business loan CORs. See Tables A7 and A8 in the Appendix for more detailed results.

Figure 16 around here

Figure 17 confirms our earlier findings regarding the difficulty to obtain substantial predictability gains from factors for consumer loan CORs. It should be noted that neither our factor-augmented forecasting models nor the stationary AR benchmark model consistently outperform the nonstationary RW model ($\widehat{co}_{t+j|t}^{RW}$). These findings are again consistent with substantially persistent dynamics (close to a unit root process) of consumer loan CORs. See Tables A9 and A10 in the Appendix for more detailed results.

Figure 17 around here

[25] Similarly, Behera, Kim, and Kim (2023) show that only U.S. factors play an important role in out-of-sample forecasting the KRW-USD real exchange rate, while Korean factors tend to serve as noise in forecasting. They explain such superior predictability of U.S. factors using high degree co-movement behavior of many bilateral exchange rates relative to the U.S.

[26] PC Factor-augmented models perform similarly. Results are available upon request.

5 Concluding Remarks

This paper proposes factor-augmented forecasting models for the net charge-off rate (COR) of the top 10 largest U.S. bank holding companies (BHCs) in a data rich environment. The COR serves as a crucial indicator of the riskiness of loan portfolios in the banking system, with potential spillover effects on both financial markets and the real economy. One of the primary contributions of our research lies in the ability of our forecasting models to serve as Early Warning Signals (EWS), offering timely insights into signs of financial market instability. By accurately predicting the COR for the top 10 BHCs, our models provide valuable information for monitoring and managing risks in the banking sector.

By leveraging individual CORs for disaggregated loans, our models aim to mitigate potential inaccuracies arising from the exit of small and intermediate banks from the financial system. Furthermore, our analysis reveals notable variations in the predictability of macroeconomic factors for CORs across different loan categories. This finding underscores the importance of considering loan disaggregation when examining the relationship between macroeconomic indicators and the COR, highlighting the need for a nuanced understanding of risk dynamics in specific loan segments.[27]

In this study, we employ various data dimensionality reduction methods on a large panel of 237 quarterly frequency macroeconomic variables from 1986:III to 2021:I. By applying Principal Component (PC) and Partial Least Squares (PLS) techniques, we extract latent common factors, which are employed to augment the benchmark model to improve the out-of-sample predictability of CORs.

We assess the prediction accuracy of our models relative to two benchmark models: the stationary autoregressive and the nonstationary random walk models. Our factor-augmented models consistently outperform these benchmarks, particularly in forecasting CORs for business loans, real estate loans, and all loans combined. Forecasting consumer loan CORs remains challenging. These findings suggest that latent factors derived from the underlying forces driving the business cycle dynamics strongly influence business loan CORs. In contrast, consumer loan CORs exhibit more persistent dynamics, potentially due to the Martingale property of consumption, which limits the gains from incorporating latent factors.

Additionally, we find that factors derived from a subset of macro predictors, specifically real activity predictors, significantly enhance the out-of-sample predictability of business loan CORs. While finance factors also offer useful predictive content for CORs, they often do not provide additional contributions when real factors are present, although they also contain stand-alone useful predictive information for CORs. These findings align with the work of Boivin and Ng (2006) who demonstrated the importance of relevant common factors for the target variable. Overall, our study demonstrates the effectiveness of factor-augmented models in forecasting the riskiness of

[27] It's notable that individual BHCs present varying loan structures, which are also subject to time variations, as depicted in Figure 2. For instance, PNC demonstrated approximately 40% and 30% shares of business and real estate loans in 2021:I, respectively, with a mere 10% share in consumer loans. This highlights the pivotal role of macroeconomic factors. On the other hand, if a bank's primary business is consumer loans, as seen in the cases of many BHCs before the Great Recession, such as JPM and BAC, the usefulness of macroeconomic factors is limited.

loan portfolios. These models have implications for financial market stability and risk management within the banking sector, offering valuable insights for informed decision-making.

References

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

Bai, J., and S. Ng (2002): "Determining the Number of Factors in Approximate Factor Models," *Econometrica*, 70(1), 191–221.

——— (2004): "A PANIC Attack on Unit Roots and Cointegration," *Econometrica*, 72(4), 1127–1177.

——— (2008): "Forecasting Economic Time Series Using Targeted Predictors," *Journal of Econometrics*, 146(2), 304–317.

Barth, J. R., S. Joo, H. Kim, K. B. Lee, S. Maglic, and X. Shen (2020): "Forecasting Net Charge-Off Rates of Banks: A PLS Approach," in *HANDBOOK OF FINANCIAL ECONOMETRICS, MATHEMATICS, STATISTICS, AND MACHINE LEARNING*, ed. by C. F. Lee, and J. C. Lee, World Scientific Book Chapters, chap. 63, pp. 2265–2301. World Scientific Publishing Co. Pte. Ltd.

Behera, S., and H. Kim (2019): "Forecasting Dollar Real Exchange Rates and the Role of Real Activity Factors," Auburn Economics Working Paper Series auwp2019-04, Department of Economics, Auburn University.

Behera, S., H. Kim, and S. Kim (2023): "Superior Predictability of American Factors of the Dollar/Won Real Exchange Rate," Auburn Economics Working Paper Series auwp2023-05, Department of Economics, Auburn University.

Berg, A., E. Borensztein, and C. Pattillo (2005): "Assessing Early Warning Systems: How Have They Worked in Practice?," IMF Working Paper 04/52. International Monetary Fund.

Berg, A., and C. Pattillo (1999): "Predicting Currency Crises: The Indicators Approach and an Alternative," *Journal of International Money and Finance*, 18(4), 561–586.

Boivin, J., and S. Ng (2006): "Are More Data Always Better for Factor Analysis?," *Journal of Econometrics*, 132(1), 169–194.

Brüggemann, A., and T. Linne (1999): "How Good are Leading Indicators for Currency and Banking Crises in Central and Eastern Europe? An Empirical Test," IWH Discussion Papers, 95.Halle Institute for Economic Research.

Bussiere, M., and C. Mulder (1999): "External Vulnerability in Emerging Market Economies - How High Liquidity Can Offset Weak Fundamentals and the Effects of Contagion," IMF Working Papers 99/88.

CHEN, S.-L., J. D. JACKSON, H. KIM, AND P. RESIANDINI (2014): "What Drives Commodity Prices?," *American Journal of Agricultural Economics*, 96(5), 1455–1468.

CHEUNG, Y.-W., AND K. S. LAI (1995): "Lag Order and Critical Values of the Augmented Dickey-Fuller Test," *Journal of Business & Economic Statistics*, 13(3), 277–80.

CHIAIE, S. D., L. FERRARA, AND D. GIANNONE (2022): "Common Factors of Commodity Prices," *Journal of Applied Econometrics*, 37(3), 461–476.

CHRISTENSEN, I., AND F. LI (2014): "Predicting Financial Stress Events: A Signal Extraction Approach," *Journal of Financial Stability*, 14(C), 54–65.

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.

DEB, P., AND M. SEFTON (1996): "The Distribution of a Lagrange Multiplier Test of Normality," *Economics Letter*, 51, 123–130.

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.

ELLIOTT, G., T. J. ROTHENBERG, AND J. H. STOCK (1996): "Efficient Tests for an Autoregressive Unit Root," *Econometrica*, 64(4), 813–836.

ENGEL, C., N. MARK, AND K. WEST (2015): "Factor Model Forecasts of Exchange Rates," *Econometric Reviews*, 34(1-2), 32–55.

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.

GREENAWAY-MCGREVY, R., N. C. MARK, D. SUL, AND J. WU (2018): "Identifying Exchange Rate Common Factors," *International Economic Review*, 59(4), 2193–2218.

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.

Jarque, C. M., and A. K. Bera (1980): "Efficient Tests for Normality, Homoscedasticity and Serial Independence of Regression Residuals," *Economics Letters*, 6(3), 255–259.

——— (1987): "A Test for Normality of Observations and Regression Residuals," *International Statistical Review*, 55, 163–172.

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., and K. Ko (2020): "Improving Forecast Accuracy of Financial Vulnerability: PLS Factor Model Approach," *Economic Modelling*, 88(C), 341–355.

Kim, H., and W. Shi (2021): "Forecasting Financial Vulnerability in the USA: A Factor Model Approach," *Journal of Forecasting*, 40(3), 439–457.

Kim, H., W. Shi, and H. H. Kim (2020): "Forecasting Financial Stress Indices in Korea: A Factor Model Approach," *Empirical Economics*, 59(6), 2859–2898.

Kim, Y., and C. Park (2020): "Are Exchange Rates Disconnected from Macroeconomic Variables? Evidence from the Factor Approach," *Empirical Economics*, 58(4), 1713–1747.

Kliesen, K. L., M. T. Owyang, and E. K. Vermann (2012): "Disentangling Diverse Measures: A Survey of Financial Stress Indexes," *Review*, (Sep), 369–398.

Liu, L., H. R. Moon, and F. Schorfheide (2023): "Forecasting with a Panel Tobit Model," *Quantitative Economics*, 14(1), 117–159.

Ludvigson, S., and S. Ng (2009): "Macro Factors in Bond Risk Premia," *Review of Financial Studies*, 22(12), 5027–5067.

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.

Pesaran, M. H. (2021): "General Diagnostic Tests for Cross-Sectional Dependence in Panels," *Empirical Economics*, 60(1), 13–50.

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.

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.

TANNER, E. (2002): "Exchange Market Pressure, Currency Crises, and Monetary Policy: Additional Evidence From Emerging Markets," IMF Working Papers No. 02/14.

WEST, K. D., AND K.-F. WONG (2014): "A Factor Model for Co-Movements of Commodity Prices," *Journal of International Money and Finance*, 42(C), 289–309.

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. Top 10 Bank Holding Companies

Name	ID	RSSDID	Location	Asset ($ Mil)	BL (%)	CL (%)	RL (%)
JPMorgan	JPM	1039502	New York, NY	3,757,576	26.3	20.1	32.8
Bank of America	BAC	1073757	Charlotte, NC	3,085,446	26.9	17.8	42.9
Wells Fargo	WFC	1120754	San Francisco, CA	1,954,901	19.7	20.9	48.0
U.S. Bancorp	USB	1119794	Minneapolis, MN	567,495	28.5	18.2	38.7
PNC	PNC	1069778	Pittsburgh, PA	554,457	32.1	13.8	39.7
Truist	TFC	1074156	Charlotte, NC	529,884	16.5	13.5	62.5
Fifth Third	FITB	1070345	Cincinnati, OH	207,731	27.7	18.6	39.5
BMO	BMO	1245415	Wilmington, DE	195,146	34.6	13.4	35.2
Citizens	CFG	1132449	Providence, RI	187,549	19.6	17.7	51.4
Keycorp	KEY	1068025	Cleveland, OH	187,198	28.8	17.4	38.4

Note: The top 10 bank holding companies (BHCs) are selected based on the dollar value of total assets as of September 30, 2021 among the largest BHCs with balanced available loan data we are interested in. Some large BHCs such as Goldman Sachs, Morgan Stanley, and Charles Schwab were excluded due to lack of sufficient business and consumer loan data. BL and CL denote the average shares of business loans and consumer loans, respectively, of each BHC during the sample period 1986:I to 2021:I. RL denotes the real estate loans during 1991:1 to 2021:1.

Table 2. Summary Statistics: Top 10 Charge-Off Rates

			All Loans COR					
ID	Mean	Median	Std Dev	Min	Max	Skew	Kurt	JB
JPM	0.710	0.519	0.474	0.251	2.447	2.284	19.821	1747
BAC	0.561	0.386	0.491	0.185	2.568	0.300	9.643	256
WFC	0.553	0.461	0.371	0.135	1.932	-0.018	7.523	118
USB	0.574	0.485	0.330	0.179	2.060	0.868	15.449	908
PNC	0.382	0.284	0.359	-0.098	2.086	1.792	16.630	1142
TFC	0.318	0.220	0.287	0.061	1.565	0.368	21.438	1958
FITB	0.423	0.276	0.422	0.100	2.075	2.990	23.269	2568
BMO	0.385	0.284	0.385	-0.162	1.889	0.776	11.956	475
CFG	0.350	0.218	0.308	0.057	1.545	1.308	10.325	348
KEY	0.456	0.302	0.458	0.099	2.379	2.105	18.596	1501
Top 10	0.568	0.419	0.391	0.200	2.313	2.821	15.998	1154
Top 100	1.025	0.740	0.671	0.390	3.360	1.386	10.472	365
All Banks	0.912	0.650	0.577	0.330	3.020	0.989	7.502	139
			Business Loans COR					
ID	Mean	Median	Std Dev	Min	Max	Skew	Kurt	JB
JPM	0.547	0.324	0.544	0.035	2.926	1.790	10.565	403
BAC	0.413	0.251	0.457	-0.260	2.718	0.733	9.423	250
WFC	0.499	0.378	0.411	0.027	2.213	0.409	6.404	70
USB	0.447	0.247	0.515	-0.322	2.894	0.109	6.033	53
PNC	0.470	0.248	0.692	-0.390	5.534	3.864	35.348	6360
TFC	0.293	0.214	0.249	-0.052	1.197	-0.386	9.403	239
FITB	0.397	0.301	0.353	-0.033	2.263	1.481	12.941	619
BMO	0.491	0.298	0.688	-0.531	4.916	2.512	31.441	4796
CFG	0.363	0.227	0.453	-0.364	2.334	0.358	15.148	851
KEY	0.445	0.200	0.610	-0.281	3.939	-0.454	15.197	860
Top 10	0.466	0.317	0.379	0.045	1.638	0.637	5.961	60
Top 100	0.745	0.520	0.615	0.030	2.660	1.342	7.297	148
All Banks	0.784	0.510	0.608	0.120	2.650	0.569	5.172	35
			Consumer Loans COR					
ID	Mean	Median	Std Dev	Min	Max	Skew	Kurt	JB
JPM	1.698	1.545	0.842	0.580	5.056	0.890	16.411	1052
BAC	1.649	1.287	0.991	0.782	5.550	-1.297	16.153	1033
WFC	1.344	1.229	0.546	0.560	4.098	1.662	21.962	2131
USB	1.396	1.309	0.546	0.298	2.882	-0.284	8.350	166
PNC	0.689	0.580	0.539	0.113	3.446	0.293	17.694	1243
TFC	0.965	0.990	0.438	0.202	2.515	0.424	15.801	946
FITB	0.708	0.591	0.367	0.223	2.080	2.243	15.057	952
BMO	0.580	0.377	0.508	0.109	2.122	-5.333	55.606	16567
CFG	0.616	0.582	0.361	0.036	1.972	-0.605	8.526	184
KEY	0.932	0.944	0.467	0.291	2.614	0.619	10.300	315
Top 10	1.424	1.329	0.708	0.634	4.143	-1.563	21.632	2052
Top 100	2.643	2.400	1.023	1.500	7.080	-0.370	7.888	141
All Banks	2.452	2.280	1.001	1.350	6.700	-0.369	7.690	130

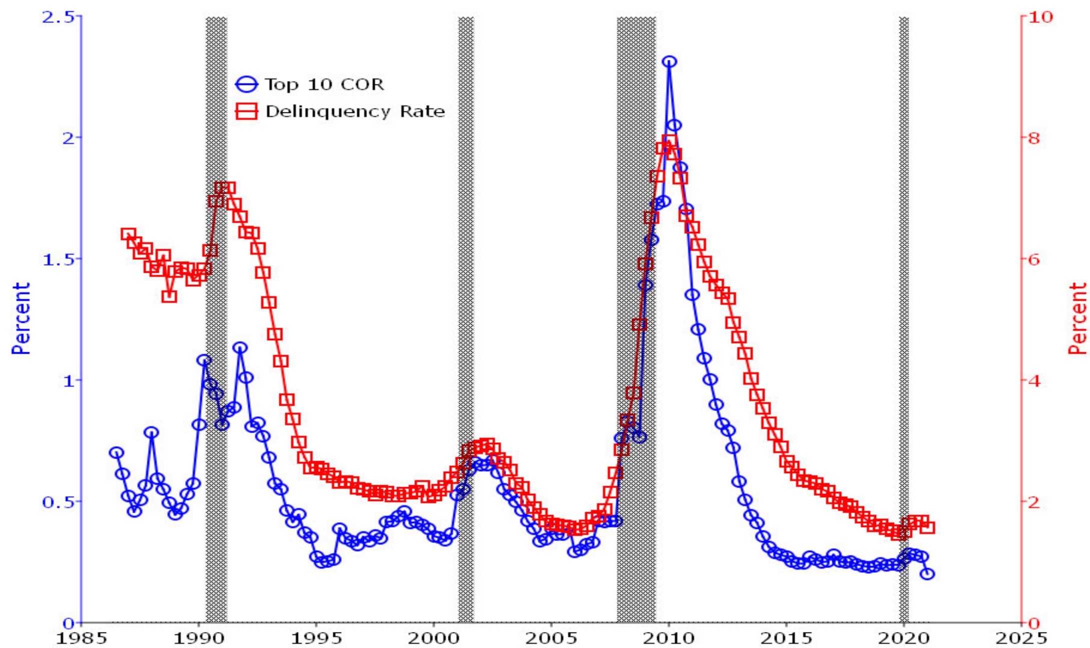
Note: Skew and Kurt denote skewness and kurtosis, respectively. Results overall imply an asymmetric and fat-tailed distribution of COR. JB denotes the Jarque-Bera statistics (Jarque and Bera, 1980, 1987; Deb and Sefton, 1996). The test rejects the null hypothesis of normality for all cases at any conventional significance level when the critical values from Deb and Sefton (1996).

Table 3. Cross-Section Dependence in the Top 10 Charge-Off Rates

	Average Cross-Correlations ($\hat{\rho}_i$)		
	$co_{i,All,t}$	$co_{i,Bus,t}$	$co_{i,Con,t}$
JPM	0.345	0.193	0.397
BAC	0.409	0.465	0.414
WFC	0.259	0.204	0.186
USB	0.238	0.466	0.333
PNC	0.234	0.381	0.206
TFC	0.323	0.431	0.290
FITB	0.213	0.383	0.327
BMO	0.178	0.399	0.110
CFG	0.309	0.271	0.317
KEY	0.343	0.469	0.294
Top 10	0.398	0.372	0.463
Top 100	0.278	0.449	0.363
All Banks	0.214	0.475	0.362
Average $\hat{\rho}_i$	0.228	0.330	0.255
CD	$23.515^{\ddagger}$	$33.968^{\ddagger}$	$26.279^{\ddagger}$

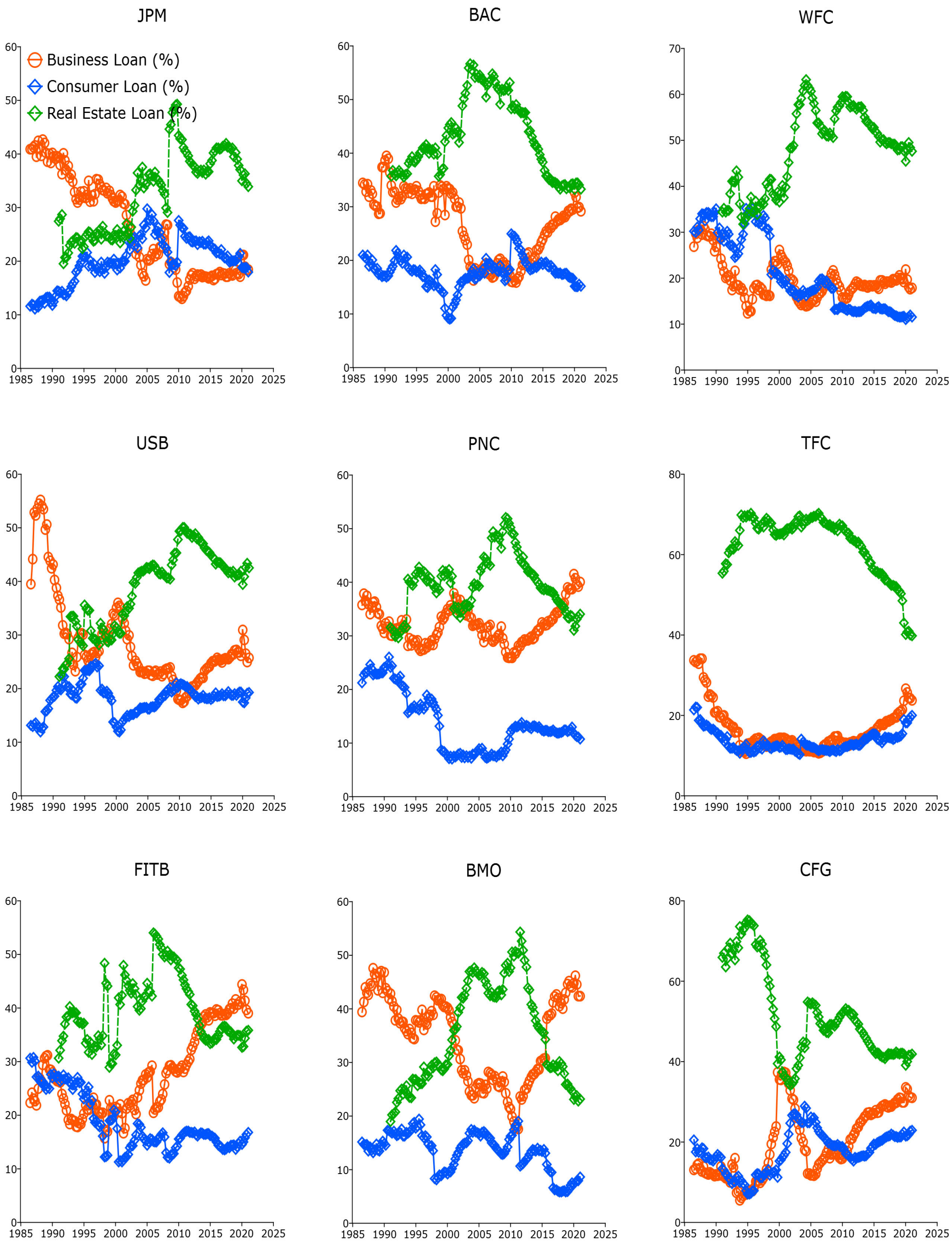
Note: $\hat{\rho}_{i,j}$ denotes the cross-correlations of the residuals $\varepsilon_{i,t}$ and $\varepsilon_{j,t}$ from the ADF regressions. We report the average cross-correlations of each CORs, $\hat{\rho}_i = N^{-1}\sum_{i\neq j}\hat{\rho}_{i,j}$. Average $\hat{\rho}_i$ is the average value of all CORs' average cross-correlations. CD denotes the cross-section dependence statistics from Pesaran (2021). The superscript ‡ denotes a rejection at the 1% signficance level.

Figure 1. Countercyclicality Dynamics of Charge-off-Rates



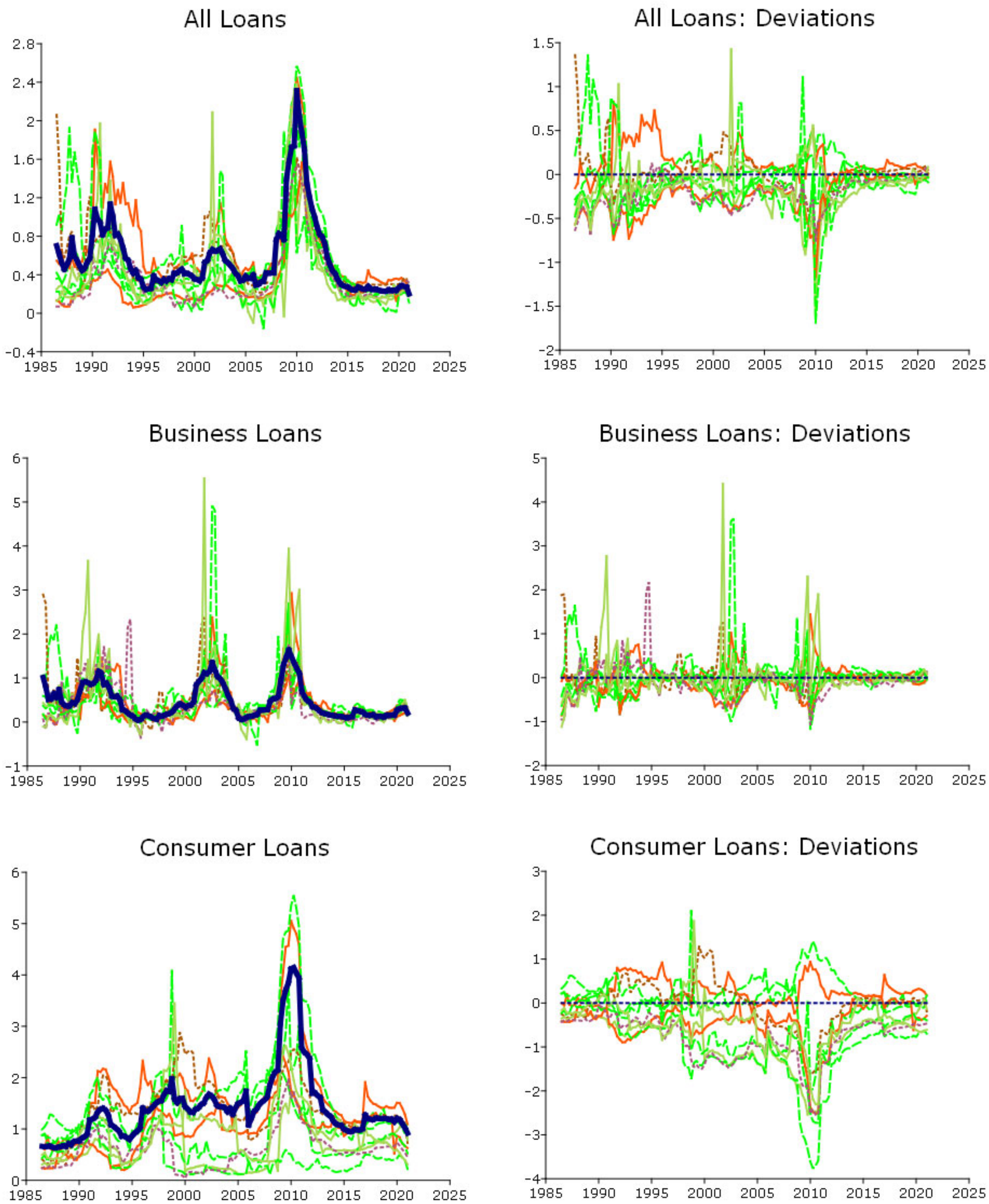
Note: We report the average COR of all loans of the top 10 BHCs in the U.S. and the delinquency rate of the top 100 U.S. banks. Shaded areas denote recessions.

Figure 2. Top 10 Business, Consumer, and Other Loan Shares



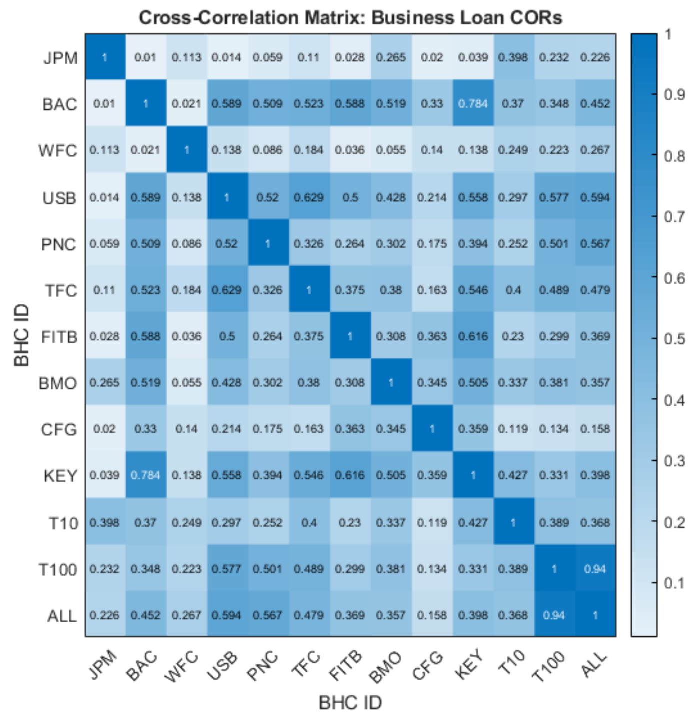
Note: All loan shares are as the percent (%) of total outstanding loan amounts of each BHC. Real estate loan shares lack 18 quarterly observations, starting from 1991:I to 2021:I.

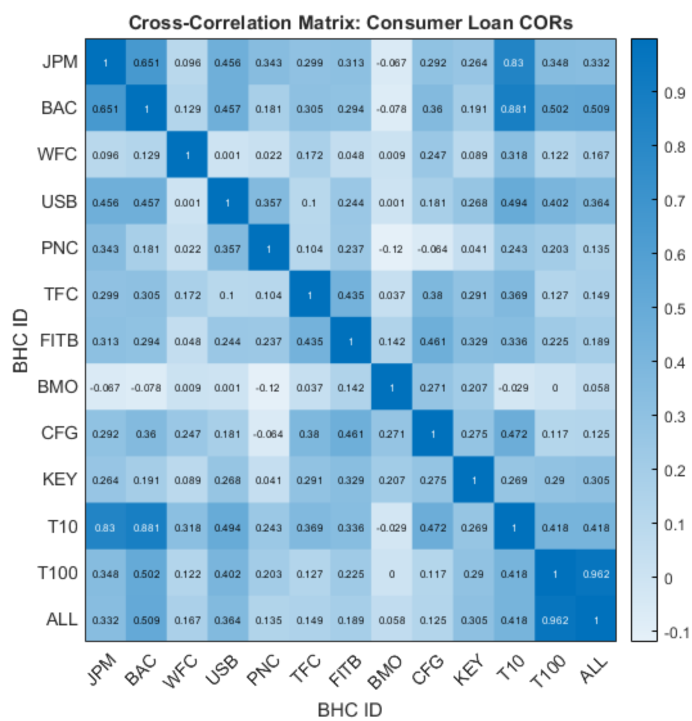
Figure 3. Top 10 Net Charge-Off Rates



Note: The solid bold lines in the first column are the average net charge-off rates (CORs) of the top 10 BHCs, whereas individual CORs are lighter lines. The figures in the second column are deviations of individual CORs from the average rates.

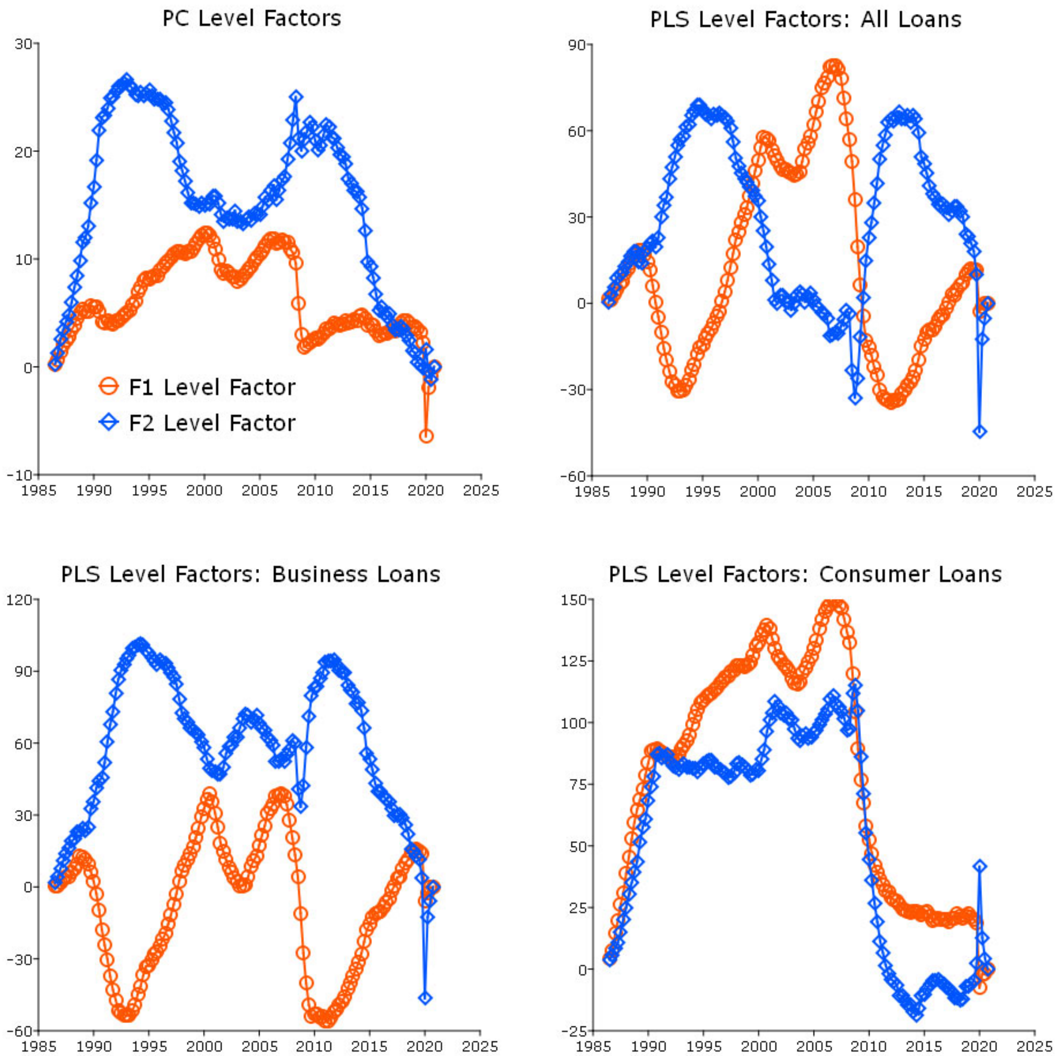
Figure 4. Cross-Correlation Matrix of Net Charge-Off Rates





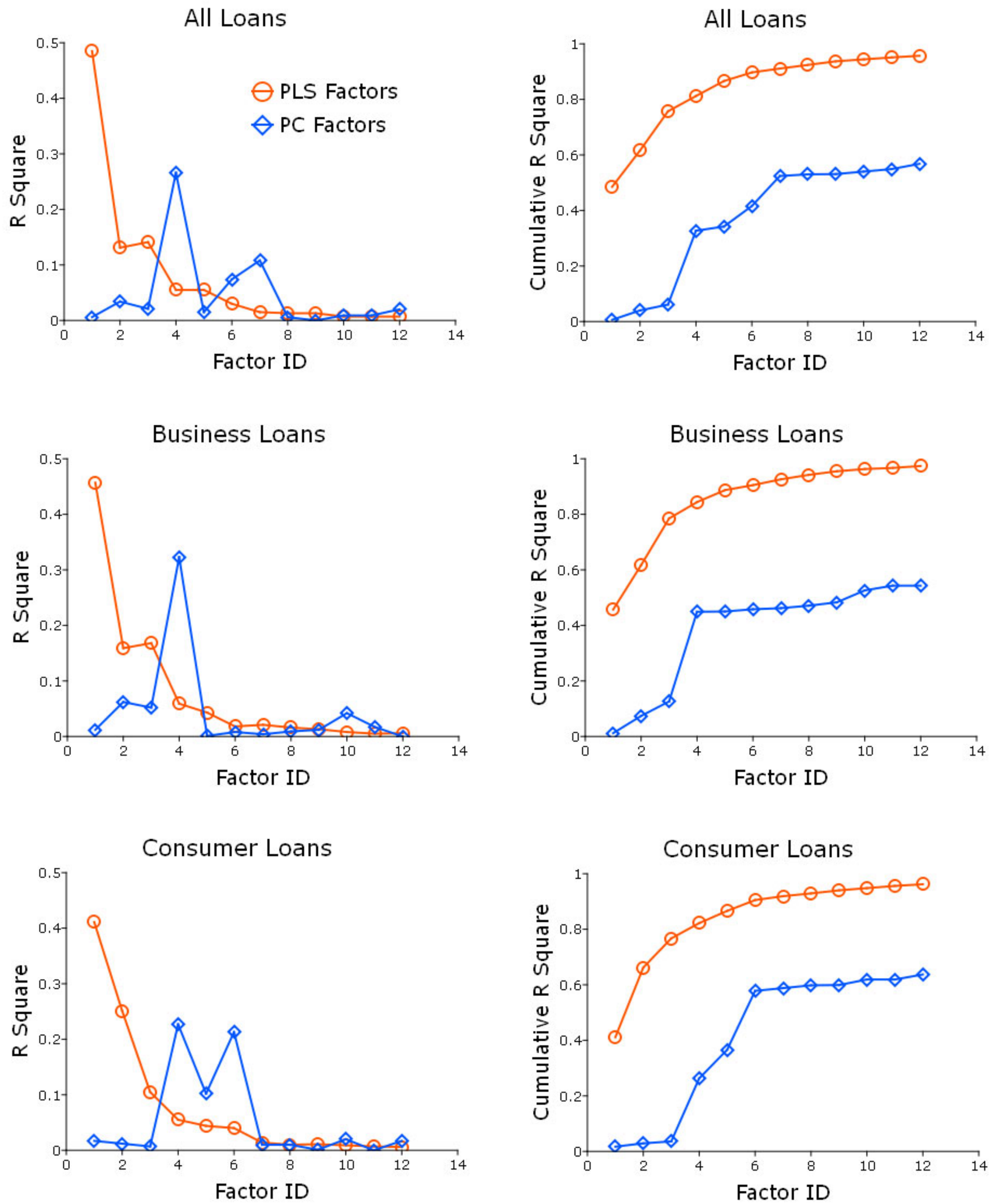
Note: The heatmap reports the cross-correlations ($\hat{\rho}_{i,j}$) of the residuals $\varepsilon_{i,t}$ and $\varepsilon_{j,t}$ from the ADF regressions of each pair of CORs.

Figure 5. Level Common Factor Estimates for Top 10 CORs



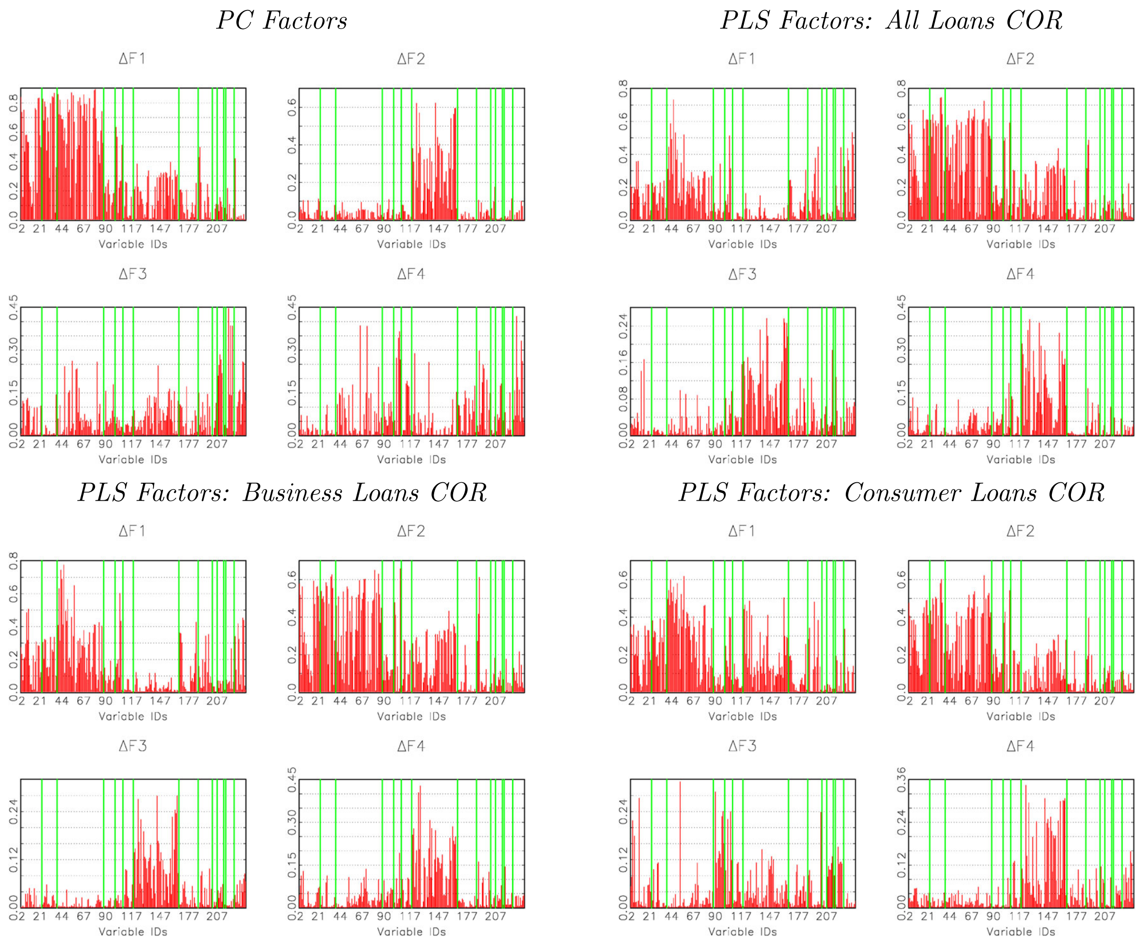
Note: We obtained up to 4 factors by applying the method of the principal components to 237 quarterly frequency macroeconomic time series variables. Level factors are obtained by re-integrating estimated common factors. PLS factors are target-specific factors for each type loans.

Figure 6. In-Sample Fit Analysis of Factor Estimates



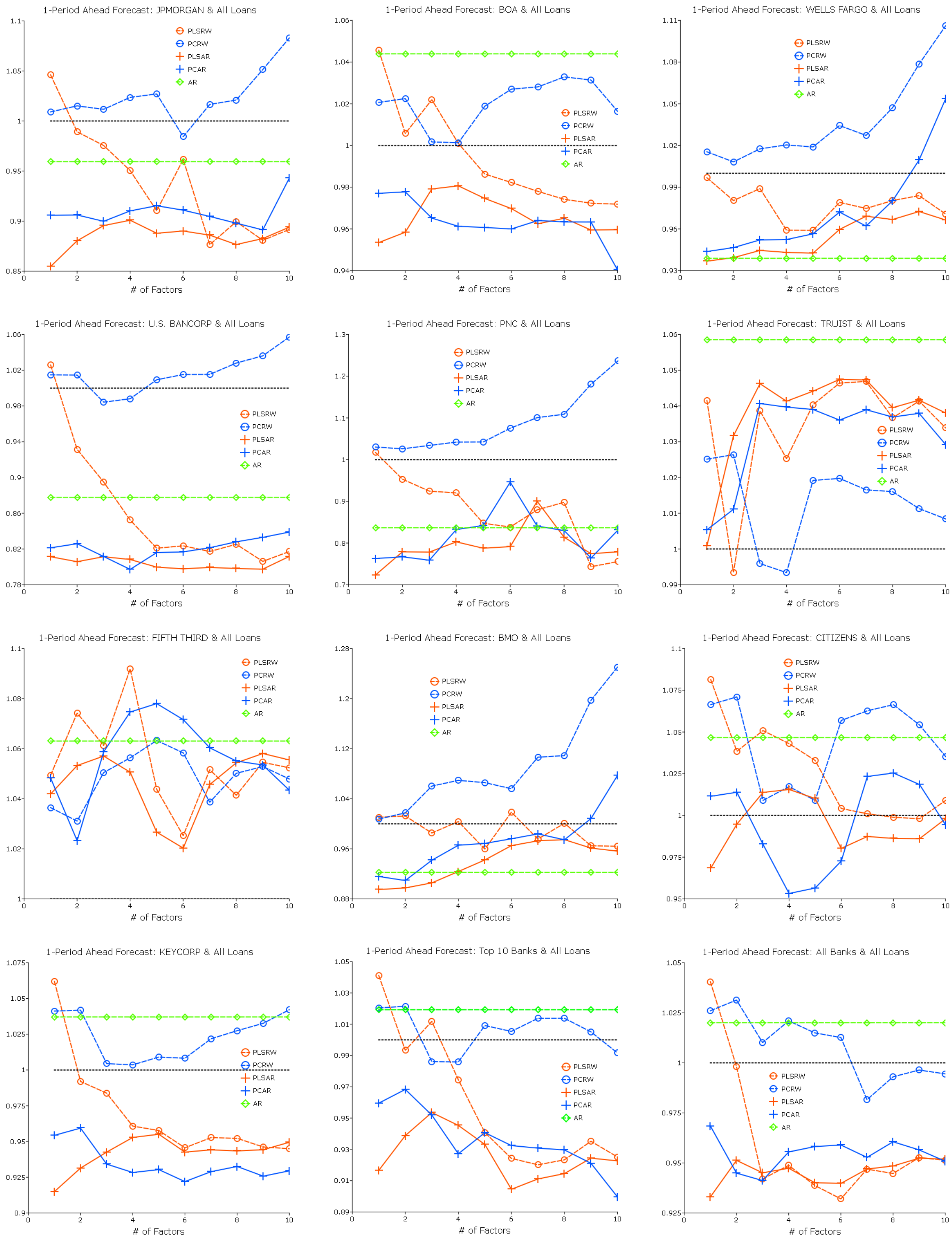
Note: Estimated R^2 are reported in the first column, while cumulative value figures are in the second column.

Figure 7. Marginal $\mathbf{R^2}$ Analysis



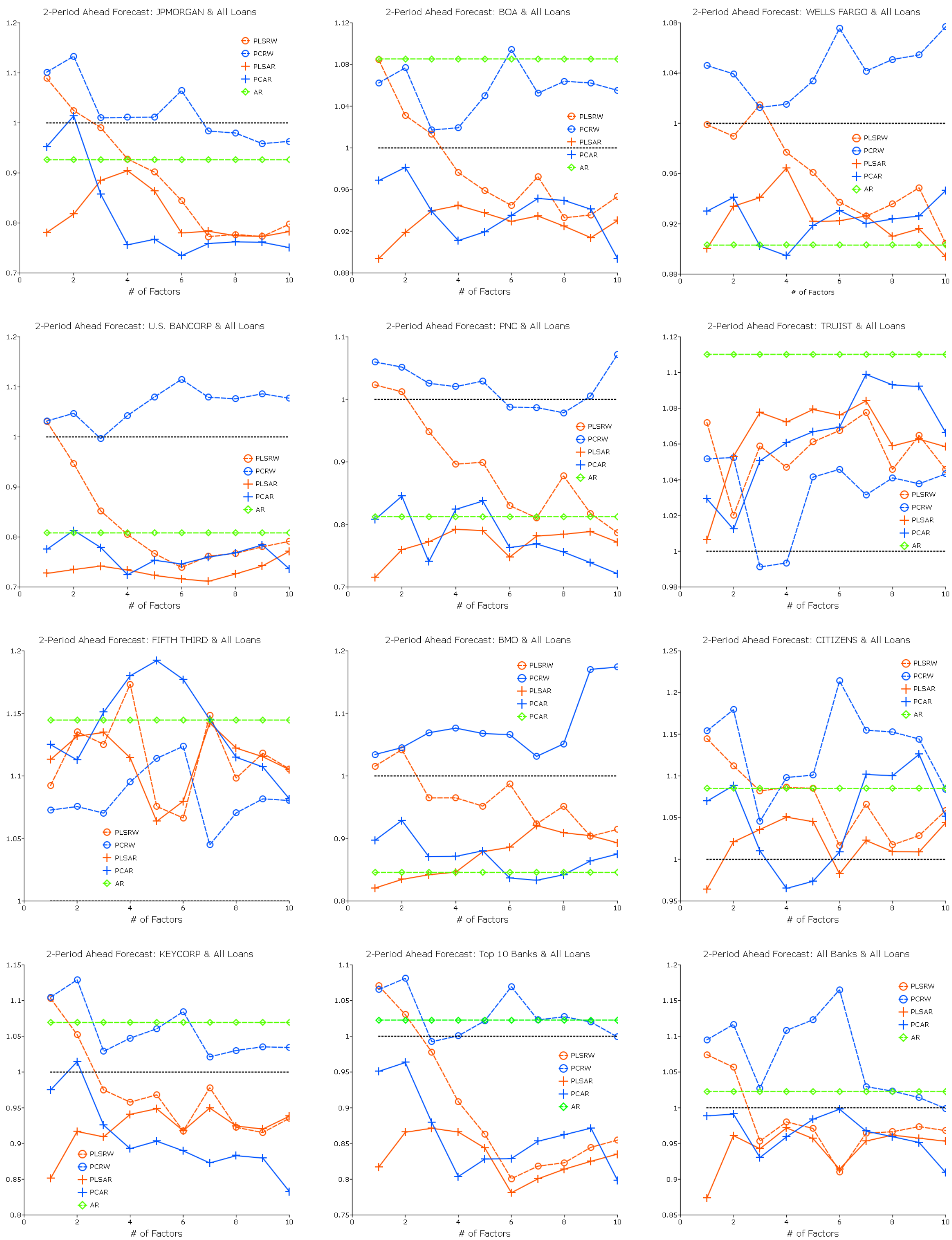
Note: The marginal R^2 is obtained by regressing each of the individual time series variables onto each estimated factor, one at a time, using the full sample of data. The individual series in each group are separated by vertical lines. The data IDs are on the x-axis.

Figure 8. 1-Period Ahead Out-of-Sample Forecast Peformance: All Loans COR



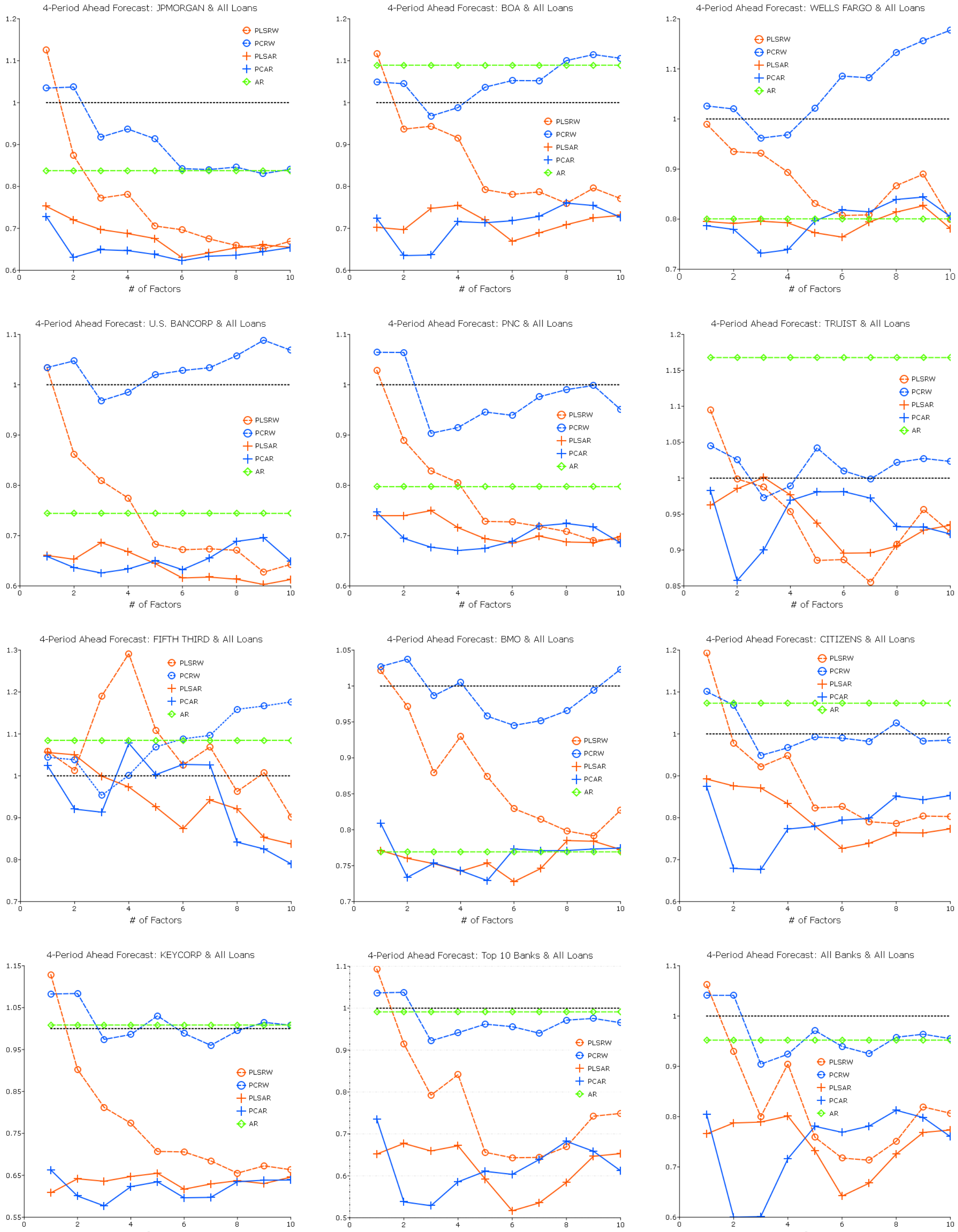
Note: We report the RRMSPE statistics with the random walk benchmark model. Therefore, lower values than 1 imply that competing models outperform the benchmark. We assess the 1-quarter ahead out-of-sample predictability of our factor models with up to 10 factors for all loan CORs of top 10 individual BHCs and the two aggregate CORs.

Figure 9. 2-Period Ahead Out-of-Sample Forecast Peformance: All Loans COR



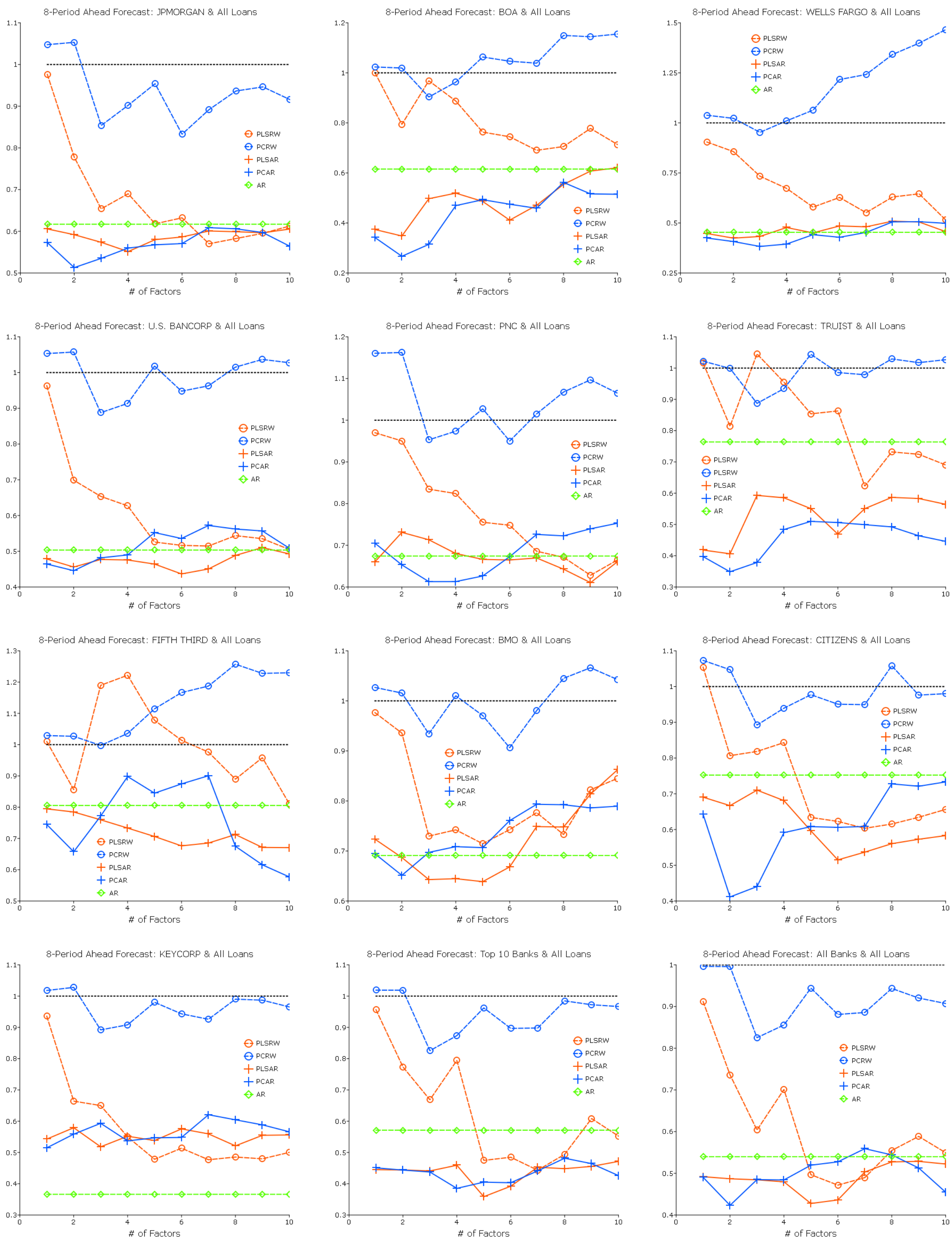
Note: We report the RRMSPE statistics with the random walk benchmark model. Therefore, lower values than 1 imply that competing models outperform the benchmark. We assess the 2-quarter ahead out-of-sample predictability of our factor models with up to 10 factors for all loan CORs of top 10 individual BHCs and the two aggregate CORs.

Figure 10. 4-Period Ahead Out-of-Sample Forecast Peformance: All Loans COR



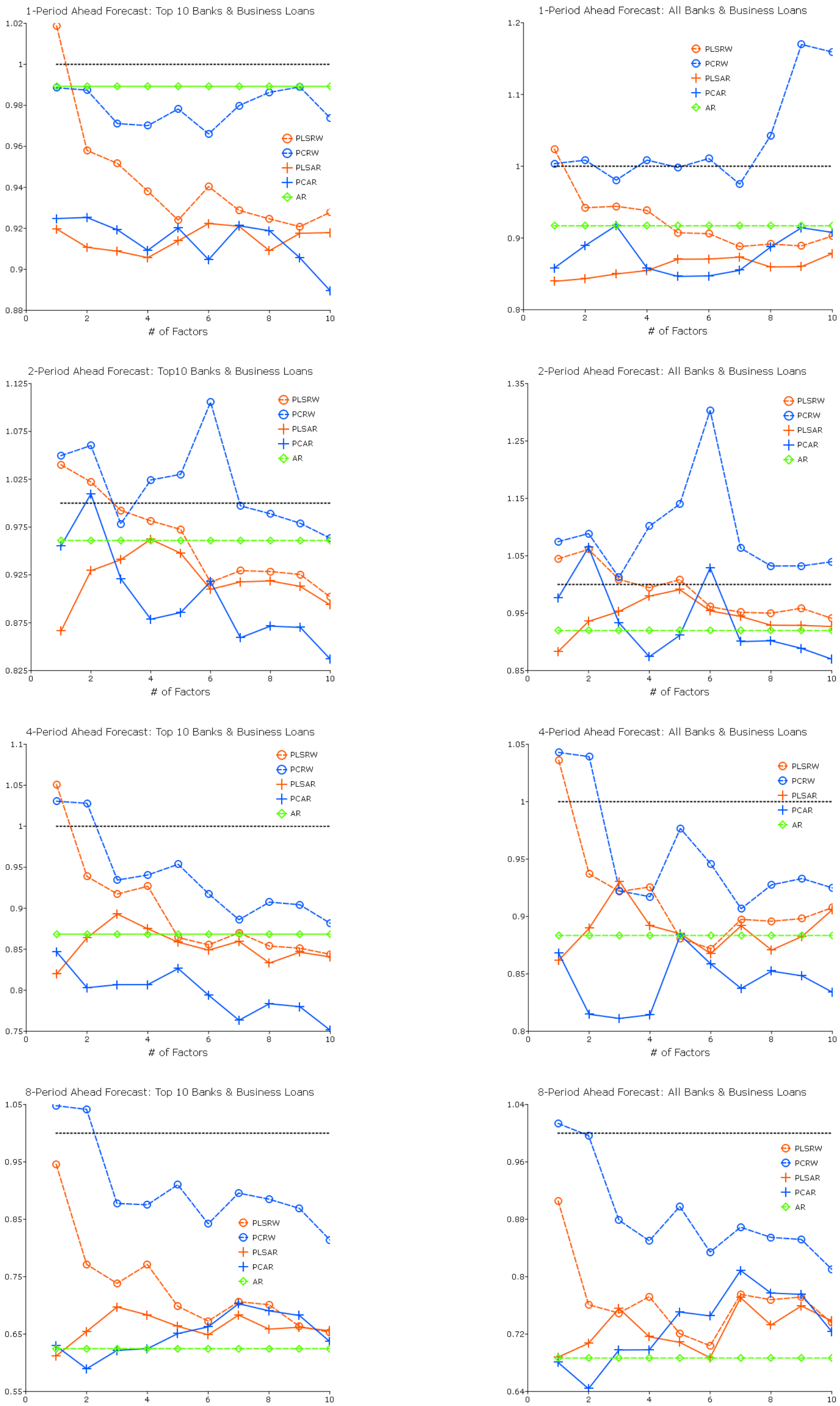
Note: We report the RRMSPE statistics with the random walk benchmark model. Therefore, lower values than 1 imply that competing models outperform the benchmark. We assess the 4-quarter (1-year) ahead out-of-sample predictability of our factor models with up to 10 factors for all loan CORs of top 10 individual BHCs and the two aggregate CORs.

Figure 11. 8-Period Ahead Out-of-Sample Forecast Peformance: All Loans COR



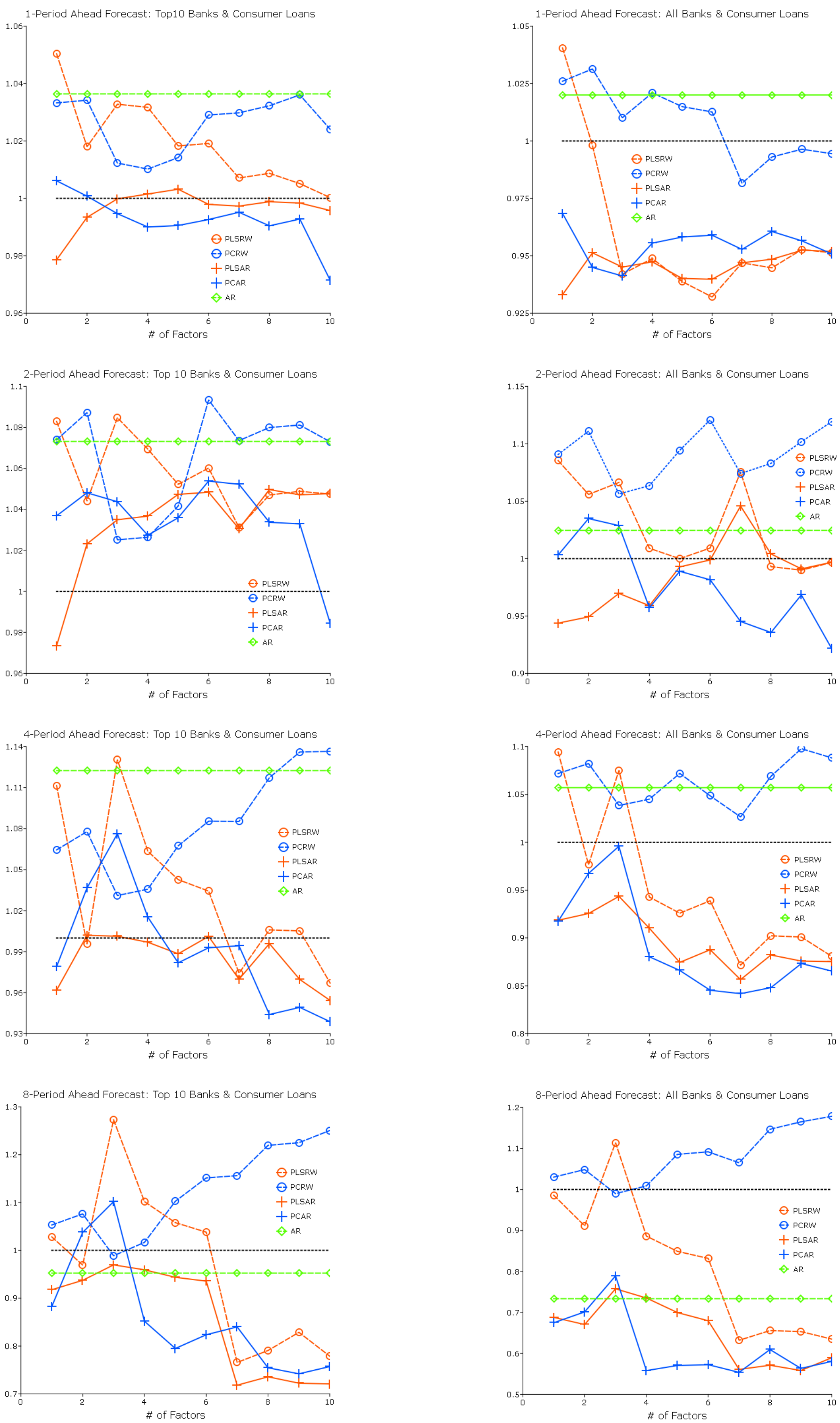
Note: We report the RRMSPE statistics with the random walk benchmark model. Therefore, lower values than 1 imply that competing models outperform the benchmark. We assess the 8-quarter (2-year) ahead out-of-sample predictability of our factor models with up to 10 factors for all loan CORs of top 10 individual BHCs and the two aggregate CORs.

Figure 12. Out of Sample Forecast Performance: Aggregate CORs of Business Loans



Note: We report the RRMSPE statistics with the random walk benchmark model. Therefore, lower values than 1 imply that competing models outperform the benchmark. We assess 1-quarter to 8-quarter ahead out-of-sample predictability of our factor models with up to 10 factors for the top 10 average COR and all banks average COR for business loans.

Figure 13. Out of Sample Forecast Performance: Aggregate CORs of Consumer Loans

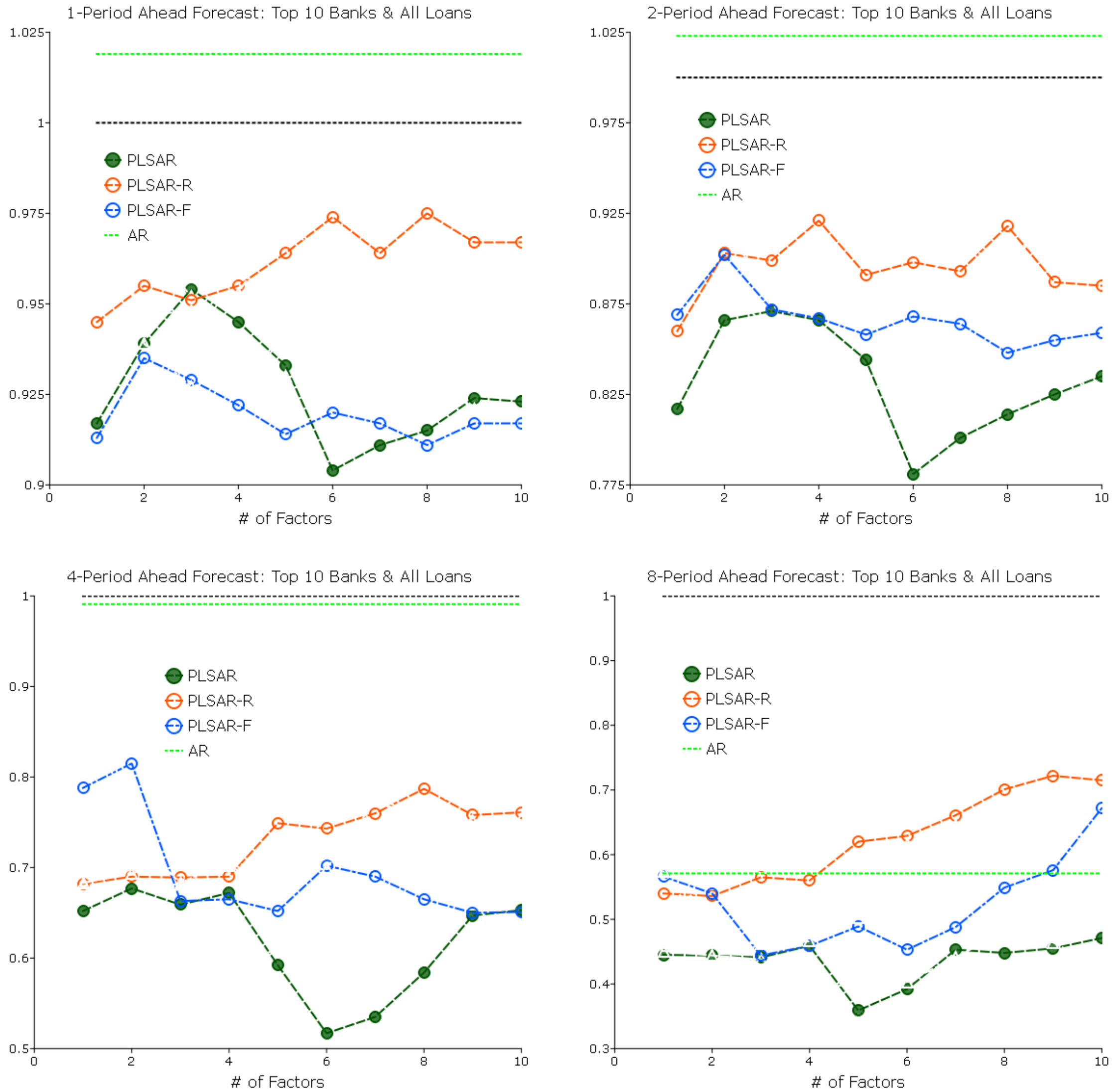


Note: We report the RRMSPE statistics with the random walk benchmark model. Therefore, lower values than 1 imply that competing models outperform the benchmark. We assess 1-quarter to 8-quarter ahead out-of-sample predictability of our factor models with up to 10 factors for the top 10 average COR and all banks average COR for consumer loans.

Figure 14. Out of Sample Forecast Performance: Aggregate CORs of Real Estate Loans

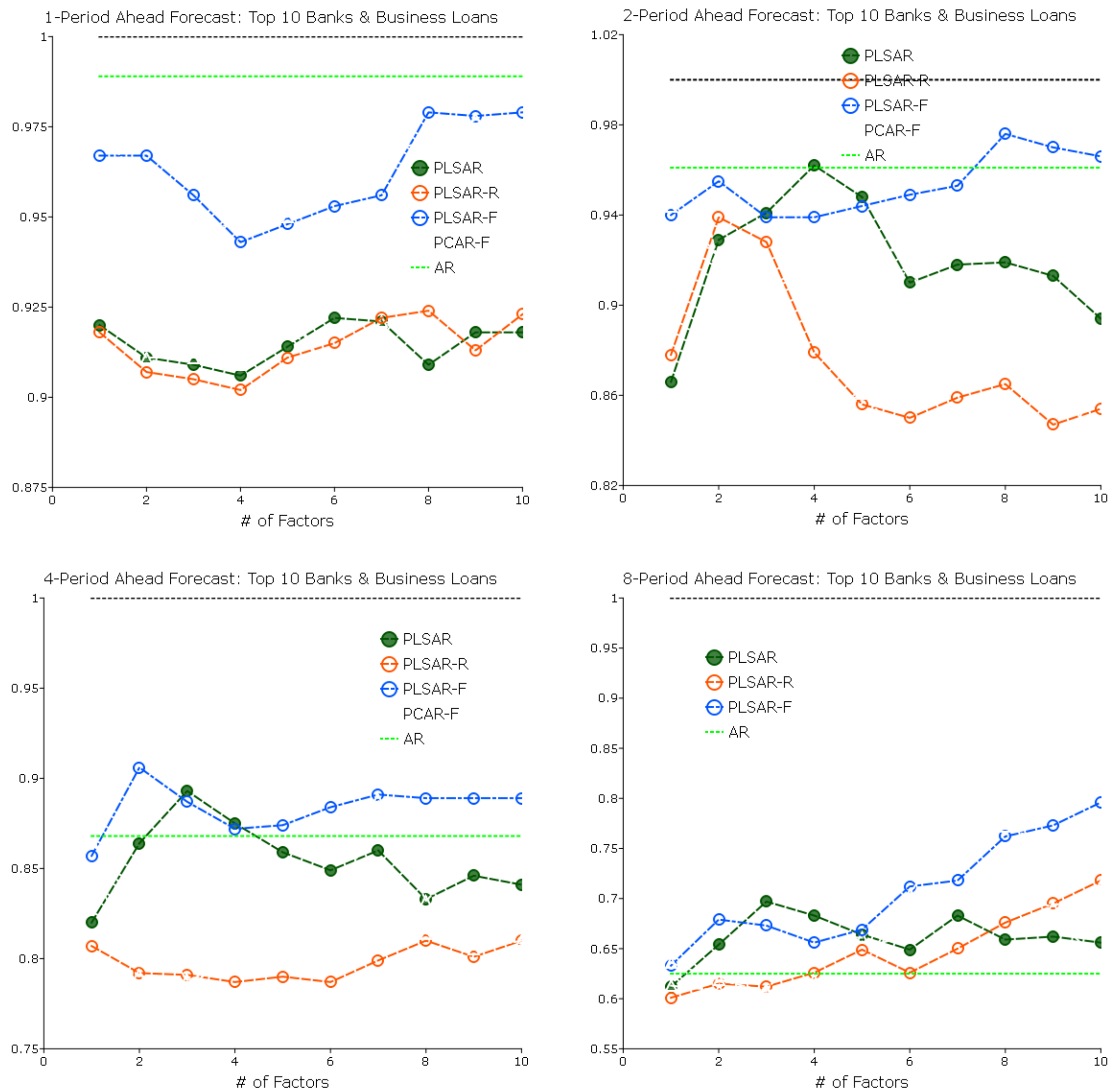
Note: We report the RRMSPE statistics with the random walk benchmark model. Therefore, lower values than 1 imply that competing models outperform the benchmark. We assess 1-quarter to 8-quarter ahead out-of-sample predictability of our factor models with up to 10 factors for the top 10 average COR and all banks average COR for real estate loans. Real estate loan shares lack 18 quarterly observations, starting from 1991:I to 2021:I.

Figure 15. Real vs. Finance Factors: Top 10 Banks Average COR of All Loans



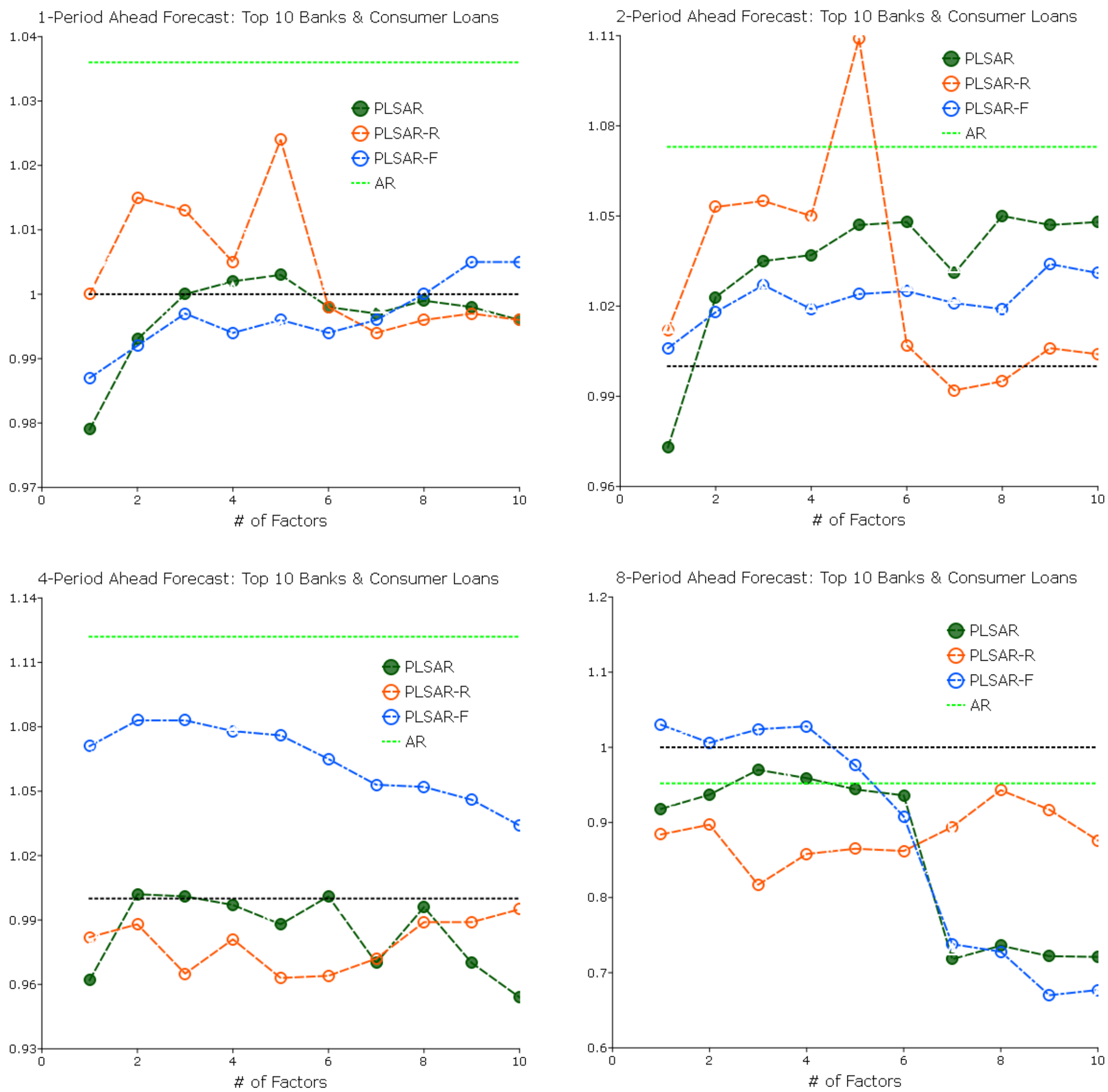
Note: We report the RRMSPE statistics with the random walk benchmark model. Therefore, lower values than 1 imply that competing models outperform the benchmark. We assess 1-quarter to 8-quarter ahead out-of-sample predictability of our factor models with up to 10 factors for the top 10 average all loan COR when factors are estimated via PLS utilizing real activity variables, financial sector variables, and all variables.

Figure 16. Real vs. Finance Factors: Top 10 Banks Average COR of Business Loans



Note: We report the RRMSPE statistics with the random walk benchmark model. Therefore, lower values than 1 imply that competing models outperform the benchmark. We assess 1-quarter to 8-quarter ahead out-of-sample predictability of our factor models with up to 10 factors for the top 10 average business loan COR when factors are estimated via PLS utilizing real activity variables, financial sector variables, and all variables.

Figure 17. Real vs. Finance Factors: Top 10 Banks Average COR of Consumer Loans



Note: We report the RRMSPE statistics with the random walk benchmark model. Therefore, lower values than 1 imply that competing models outperform the benchmark. We assess 1-quarter to 8-quarter ahead out-of-sample predictability of our factor models with up to 10 factors for the top 10 average consumer loan COR when factors are estimated via PLS utilizing real activity variables, financial sector variables, and all variables.

Appendix

Table A1. Macroeconomic Data Descriptions

Classifications	Group ID	Data ID	Data Descriptions
Real Activity	#1	1-22	NIPA
	#2	23-38	Industrial Production
	#3	39-87	Employment and Unemployment
	#4	88-99	Housing
	#5	100-107	Inventories, Orders, and Sales
	#6	108-118	Earnings and Productivity
Nominal/Financial	#7	119-166	Prices
	#8	167-186	Interest Rates
	#9	187-201	Money and Credit
	#10	202-206	Exchange Rates
	#11	207-213	Stock Markets
	#12	214-215	Others
	#13	216-224	Household Balance Sheets
	#14	225-237	Non-Household Balance Sheets

Note: We obtained all data from the FRED-QD (https://research.stlouisfed.org/econ/mccracken/fred-databases/). Quantity variables are log-transformed, while percent variables are divided by 100.

Table A2. DFGLS Unit Root Test Results

ID	$co_{i,A,t}$	$co_{i,B,t}$	$co_{i,C,t}$
JPM	-2.156†	-2.235†	-1.443
BAC	-1.470	-2.153†	-1.283
WFC	-1.408	-1.677*	-1.637
USB	-0.339	-0.574	-1.044
PNC	-2.247†	-2.970‡	-2.287†
TFC	-1.360	-2.063†	-1.216
FITB	-1.621	-1.953*	-2.105†
BMO	-3.003‡	-3.114‡	-1.888*
CFG	-1.306	-2.465†	-1.456
KEY	-1.457	-1.759*	-1.755*
Top 10	-1.516	-1.688*	-1.255
Top 100	-2.168†	-1.942*	-1.700*
All Banks	-1.929*	-2.796‡	-1.505

Note: We report the DFGLS unit root test statistics by Elliott, Rothenberg, and Stock (1996). The number of lags was selected by the Bayesian Information Criteria. *, †, and ‡ denote rejections of the unit root null hypothesis at the 10%, 5%, and 1% significance level, respectively.

Table A3: Out-of-Sample Forecast Results: All Loans COR

		Top10 Banks RRMSPE					*All Banks RRMSPE*				
$\#H$	$\#F$	PLS_{RW}	PC_{RW}	PLS_{AR}	PC_{AR}	AR	PLS_{RW}	PC_{RW}	PLS_{AR}	PC_{AR}	AR
1	1	1.041	1.020	0.917	0.960	1.019	1.041	1.026	0.933	0.968	1.020
	2	0.993	1.021	0.939	0.968	1.019	0.998	1.031	0.951	0.945	1.020
	3	1.012	0.986	0.954	0.952	1.019	0.942	1.010	0.945	0.941	1.020
	4	0.974	0.986	0.945	0.927	1.019	0.949	1.021	0.947	0.955	1.020
	5	0.941	1.009	0.933	0.941	1.019	0.939	1.015	0.940	0.958	1.020
	6	0.924	1.005	0.904	0.932	1.019	0.932	1.013	0.940	0.959	1.020
	7	0.920	1.014	0.911	0.931	1.019	0.947	0.982	0.947	0.953	1.020
	8	0.923	1.014	0.915	0.930	1.019	0.945	0.993	0.948	0.961	1.020
	9	0.935	1.005	0.924	0.921	1.019	0.953	0.996	0.952	0.957	1.020
	10	0.925	0.992	0.923	0.900	1.019	0.951	0.994	0.952	0.951	1.020
2	1	1.071	1.065	0.817	0.951	1.023	1.074	1.095	0.874	0.989	1.023
	2	1.030	1.081	0.866	0.963	1.023	1.057	1.117	0.961	0.991	1.023
	3	0.978	0.992	0.871	0.879	1.023	0.953	1.027	0.943	0.930	1.023
	4	0.908	1.001	0.866	0.804	1.023	0.980	1.108	0.972	0.959	1.023
	5	0.863	1.022	0.844	0.828	1.023	0.971	1.123	0.957	0.984	1.023
	6	0.801	1.069	0.781	0.829	1.023	0.910	1.165	0.914	0.998	1.023
	7	0.818	1.023	0.801	0.853	1.023	0.965	1.030	0.953	0.967	1.023
	8	0.823	1.027	0.814	0.862	1.023	0.966	1.023	0.962	0.960	1.023
	9	0.845	1.020	0.825	0.872	1.023	0.974	1.014	0.957	0.951	1.023
	10	0.855	0.999	0.835	0.798	1.023	0.968	0.999	0.953	0.909	1.023
4	1	1.093	1.036	0.652	0.735	0.991	1.062	1.041	0.766	0.805	0.952
	2	0.915	1.038	0.677	0.538	0.991	0.930	1.041	0.787	0.600	0.952
	3	0.792	0.922	0.659	0.529	0.991	0.800	0.904	0.789	0.601	0.952
	4	0.841	0.941	0.672	0.586	0.991	0.904	0.924	0.801	0.716	0.952
	5	0.656	0.962	0.592	0.611	0.991	0.759	0.971	0.732	0.781	0.952
	6	0.643	0.956	0.517	0.603	0.991	0.718	0.940	0.642	0.769	0.952
	7	0.644	0.940	0.535	0.639	0.991	0.713	0.925	0.668	0.781	0.952
	8	0.669	0.971	0.584	0.682	0.991	0.751	0.958	0.725	0.813	0.952
	9	0.742	0.976	0.647	0.659	0.991	0.819	0.964	0.768	0.798	0.952
	10	0.749	0.965	0.653	0.612	0.991	0.807	0.955	0.773	0.760	0.952
8	1	0.956	1.019	0.445	0.451	0.571	0.912	0.996	0.492	0.490	0.540
	2	0.773	1.019	0.444	0.444	0.571	0.736	0.996	0.487	0.423	0.540
	3	0.668	0.825	0.441	0.438	0.571	0.604	0.825	0.484	0.485	0.540
	4	0.794	0.874	0.459	0.385	0.571	0.701	0.856	0.480	0.484	0.540
	5	0.475	0.962	0.359	0.405	0.571	0.497	0.944	0.428	0.519	0.540
	6	0.485	0.897	0.392	0.403	0.571	0.472	0.881	0.436	0.527	0.540
	7	0.443	0.898	0.453	0.441	0.571	0.490	0.886	0.503	0.559	0.540
	8	0.494	0.985	0.448	0.482	0.571	0.555	0.943	0.527	0.544	0.540
	9	0.609	0.972	0.455	0.464	0.571	0.589	0.920	0.529	0.513	0.540
	10	0.552	0.967	0.471	0.426	0.571	0.549	0.907	0.522	0.455	0.540

Note: We report the RRMSPE statistics with the random walk benchmark model. Therefore, lower values than 1 imply that competing models outperform the benchmark. We assess the 1-quarter to 8-quarter ahead out-of-sample predictability of our factor models with up to 10 factors for all loan average CORs of the top 10 BHCs and all banks in the U.S.

Table A4: Out-of-Sample Forecast Results: Business Loans COR

		Top10 Banks RRMSPE					*All Banks RRMSPE*				
$\#H$	$\#F$	PLS_{RW}	PC_{RW}	PLS_{AR}	PC_{AR}	AR	PLS_{RW}	PC_{RW}	PLS_{AR}	PC_{AR}	AR
1	1	1.019	0.989	0.920	0.925	0.989	1.024	1.004	0.840	0.858	0.917
	2	0.958	0.987	0.911	0.925	0.989	0.942	1.008	0.843	0.889	0.917
	3	0.952	0.971	0.909	0.919	0.989	0.944	0.981	0.850	0.917	0.917
	4	0.938	0.970	0.906	0.909	0.989	0.938	1.009	0.855	0.858	0.917
	5	0.924	0.978	0.914	0.920	0.989	0.907	0.998	0.870	0.846	0.917
	6	0.940	0.966	0.922	0.905	0.989	0.906	1.011	0.871	0.847	0.917
	7	0.929	0.980	0.921	0.921	0.989	0.888	0.975	0.873	0.855	0.917
	8	0.925	0.986	0.909	0.919	0.989	0.891	1.043	0.860	0.887	0.917
	9	0.921	0.989	0.918	0.906	0.989	0.889	1.170	0.860	0.914	0.917
	10	0.928	0.974	0.918	0.890	0.989	0.903	1.159	0.878	0.908	0.917
2	1	1.040	1.050	0.866	0.955	0.961	1.045	1.075	0.883	0.977	0.920
	2	1.022	1.060	0.929	1.010	0.961	1.061	1.089	0.936	1.065	0.920
	3	0.992	0.978	0.941	0.921	0.961	1.009	1.012	0.953	0.933	0.920
	4	0.981	1.024	0.962	0.879	0.961	0.994	1.102	0.980	0.874	0.920
	5	0.972	1.030	0.948	0.886	0.961	1.008	1.140	0.991	0.911	0.920
	6	0.917	1.106	0.910	0.918	0.961	0.961	1.303	0.954	1.029	0.920
	7	0.929	0.997	0.918	0.859	0.961	0.952	1.063	0.944	0.900	0.920
	8	0.928	0.989	0.919	0.871	0.961	0.950	1.032	0.929	0.902	0.920
	9	0.925	0.979	0.913	0.870	0.961	0.959	1.032	0.929	0.889	0.920
	10	0.902	0.963	0.894	0.837	0.961	0.941	1.040	0.927	0.870	0.920
4	1	1.051	1.031	0.820	0.847	0.868	1.036	1.043	0.862	0.868	0.883
	2	0.939	1.028	0.864	0.803	0.868	0.937	1.039	0.890	0.815	0.883
	3	0.917	0.934	0.893	0.807	0.868	0.922	0.922	0.930	0.811	0.883
	4	0.927	0.941	0.875	0.807	0.868	0.926	0.917	0.892	0.814	0.883
	5	0.864	0.954	0.859	0.826	0.868	0.881	0.977	0.885	0.884	0.883
	6	0.855	0.918	0.849	0.794	0.868	0.872	0.946	0.868	0.859	0.883
	7	0.870	0.886	0.860	0.763	0.868	0.897	0.907	0.892	0.837	0.883
	8	0.854	0.908	0.833	0.784	0.868	0.896	0.928	0.871	0.852	0.883
	9	0.851	0.904	0.846	0.780	0.868	0.898	0.933	0.883	0.848	0.883
	10	0.844	0.882	0.841	0.751	0.868	0.908	0.925	0.906	0.834	0.883
8	1	0.946	1.048	0.612	0.629	0.625	0.905	1.013	0.688	0.680	0.687
	2	0.772	1.041	0.654	0.590	0.625	0.761	0.997	0.707	0.644	0.687
	3	0.738	0.878	0.697	0.622	0.625	0.749	0.879	0.756	0.698	0.687
	4	0.771	0.875	0.683	0.624	0.625	0.772	0.850	0.716	0.698	0.687
	5	0.698	0.910	0.664	0.651	0.625	0.721	0.898	0.708	0.751	0.687
	6	0.673	0.842	0.649	0.663	0.625	0.704	0.834	0.687	0.745	0.687
	7	0.706	0.896	0.683	0.703	0.625	0.775	0.869	0.771	0.809	0.687
	8	0.701	0.885	0.659	0.691	0.625	0.768	0.855	0.733	0.777	0.687
	9	0.663	0.869	0.662	0.682	0.625	0.772	0.852	0.759	0.775	0.687
	10	0.653	0.814	0.656	0.637	0.625	0.736	0.810	0.738	0.723	0.687

Note: We report the RRMSPE statistics with the random walk benchmark model. Therefore, lower values than 1 imply that competing models outperform the benchmark. We assess the 1-quarter to 8-quarter ahead out-of-sample predictability of our factor models with up to 10 factors for business loan average CORs of the top 10 BHCs and all banks in the U.S.

Table A5: Out-of-Sample Forecast Results: Consumer Loans COR

		Top10 Banks RRMSPE					*All Banks RRMSPE*				
$\#H$	$\#F$	PLS_{RW}	PC_{RW}	PLS_{AR}	PC_{AR}	AR	PLS_{RW}	PC_{RW}	PLS_{AR}	PC_{AR}	AR
1	1	1.050	1.033	0.979	1.006	1.036	1.049	1.036	0.954	0.977	1.012
	2	1.018	1.034	0.993	1.001	1.036	1.012	1.042	0.952	0.982	1.012
	3	1.033	1.012	1.000	0.995	1.036	1.005	1.038	0.958	1.004	1.012
	4	1.032	1.010	1.002	0.990	1.036	0.973	1.038	0.953	0.960	1.012
	5	1.018	1.014	1.003	0.991	1.036	0.964	1.028	0.960	0.962	1.012
	6	1.019	1.029	0.998	0.993	1.036	0.966	1.033	0.958	0.952	1.012
	7	1.007	1.030	0.997	0.995	1.036	0.962	1.017	0.948	0.945	1.012
	8	1.009	1.032	0.999	0.990	1.036	0.957	1.024	0.962	0.943	1.012
	9	1.005	1.036	0.998	0.993	1.036	0.963	1.031	0.968	0.955	1.012
	10	1.000	1.024	0.996	0.971	1.036	0.971	1.024	0.972	0.937	1.012
2	1	1.083	1.074	0.973	1.037	1.073	1.085	1.091	0.944	1.003	1.025
	2	1.044	1.087	1.023	1.048	1.073	1.056	1.111	0.949	1.035	1.025
	3	1.085	1.025	1.035	1.044	1.073	1.066	1.056	0.970	1.029	1.025
	4	1.069	1.026	1.037	1.027	1.073	1.009	1.063	0.959	0.957	1.025
	5	1.052	1.042	1.047	1.036	1.073	1.000	1.094	0.993	0.989	1.025
	6	1.060	1.093	1.048	1.054	1.073	1.009	1.121	0.999	0.982	1.025
	7	1.031	1.074	1.031	1.052	1.073	1.076	1.074	1.046	0.945	1.025
	8	1.047	1.080	1.050	1.034	1.073	0.993	1.083	1.004	0.936	1.025
	9	1.049	1.081	1.047	1.033	1.073	0.990	1.102	0.991	0.968	1.025
	10	1.047	1.073	1.048	0.984	1.073	0.997	1.119	0.997	0.921	1.025
4	1	1.111	1.064	0.962	0.979	1.122	1.094	1.072	0.919	0.917	1.057
	2	0.995	1.078	1.002	1.037	1.122	0.977	1.082	0.926	0.967	1.057
	3	1.130	1.031	1.001	1.076	1.122	1.075	1.038	0.944	0.996	1.057
	4	1.063	1.036	0.997	1.015	1.122	0.943	1.045	0.910	0.880	1.057
	5	1.043	1.067	0.988	0.982	1.122	0.926	1.072	0.874	0.866	1.057
	6	1.034	1.085	1.001	0.993	1.122	0.939	1.049	0.888	0.845	1.057
	7	0.974	1.085	0.970	0.994	1.122	0.872	1.027	0.856	0.842	1.057
	8	1.006	1.117	0.996	0.944	1.122	0.902	1.069	0.882	0.848	1.057
	9	1.005	1.136	0.970	0.949	1.122	0.901	1.098	0.876	0.873	1.057
	10	0.967	1.137	0.954	0.939	1.122	0.881	1.088	0.875	0.865	1.057
8	1	1.028	1.053	0.918	0.883	0.952	0.986	1.030	0.688	0.675	0.734
	2	0.969	1.076	0.937	1.038	0.952	0.911	1.049	0.671	0.701	0.734
	3	1.273	0.989	0.970	1.102	0.952	1.114	0.990	0.758	0.789	0.734
	4	1.102	1.017	0.959	0.852	0.952	0.886	1.009	0.735	0.558	0.734
	5	1.058	1.103	0.944	0.795	0.952	0.849	1.086	0.700	0.571	0.734
	6	1.038	1.151	0.936	0.824	0.952	0.832	1.092	0.680	0.573	0.734
	7	0.766	1.156	0.718	0.840	0.952	0.632	1.065	0.561	0.554	0.734
	8	0.791	1.219	0.736	0.755	0.952	0.656	1.147	0.571	0.610	0.734
	9	0.829	1.225	0.722	0.742	0.952	0.653	1.165	0.558	0.564	0.734
	10	0.779	1.251	0.721	0.757	0.952	0.635	1.178	0.590	0.581	0.734

Note: We report the RRMSPE statistics with the random walk benchmark model. Therefore, lower values than 1 imply that competing models outperform the benchmark. We assess the 1-quarter to 8-quarter ahead out-of-sample predictability of our factor models with up to 10 factors for consumer loan average CORs of the top 10 BHCs and all banks in the U.S.

Table A6: Out-of-Sample Forecast Results: Real Factors for All Loans COR

		Top10 Banks RRMSPE					*All Banks RRMSPE*				
$\#H$	$\#F$	PLS_{RW}	PC_{RW}	PLS_{AR}	PC_{AR}	AR	PLS_{RW}	PC_{RW}	PLS_{AR}	PC_{AR}	AR
1	1	1.030	1.018	0.945	0.969	1.019	1.035	1.029	0.950	0.970	1.020
	2	0.990	0.996	0.955	0.947	1.019	0.984	0.993	0.961	0.942	1.020
	3	0.976	1.001	0.951	0.948	1.019	0.965	0.991	0.956	0.946	1.020
	4	0.973	1.002	0.955	0.953	1.019	0.965	0.993	0.962	0.946	1.020
	5	0.979	1.012	0.964	0.974	1.019	0.976	0.987	0.972	0.937	1.020
	6	0.985	0.993	0.974	0.957	1.019	0.982	1.002	0.977	0.970	1.020
	7	0.987	0.981	0.964	0.952	1.019	0.981	0.998	0.976	0.974	1.020
	8	0.997	0.980	0.975	0.951	1.019	0.971	0.995	0.970	0.970	1.020
	9	0.993	0.977	0.967	0.948	1.019	0.970	0.986	0.957	0.965	1.020
	10	1.002	0.977	0.967	0.946	1.019	0.968	0.985	0.955	0.968	1.020
2	1	1.058	1.090	0.860	0.990	1.023	1.072	1.133	0.920	1.044	1.023
	2	1.040	1.014	0.903	0.886	1.023	1.077	1.036	1.007	0.934	1.023
	3	0.998	0.999	0.899	0.859	1.023	1.040	0.990	1.000	0.897	1.023
	4	1.060	1.013	0.921	0.882	1.023	1.024	0.998	0.951	0.915	1.023
	5	0.959	1.040	0.891	0.878	1.023	0.984	0.988	0.948	0.891	1.023
	6	0.962	1.018	0.898	0.868	1.023	0.988	1.010	0.970	0.939	1.023
	7	0.959	0.996	0.893	0.881	1.023	0.991	1.010	0.964	0.963	1.023
	8	0.979	0.990	0.918	0.874	1.023	0.980	0.991	0.959	0.936	1.023
	9	0.974	0.984	0.887	0.870	1.023	0.988	0.996	0.941	0.949	1.023
	10	0.980	0.988	0.885	0.865	1.023	0.978	1.014	0.924	0.967	1.023
4	1	1.085	1.032	0.682	0.683	0.991	1.072	1.050	0.790	0.769	0.952
	2	0.863	0.933	0.690	0.694	0.991	0.870	0.933	0.783	0.767	0.952
	3	0.895	0.949	0.689	0.704	0.991	0.909	0.968	0.759	0.795	0.952
	4	0.817	0.947	0.690	0.702	0.991	0.851	0.967	0.776	0.805	0.952
	5	0.889	0.952	0.749	0.673	0.991	0.948	0.962	0.856	0.758	0.952
	6	0.875	1.006	0.743	0.760	0.991	0.943	1.020	0.864	0.858	0.952
	7	0.870	0.986	0.760	0.777	0.991	0.959	1.003	0.894	0.866	0.952
	8	0.887	0.980	0.787	0.757	0.991	0.923	0.998	0.870	0.848	0.952
	9	0.886	0.972	0.758	0.752	0.991	0.931	0.986	0.842	0.845	0.952
	10	0.868	0.981	0.761	0.754	0.991	0.910	0.994	0.840	0.851	0.952
8	1	1.055	1.030	0.540	0.515	0.571	0.977	1.017	0.486	0.464	0.540
	2	0.802	0.902	0.536	0.554	0.571	0.710	0.845	0.496	0.478	0.540
	3	0.894	0.950	0.565	0.580	0.571	0.771	0.910	0.513	0.535	0.540
	4	0.784	0.952	0.560	0.596	0.571	0.695	0.909	0.513	0.563	0.540
	5	0.872	0.974	0.620	0.553	0.571	0.802	0.949	0.598	0.526	0.540
	6	0.841	1.053	0.629	0.616	0.571	0.776	1.070	0.595	0.601	0.540
	7	0.838	1.018	0.661	0.646	0.571	0.823	1.054	0.653	0.645	0.540
	8	0.873	1.011	0.701	0.592	0.571	0.800	1.052	0.632	0.607	0.540
	9	0.898	1.004	0.722	0.588	0.571	0.832	1.037	0.625	0.604	0.540
	10	0.835	1.028	0.715	0.589	0.571	0.782	1.058	0.628	0.618	0.540

Note: We report the RRMSPE statistics with the random walk benchmark model. Therefore, lower values than 1 imply that competing models outperform the benchmark. We assess the 1-quarter to 8-quarter ahead out-of-sample predictability of our factor models with up to 10 factors for all loan average CORs of the top 10 BHCs and all banks in the U.S using real activity variables only.

Table A7: Out-of-Sample Forecast Results: Finance Factors for All Loans COR

		Top10 Banks RRMSPE					*All Banks RRMSPE*				
$\#H$	$\#F$	PLS_{RW}	PC_{RW}	PLS_{AR}	PC_{AR}	AR	PLS_{RW}	PC_{RW}	PLS_{AR}	PC_{AR}	AR
1	1	1.048	1.021	0.913	1.001	1.019	1.032	1.013	0.968	1.009	1.020
	2	1.000	1.019	0.935	0.939	1.019	1.015	1.016	0.974	0.977	1.020
	3	0.968	0.999	0.929	0.927	1.019	0.971	1.016	0.956	0.958	1.020
	4	0.965	1.007	0.922	0.931	1.019	0.957	1.021	0.957	0.962	1.020
	5	0.929	1.006	0.914	0.930	1.019	0.963	1.029	0.961	0.963	1.020
	6	0.937	1.017	0.920	0.928	1.019	0.985	1.031	0.972	0.972	1.020
	7	0.933	1.016	0.917	0.926	1.019	0.984	1.022	0.975	0.980	1.020
	8	0.929	1.008	0.911	0.921	1.019	0.980	1.015	0.974	0.972	1.020
	9	0.935	1.015	0.917	0.929	1.019	0.980	1.014	0.973	0.976	1.020
	10	0.934	1.010	0.917	0.931	1.019	0.977	1.011	0.972	0.983	1.020
2	1	1.064	1.029	0.869	0.993	1.023	1.053	1.024	0.920	0.993	1.023
	2	0.996	1.030	0.902	0.915	1.023	1.007	1.027	0.935	0.893	1.023
	3	0.922	1.002	0.872	0.883	1.023	0.889	0.991	0.870	0.877	1.023
	4	0.899	1.013	0.867	0.888	1.023	0.872	1.005	0.867	0.890	1.023
	5	0.858	1.013	0.858	0.882	1.023	0.878	1.016	0.873	0.889	1.023
	6	0.879	1.024	0.868	0.884	1.023	0.920	1.017	0.892	0.899	1.023
	7	0.876	1.017	0.864	0.907	1.023	0.921	1.008	0.899	0.926	1.023
	8	0.869	1.010	0.848	0.897	1.023	0.913	0.995	0.893	0.906	1.023
	9	0.875	1.029	0.855	0.922	1.023	0.915	1.005	0.896	0.929	1.023
	10	0.879	1.020	0.859	0.936	1.023	0.906	1.002	0.894	0.956	1.023
4	1	1.065	1.043	0.788	0.963	0.991	1.023	1.026	0.868	0.929	0.952
	2	1.006	1.046	0.815	0.736	0.991	0.979	1.021	0.873	0.698	0.952
	3	0.795	0.984	0.663	0.683	0.991	0.727	0.916	0.682	0.688	0.952
	4	0.776	1.015	0.665	0.708	0.991	0.723	0.962	0.673	0.743	0.952
	5	0.700	1.041	0.652	0.690	0.991	0.708	1.001	0.693	0.776	0.952
	6	0.814	1.044	0.702	0.699	0.991	0.826	0.994	0.767	0.772	0.952
	7	0.803	1.028	0.690	0.774	0.991	0.831	0.977	0.788	0.812	0.952
	8	0.797	0.997	0.665	0.752	0.991	0.824	0.944	0.770	0.774	0.952
	9	0.794	1.053	0.650	0.806	0.991	0.821	0.982	0.773	0.810	0.952
	10	0.780	1.025	0.651	0.867	0.991	0.792	0.957	0.761	0.840	0.952
8	1	0.879	1.019	0.566	0.574	0.571	0.840	1.009	0.541	0.514	0.540
	2	0.878	1.020	0.540	0.467	0.571	0.805	0.978	0.529	0.510	0.540
	3	0.560	0.900	0.444	0.462	0.571	0.515	0.812	0.540	0.519	0.540
	4	0.557	0.927	0.459	0.466	0.571	0.515	0.836	0.601	0.513	0.540
	5	0.458	0.925	0.489	0.502	0.571	0.502	0.853	0.610	0.508	0.540
	6	0.537	0.854	0.453	0.605	0.571	0.534	0.777	0.606	0.671	0.540
	7	0.507	0.858	0.488	0.592	0.571	0.555	0.811	0.602	0.705	0.540
	8	0.511	0.823	0.549	0.627	0.571	0.551	0.768	0.625	0.745	0.540
	9	0.520	0.964	0.576	0.657	0.571	0.549	0.832	0.623	0.723	0.540
	10	0.544	0.972	0.672	0.609	0.571	0.614	0.850	0.706	0.668	0.540

Note: We report the RRMSPE statistics with the random walk benchmark model. Therefore, lower values than 1 imply that competing models outperform the benchmark. We assess the 1-quarter to 8-quarter ahead out-of-sample predictability of our factor models with up to 10 factors for all loan average CORs of the top 10 BHCs and all banks in the U.S. using finance factors only.

Table A8: Out-of-Sample Forecast Results: Real Factors for Business Loans COR

		Top10 Banks RRMSPE					*All Banks RRMSPE*				
$\#H$	$\#F$	PLS_{RW}	PC_{RW}	PLS_{AR}	PC_{AR}	AR	PLS_{RW}	PC_{RW}	PLS_{AR}	PC_{AR}	AR
1	1	1.009	0.989	0.918	0.928	0.989	1.020	1.010	0.856	0.883	0.917
	2	0.959	0.979	0.907	0.911	0.989	0.957	0.972	0.861	0.849	0.917
	3	0.940	0.972	0.905	0.911	0.989	0.923	0.968	0.859	0.837	0.917
	4	0.932	0.988	0.902	0.915	0.989	0.977	0.965	0.848	0.857	0.917
	5	0.926	0.994	0.911	0.927	0.989	0.883	0.957	0.838	0.844	0.917
	6	0.938	0.979	0.915	0.902	0.989	0.879	0.962	0.839	0.828	0.917
	7	0.942	0.952	0.922	0.883	0.989	0.876	0.963	0.843	0.849	0.917
	8	0.941	0.948	0.924	0.882	0.989	0.871	0.963	0.833	0.849	0.917
	9	0.935	0.943	0.913	0.877	0.989	0.871	0.960	0.834	0.847	0.917
	10	0.940	0.946	0.923	0.878	0.989	0.863	0.961	0.833	0.852	0.917
2	1	1.029	1.085	0.878	1.019	0.961	1.049	1.119	0.914	1.067	0.920
	2	1.070	1.026	0.939	0.916	0.961	1.142	1.077	0.976	0.908	0.920
	3	0.983	0.957	0.928	0.841	0.961	1.005	0.971	0.928	0.847	0.920
	4	0.992	0.970	0.879	0.847	0.961	0.936	1.043	0.851	0.880	0.920
	5	0.898	0.981	0.856	0.861	0.961	0.922	1.008	0.840	0.885	0.920
	6	0.897	0.968	0.850	0.837	0.961	0.911	1.016	0.831	0.879	0.920
	7	0.898	0.952	0.859	0.823	0.961	0.898	1.020	0.879	0.851	0.920
	8	0.895	0.947	0.865	0.827	0.961	0.929	0.996	0.882	0.838	0.920
	9	0.891	0.930	0.847	0.821	0.961	0.942	0.998	0.889	0.856	0.920
	10	0.892	0.970	0.854	0.873	0.961	0.937	1.000	0.884	0.874	0.920
4	1	1.054	1.029	0.807	0.798	0.868	1.053	1.052	0.864	0.850	0.883
	2	0.861	0.908	0.792	0.791	0.868	0.872	0.906	0.834	0.816	0.883
	3	0.862	0.887	0.791	0.782	0.868	0.879	0.905	0.827	0.820	0.883
	4	0.867	0.920	0.787	0.807	0.868	0.881	0.955	0.838	0.869	0.883
	5	0.857	0.934	0.790	0.799	0.868	0.878	0.952	0.841	0.860	0.883
	6	0.855	0.965	0.787	0.825	0.868	0.874	0.989	0.836	0.885	0.883
	7	0.837	0.938	0.799	0.818	0.868	0.894	0.970	0.869	0.883	0.883
	8	0.851	0.931	0.810	0.814	0.868	0.906	0.967	0.885	0.878	0.883
	9	0.846	0.914	0.801	0.805	0.868	0.910	0.949	0.880	0.868	0.883
	10	0.857	0.926	0.810	0.810	0.868	0.915	0.957	0.884	0.874	0.883
8	1	0.997	1.040	0.601	0.613	0.625	0.961	1.019	0.683	0.672	0.687
	2	0.753	0.842	0.615	0.608	0.625	0.747	0.813	0.681	0.642	0.687
	3	0.743	0.847	0.612	0.609	0.625	0.757	0.825	0.681	0.663	0.687
	4	0.749	0.887	0.626	0.662	0.625	0.768	0.890	0.704	0.745	0.687
	5	0.758	0.944	0.649	0.661	0.625	0.774	0.922	0.725	0.757	0.687
	6	0.751	1.019	0.626	0.700	0.625	0.767	0.994	0.705	0.796	0.687
	7	0.741	0.984	0.650	0.718	0.625	0.822	0.982	0.774	0.820	0.687
	8	0.746	0.981	0.676	0.702	0.625	0.818	0.972	0.784	0.805	0.687
	9	0.766	0.960	0.695	0.692	0.625	0.840	0.957	0.799	0.797	0.687
	10	0.796	0.993	0.718	0.713	0.625	0.871	0.980	0.823	0.822	0.687

Note: We report the RRMSPE statistics with the random walk benchmark model. Therefore, lower values than 1 imply that competing models outperform the benchmark. We assess the 1-quarter to 8-quarter ahead out-of-sample predictability of our factor models with up to 10 factors for business loan average CORs of the top 10 BHCs and all banks in the U.S. using real activity variables only.

Table A9: Out-of-Sample Forecast Results: Finance Factors for Business Loans COR

		Top10 Banks RRMSPE					*All Banks RRMSPE*				
$\#H$	$\#F$	PLS_{RW}	PC_{RW}	PLS_{AR}	PC_{AR}	AR	PLS_{RW}	PC_{RW}	PLS_{AR}	PC_{AR}	AR
1	1	1.022	1.000	0.967	0.974	0.989	1.010	1.009	0.889	0.902	0.917
	2	0.993	1.004	0.967	0.958	0.989	0.965	1.016	0.888	0.887	0.917
	3	0.981	0.990	0.956	0.945	0.989	0.943	1.026	0.888	0.898	0.917
	4	0.963	0.997	0.943	0.949	0.989	0.955	1.026	0.915	0.895	0.917
	5	0.967	1.000	0.948	0.950	0.989	0.972	1.049	0.941	0.910	0.917
	6	0.972	1.011	0.953	0.962	0.989	0.973	1.104	0.941	0.962	0.917
	7	0.972	1.014	0.956	0.968	0.989	0.989	1.101	0.952	0.982	0.917
	8	0.991	1.009	0.979	0.971	0.989	0.998	1.107	0.967	1.002	0.917
	9	0.988	1.012	0.978	0.977	0.989	0.994	1.078	0.964	0.986	0.917
	10	0.995	1.010	0.979	0.974	0.989	1.042	1.088	0.995	1.005	0.917
2	1	1.030	1.008	0.940	0.944	0.961	1.016	1.012	0.909	0.907	0.920
	2	0.983	1.018	0.955	0.922	0.961	0.993	1.023	0.929	0.898	0.920
	3	0.952	0.995	0.939	0.908	0.961	0.946	1.019	0.924	0.904	0.920
	4	0.940	1.001	0.939	0.913	0.961	0.934	1.041	0.931	0.921	0.920
	5	0.954	1.004	0.944	0.916	0.961	0.945	1.031	0.935	0.899	0.920
	6	0.962	1.015	0.949	0.929	0.961	0.939	1.042	0.929	0.913	0.920
	7	0.965	1.019	0.953	0.942	0.961	0.959	1.042	0.949	0.939	0.920
	8	0.981	0.994	0.976	0.936	0.961	0.970	1.029	0.962	0.939	0.920
	9	0.973	0.998	0.970	0.955	0.961	0.949	1.007	0.934	0.934	0.920
	10	0.978	0.992	0.966	0.954	0.961	0.946	1.014	0.930	0.946	0.920
4	1	1.007	1.028	0.857	0.865	0.868	0.992	1.025	0.877	0.877	0.883
	2	0.961	1.045	0.906	0.825	0.868	0.923	1.026	0.879	0.831	0.883
	3	0.901	0.955	0.887	0.788	0.868	0.861	0.916	0.847	0.770	0.883
	4	0.869	0.972	0.872	0.797	0.868	0.820	0.937	0.818	0.786	0.883
	5	0.869	0.977	0.874	0.802	0.868	0.822	0.959	0.823	0.808	0.883
	6	0.896	0.984	0.884	0.817	0.868	0.827	0.965	0.820	0.830	0.883
	7	0.902	1.033	0.891	0.888	0.868	0.843	0.964	0.831	0.861	0.883
	8	0.896	0.948	0.889	0.832	0.868	0.877	0.918	0.867	0.829	0.883
	9	0.897	0.973	0.889	0.860	0.868	0.880	0.929	0.866	0.843	0.883
	10	0.900	0.972	0.889	0.873	0.868	0.880	0.942	0.872	0.862	0.883
8	1	0.816	1.048	0.633	0.636	0.625	0.815	1.027	0.682	0.679	0.687
	2	0.795	1.068	0.679	0.619	0.625	0.774	0.999	0.704	0.682	0.687
	3	0.724	0.967	0.673	0.607	0.625	0.714	0.889	0.679	0.637	0.687
	4	0.695	0.957	0.656	0.619	0.625	0.695	0.873	0.664	0.642	0.687
	5	0.677	0.905	0.669	0.631	0.625	0.691	0.832	0.685	0.650	0.687
	6	0.719	0.873	0.712	0.704	0.625	0.713	0.833	0.707	0.781	0.687
	7	0.719	0.965	0.718	0.781	0.625	0.724	0.938	0.712	0.880	0.687
	8	0.766	0.900	0.762	0.772	0.625	0.778	0.892	0.764	0.874	0.687
	9	0.788	0.932	0.773	0.752	0.625	0.803	0.882	0.789	0.818	0.687
	10	0.809	0.980	0.796	0.762	0.625	0.800	0.932	0.791	0.840	0.687

Note: We report the RRMSPE statistics with the random walk benchmark model. Therefore, lower values than 1 imply that competing models outperform the benchmark. We assess the 1-quarter to 8-quarter ahead out-of-sample predictability of our factor models with up to 10 factors for business loan average CORs of the top 10 BHCs and all banks in the U.S. using finance variables only.

Table A10: Out-of-Sample Forecast Results: Real Factors for Consumer Loans COR

		Top10 Banks RRMSPE					*All Banks RRMSPE*				
$\#H$	$\#F$	PLS_{RW}	PC_{RW}	PLS_{AR}	PC_{AR}	AR	PLS_{RW}	PC_{RW}	PLS_{AR}	PC_{AR}	AR
1	1	1.042	1.034	1.000	1.014	1.036	1.046	1.041	0.965	0.982	1.012
	2	1.017	1.013	1.015	1.010	1.036	1.013	1.018	0.969	0.969	1.012
	3	1.031	1.008	1.013	1.005	1.036	1.010	1.012	0.974	0.964	1.012
	4	1.026	1.011	1.005	1.013	1.036	1.011	1.014	0.967	0.987	1.012
	5	1.025	1.019	1.024	0.999	1.036	0.982	1.008	0.967	0.995	1.012
	6	1.017	1.005	0.998	1.006	1.036	0.976	1.020	0.961	1.002	1.012
	7	1.006	1.003	0.994	1.003	1.036	0.981	1.023	0.959	0.996	1.012
	8	1.006	1.003	0.996	1.001	1.036	0.978	1.023	0.967	0.995	1.012
	9	1.005	1.003	0.997	1.000	1.036	0.975	1.017	0.966	0.984	1.012
	10	0.998	1.005	0.996	0.992	1.036	0.976	1.014	0.963	0.984	1.012
2	1	1.075	1.093	1.012	1.069	1.073	1.086	1.114	0.981	1.044	1.025
	2	1.063	1.036	1.053	1.032	1.073	1.079	1.065	0.982	0.974	1.025
	3	1.096	1.029	1.055	1.021	1.073	1.088	1.059	1.006	0.958	1.025
	4	1.093	1.044	1.050	1.047	1.073	1.105	1.076	0.986	1.047	1.025
	5	1.078	1.090	1.109	1.020	1.073	1.017	1.064	0.953	1.039	1.025
	6	1.039	1.066	1.007	1.021	1.073	0.995	1.066	0.938	1.067	1.025
	7	1.031	1.046	0.992	1.033	1.073	0.995	1.057	0.943	1.118	1.025
	8	1.034	1.039	0.995	1.023	1.073	0.975	1.056	0.944	1.088	1.025
	9	1.041	1.038	1.006	1.017	1.073	0.966	1.055	0.943	1.080	1.025
	10	1.023	1.026	1.004	1.011	1.073	0.988	1.046	0.955	1.076	1.025
4	1	1.106	1.064	0.982	0.977	1.122	1.096	1.074	0.901	0.896	1.057
	2	0.985	1.012	0.988	1.019	1.122	0.961	1.012	0.917	0.953	1.057
	3	1.087	1.049	0.965	1.061	1.122	1.052	1.029	0.930	0.973	1.057
	4	1.064	1.055	0.981	1.066	1.122	1.046	1.033	0.924	0.977	1.057
	5	1.027	1.070	0.963	1.074	1.122	1.000	1.021	0.909	0.996	1.057
	6	1.003	1.096	0.964	1.046	1.122	0.972	1.063	0.922	0.980	1.057
	7	1.028	1.098	0.972	0.981	1.122	0.996	1.071	0.925	0.926	1.057
	8	1.044	1.100	0.989	0.964	1.122	1.016	1.075	0.951	0.916	1.057
	9	1.025	1.096	0.989	0.947	1.122	0.967	1.068	0.928	0.905	1.057
	10	1.026	1.102	0.995	0.948	1.122	0.952	1.077	0.925	0.908	1.057
8	1	1.061	1.067	0.884	0.864	0.952	1.019	1.045	0.664	0.647	0.734
	2	0.943	0.973	0.897	0.968	0.952	0.874	0.935	0.681	0.777	0.734
	3	1.165	1.092	0.817	1.113	0.952	1.082	0.999	0.725	0.840	0.734
	4	1.125	1.095	0.858	1.108	0.952	1.069	1.003	0.690	0.840	0.734
	5	1.106	1.118	0.865	1.128	0.952	1.010	0.997	0.690	0.821	0.734
	6	1.031	1.205	0.862	1.067	0.952	0.957	1.101	0.711	0.810	0.734
	7	1.119	1.196	0.894	0.885	0.952	1.008	1.104	0.713	0.690	0.734
	8	1.154	1.207	0.943	0.867	0.952	1.017	1.121	0.745	0.687	0.734
	9	1.074	1.203	0.917	0.825	0.952	0.907	1.114	0.705	0.676	0.734
	10	0.979	1.232	0.876	0.853	0.952	0.852	1.147	0.703	0.696	0.734

Note: We report the RRMSPE statistics with the random walk benchmark model. Therefore, lower values than 1 imply that competing models outperform the benchmark. We assess the 1-quarter to 8-quarter ahead out-of-sample predictability of our factor models with up to 10 factors for consuer loan average CORs of the top 10 BHCs and all banks in the U.S. using real activity variables only.

Table A11: Out-of-Sample Forecast Results: Finance Factors for Consumer Loans COR

		Top10 Banks RRMSPE					*All Banks RRMSPE*				
$\#H$	$\#F$	PLS_{RW}	PC_{RW}	PLS_{AR}	PC_{AR}	AR	PLS_{RW}	PC_{RW}	PLS_{AR}	PC_{AR}	AR
1	1	1.049	1.021	0.987	1.002	1.036	1.036	1.017	0.982	0.989	1.012
	2	1.016	1.023	0.992	1.011	1.036	1.018	1.021	0.977	0.995	1.012
	3	1.036	1.021	0.997	1.005	1.036	1.010	1.024	0.980	0.991	1.012
	4	1.020	1.020	0.994	1.001	1.036	0.992	1.021	0.977	0.983	1.012
	5	1.016	1.021	0.996	0.995	1.036	0.986	1.030	0.978	0.968	1.012
	6	1.010	1.040	0.994	1.002	1.036	0.984	1.046	0.973	0.979	1.012
	7	1.012	1.038	0.996	0.998	1.036	0.982	1.040	0.973	0.982	1.012
	8	1.013	1.034	1.000	0.994	1.036	0.983	1.036	0.978	0.982	1.012
	9	1.017	1.035	1.005	0.991	1.036	0.979	1.030	0.972	0.980	1.012
	10	1.017	1.030	1.005	0.985	1.036	0.992	1.029	0.982	0.975	1.012
2	1	1.073	1.032	1.006	1.013	1.073	1.064	1.033	0.976	0.979	1.025
	2	1.010	1.038	1.018	1.036	1.073	1.023	1.046	0.960	0.999	1.025
	3	1.055	1.019	1.027	1.027	1.073	1.025	1.029	0.965	0.983	1.025
	4	1.025	1.025	1.019	1.020	1.073	0.981	1.049	0.953	0.967	1.025
	5	1.024	1.024	1.024	1.015	1.073	0.980	1.051	0.962	0.935	1.025
	6	1.022	1.049	1.025	1.027	1.073	0.993	1.078	0.964	0.951	1.025
	7	1.019	1.046	1.021	1.023	1.073	0.953	1.075	0.936	0.970	1.025
	8	1.017	1.047	1.019	1.018	1.073	0.944	1.079	0.928	0.974	1.025
	9	1.027	1.053	1.034	1.013	1.073	0.946	1.064	0.928	0.959	1.025
	10	1.022	1.048	1.031	1.012	1.073	0.960	1.057	0.940	0.950	1.025
4	1	1.086	1.046	1.071	1.065	1.122	1.058	1.036	1.033	1.031	1.057
	2	1.058	1.060	1.083	1.135	1.122	1.070	1.055	1.034	1.066	1.057
	3	1.144	1.017	1.083	1.113	1.122	1.079	1.000	1.026	1.019	1.057
	4	1.079	1.045	1.078	1.081	1.122	0.987	1.031	0.989	0.987	1.057
	5	1.070	1.068	1.076	1.039	1.122	1.001	1.066	1.002	0.937	1.057
	6	1.072	1.086	1.065	1.042	1.122	1.013	1.081	0.985	0.920	1.057
	7	1.055	1.082	1.053	1.037	1.122	1.008	1.067	0.987	0.943	1.057
	8	1.060	1.079	1.052	1.028	1.122	1.005	1.041	0.984	0.927	1.057
	9	1.065	1.100	1.046	0.980	1.122	0.975	1.049	0.931	0.902	1.057
	10	1.028	1.090	1.034	0.982	1.122	0.937	1.035	0.913	0.924	1.057
8	1	1.004	1.044	1.030	1.009	0.952	0.979	1.037	0.750	0.763	0.734
	2	1.050	1.075	1.006	1.034	0.952	0.987	1.067	0.761	0.731	0.734
	3	1.109	0.980	1.024	1.008	0.952	0.925	0.963	0.761	0.693	0.734
	4	1.055	1.013	1.028	0.937	0.952	0.850	1.002	0.735	0.657	0.734
	5	0.956	1.030	0.976	0.793	0.952	0.775	1.026	0.729	0.572	0.734
	6	0.964	1.013	0.908	0.709	0.952	0.809	1.012	0.678	0.468	0.734
	7	0.797	1.023	0.738	0.731	0.952	0.652	1.022	0.553	0.572	0.734
	8	0.780	1.010	0.728	0.700	0.952	0.656	0.980	0.561	0.549	0.734
	9	0.773	1.065	0.670	0.628	0.952	0.619	1.012	0.481	0.554	0.734
	10	0.734	1.085	0.677	0.673	0.952	0.653	1.048	0.567	0.593	0.734

Note: We report the RRMSPE statistics with the random walk benchmark model. Therefore, lower values than 1 imply that competing models outperform the benchmark. We assess the 1-quarter to 8-quarter ahead out-of-sample predictability of our factor models with up to 10 factors for consumer loan average CORs of the top 10 BHCs and all banks in the U.S. using finance variables only.