DISCUSSION PAPER SERIES

DP13034

THE FORCASTING PERFORMANCE OF DYNAMIC FACTOR MODELS WITH VINTAGE DATA

Luca Di Bonaventura, Mario Forni and Francesco Pattarin

MONETARY ECONOMICS AND FLUCTUATIONS

ISSN 0265-8003

THE FORCASTING PERFORMANCE OF DYNAMIC FACTOR MODELS WITH VINTAGE DATA

Luca Di Bonaventura, Mario Forni and Francesco Pattarin

Discussion Paper DP13034
Published 03 July 2018
Submitted 03 July 2018

Centre for Economic Policy Research
33 Great Sutton Street, London EC1V 0DX, UK
Tel: +44 (0)20 7183 8801
www.cepr.org

THE FORCASTING PERFORMANCE OF DYNAMIC FACTOR MODELS WITH VINTAGE DATA

Abstract

We present a comparative analysis of the forecasting performance of two dynamic factor models, the Stock and Watson (2002a, b) model and the Forni, Hallin, Lippi and Reichlin (2005) model, based on vintage data. Our dataset contains 107 monthly US “first release” macroeconomic and financial vintage time series, spanning the 1996:12 to 2017:6 period with monthly periodicity, extracted from the Bloomberg database†. We compute real-time one-month-ahead forecasts with both models for four key macroeconomic variables: the month-on-month change in industrial production, the unemployment rate, the core consumer price index and the ISM Purchasing Managers’ Index. First, we find that both the Stock and Watson and the Forni, Hallin, Lippi and Reichlin models outperform simple autoregressions for industrial production, unemployment rate and consumer prices, but that only the first model does so for the PMI. Second, we find that neither models always outperform the other. While Forni, Hallin, Lippi and Reichlin’s beats Stock and Watson’s in forecasting industrial production and consumer prices, the opposite happens for the unemployment rate and the PMI.

JEL Classification: C01, C32, C52, C53, E27, E37

Keywords: Dynamic factor models, Forecasting, Forecasting Performance, Vintage data, First release data

Luca Di Bonaventura - 71661@studenti.unimore.it
Credito Emiliano Spa

Mario Forni - mario.forni@unimore.it
Dipartimento di Economia "Marco Biagi" and CEPR

Francesco Pattarin - francesco.pattarin@unimore.it
Dipartimento di Economia "Marco Biagi"

The Forcasting Performance of Dynamic Factor Models with Vintage Data

Luca Di Bonaventura
Credito Emiliano Spa

Mario Forni
Department of Economics
RECent – Center for Economomic Research
University of Modena and Reggio Emilia
CEPR – Center for Economic Policy Research

Francesco Pattarin*
Department of Economics
CEFIN – Centro Studi Banca e Finanza
University of Modena and Reggio Emilia

Abstract

We present a comparative analysis of the forecasting performance of two dynamic factor models, the Stock and Watson (2002a, b) model and the Forni, Hallin, Lippi and Reichlin (2005) model, based on vintage data. Our dataset contains 107 monthly US "first release" macroeconomic and financial vintage time series, spanning the 1996:12 to 2017:6 period with monthly periodicity, extracted from the Bloomberg database†. We compute real-time one-month-ahead forecasts with both models for four key macroeconomic variables: the month-on-month change in industrial production, the unemployment rate, the core consumer price index and the ISM Purchasing Managers' Index. First, we find that both the Stock and Watson and the Forni, Hallin, Lippi and Reichlin models outperform simple autoregressions for industrial production, unemployment rate and consumer prices, but that only the first model does so for the PMI. Second, we find that neither models always outperform the other. While Forni, Hallin, Lippi and Reichlin's beats Stock and Watson's in forecasting industrial production and consumer prices, the opposite happens for the unemployment rate and the PMI.

Keywords: Dynamic factor models, Forecasting, Forecasting Performance, Vintage data, First release data

JEL Codes: C01, C32, C52, C53, E27, E37

** Corresponding author: francesco.pattarin@unimore.it
†† We thank Jetsone Asset Management PLC, London for providing us the Bloomberg data.

Introduction

Dynamic factor models represent each element of a multiple set of time series as the sum of two orthogonal terms: the common component and the idiosyncratic component. The common component represents the part of each series which determines its co-movements with the other series, while the idiosyncratic component represents the residual part of its variability over time. The common component is driven by relatively few factors, as compared to the many series of interest.

Among the many versions of dynamic factor models, we compare the forecasting performance of two well-known models: the model proposed by Stock and Watson (2002a, b) SW hereafter, and that proposed by Forni, Hallin, Lippi and Reichlin (2005) FHLR hereafter.

The SW model requires two main steps to forecast a target variable. First, common factors are estimated using a static (time domain) principal components estimator. Then, the estimated factors are used as inputs (*i.e.* regressors) in an Autoregressive Distributed Lag (ADL) model.

Like the SW model, FHLR builds forecasts of the target variable in two steps. In the first step, the frequency-domain technique proposed by Forni and co-authors (2000) is used to estimate the covariance matrices of the common and idiosyncratic components. Then, the second step relies on these estimates to get a consistent estimator of the space spanned by the static factors. This is achieved by solving the generalized principal components problem, and produces a set of dynamic principal components that, in turn, deliver a predictor of the future values of the series which is a linear projection of the future values of the common component on the factor space. Optionally, an independent forecast of the idiosyncratic component may be added.

Since the FLHR model was presented, many researchers have tackled the issue of comparing its forecasting performance against the SW model. The existing literature shows mixed results so far. Using a dataset of 147 US macroeconomic and financial series spanning the 1959:1 to 1998:12 period, Boivin and Ng (2005) forecast eight variables and found that generally SW outperforms FHLR. On the contrary, Den Reijen (2005) found that FHLR outperforms SW in forecasting the Dutch GDP on a quarterly dataset of 370 series running from 1980:Q1 to 2002:Q4. The same conclusion is drawn by Schumacher (2007) for the German GDP, who also found that the model proposed by Kapetanios and Marcellino (2009) is similar to FHLR. Schumacher's work is based on inputs from 124 quarterly series over the period 1978:Q1 – 2004:Q1. D'Agostino and Giannone (2012) compare the two models on dataset of 147 US monthly observations on macroeconomic and financial variables. They found that SW and FHLR have similar forecasting performance and produce highly collinear forecasts. More recently, Forni, Giovannelli, Lippi and Soccorsi (2016) performed a comparison on a dataset of 115 monthly US macroeconomic and financial variables observed from 1959:1 to 2014:8, extending it to a further dynamic factor model recently proposed by Forni, Hallin, Lippi and Zaffaroni (2015, 2016). They found that FHLR wins over SW in forecasting industrial production but not consumer prices, where the Forni, Hallin, Lippi and Zaffaroni model performs best. Similar results are found by Della Marra (2016) on a dataset of 176 EU monthly macroeconomic and financial variables spanning the period 1986:2 – 2015:11

The forecasting performance comparison of the present work is different from those traditionally employed in the literature, because of the type of data used. Instead of using commonly available series, that undergo revision procedures after their first release, we use genuine first release data (*i.e.* not furtherly adjusted) data, also known as "vintage data". Vintage data allow to perform *ex-post* a real-time out-of-sample forecasting exercise, because these are the data that were actually available in the past if the model was to be applied then. The fact that real-time datasets contain missing values at the end of the in-sample period (the so called "ragged-edge" problem) is solved as in Altissimo and co-authors (2010) shifting forward the series when last observations are missing. To our knowledge, there is not any other research on dynamic factor models that exploits first-release data as we do.

We compare the forecasting performances of SW and FHLR on a dataset that contains 107 monthly US first-release macroeconomic and financial time series spanning 1996:12 to 2017:6 extracted from the Bloomberg database. The real-time one month ahead forecasts are made for four key macroeconomic variables, namely industrial production month-on-month change (IP), the unemployment rate (UN), the core consumer price index (CPI) and the ISM Purchasing Managers' Index (PMI). We use three different methods to determine the number of factors: the Information Criterion developed by Bai and Ng (2002), the Edge Distribution proposed by Onatski (2010) and the Eigenvalues Ratio of Ahn and Horenstein (2013). Moreover, all the forecasts are made using either the whole dataset or excluding – in turn – various blocks of variables or dropping some variables according to two rules: the first rule is that proposed in Boivin and Ng (2006), the second rule was devised by us and is explained further.

The rest of the paper is structured as follows. Section 1 gives a brief overview of factor models; paragraphs 1.1 and 1.2 describe the Stock and Watson (2002a, b) and the Forni, Hallin, Lippi and Reichlin (2005) models; section 1.3 and 1.4 present state-of-the-art methods for estimating the number of static and dynamic factors, respectively; section 1.5 concludes with an overview of different rules to par-down the number of input series in a dataset before the estimation. In Section 2 we shortly describe our dataset and presents our results. Section 3 concludes. The details of our data set are reported in the Appendix.

1. Factor Models for Time Series

Let X be a N x T matrix that represents the panel of observed time series, where N is the cross-sectional dimension $(i=1,\ldots,N)$ and T is the time dimension $(t=1,\ldots,T)$:

$$X=\begin{bmatrix} x_{11} & \cdots & x_{1T} \\ \vdots & \ddots & \vdots \\ x_{N1} & \cdots & x_{NT} \end{bmatrix}$$

Let's also assume that each time series in the panel is covariance stationary, with zero mean and unit variance.

In dynamic factor models, the vector of N time series is represented at each point in time t as the sum of two mutually orthogonal unobservable vectors, the common components χ_t and the idiosyncratic components ξ_t. While the common components represent the part of the series which co-move, the idiosyncratic components represent the residual part of the observed series. The vector of common components is driven by a vector of factors whose size is much smaller than the number of series, which creates the co-movements between the time series. In vector notation:

$$x_t=\chi_t+\xi_t, \forall t\in\{1,\ldots,T\}; \qquad \text{with} \quad \chi_t \perp \xi_t$$

where:

(i) $x_t=[x_{1t},\ldots,x_{Nt}]'$ is the N x 1 column vector that contains the N observed time series at time *t*;

(ii) $\chi_t=[\chi_{1t},\ldots,\chi_{Nt}]'$ is the N x 1 column vector of common components at time *t*;

(iii) $\xi_t=[\xi_{1t},\ldots,\xi_{Nt}]'$ is the N x 1 column vector of idiosyncratic components at time *t* .

Different classes of factor models can be specified, according to the assumptions made for the common and the idiosyncratic components. A first distinction can be made according to the functional form assumed for the common components χ_t. With respect to this feature, two different classes of factor models are specified: static and dynamic factor models. In static factor models, the common components χ_t do not take into account the lags of the factors:

$$\chi_t=\Lambda F_t$$

where:

(i) $F_t \in R^r$ is the r x 1 column vector whose elements are the r (unobservable) static factors at time t, with $r \ll N$;

(ii) $\Lambda \in R^{N x r}$ is the factor loadings matrix (with N rows and r columns).

In dynamic factor models, the common components χ_t also take into account the lags of the factors:

$$\chi_t = \Lambda(L) f_t = \Lambda_0 f_t + \Lambda_1 f_{t-1} + \Lambda_2 f_{t-2} \ldots$$

where:

(i) $f_t \in R^q$ is the q x 1 column vector whose elements are the q (unobservable) dynamic factors at time t, with $q \ll N$;

(ii) $\Lambda(L) = \sum_{i=0}^{\infty} \Lambda_i L^i$ is the lag polynomial matrix in non-negative powers of L, where the lag operator L is such that $L^i f_t = f_{t-i}$; and each Λ_i is a N x q matrix of factor loadings.

Assuming that the maximum number of lags of the lag polynomial matrix $\Lambda(L)$ is fixed at $s < \infty$, then the dynamic factor model can be rewritten in an equivalent static form:

$$x_t = \Lambda F_t + \xi_t$$

where:

(i) $x_t = [x_{1t}, \ldots, x_{Nt}]'$ is the N x 1 column vector that contains the N observed time series at time t;

(ii) $F_t = [f_t', f_{t-1}', \ldots, f_{t-s}']' \in R^r$ is the r x 1 column vector whose elements are the r (unobservable) static factors at time t, with $r = q(s+1)$.

(iii) $\Lambda = [\Lambda_0, \Lambda_1, \ldots, \Lambda_s]$ is the factor loadings matrix (with N rows and r columns).

(iv) $\xi_t = [\xi_{1t}, \ldots, \xi_{Nt}]'$ is the N x 1 column vector of idiosyncratic components at time t.

A second distinction between factor models can be made according to the assumptions made with respect to the idiosyncratic components ξ_t. With respect to this feature, two different classes are specified: exact and approximate (or generalized) factor models. In exact factor models, the idiosyncratic components have no cross sectional and time dependence, *i.e.* both the covariance matrix and the cross- autocovariance matrices of

ξ_t are diagonal: $Cov(\xi_{it}, \xi_{j\tau}) = 0, \forall t, \tau, \forall i, j, i \neq j$. In approximate factor models the idiosyncratic components ξ_t are allowed to have mild cross-sectional dependence and also time dependence.

Among the different versions of factor models, we only consider two different approximate dynamic factor models which are commonly used for forecasting macroeconomic variables, namely the model of Stock and Watson (2002a, b) and the model of Forni, Hallin, Lippi and Reichlin (2005).

1.1 The Stock and Watson model

Stock and Watson (2002a, b) assumed an approximate dynamic factors structure with finite lags (*i.e.* a static representation) for the covariance stationary standardized panel X, and proposed a procedure based on two main steps to forecast a time series x_i, the i^{th}-variable of X. First, they estimate the factors F_t using the static (time domain) principal components estimator. Then, they use the estimated static factors $\hat{F}_t$ as inputs (*i.e.* regressors) in an Autoregressive Distributed Lag (ADL) model. Stock and Watson (2002a) showed that the principal components estimator $\hat{F}_t$ is a consistent estimator of the space spanned by the unobservable factors F_t as both $N, T \longrightarrow \infty$, and Bai and Ng (2006) showed that using the principal components factors estimates as regressors does not affect the consistency properties of the OLS coefficient estimators, as long as $\sqrt{T}/N \longrightarrow 0$ as $N, T \longrightarrow \infty$. In detail, the SW forecasting model requires the following steps:

1. Compute the sample covariance matrix $\hat{\Gamma}_x(0) = T^{-1} \sum_{t=1}^{T} x_t x_t'$.
2. Compute the N eigenvalues of $\hat{\Gamma}_x(0)$, $\hat{\mu}_1, \ldots, \hat{\mu}_N$.
3. Estimate the factors F_t through the static principal components estimator: $\hat{F}_t = (\sqrt{N})^{-1} \hat{P}' x_t$, where $\hat{P} = [\hat{p}_1, \ldots, \hat{p}_r]$ is the N x r matrix that contains the normalized (*i.e.* with unit norm) eigenvectors corresponding to the first r largest eigenvalues of the sample covariance matrix.
4. Estimate OLS (using data up to $T-h$) the constant α, the r x m coefficients of the matrix $B = [\beta_1, \ldots, \beta_m]$ and the p coefficients $\gamma_1, \ldots, \gamma_p$ of the ADL (m, p) model in which the estimated factors $\hat{F}_t$ are used as predictors together with the variable to be forecast:
$$x_{i,t+h}^h = \alpha_h + \sum_{j=1}^{m} \beta_{hj}' \hat{F}_{t-j+1} + \sum_{k=1}^{p} \gamma_{hk} x_{i,t-j+1} + \epsilon_{t+h}^h$$

where $x^h_{i,t+h}$ is the h-step ahead variable to be forecast, ϵ^h_{t+h} is the linear regression error and the subscripts denote the dependence of the estimates on the forecasting horizon h. The SW forecasting equation is then given by:

$$\hat{x}^h_{i,T+h \vee T} = \hat{\alpha}_h + \sum_{j=1}^{m} \hat{\beta}'_{hj} \hat{F}_{T-j+1} + \sum_{k=1}^{p} \hat{\gamma}_{hk} x_{i,T-j+1}$$

where $\hat{\alpha}_h, \hat{\beta}_{hj}$ and $\hat{\gamma}_{hk}$ are the OLS estimates obtained using data up to time T, and $\hat{x}^h_{i,T+h \vee T}$ is the h-step ahead forecast of $x_{i,T+h}$ based on $\{x_{i,T}, x_{i,T-1}, \ldots, \hat{F}_T, \hat{F}_{T-1}, \ldots\}$.

Stock and Watson named the above forecasting equation "DI-AR Lag" (Diffusion Index – Autoregressive, with Lags). They also proposed two alternative restricted versions of the DI-AR Lag forecasting equation: the "DI-AR" (Diffusion Index – Autoregressive) forecasts, in which the lags of the factors are dropped ¿ i.e. $\hat{x}^h_{i,T+h \vee T} = \hat{\alpha}_h + \hat{\beta}'_h \hat{F}_T + \sum_{k=1}^{p} \hat{\gamma}_{hk} x_{i,T-j+1} ¿$; and the "DI" (Diffusion Index) forecasts, in which both the autoregressive part and the lags of the factors are dropped ¿ i.e. $\hat{x}^h_{i,T+h \vee T} = \hat{\alpha}_h + \hat{\beta}'_h \hat{F}_T ¿$. Which of the three forecasting equations performs better is an empirical question that I will try to answer in the forecasting exercise of the next chapter.

The calibration of three parameters is required in order to make SW operational in its more general version (DI-AR Lag):

(i) The number of static factors (r).
(ii) The maximum number of lags (m) for the factors.
(iii) The maximum number of lags (p) for the response variable.

In the forecasting exercise of the next chapter we set the maximum number of lags m ¿ and p) using the well-known BIC information criteria, while number of static factors (r) is estimated using three different methods explained in section 1.3.

1.2 The Forni, Hallin, Lippi and Reichlin model

Forni, Hallin, Lippi and Reichlin (2005) assumed an approximate dynamic factors structure with finite lags (*i.e.* static representation) for the covariance stationary standardized panel X, and proposed a procedure based on two main steps to forecast x_i. In the first step, they use the frequency-domain technique (*i.e.* dynamic PCA) of Forni and co-authors (2000) to estimate covariance matrices of the common and idiosyncratic components. In the second step, they use these estimates to obtain a consistent estimator of the space spanned by the static factors F_t solving the generalized principal components problem, and propose a predictor obtained as the linear projection of future values of the common component on the factor space. In detail, the FHLR model requires the following steps:

1. Compute $(2M+1)$ k-lag sample covariance matrices $\hat{\Gamma}_x(k)=T^{-1}\sum_{t=1}^{T} x_t x_{t-k}'$, where $k\in\{-M,\ldots,M\}$ and M is the truncation parameter used in the next step to estimate the spectral density matrix of x_t .
2. Use the estimates above to compute the sample spectral density matrix $\hat{\Sigma}_x(\theta_h)$ for each frequency $\theta_h=\frac{2\pi h}{2H}$, where $h\in\{0,\ldots,2H\}$:

$$\hat{\Sigma}_x(\theta_h)=\sum_{k=-M}^{M} w(M^{-1}k)\hat{\Gamma}_x(k)e^{-ik\theta_h}$$

   where $w(M^{-1}k)$ is a positive even weight function and $i=\sqrt{-1}$ is the imaginary unit. The above estimator is consistent as long as the following two conditions are meet: 1) $M>0$, $M\longrightarrow\infty$ and $M/T\longrightarrow 0$ as $T\longrightarrow\infty$; 2) $w(0)=1$, $|w(\alpha)|\leq 1\,\forall\alpha,$ and $w(\alpha)=0\,\forall|\alpha|>1$. The above estimator is called "lag window estimator" of $\Sigma_x(\theta)$.
3. Compute the dynamic eigenvalues of the sample spectral density matrix $\hat{\Sigma}_x(\theta_h)$ for each frequency θ_h ; i.e. compute $\hat{\lambda}_1(\theta_h),\ldots,\hat{\lambda}_n(\theta_h)$. Since the spectral density is a Hermitian matrix, its eigenvalues are real.
4. For each frequency θ_h , use the q-largest dynamic eigenvalues of $\hat{\Sigma}_x(\theta_h)$ and the corresponding eigenvectors to estimate the spectral density matrix of the common component χ_t :

$$\hat{\Sigma}_\chi(\theta_h)=\hat{P}(\theta_h)\hat{\Lambda}(\theta_h)\hat{P}'(\theta_h)$$

   where $\hat{\Lambda}(\theta_h)$ is the q x q diagonal matrix of the dynamic eigenvalues (arranged in decreasing order) of $\hat{\Sigma}_x(\theta_h)$ at frequency θ_h , $\hat{P}(\theta_h)=[\hat{p}_1(\theta_h),\ldots,\hat{p}_q(\theta_h)]$ is the N x q unitary matrix that contains the normalized complex eigenvectors corresponding to the q-largest dynamic eigenvalues of $\hat{\Sigma}_x(\theta_h)$ at frequency θ_h , i.e. $\hat{P}(\theta_h)\hat{P}^{¿}(\theta_h)=\hat{P}^{¿}(\theta_h)\hat{P}(\theta_h)=I$, where $\hat{P}^{¿}(\theta_h)$ is the complex conjugate of $\hat{P}(\theta_h)$.

5. Estimate the spectral density of the idiosyncratic component ξ_t by $\hat{\Sigma}_\xi(\theta_h)=\hat{\Sigma}_x(\theta_h)-\hat{\Sigma}_\chi(\theta_h)$, $\forall\theta_h$. Note that the above estimator exploits the fact that the spectral density of x_t can be decomposed at each frequency as $\hat{\Sigma}_x(\theta_h)=\hat{\Sigma}_\chi(\theta_h)+\hat{\Sigma}_\xi(\theta_h)$, since the common and the idiosyncratic components are assumed to be orthogonal.
6. Estimate the *k*-lag sample covariance matrices of the common and idiosyncratic components through the following Inverse (Discrete) Fourier Transforms:

$$\hat{\Gamma}_\chi(k)=(2H+1)^{-1}\sum_{h=-H}^{H}\hat{\Sigma}_\chi(\theta_h)e^{-ik\theta_h},k\in\{-M,\dots,M\}$$

$$\hat{\Gamma}_\xi(k)=(2H+1)^{-1}\sum_{h=-H}^{H}\hat{\Sigma}_\xi(\theta_h)e^{-ik\theta_h},k\in\{-M,\dots,M\}$$

7. Obtain the N complex numbers $\hat{v}_j$ solving $det\left(\hat{\Gamma}_\chi(0)-\hat{v}_j\hat{\hat{\Gamma}}_\xi(0)\right)=0$, where $\hat{\Gamma}_\chi(0)$ is the $N\,x\,N$ estimated covariance matrix of the common component and $\hat{\hat{\Gamma}}_\xi(0)$ is the $N\,x\,N$ diagonal matrix obtained setting equal to 0 all the off-diagonal entries of the estimated covariance matrix of the idiosyncratic component $\hat{\Gamma}_\xi(0)$. The N complex numbers $\hat{v}_1,\dots,\hat{v}_N$ are then called *generalized eigenvalues*.
8. Use the generalized eigenvalues computed above to obtain the corresponding N *generalized eigenvectors* $\hat{Z}_1,\dots,\hat{Z}_N$, i.e. the N column vectors satisfying $\hat{Z}_j\hat{\Gamma}_\chi(0)=\hat{v}_j\hat{Z}_j\hat{\hat{\Gamma}}_\xi(0)$, s.t. $\hat{Z}_j'\hat{\hat{\Gamma}}_\xi(0)\hat{Z}_i=1$ for $i=j$, and 0 otherwise.
9. Estimate the static factors F_t through the *generalized principal components* (GPC) estimator: $\hat{F}_t=\hat{Z}'x_t$, where $\hat{Z}=\left[\hat{Z}_1,\dots,\hat{Z}_r\right]$ is the N x r matrix that contains the normalized (i.e. with norm 1) generalized eigenvectors corresponding to the first r largest generalized eigenvalues of the pair of matrices $\left(\hat{\Gamma}_\chi(0),\hat{\hat{\Gamma}}_\xi(0)\right)$.
10. Finally, the FHLR *h*-step ahead forecasting equation is then obtained projecting future values of the common component on the factor space:

$$\hat{x}^h_{i,T+h\vee T}=\hat{\chi}^h_{i,T+h\vee T}=\hat{\Gamma}_{\chi,i.}(h)\hat{Z}\left(\hat{Z}'\hat{\Gamma}_x(0)\hat{Z}\right)\hat{F}_t$$

where *h* is the forecasting horizon, $\hat{\Gamma}_{\chi,i.}(h)$ is the *i*-th row of the *k*-lag sample covariance matrices of the common components, $\hat{Z}$ is the N x r matrix that contains the generalized eigenvectors, $\hat{\Gamma}_x(0)$ is the sample covariance matrix, i.e. $\hat{\Gamma}_x(0)=T^{-1}\sum_{t=1}^{T} x_t x_t^{'}$ and $\hat{F}_t$ are the estimated static factors.

We name the above forecasting equation FHLR_1. Another forecasting equation could be obtained forecasting also the idiosyncratic component. This could be done separately since it is assumed that ξ_t is orthogonal to χ_t. In this case, the FHLR *h*-step ahead forecasting equation is obtained adding the projection of future values of the idiosyncratic component on the present and past values of the response variable x_i to the projection of future values of the common component on the factor space: $\hat{x}^h_{i,T+h\vee T}=\hat{\chi}^h_{i,T+h\vee T}+\hat{\xi}^h_{i,T+h\vee T}$, where:

$$\hat{\chi}^h_{i,T+h\vee T}=\hat{\Gamma}_{\chi,i.}(h)\hat{Z}\left(\hat{Z}'\hat{\Gamma}_x(0)\hat{Z}\right)\hat{F}_t$$

and

$$\hat{\xi}^h_{i,T+h\vee T}=\left[\hat{\Gamma}_{\xi,ii}(h),\ldots,\hat{\Gamma}_{\xi,ii}(h+p-1)\right]W_i^{-1}(p)\left[x_{iT},\ldots,x_{iT-p+1}\right]'$$

where the idiosyncratic component forecasting equation is taken following D'Agostino and Giannone (2012), $\hat{\Gamma}_{\xi,ii}(h)$ denotes the entry of the *i*-th row and the *i*-th column of the matrix $\hat{\Gamma}_\xi(h)$, *p* is the maximum number of lag for the idiosyncratic component and

$$W_i(p)=\begin{bmatrix}\hat{\Gamma}_{x,ii}(0) & \cdots & \hat{\Gamma}_{x,ii}(-(p-1)) \\ \cdots & \cdots & \cdots \\ \hat{\Gamma}_{x,ii}(p-1) & \cdots & \hat{\Gamma}_{x,ii}(0)\end{bmatrix}$$

where $\hat{\Gamma}_{x,ii}(p)$ denotes the entry of the *i*-th row and the *i*-th column of the matrix $\hat{\Gamma}_x(p)$. We call this forecasting equation FHLR_2.

The calibration of five parameters is required in order to use FHLR_1:

(i) The number of static factors (r).
(ii) The number of dynamic factors (q).
(iii) The size of the truncation parameter *M* for the spectral density estimation.
(iv) The weight function $w(M^{-1}k)$ for the spectral density estimation.
(v) The number of frequencies $(2H)$ for the computation of the spectral density.

In the forecasting exercise of the next chapter, we set $M=floor(\sqrt{T})$ and $w(M^{-1}k)=1-\frac{|k|}{M+1}$ (*i.e.* triangular window), as suggested by Forni and co-authors.

With respect to the selection of the number of static factors (r), we use three different methods explained in section 2.4. Finally, the number of dynamic factors (q) is estimated using the well know Hallin and Liska (2007) method, which is explained in the next section.

1.3 Determining the number of static factors

In order to use the approximate static factor models presented in the previous section, the unknown true number of static factors (r) has to be estimated. Among the several techniques proposed in the literature for addressing this issue, we consider three different well-known methods: the Information Criterion developed by Bai and Ng (2002), the Edge Distribution proposed by Onatski (2010) and the Eigenvalues Ratio of Ahn and Horenstein (2013).

The pioneering method developed by Bai and Ng (2002) consistently estimate the true number of static factors by minimizing one of two alternative information criterion, called IC_p and PC_p. These two model selection functions are modifications of the well-known AIC and BIC criterion that take into account both the cross-section and the time dimension of the dataset in the overfitting penalty term. Both the IC_p and the PC_p criteria can be specified in three different forms, according to the function chosen for the penalty term. The authors also specified the conditions that the penalty function $p(N,T)$ should satisfy to consistently estimate the true number of static factors, i.e. $p(N,T)\longrightarrow 0$ and $\left(min\left(\sqrt{N},\sqrt{T}\right)\right)^2 p(N,T)\longrightarrow\infty$ as $N,T\longrightarrow\infty$.

In our analysis we consider only one of the six different information criteria proposed by Bai and Ng (2002) for setting the number of static factors (r). Specifically, we use the function that the authors have called IC_{p2}, which performed well in their Monte Carlo simulations:

$$IC_{p2}=\ln\left(V\left(k,\hat{\Lambda},\hat{F}_t\right)\right)+kp(N,T)$$

where

$$V\left(k,\hat{\Lambda},\hat{F}_t\right)=(NT)^{-1}\sum_{t=1}^{T}\left(x_t-\hat{\Lambda}(k)\hat{F}_t(k)\right)'\left(x_t-\hat{\Lambda}(k)\hat{F}_t(k)\right)=(NT)^{-1}\sum_{t=1}^{T}\hat{\xi}_t'\hat{\xi}_t$$

is the cross-sectional average variance of the idiosyncratic component when the model is specified using k static factors,

$$\hat{\Lambda}(k)=\sqrt{N}\hat{P}_t \quad\text{and}\quad \hat{F}_t(k)=\left(\sqrt{N}\right)^{-1}\hat{P}_t'x_t$$

are, respectively, the factor loadings and factors static principal components estimators using the largest k eigenvectors of the sample covariance matrix $\hat{\Gamma}_x(0)$, and

$$p(N,T)=\frac{N+T}{NT}\ln(min\{N,T\})$$

is the penalty term.

The number of static factors (r) is obtained as the value of k that minimize the function IC_{p2} , for $k\in\{0,1,\ldots,k_{max}\}$, where the upper bound k_{max} represents the maximum number of static factor assumed plausible *a priori* by the researcher:

$$\hat{r}=\underset{0\leq k\leq k_{max}}{argmin}\, IC_{p2}(k) \quad .$$

The authors suggested to consider $k_{max}=8\int\left[\left(\frac{min\{N,T\}}{100}\right)^{0.25}\right]$ as a possible heuristic choice of the maximum number of factors to be tested.

Bai and Ng (2002) estimators are often been criticized because of two main reasons. First, these estimators are quite sensitive on the choice of the parameter k_{max} . Second, as long as the choice of the penalty function meets the consistency conditions presented above, multiplying the penalty term by any positive scalar allows to estimate consistently the true number of static factors. Therefore, finite sample performances of the estimator critically depend on the choice made. As a result, it is well known in the literature that Bai and Ng (2002) criteria is non-robust, in the sense that it often overestimates or underestimates the true number of factors in finite samples[1]. Alessi, Barigozzi and Capasso (2010) tackled this problem proposing a data-driven technique (based on the influential work of Hallin and Liska (2007) presented in Section 2.5) to tune the choice of the penalty term. The resulting criteria maintains the same consistency properties as those of the Bai and Ng (2002) criterion, but it has better finite sample performance.

Onatski (2010) developed an alternative estimator of the number of static factors (r) based on the empirical distribution of the eigenvalues of the sample covariance matrix $\hat{\Gamma}_x(0)$. The proposed estimator, named “Edge Distribution” is based on the fact that, for any $k>r$, the difference between two adjacent eigenvalues (of the sample covariance matrix $\hat{\Gamma}_x(0)$) $\hat{\mu}_k-\hat{\mu}_{k+1}$ converges to 0, and $\hat{\mu}_r-\hat{\mu}_{r+1}$ diverge to ∞ as both $N,T\longrightarrow\infty$. The Onatski (2010) method requires the calibration of a

1 About this, see the empirical application results in Forni and co-aouthors (2009) or the Monte Carlo results in Alessi, Barigozzi and Capasso (2010), Onatski (2010) and Ahn and Horenstein (2013))

threshold δ that separates the diverging differenced eigenvalues from the bounded ones. This threshold δ is sharp, in the sense that is cannot be arbitrarily scaled without compromise the consistency results.

The number of static factors (r) is obtained through the following algorithm:

1. Set the maximum number of iteration and the maximum number of static factors to be tested (k_{max}¿;
2. Set $j=k_{max}+1$ and obtain the N eigenvalues $\hat{\mu}_1,\ldots,\hat{\mu}_N$ of the sample covariance matrix $\hat{\Gamma}_x(0)$;
3. Run a simple linear regression in which $\{\hat{\mu}_j,\ldots,\hat{\mu}_{j+4}\}$ is regressed on $\{(j-1)^{2/3},\ldots,(j+3)^{2/3}\}$ and the constant. Then, retrieve the OLS slope estimate $\hat{\beta}$. Set $\delta=2|\hat{\beta}|$;
4. If $\hat{\mu}_k-\hat{\mu}_{k+1}<\delta\ \forall k\leq k_{max}$, then $\hat{r}(\delta)=0$; otherwise $\hat{r}(\delta)=max\{k\leq k_{max}:\hat{\mu}_k-\hat{\mu}_{k+1}\geq\delta\}$;
5. Set $j=\hat{r}(\delta)+1$ and repeat steps 2 and 3 until convergence, i.e. until you get the same $\hat{r}(\delta)$ in two consecutive iterations.

This procedure consistently estimates r under the assumption of normality for the idiosyncratic component. Without this assumption, the consistency is reached if the residuals are either serially or cross-sectionally correlated, but not both. However, in his Monte Carlo exercise the author showed that the *ED* estimator has good finite sample performance even in the case in which the residuals are both cross-sectionally and auto-correlated.

Ahn and Horenstein (2013) proposed two alternative consistent estimators of the number of static factors, named “Eigenvalue Ratio” and “Growth Ratio” estimators. These estimators are based on the approximate factor model assumption that the r largest eigenvalues of the sample covariance matrix $\hat{\Gamma}_x(0)$ diverge to infinity as the number of series (N) increases whereas the remaining ($N-r$¿ eigenvalues do not – they are bounded. With respect to the Onastki (2010) estimator, Ahn and Horenstein (2013) have the comparative advantage of not requiring the use of a pre-specified threshold; moreover, they have shown by Monte Carlo simulations that their estimators are robust with respect to the choice of the upper bound k_{max} .

In this work we consider only the Eigenvalue Ratio estimator, in which the number of static factors (r) is obtained as the value of k that maximize the ratio of two adjacent eigenvalues (arranged in decreasing order) of the sample covariance matrix $\hat{\Gamma}_x(0)$, for $k\in\{0,1,\ldots,k_{max}\}$:

$$\hat{r} = \underset{0 \le k \le k_{max}}{argmax} \frac{\hat{\mu}_k}{\hat{\mu}_{k+1}} \quad ,$$

where $\hat{\mu}_k$ is the k -th largest eigenvalue of $\hat{\Gamma}_x(0)$ and $\hat{\mu}_0 = (min\{N,T\})^{-1} \sum_{k=1}^{min\{N,T\}} \hat{\mu}_k \ / \ \ln(min\{N,T\})$ is the authors' suggested choice for the "mock eigenvalue". Different choices of $\hat{\mu}_0$ could be made, as long as they satisfy the consistency conditions $\hat{\mu}_0 \longrightarrow 0$ and $min\{N,T\}\hat{\mu}_0 \longrightarrow \infty$ as $N,T \longrightarrow \infty$.

1.4 Determining the number of dynamic factors

The FHLR forecasting model requires to estimate the unknown number of dynamic factors (q) to be made operational. Here we consider the popular approach proposed by Hallin and Liska (2007). This method is based on the generalized dynamic factor model assumption that the q largest dynamic eigenvalues of the sample spectral matrix of the observations $\hat{\Sigma}_x(\theta_h)$ diverge to infinity as the number of series N increases, whereas the remaining $N-q$ eigenvalues are bounded.

Hallin and Liska (2007) consistently estimate the true number of dynamic factors by minimizing one of two alternative information criterion, called IC_1 and IC_2 . Each of these criteria requires to consistently estimate the spectral density matrix of X . This could be done either using the lag window estimator presented in Section 1.3 or the periodogram-smoothing estimator. Therefore, both the IC_1 and the IC_2 criteria could be specified in two different forms, according to the spectral density estimator chosen. As recommended by the authors, we use only the criterion based on the lag window estimator. Then, the information criteria to be minimized takes one of the two the following forms:

$$IC_1(k) = (N)^{-1} \sum_{i=k+1}^{N} (2M+1)^{-1} \sum_{h=-M}^{M} \hat{\lambda}_i(\theta_h) + kcp(N,T)$$

or

$$IC_2(k) = \ln\left[(N)^{-1} \sum_{i=k+1}^{N} (2M+1)^{-1} \sum_{h=-M}^{M} \hat{\lambda}_i(\theta_h)\right] + kcp(N,T)$$

where M is a truncation parameter used to estimate the spectral density matrix of the observations, $\theta_h = \frac{\pi h}{M+0.5}$ with $h \in \{-M, \ldots, M\}$ are the $2M+1$ grid points at which the spectral density is estimated, $p(N,T)$ is an appropriate penalty function such that $p(N,T) \longrightarrow 0$ and $min\{N, M^2, \sqrt{T/M}\} p(N,T) \longrightarrow \infty$ as $N, T \longrightarrow \infty$ and c is an arbitrary positive real number present in the equation because, as for the Bai and Ng (2002) information criteria, if the penalty $p(N,T)$ satisfies the consistency conditions above, then also any penalty $cp(N,T)$ does.

We use IC_2, which performed better in their Monte Carlo simulations. Based on their results, we also set

$$M = floor[0.75\sqrt{T}]$$

and

$$p(N,T) = (M^{-2} + \sqrt{M/T} + N^{-1}) \ln(min\{N, M^2, \sqrt{T/M}\}) \quad .$$

The number of dynamic factor is set to the value of *k* that minimize the chosen information criteria $IC_2(k)$ for $k \in \{0, 1, \ldots, k_{max}\}$, where the upper bound k_{max} represents the maximum number of dynamic factor assumed plausible *a priori*:

$$\hat{q} = \underset{0 \le k \le k_{max}}{argmin} IC_2(k) \quad .$$

Notice that in this procedure the positive constant c is considered to be known. However, in practical applications the value of c is not known and has to be calibrated. This is a delicate step because the estimated number of factors is a function of c : if the value of the constant is set too small, then there is underpenalization – and therefore overestimation – of q, while if the value of the constant is set too large, then there is overpenalization – and therefore underestimation – of q. To solve this problem, the authors also propose a data-driven methodology to calibrate the positive constant c and therefore estimate q. In detail, their algorithm requires the following steps:

1. Choose the penalty term $p(N,T)$ and set the values of M, k_{max}, c_{min}, c_{max} and the size of the sub-blocks.
2. Define $\tilde{p}(c,N,T) = c\, p(N,T)$, with $c \in C \subset [c_{min}, c_{max}] \subset R^{+¿}$.
3. Perform a random permutation of the positions of the N series available (optional) and choose J sub-blocks of the series of size (n_j, τ_j) with $j \in \{1, \ldots, J\}$ such that $0 < n_1 < \ldots < n_J = N$ and $0 \le \tau_1 \le \ldots \le \tau_J = T$.
4. Repeat the following step for every $c \in C$:

   Compute the J minimizers

$$\hat{q}_{n_j,\tau_j}(c) = \underset{0\le k\le k_{max}}{argmin} \ln\left[(N)^{-1}\sum_{i=k+1}^{N}(2M+1)^{-1}\sum_{h=-M}^{M}\hat{\lambda}_i(\theta_h)\right] + kcp(N,T) \quad ,$$

that is one $\hat{q}_{n_j,\tau_j}(c)$ for each of the sub-blocks (n_j,τ_j), with $j\in\{1,\dots,J\}$. Notice that the dependence of the number of dynamic factors on c is now explicit. Then, compute the empirical variance of the J minimizers $\hat{q}_{n_j,\tau_j}(c)$:

$$S^2(c) = J^{-1}\sum_{j=1}^{J}\left[\hat{q}_{n_j,\tau_j}(c) - J^{-1}\sum_{j=1}^{J}\hat{q}_{n_j,\tau_j}(c)\right]^2$$

5. Obtain $\hat{c}$ as the smallest value of $c\in C$ such that $\hat{q}_{n_j,\tau_j}(c)$ is a constant function of j (i.e. the empirical variance $S^2(c)=0$) and such that $\hat{q}_{n_j,\tau_j}(c)<k_{max}$ (*i.e.* do not consider the so called "first stability interval").

We set $k_{max}=12$, $n_j = floor\left(\frac{N}{4}\right)+j-1$ and I keep fixed the time dimension of the sub-blocks, i.e. $\tau_j = T\,\forall j$.

1.5 Estimating the factors using a subset of the data available

In approximate (static or dynamic) factor models, all the consistency results require the divergence of both the panel cross-sectional dimension N and time dimension T. Therefore, in empirical applications researchers have always used as much series as were available to estimate the (space spanned by the) factors.

However, Boivin and Ng (2006) found that more data is not always better for estimating the factors and for forecasting using factors estimates. In details, they found that using only a subset of the available series might enhance the precision of the factor estimates and the forecasts accuracy when:

(i) The series to be forecasted depends on "dominated" factors, *i.e.* factors which have a relatively small importance in the available dataset.
(ii) The idiosyncratic components of the available series are strongly cross-correlated.
(iii) Some available variables are too noisy, *i.e.* they have a relatively large idiosyncratic component.

Bai and Ng (2008) exploited the case (i) proposing to estimate the factors using only the subset of the available dataset that contains series which have been tested to have predictive power for the variable to be forecasted. They call these selected series "targeted predictors"[2].

Boivin and Ng (2006) focused mainly on the case (ii) and proposed two rules to select the series to be dropped, based on the correlation between the idiosyncratic components of the available series:

Rule 1: For each of the N available series, compute the sample correlation coefficient between the idiosyncratic component of the variable i and j ; i.e. compute $\hat{\rho}_{ij}$ $\forall i, j$ with $i \neq j$. Then, retrieve $\hat{\rho}_1^{¿}(i) = \max_j |\hat{\rho}_{ij}| \forall i$; i.e. the highest absolute sample correlation coefficient between the idiosyncratic component of the variable i and j . Identify as j_i^1 the variable whose idiosyncratic component is the most correlated with those of series i . Then, obtain the set of the series whose idiosyncratic component is most correlated with the idiosyncratic component of some other series: $j^1 = \{j_i^1\} = \{j_1^1, \ldots, j_N^1\}$ and drop all these variables. If two series are most correlated with each other, do not drop both, but only the one that has the smallest common component's relative importance, where:

$$R_i^2 = \sum_{t=1}^{T} \chi_{it}^2 / \sum_{t=1}^{T} x_{it}^2$$

is the relative importance of the common component in the i -th variable.

Rule 2: Drop the series following *Rule 1*. Then, drop also all the series whose idiosyncratic component is second most correlated with the idiosyncratic component of some other series. Similarly, to Rule 1, if two series are secondly most correlated with each other, do not drop both, but only the one that has the smallest common component's relative importance.

We have tried both Rule 1 and Rule 2 to test if the forecast accuracy could be improved dropping series based on the correlation between their idiosyncratic components. Furthermore, we devised another rule to select the series to be dropped, based on relative importance of the common component of the available series. Our proposed "Rule 3" tries to exploit Bai and Ng (2006) result that it might be possible to improve both the factor estimates' precision and forecasts accuracy dropping series that have a relatively large idiosyncratic component; i.e. case (iii). The implementation details of Rule 3 are:

2 WeI have not implemented this promising approach which we leave as a future research topic.

Rule 3: For each of the N available series, compute the relative importance of the common component statistic, i.e. compute $R_i^2 = \sum_{t=1}^{T} \chi_{it}^2 / \sum_{t=1}^{T} x_{it}^2$ $, \forall i \in \{1, \ldots, N\}$. Then, sort them in increasing order and compute the *p*-th percentile. Drop all the series below the *p*-th percentile. We set $p = 25$.

Finally, we also test a fourth rule in which the series to be dropped have both large and strongly correlated idiosyncratic components, *i.e.* case (ii) and (iii). Specifically, this combines Rule 1 with Rule 3: first, drop the series according to Rule 1; then, apply Rule 3 on the remaining series and drop the variables accordingly. We call this “Rule 4”.

2. Comparisons of Forecasting Performances

In order to run a real-time out-of-sample forecasting exercise, it is necessary to use data actually available at each point in time in which the forecasts are made. Therefore, we create a vintage dataset of "first-release" macroeconomic data, which are the first preliminary data released by official statistical agencies. This is in contrast with what typically happens in the dynamic factor models forecasting literature, where it is common to perform a pseudo real-time forecasting exercise using the last revised data available at when the dataset is built. These data are almost always different from those that are first published and actually known in real time.

Our dataset contains 107 monthly US first-release macroeconomic and financial monthly time series, spanning the 1996:12 to 2017:6 period. All series come from the Bloomberg database. We classify these series into seven broad groups:

1. Output (series 1-15),
2. Labor Market (series 16-28),
3. Housing (series 29-38),
4. Consumption, Orders and Inventories (series 39-47),
5. Money&Credit (series 48-55),
6. Financials (series 56-86),
7. Prices (series 87-107).

All series are transformed to achieve stationarity and no treatment for outliers is applied. A complete description of the dataset and further information on how the variables are made stationary can be found in the Appendix.

A further issue in real-time macroeconomic forecasting is how to deal with the so-called "ragged-edge" problem: real-time datasets contain missing values at the end of the in-sample period. This happens because macroeconomic data are not released at the same time, and therefore new observations on some series are released later than those on other series. As in Altissimo and co-authors (2010), we deal with this problem through the "forward realignment" method: we shift forward the series whose last observations are missing. As a result, we lose some observations at the beginning of the sample period, which is then cut in order to have a balanced dataset. After this realignment, our actual dataset begins in 1998:2, fourteen months later than the beginning of the original raw series.

The forecasting exercise simulates real-time forecasts of four key macroeconomic variables, namely: the industrial production month-on-month change (IP), the unemployment rate, the Core Consumer Price Index (*i.e.* all CPI items less food and energy) and the ISM Purchasing Managers' Index (PMI). It is important to notice that, even if the four series of interest belong to a single dataset, the forward realignment procedure requires peculiar shifting depending on which variable is forecasted. This is because the target variables are not released at the same time. Therefore, in making forecasts the day before of the first release, four different realignments are made depending on which is the target variable.

The SW and FHLR models are first estimated over the seven-years period beginning on 1998:2 and ending on 2005:1. Then both models are used to compute the out-of-sample one-step-ahead forecast, that is the forecast for the period 2005:2. Then the window is rolled on by one month (*i.e.* an additional period of data is added and the first period of the first sample is deleted) and the models are re-estimated. Therefore, the next forecast is based on models estimated on a new sample that differs from the former by one data point. This exercise is repeated until the models are estimated in the last sample, which spans the period 2010:6 -2017:5, and the last forecast is made for 2017:6. Overall, 149 forecasts are computed for each target variable.

All variables are normalized to have zero sample mean and unit sample variance, so that the forecasting exercise is run on covariance-stationary and standardized series. The standardized forecasts are then inverted to their original scales, and the transformations made to achieve stationarity are reverse engineered. This process is carried out for:

1. a univariate AR model of order four,
2. three specifications of SW (DI-AR Lag, DI-AR and DI),
3. two specifications of FHLR (FHLR_1 and FHLR_2.

The forecasting performance of each model is measured by the relative mean squared forecasting error (RMSFE):

$$RMSFE(model)=\frac{\sum_{\tau=\tau_0}^{T-h}\left(y_{\tau+h}-\hat{y}_{\tau+h\vee\tau}^{model}\right)^2/(T-h-\tau_0)}{\sum_{\tau=\tau_0}^{T-h}\left(y_{\tau+h}-\hat{y}_{\tau+h\vee\tau}^{bench}\right)^2/(T-h-\tau_0)}$$

where T is the last date of the sample, corresponding to 2017:6; $h=1$ is the forecasting horizon; τ_0 is the last date of the first rolling window, corresponding to 2005:1; $model$ is the forecasting model considered and ranges over DI-AR Lag, DI-AR, DI, FHLR_1 and FHLR_2; $bench$ is the benchmark model, which is the AR model of order 4; $y_{\tau+h}$ is the value of the variable to be predicted at time $\tau+h$ and $\hat{y}_{\tau+h\vee\tau}$ is the corresponding h -step ahead forecast. A value of the RMSFE below one indicates that the forecasting model considered has had, on average, lower squared forecasting errors than those of the benchmark AR model.

Table 1

Description of the datasets of predictors

Data Sets	Series	Series Numbers
D1	All time series	1 – 107
D2	All time series except those in the Output group	16 – 107
D3	All time series except those in the Labor Market group	1 – 15 and 29 – 107
D4	All time series except those in the Housing group	1 – 28 and 39 – 107
D5	All time series except those in the Consumption, Orders and Inventories group	1 – 38 and 48 – 107
D6	All time series except those in the Money&Credit" group	1 – 47 and 56 – 107
D7	All time series except those in the Financials group	1 – 55 and 87 – 107
D8	All time series except those in the Prices group	1 – 86
D9	All time series except those deleted according to Rule 1	Variable
D10	All time series except those deleted according to Rule 2	Variable
D11	All time series except those deleted according to Rule 3	Variable
D12	All time series except those deleted according to Rule 4	Variable

For each of the four key target variables, the forecasting models are estimated considering twelve different datasets, as detailed in Table 1. Notice that the series deleted according to Rules 1 to 4 can be different at each time window.

2.1 Forecasting Performance Results

The RMSFE results for Industrial Production are shown in Table 2. The best forecasting results of each model are obtained estimating the number of static factors through the Bai and Ng (2002) method. In general, FHLR outperforms SW. Moreover, FHLR outperforms the benchmark almost always, while for SW this outcome depends on both the datasets used to estimate the factors and on the methods applied to determine the number of factors.

Comparing the different specifications of the two dynamic factor models, it can be seen that the three SW specifications provide very similar results and most of the times the simple DI model has a slightly better forecasting performance. Also, the two FHLR specifications yield almost the same results, so forecasting only the common component provides results that are comparable – and many times slightly better – than those obtained forecasting also the idiosyncratic component.

With respect to the dataset used, an interesting result that holds for all models' specifications and for all the methods used to estimate the number of factors is that excluding the Financial or the Prices blocks of variables (*i.e.* using dataset D7 or D8) improves the forecasting performance. On the contrary, excluding the Labor Market block (*i.e.* using dataset D3) deteriorates the forecasting performance. Results on other blocks of variables are not easily interpretable, because they depend either on the forecasting model used or on the method chosen to determine the number of factors. Nevertheless, it is always possible to improve the forecasting accuracy of the five models' specifications by dropping some block of variables, and sometimes these improvements are substantial. These results provide evidence that supports the finding of Boivin and Ng (2006), who stated that using more data to estimate the factors may worsen forecasting performance.

Table 2

Comparisons of Forecasting Performances on Industrial Production

Data Sets	Models				
	SW DI-AR Lag	SW DI-AR	SW DI	FHLR_1	FHLR_2
Information Criteria (Bai and Ng, 2002)					
D1	1.05	1.06	1.06	0.90	0.90
D2	1.04	1.05	1.04	0.87	0.88
D3	1.14	1.15	1.14	0.94	0.94
D4	1.08	1.08	1.08	0.91	0.91
D5	1.05	1.07	1.05	0.93	0.93
D6	0.99	1.00	1.00	0.89	0.89
D7	**0.88**	0.95	0.92	0.86	0.85
D8	0.97	0.99	0.94	**0.80**	**0.79***
D9	0.93	0.99	0.97	0.90	0.91
D10	1.12	1.14	1.12	1.04	1.05
D11	0.96	0.96	**0.89**	0.85	0.84
D12	0.97	**0.88**	0.90	0.88	0.87
Edge Distribution (Onatski, 2010)					
D1	1.06	0.99	0.99	0.97	0.98
D2	1.03	1.06	0.99	0.99	1.00
D3	1.07	1.02	1.00	0.98	0.98
D4	0.97	1.02	1.01	0.89	0.90
D5	1.04	1.02	1.00	0.99	0.99
D6	1.03	1.00	0.97	0.97	0.97
D7	1.00	0.99	0.96	0.97	0.98
D8	1.00	0.93	0.92	0.91	0.91
D9	1.06	1.05	1.03	0.97	0.97
D10	0.95	0.96	0.93	0.94	0.94
D11	1.04	1.00	0.98	0.97	0.98
D12	1.07	1.05	1.03	0.97	0.99
Eigenvalue Ratio (Ahn & Horenstein, 2013)					
D1	1.00	0.93	0.93	0.92	0.93
D2	0.96	0.99	0.93	0.92	0.93
D3	1.00	0.99	0.97	0.94	0.96
D4	0.95	0.99	0.96	0.86	0.87
D5	0.99	0.96	0.96	0.94	0.95
D6	0.98	0.92	0.92	0.91	0.92
D7	0.89	0.91	0.89	0.90	0.90
D8	0.99	0.93	0.92	0.91	0.91
D9	0.99	1.00	0.97	0.92	0.92
D10	1.04	1.03	0.99	0.90	0.91
D11	0.98	0.93	0.93	0.92	0.93
D12	1.01	0.99	0.98	0.91	0.92

Note: Relative Root Mean Squared Error against the AR benchmark. The best forecasting performance of each model marked in boldface. Among these five top forecasting performances, the best one is starred.

The performance comparisons on the unemployment rate are shown in Table 3. Both SW and FHLR outperform the benchmark, no matter what is the model specification and which dataset or which method to select the number of factors are used. Overall, SW tends to outperform FHLR. Comparing the different specifications, the simple FHLR_1 performs better than the more complex FHLR_2, in which both the common and the idiosyncratic component are forecasted.

Both SW and FHLR lead to similar conclusions about the informative contents of the different datasets. In detail, excluding the Output, Labor Market, Consumption, Orders and Inventories or the Financials blocks (*i.e.* using dataset D2, D3, D5 or D7 respectively) generally improves the forecasting performance. The same happens dropping variables according to Rule 2 or Rule 3 (*i.e.* using dataset D10 or D11), while excluding the Prices block (*i.e.* using dataset D8) improves the forecasting performance. Indeed, the best forecasting results of all the five models are obtained using dataset D8. While for all SW specifications these performances are achieved estimating the number of factors through the Onatski (2010) method, the best performances for the two FHLR specification are obtained using the Bai and Ng (2002) method. The RMSFE improvements excluding the price variables are considerable. For example, the mean squared forecasted errors of the DI models is 84% that of the autoregressive model when using the largest dataset (i.e. D1), while it is only 76% that of the benchmark model when excluding the prices block of variables. Therefore, the Boivin and Ng (2006) result that using more data to estimate the factors may provide worse forecasting performance is supported by our analysis on the unemployment rate forecasts. The simple DI specification of SW has better forecasting performance than the two alternative specifications, where lags on the response variable or on the factors are allowed. For FHLR a similar conclusion applies: the simple FHLR_1 performs better than FHLR_2 – which includes an idiosyncratic component.

Table 3

Comparisons of Forecasting Performances on Unemployment

Data Sets	Models				
	SW DI-AR Lag	SW DI-AR	SW DI	FHLR_1	FHLR_2
Information Criteria (Bai and Ng, 2002)					
D1	0.88	0.88	0.87	0.88	0.9
D2	0.9	0.88	0.87	0.9	0.92
D3	0.92	0.92	0.91	0.94	0.96
D4	0.86	0.85	0.84	0.85	0.9
D5	0.92	0.9	0.9	0.9	0.92
D6	0.9	0.9	0.88	0.86	0.89
D7	0.88	0.88	0.84	0.86	0.9
D8	0.8	0.8	0.77	**0.81**	**0.83**
D9	0.85	0.87	0.85	0.86	0.88
D10	0.9	0.9	0.9	0.96	0.98
D11	0.91	0.92	0.9	0.91	0.92
D12	0.83	0.88	0.89	0.93	0.95
Edge Distribution (Onatski, 2010)					
D1	0.88	0.85	0.84	0.88	0.91
D2	0.9	0.89	0.88	0.94	0.96
D3	0.9	0.87	0.86	0.92	0.94
D4	0.84	0.83	0.82	0.86	0.9
D5	0.89	0.86	0.85	0.9	0.92
D6	0.91	0.85	0.84	0.89	0.92
D7	0.85	0.82	0.82	0.88	0.91
D8	**0.79**	**0.78**	**0.76***	0.82	0.84
D9	0.87	0.84	0.84	0.89	0.9
D10	0.91	0.91	0.9	0.93	0.95
D11	0.87	0.85	0.84	0.89	0.91
D12	0.87	0.82	0.82	0.88	0.9
Eigenvalue Ratio (Ahn & Horenstein, 2013)					
D1	0.85	0.84	0.81	0.86	0.88
D2	0.86	0.87	0.84	0.91	0.93
D3	0.89	0.88	0.85	0.91	0.94
D4	0.84	0.84	0.82	0.86	0.89
D5	0.87	0.86	0.83	0.88	0.91
D6	0.88	0.83	0.81	0.87	0.9
D7	0.84	0.81	0.8	0.85	0.88
D8	0.82	0.81	0.78	0.85	0.87
D9	0.87	0.84	0.83	0.89	0.91
D10	0.88	0.88	0.88	0.91	0.94
D11	0.85	0.84	0.81	0.87	0.9
D12	0.87	0.83	0.83	0.88	0.92

Note: Relative Root Mean Squared Error against the AR benchmark. The best forecasting performances of each model are marked in boldface. Among the top forecasting performances, the best one is starred.

Table 4 shows our results for the CPI. The table shows that the best forecasting results of each model are obtained estimating the number of static factors through the Onatski (2010) method, and that in general FHLR model outperforms the SW model. However, the two dynamic factor models rarely beat the benchmark, and when they do the relative forecasting improvements are quite small. More precisely, SW beats the AR model only 10% of the time, while FHLR outperforms the benchmark approximately 30% of the time. The best performance is achieved by the FHLR_2 specification in which both the common and the idiosyncratic component are forecast on the dataset that exclude the Output block of variables. In this case, the mean squared forecast errors is 96% relative to the benchmark autoregressive model. Comparing the different specifications of the two dynamic factor models (SW and FHLR), it can be seen that the DI-AR Lag specification of SW never beats the autoregressive model, and in general the DI-AR specification has better forecast performance than the other two alternative specifications (DI and DI-AR Lag). For FHLR, the FHLR_2 specification performs better than FHLR_1 where only the common component is forecast. A further interesting result that holds for all the five forecasting models and for all the methods used to estimate the number of factors is that dropping the variables according to Rule 1 (*i.e.* using dataset D9) improves the forecasting performance for CPI. Results on other block of variables are not easily interpretable because depend either on the forecasting model used or on the method chosen to determine the number of factors. Nevertheless, Boivin and Ng (2006) result are still supported, even if the improvements here are marginal.

Table 4

Comparisons of Forecasting Performances on Consumer Prices

Data Sets	Models				
	SW DI-AR Lag	SW DI-AR	SW DI	FHLR_1	FHLR_2
Information Criteria (Bai and Ng, 2002)					
D1	1.15	1.12	1.16	1.03	1.01
D2	1.14	1.10	1.15	1.03	0.98
D3	1.12	1.09	1.13	1.05	1.04
D4	1.11	1.08	1.12	1.01	0.98
D5	1.14	1.10	1.15	1.05	1.03
D6	1.14	1.11	1.13	1.04	1.01
D7	1.07	1.02	1.06	1.02	1.00
D8	1.17	1.15	1.21	1.11	1.04
D9	1.09	1.06	1.09	1.03	0.99
D10	1.07	1.02	1.07	1.08	1.02
D11	1.18	1.14	1.18	1.05	1.01
D12	1.17	1.11	1.15	1.02	0.98
Edge Distribution (Onatski, 2010)					
D1	1.04	1.00	1.03	1.02	0.97
D2	1.00	0.98	**0.98**	**0.99**	**0.96***
D3	1.01	**0.97**	1.00	**0.99**	0.98
D4	1.01	0.99	1.03	1.01	0.98
D5	1.03	0.99	1.00	1.00	0.98
D6	1.05	1.00	1.04	1.03	1.00
D7	1.04	1.00	1.02	1.01	1.01
D8	**1.00**	0.99	1.01	1.01	0.97
D9	**1.00**	**0.97**	1.00	1.00	0.97
D10	1.02	1.00	1.00	1.03	0.99
D11	1.04	1.00	1.03	1.02	0.97
D12	1.03	0.99	1.02	1.03	0.98
Eigenvalue Ratio (Ahn & Horenstein, 2013)					
D1	1.04	1.00	1.04	1.04	0.98
D2	1.05	1.01	1.07	1.06	1.02
D3	1.03	1.00	1.03	1.04	1.02
D4	1.04	1.01	1.06	1.01	0.97
D5	1.03	1.00	1.04	1.03	1.00
D6	1.04	1.01	1.04	1.04	0.99
D7	1.02	0.98	1.03	1.03	0.98
D8	1.04	1.02	1.05	1.06	1.01
D9	1.02	0.99	1.03	1.03	0.98
D10	1.05	1.02	1.07	1.08	1.02
D11	1.04	1.00	1.04	1.03	0.98
D12	1.02	0.99	1.03	1.02	0.97

Note: Relative Root Mean Squared Error against the AR benchmark. The best forecasting performances of each model are marked in boldface. Among the top forecasting performances, the best one is starred.

Table 5 shows results for the ISM-PMI index. SW outperforms FHLR which never beats the AR benchmark. Comparing the different specifications of SW, it can be seen that, in general, the DI-AR Lag specification has better forecasting performance than the alternative specifications (DI and DI-AR) and beats the benchmark almost always. The opposite happens for the simple DI specification, which never outperforms the autoregressive model. With DI-AR Lag and DI-AR the best performances are achieved estimating the number of factors through the Eigenvalue Ratio method, while the best performances for the two FHLR specifications and for the SW DI specification are achieved using the Edge Distribution method. An interesting result that holds for all the five forecasting models and for all the methods used to estimate the number of factors is that excluding the Housing block of variables (*i.e.* using dataset D4) substantially improves the forecasting performance. For example, the mean squared forecasted errors of the DI-AR Lag model is 92% of the AR model when using the largest dataset (i.e. D1), while it is only 85% that of the benchmark model when excluding the Prices block of variables – and estimating the number of factors through the Eigenvalue Ratio method. On the contrary, excluding the Labor Market block of variables (*i.e.* using dataset D3) deteriorates the forecasts. Results on other block of variables are not easily interpretable because depend either on the forecasting model used or on the method chosen to determine the number of factors. Nevertheless, it is always possible to improve the forecasting accuracy of the five models dropping some block of variables. Therefore, the Boivin and Ng (2006) result that using more data to estimate the factors may provide worse forecasting performance supported also for the ISM PMI forecasts.

Table 5

Comparisons of Forecasting Performances on ISM-PMI

Data Sets	Models				
	SW DI-AR Lag	SW DI-AR	SW DI	FHLR_1	FHLR_2
Information Criteria (Bai and Ng, 2002)					
D1	0.94	0.95	2.17	2.63	2.46
D2	0.96	0.98	2.27	2.66	2.23
D3	0.99	1.00	2.41	2.96	2.62
D4	0.92	0.93	1.71	2.13	1.75
D5	0.94	0.97	2.54	2.87	2.33
D6	0.95	0.97	2.23	2.52	2.20
D7	1.00	1.01	2.51	2.55	2.33
D8	0.93	0.95	2.28	2.81	2.38
D9	0.87	0.88	2.28	2.18	2.02
D10	1.00	1.01	2.67	2.19	1.95
D11	0.92	0.94	2.20	2.70	2.35
D12	0.99	1.02	2.45	2.27	1.98
Edge Distribution (Onatski, 2010)					
D1	0.95	0.99	2.49	3.15	2.92
D2	0.94	1.00	2.97	3.72	3.44
D3	0.99	1.01	2.79	3.39	3.12
D4	0.87	0.87	1.90	1.90	**1.71**
D5	0.96	0.98	2.66	3.28	2.98
D6	0.96	0.99	2.48	3.13	2.98
D7	0.99	0.98	2.31	3.23	3.00
D8	0.97	0.95	2.87	3.59	3.35
D9	0.90	0.94	**1.57**	2.45	2.34
D10	1.04	1.02	2.22	**1.88**	1.72
D11	0.89	0.96	2.32	2.66	2.48
D12	0.95	0.99	2.52	3.18	2.94
Eigenvalue Ratio (Ahn & Horenstein, 2013)					
D1	0.92	0.97	2.76	3.41	3.19
D2	0.92	0.98	3.46	3.46	3.84
D3	0.95	0.97	3.16	3.91	3.65
D4	**0.85***	**0.86**	2.10	2.04	1.81
D5	0.94	0.97	3.07	3.70	3.33
D6	0.94	0.98	2.79	3.39	3.15
D7	0.96	0.97	2.61	3.41	3.21
D8	0.99	0.98	3.00	3.71	3.44
D9	0.86	0.92	2.59	2.77	2.55
D10	0.96	1.00	3.47	2.56	2.24
D11	0.87	0.94	2.63	2.99	2.76
D12	0.92	0.97	2.77	3.49	3.22

Note: Relative Root Mean Squared Error against the AR benchmark. The best forecasting performances of each model are marked in boldface. Among the top forecasting performances, the best one is starred.

3. Concluding Remarks

In this paper we compare the forecasting performance of two popular dynamic factor models on vintage data: Stock and Watson (2002a,b) and Forni, Hallin, Lippi and Reichlin (2005). Our first-release dataset has 107 US macroeconomic and financial monthly time series, spanning the 1996:12 – 2017:6 period. Data were collected from the Bloomberg database. Vintage data allow to perform a real-time out-of-sample forecasting exercise, because these data are those that were actually available at each point in time in which the forecasts could have been made. The fact that real-time datasets contain missing values at the end of the in-sample period ("ragged-edge" problem) is solved as in Altissimo co-authors (2010) shifting forward the series whose last observations are missing.

The real-time one-month ahead forecasts are made for four key macroeconomic variables: Industrial Production month-on-month change, unemployment rate, Core Consumer Price Index and the ISM Purchasing Managers' Index. The comparison is run using three different methods to determine the number of static factors: the Information Criterion developed by Bai and Ng (2002), the Edge Distribution proposed by Onatski (2010) and the Eigenvalues Ratio of Ahn and Horenstein (2013). Moreover, all the forecasts are made using either the whole dataset, or excluding, in turn, various block of variables, or dropping from the whole dataset some variables according to the rule proposed by either Boivin and Ng (2006) or by us.

Four main results emerge from our analysis. First, both the Stock&Watson and the Forni and co-authors models outperform a simple autoregressive model for industrial production, unemployment rate and consumer prices. For the ISM-PMI, only Stock&Watsons' model outperforms the autoregression. Second, there is not a model that always outperform the other: while Forni and co-authors' beats Stock&Watson's for industrial production consumer prices, the opposite happens for the unemployment rate and ISM-PMI. Third, the best forecasting performance are obtained estimating the number of static factors with different methods: for the industrial production, the best is achieved using the Bai and Ng (2002) method; for both the unemployment rate and consumer prices the Onatski (2010) method is best; ISM-PMI the Ahn and Horenstein (2013) method is preferred. Finally, it is always possible to improve the forecasting accuracy of the predictive models dropping some block of variables, and sometimes these improvements are substantial. Therefore, as found also by Boivin and Ng (2006), using more data to estimate the factors worsen forecasting performance. Overall, our results suggest that different model and specification choices, methods for pinning down the number of factors and choices of the set of predictors interplay significantly in determining out-of-sample forecasting performances. Our results are robust to the spurious effect that relying on revised data may induce, thanks to us using first-release vintage data. Contrary to what is commonly found in the existing literature on model comparisons, this fact makes our analysis realistic in the sense that it is grounded on the information that was actually available when the forecasts could have been produced.

References

AHN S. C. AND A.R. HORENSTEIN (2013): “Eigenvalue ratio test for the number of factors”, *Econometrica*, **81**, 1203-1227.

ALESSI L., M. BARIGOZZI AND M. CAPASSO (2010): “Improved penalization for determining the number of factors in approximate factor models”, *Statistics and Probability Letters*, **80**, 1806-1813.

ALTISSIMO F., R. CRISTADORO, M. FORNI, M. LIPPI AND G. VERONESE (2010): “New Eurocoin: tracking economic growth in real time”, *Review of Economics and Statistics*, **92**, 1024-1034.

BAI J. AND S. NG (2002): “Determining the number of factors in approximate factor models”, *Econometrica*, **70,** 191-221.

BAI J. AND S. NG (2006): “Confidence intervals for diffusion index forecasts and inference for factor-augmented regressions”, *Econometrica*, **74**, 1133-1150.

BAI J. AND S. NG (2008): “Forecasting economic time series using targeted predictors”, *Journal of Econometrics*, **146**, 304-317.

BOIVIN J. AND S. NG (2006): “Understanding and comparing factor-based forecasts”, *International Journal of Central Banking*, **3**, 117-151.

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

D’AGOSTINO A. AND D. GIANNONE (2012): “Comparing alternative predictors based on large panel factor models”, *Oxford bulletin of economics and statistics*, **74**, 2012

DELLA MARRA F. (2017): “A forecasting performance comparison of dynamic factor models based on static and dynamic methods”, *Communications in Applied and Industrial Mathematics*, **8**, 43-66.

DEN REIJER A.H.J. (2005): “Forecasting Dutch GDP using large scale factor models”, DNB Working Paper 28, DNB.

FORNI M., D. GIANNONE, M. LIPPI AND L. REICHLIN (2009): “Opening the black box: structural factor models versus structural VARs”, *Econometric Theory*, **25**, 1319-1347.

FORNI M., A. GIOVANNELLI, M. LIPPI AND S. SOCCORSI (2016): “Dynamic factor model with infinite dimensional factor space: forecasting”, tech.rep., CEPR Discussion Paper, No DP11161.

FORNI M., M. HALLIN, M. LIPPI AND L. REICHLIN (2000): “The generalized dynamic factor model: identification and estimation”, *Review of Economics and Statistics*, **82**, 540-554.

FORNI M., M. HALLIN, M. LIPPI AND L. REICHLIN (2005): “The generalized dynamic factor model: one-sided estimation and forecasting”, *Journal of the American Statistical Association*, **100**, 830-840.

FORNI M., M. HALLIN, M. LIPPI AND P. ZAFFARONI (2015): “Dynamic factor model with infinite dimensional factor space: representation”, *Journal of Econometrics*, **185**, 359-371.

FORNI M., M. HALLIN, M. LIPPI AND P. ZAFFARONI (2016): “Dynamic factor model with infinite dimensional factor space: asymptotic analysis”, tech. rep., EIEF Working Paper 2016/07.

HALLIN M. AND R. LISKA (2007): “Determining the number of factors in the general dynamic factor model”, *Journal of the American Statistical Association*, **102**, 603-617.

KAPETANIOS G. AND M. MARCELLINO (2009): “A parametric estimation method for dynamic factor models of large dimensions”, *Journal of Time Series Analysis*, **30**, 208-238.

McCRACKEN M. W. AND S. NG (2016): “FRED-MD: a monthly database for macroeconomic research”, *Journal of Business & Economic Statistics*, **34**, 574-589.

ONATSKI A. (2010): “Determining the number of factors from empirical distribution of eigenvalues”, *Review of Economic and Statistics*, **92**, 1004-1016.

SCHUMACHER C. (2007): “Forecasting German GDP using alternative factor models based on large datasets, *Journal of Forecasting*, **26**, 271-302.

STOCK, J. H. and M. W. WATSON (2002a): “Forecasting using principal components from a large number of predictors”, *Journal of the American Statistical Association*, **97**, 1167-1179.

STOCK, J. H. and M. W. WATSON (2002b): “Macroeconomic forecasting using diffusion indexes”, *Journal of Business & Economic Statistics*, **20**, 147-162.

Appendix

In this appendix we give a description of the series that compose the balanced dataset used in our analysis. The dataset contains 107 monthly US "first-release" macroeconomic and financial time series spanning 1996:12 to 2017:6 belonging to seven broad groups:

1. Output (series 1-15),
2. Labor Market (series 16-28),
3. Housing (series 29-38),
4. Consumption, Orders and Inventories (series 39-47),
5. Money&Credit (series 48-55),
6. Financials (series 56-86),
7. Prices (series 87-107).

The structure of dataset is similar to the FRED-MD database, described in McCracken and Ng (2016), which is a publicly available dataset maintained and updated with monthly periodicity by the Federal Reserve Bank of St. Louis. FRED-MD contains an unbalanced panel of 134 US monthly *revised* macroeconomic and financial series, starting from 1959:1. Our dataset differs from FRED-MD because 18 series have been added and 43 have been excluded.

The 18 added series are: (1) US Continuing Jobless Claims SA, (2) US Manufacturers New Orders Total MoM SA, (3) US Aggregate Reserves Depository Institutions plus Vault cash NSA, (4) US/ Russia Foreign exchange rate, (5) US/Euro Foreign exchange rate, (6) US/Mexico Foreign exchange rate, (7) US/ Chinese Renminbi exchange rate, (8) Dow Jones Industrial average Index, (9) Dow Jones Industrial average Price Earnings Ratio, (10) Russell 2000 Index, (11) VIX Index, (12) US Personal Consumption Expenditures less Food and Energy ("Core PCE") SA MoM, (13) US CPI: All items less Energy SA, (14) US CPI: All Items Less Food and Energy ("Core CPI") SA, (15) US CPI: Fuels and Utilities SA, (16) US CPI: Food and beverages SA, (17) US CPI: Recreation SA, (18) Gold Price.

We dropped the 43 excluded variables either because they are not in the Bloomberg database or because the first release economic dates are missing – these dates are necessary for the real-time forecasting exercise in order to shift the series according to the timeliness of the variables – or because some data are missing. The dropped variables are: (1) US Real Personal Income, (2) US Real personal income ex transfer receipts, (3) US Help-Wanted Index for United States, (4) US Ratio of Help Wanted/No. Unemployed, (5) US Civilian Employment (6) US Civilians Unemployed for 15–26 Weeks (7) US Civilians Unemployed for 27 Weeks and Over, (8) US All Employees: Goods-Producing Industries (9) US All Employees: Mining and Logging: Mining (10) US All Employees: Construction, (11) US All Employees: Manufacturing, (12) US All Employees: Durable goods, (13) US All Employees: Nondurable goods, (14) US All Employees: Service-

Providing Industries, (15) US All Employees: Trade, Transportation & Utilities, (16) US All Employees: Wholesale Trade, (17) US All Employees: Retail Trade, (18) US All Employees: Financial Activities, (19) US All Employees: Government, (20) US Average Weekly Hours: Goods-Producing, (21) US Average Weekly Overtime Hours: Manufacturing, (22) US Average Weekly Hours: Manufacturing, (23) US Real personal consumption expenditures, (24) ISM: US PMI Composite Index, (25) US New Orders for Consumer Goods, (25) US New Orders for Nondefense Capital Goods, (26) US Total Business Inventories, (27) US Total Business: Inventories to Sales Ratio, (28) US Real M2 Money Stock, (29) St. Louis Adjusted US Monetary Base, (30) US Real Estate Loans at All Commercial Banks, (31) US Nonrevolving consumer credit to Personal Income, (32) US Consumer Motor Vehicle Loans Outstanding, (33) US Total Consumer Loans and Leases Outstanding, (34) US Securities in Bank Credit at All Commercial Banks, (35) Trade Weighted U.S. Dollar Index: Major Currencies, (36) US PPI: Finished Goods, (37) US PPI: Finished Consumer Goods, (38) US PPI: Intermediate Materials, (39) US PPI: Crude Materials, (40) US PPI: Metals and metal products, (41) US CPI: Apparel, (42) US CPI: Medical Care, (43) S&P's Composite Common Stock: Dividend Yield.

Our dataset starts is shorter than FRED-MD as it starts in December 1996 because first-release data for early periods are not available in the Bloomberg database for any series.

Table A provides a description of our dataset, of the transformations applied to each variable to achieve stationarity and of the correspondence of each series with similar series in FRED-MD. The "In FRED-MD" column gives mnemonics used to name similar series in the FRED-MD database. The "Bloomberg Ticker" column gives the Bloomberg ticker used to download each series. Finally, the entries in column "TCode" refers the transformations applied to each raw time series in order to meet the stationarity condition:

1. X_t no transformation,
2. $(1-L)X_t$,
3. $(1-L)(X_t/X_{t-1}-1)$,
4. $\ln(X_t)100$,
5. $(1-L)\ln(X_t)100$,
6. $(1-L^2)\ln(X_t)100$,
7. $(1-L)(1-L^{12})\ln(X_t)$ 100.

Table A
Description of the Data Set

Count	TCode	Bloomberg Ticker	Description	In FRED-MD
Output Group				
1	1	IP CHNG Index	IP Index MoM SA	INDPRO
2	5	IPTLTOTL Index	IP: Final products (market group) SA	IPFINAL
3	1	IPTLNOCH Index	IP: Final products Nonindustrial supplies MoM SA	IPFPNSS
4	5	IPTLCG Index	IP: Consumer Goods SA	IPCONGD

5	5	ICGDDCGS Index	IP: Durable Consumer Goods SA	IPDCONGD
6	5	IPNDTOTL Index	IP: Nondurable Consumer Goods	IPNCONGD
7	5	IPEQBUS Index	IP: Business Equipment SA	IPBUSEQ
8	5	IPTLMATS Index	IP: Materials SA	IPMAT
9	5	IDGMTOT Index	IP: Durable Materials SA	IPDMAT
10	5	INDMTOT Index	IP: Nondurable Materials SA	IPNMAT
11	5	IPMG Index	IP: Manufacturing SA	IPMANSIC
12	5	IPMUUTIL Index	IP: Utilities SA	IPB51222s
13	5	IPTSENRG Index	IP: Energy SA	IPFUELS
14	1	NAPMPMI Index	ISM Manufacturing PI SA	NAPMPI
15	2	CPMFTOT Index	US Capacity utilization: Manufacturing Total SIC SA	CUMFNS
			Labor Market Group	
16	1	NAPMEMPL Index	ISM Manufacturing: Employment Index SA	NAPMEI
17	2	USURTOT Index	US Unemployment Rate SA	UNRATE
18	2	INJCJC Index	US Initial Jobless Claims SA	CLAIMSx
19	2	INJCSP Index	US Continuing Jobless Claims SA	
20	5	USLFTOT Index	US Labor force SA	CLF16OV
21	2	NFP TCH Index	US Empl Nonfarm Payrolls Tot MoM Net Change SA	PAYEMS
22	2	USDUMEAN Index	US Average Unemployment Duration (weeks) SA	UEMPMEAN
23	5	USDULSFV Index	US Civilians Unempl: Less than 5 weeks (thous) SA	UEMPLT5
24	5	USDUFVFR Index	US Civilians Unemployed: 5-14 weeks (thousands) SA	UEMP5TO14
25	5	USDUFIFT Index	US Civilians Unemployed: over 15 weeks (thous) SA	UEMP15OV
26	5	USAPGOOD Index	US avg weekly Payrolls: Goods-Producing SA	CES0600000008
27	5	USECTOT Index	US avg weekly Payrolls: Construction SA	CES2000000008
28	5	USAWMANU Index	US avg weekly Payrolls: Manufacturing SA	CES3000000008
			Housing Group	
29	4	NHSPSTOT Index	US New Privately Owned Housing Units Started: Total SA	HOUST
30	4	NHSPSNE Index	US New Privately Owned Housing Units Start: Northeast SA	HOUSTNE
31	4	NHSPSSO Index	US New Privately Owned Housing Units Started: South SA	HOUSTMW
32	4	NHSPSMW Index	US New Privately Owned Housing Units Started: Midwest SA	HOUSTS
33	4	NHSPSWE Index	US New Privately Owned Housing Units Started: West SA	HOUSTW
34	4	NHSPATOT Index	US New Private Housing Permits (thousands): Total (SAAR)	PERMIT
35	4	NHSPANE Index	US New Private Housing Permits (thous): Northeast (SAAR)	PERMITNE
36	4	NHSPAMW Index	US New Private Housing Permits (thous): Midwest (SAAR)	PERMITMW
37	4	NHSPASO Index	US New Private Housing Permits (thousands): South (SAAR)	PERMITS
38	4	NHSPAWE Index	US New Private Housing Permits (thousands): West(SAAR)	PERMITW
			Consumption, Orders and Inventories Group	
39	2	CONSSENT Index	University of Michigan Consumer Sentiment Index NSA	UMCSENTx
40	1	NAPMNEWO Index	ISM Manufacturing: New Orders Index SA	NAPMNOI
41	1	NAPMINV Index	ISM Manufactoring: Inventories NSA	NAPMII
42	1	NAPMSUPL Index	ISM Manufactoring: Supplier Deliveries SA	NAPMSDI
43	1	DGNOCHNG Index	US New Orders for Durable goods SA	AMDMNOx
44	1	TMNOCHNG Index	US Manufacturers New Orders Total MoM SA	
45	1	DGUOTOT Index	US Unfilled Orders for Durable Goods	AMDMUOx
46	1	MTIBCHNG Index	US Real Manufactoring & Trade Sales (Millions) SA	CMRMTSPLx
47	1	RSTAMOM Index	US Retail & Food Services sales SA Tot Monthly % change SA	RETAILx
			Money&Credit Group	
48	6	M1 Index	M1 Money stock SA	M1SL
49	6	M2 Index	M2 Money stock SA	M2SL
50	6	ARDIRRNA Index	US Aggr Reserves Depository Institutions Required NSA	TOTRESNS
51	6	ARDITOTN Index	US Aggr Reserves Depository Instit plus Vault cash NSA	
52	2	ARDINBRN Index	US Aggr Reserves Depository Instit Non Borrowed NSA	NONBORRES
53	6	MZM Index	MZM Money supply SA	MZMSL
54	6	ALCBC&IL INDEX	US C&I Loans SA	BUSLOANS
55	6	CCOSNREV Index	US Total Nonrevolving Credit	NONREVSL
			Financials Group	
56	2	FED30D Index	Effective Federal Funds Rate 30 Day	FEDFUNDS
57	2	CPDR9AFC Index	FED 3 months AA Fin Commercial Paper Interest Rate	CP3Mx
58	2	H15T3M Index	US 3-month Treasury Bill yield	TB3MS
59	2	H15T6M Index	US 6-month Treasury Bill yield	TB6MS
60	2	H15T1Y Index	US 1-year Treasury rate	GS1
61	2	H15T5Y Index	US 5-year Treasury rate	GS5
62	2	H15T10Y Index	US 10-year Treasury rate	GS10
63	2	MOODCAAA Index	Moody's Seasoned AAA Corporate Bond Yield	AAA
64	2	MOODCBAA Index	Moody's Seasoned BAA Corporate Bond Yield	BAA
65	5	USDJPY Curncy	US/Japan Foreign exchange rate	EXJPUSx

66	5	USDCAD Curncy	US/Canada Foreign exchange rate	EXCAUSx
67	5	USDRUB Curncy	US/ Russia Foreign exchange rate	
68	5	USDGBP Curncy	US/UK Foreign exchange rate	EXUSUKx
69	5	USDCHF Curncy	US/ Switzerland Foreign exchange rate	EXSZUSx
70	5	USDEUR Curncy	US/Euro Foreign exchange rate	
71	5	USDMXN Curncy	US/ Mexico Foreign exchange rate	
72	5	USDCNY Curncy	US/ Chinese Reminbi exchange rate	
73	5	SPX Index	SP500 Index	S&P500
74	5	SPX Index	SP500 Price Earnings Ratio	S&P500 PE ratio
75	5	INDU Index	Dow Jones Industrial average Index	
76	5	INDU Index	Dow Jones Industrial average Price Earnings Ratio	
77	5	RTY Index	Russell 2000 Index	
78	1		3-Month Commercial Paper Minus FEDFUNDS	COMPAPFFx
79	1		3-Month Treasury C Minus FEDFUNDS	TB3SMFFM
80	1		6-Month Treasury C Minus FEDFUNDS	TB6SMFFM
81	1		1-Year Treasury C Minus FEDFUNDS	T1YFFM
82	1		5-Year Treasury C Minus FEDFUNDS	T5YFFM
83	1		10-Year Treasury C Minus FEDFUNDS	T10YFFM
84	1		Moody's Aaa Corporate Bond Minus FEDFUNDS	AAAFFM
85	1		Moody's Baa Corporate Bond Minus FEDFUNDS	BAAFFM
86	1	VIX Index	VIX Index	
			Financials Group	
87	1	PCE DEFM Index	US Pers Consump Expend: Chain Index MoM SA	PCEPI
88	1	PCE CMOM Index	US PCE less Food and Energy ("Core PCE") SA MoM	
89	6	PCE DRBD Index	US Pers Cons Expend: Durable Goods Price deflator SA	DDURRG3M086SBEA
90	6	PCE NDRD Index	US PCE: Non Durable Goods Price deflator SA	DNDGRG3M086SBEA
91	6	PCE SRVD Index	US Pers Consum Expend: Services Price deflator SA	DSERRG3M086SBEA
92	1	CPI CHNG Index	US CPI: All items SA	CPIAUCSL
93	6	CPUPAXFD Index	US CPI: All items less Food SA	CPIULFSL
94	6	CPUPAXSH Index	US CPI: All items less Shelter SA	CUUR0000SA0L2
95	6	CPUPAXMC Index	US CPI: All items less Medical Care SA	CUSR0000SA0L5
96	6	CPUPAXEN Index	US CPI: All items less Energy SA	
97	6	CPUPAXFE Index	US CPI: All items less Food and Energy ("Core CPI") SA	
98	6	CPCATOT Index	US CPI: Commodities SA	CUSR0000SAC
99	6	CPSSTOT Index	US CPI: Services SA	CUSR0000SAS
100	6	CPCADUR Index	US CPI: Durables SA	CUUR0000SAD
101	6	CPSTPRV Index	US CPI: Private transportation SA	CPITRNSL
102	6	CPSHFU Index	US CPI: Fuels and Utilities SA	
103	6	CPSFTOT Index	US CPI: Food and beverages SA	
104	6	CPSRTOT Index	US CPI: Recreation SA	
105	1	NAPMPRIC Index	ISM Manufacturing: Price Index SA	NAPMPRI
106	6	CL1 Comdty	Crude Oil Price	OILPRICEx
107	6	GC1 Comdty	Gold Price	