UNIVERSITY OF PARMA

DOCTORAL THESIS IN ECONOMICS AND SOCIAL SCIENCES

XXVIII PH.D. CYCLE

On the macroeconomic forecasting performance of selected dynamic factor models

Supervisor:
Prof. Simona SANFELICI
Co-Supervisor:
Prof. Marco LIPPI

Author:
Fabio DELLA MARRA

Coordinator:
Prof. Mario MENEGATTI

S.S.D.: SECS-S/06

Declaration of Authorship

Hereby, I testify that I did this Ph.D. thesis independently and without accessing unathourized help. Apart form the ones mentioned in this document, no further resources have been used.

Reggio Emilia, 10.06.2017

Fabio Della Marra.

"Forecasting is the art of saying what will happen, and then explaining why it didn't!"

Anonymous (communicated by Balaji Rajagopalan)

University of Parma

Abstract

Department of Economics and Management

Doctor of Philosophy in Economics and Social Sciences

On the macroeconomic forecasting performance of selected dynamic factor models

by Fabio DELLA MARRA

In this doctoral thesis, we compare the forecasting performance of three dynamic factor models on macroeconomic and financial datasets. The purpose of the first two chapters is to provide an incremental contribution with respect to the body of literature comparing static versus dynamic factor models. Previous literature compares the forecasting performance of the static factor model SW (see Stock and M. W. Watson, 2002a, Stock and M. W. Watson, 2002b) against those of the dynamic factor model FHLR (see Forni, Hallin, Lippi, and Reichlin, 2000, Forni, Hallin, Lippi, and Reichlin, 2005). This work adds a third dynamic factor model, which is the recently published FHLZ (see Forni, Hallin, Lippi, and Zaffaroni, 2015, Forni, Hallin, Lippi, and Zaffaroni, 2016). In the third chapter, we compare the forecasting performance of two static factor models cast in a state-space form. In the first one, the conditional moments of the factors are estimated under proper hypothesis of linearity and gaussianity of the data. In the second one, the assumptions of linearity and gaussianity are relaxed for the estimation of the conditional moments of the factors.
Chapter 1 presents an application of the three factor models (SW, FHLR and FHLZ) for forecasting purposes. It compares the pseudo real-time forecast performances of the three factor models against a benchmark AR(4) (an autoregressive process of order 4) over a dataset of 176 EU macroeconomic and financial time series. In this exercise, FHLZ generally outperforms all methods on the forecasting of the Consumer Price Index (CPI). Instead, no method seems to outperform the others in forecasting the Industrial Production (IP), but all dynamic factor models outperform the benchmark AR(4).
Chapter 2 presents two applications on the same topic of the previous chapter. The most innovative part of these applications is that a genetic algorithm is employed to calibrate the three dynamic factor models. The first application exposed in this chapter employs the same dataset of Chapter 1. Instead, in the second application a dataset of 115 US macroeconomic and financial time series is employed. In this chapter, we show that FHLR tends globally to outperform the other methods on the real variables and that FHLZ tends globally to outperform the other methods on the nominal variables. As to EU dataset, in chapter 1 we found similar results for the CPI, but mixed evidences appeared for the IP. As to the US dataset, Forni, Giovannelli, et al., 2016 found similar but less significant results.
Chapter 3 extends a previous study from Banbura and Modugno, 2014, by comparing the forecasting performance of a dynamic factor model cast in state-space form in which the conditional moments relative to the factors are estimated by means of the two following techniques:

(i) *Kalman filter:* as in Banbura and Modugno, 2014, the conditional moments relative to the factors are estimated under the hypothesis that the data generating process (DGP) is linear and that the error terms follow a Gaussian distribution;

(ii) *Paticle Filter:* in this case, the conditional moments relative to the factors are estimated in a more general framework, in which the DGP may be affected by sources of nonlinearity and in which the error terms may not follow a Gaussian distribution.

Up to our knowledge, the estimation of the conditional moments of the factors by means of the Particle Filter has not been carried out yet. In this application, we employ the same Small dataset of 14 EU/US macroeconomic and financial time series from Banbura and Modugno, 2014. We show that the assumptions of linearity of the DGP and of a gaussian distribution for the error terms seems to hold in this macroeconomic setting. Hence, the estimation of the conditional moments of the factors by means of the Kalman Filter seems to be the more appropriate choice in macroeconomic forecasting. However, it is also possible that the particle filter may outperform in financial forecasting. As can be seen in Habibnia, 2017, it appears that accounting for the sources of nonlinearity in the DGP plays a more relevant role on forecasting financial variables.

Acknowledgements

I would like to thank my supervisor prof. Simona Sanfelici for her encouragement and support throughout all the years of my Ph.D. studies.
I am also grateful to my co-supervisor prof. Marco Lippi. Aside from being a top professor, he is also a great teacher who taught me all I know about time series and forecasting with great patience and dedication. I owe him far more than one.
Last but definitely not least, I would like to thank prof. Matteo Barigozzi, prof. Luca Consolini, prof. Gino Favero, dr.Alessandro Giovannelli, dr. Marco Ieva, prof. Fabrizio Laurini, prof. Paola Modesti, prof. Annamaria Olivieri, prof. Aurelio Piazzi, prof. Stefano Soccorsi, dr. Giuseppina Tomasello, Valentino Triani, the *EIEF guys* and my family for their great support and help. Without them, this work would have not been possible.

Contents

List of Figures

List of Tables

Dedicated to my family

Chapter 1

A forecasting performance comparison of dynamic factor models based on static and dynamic methods

1.1 Introduction

In this chapter, a comparative analysis of the forecasting performance of three Large-Dimensional Dynamic Factor Models is presented. As a key feature, Dynamic Factor Models represent each variable in a dataset as the sum of two orthogonal terms: a *common component* χ_t, driven by a reduced (as compared to the number of series in the dataset) number of common factors, and an *idiosyncratic component* ξ_t, which represents measurement errors or local features. Among the different versions of the Dynamic Factor Models we selected:

(i) *SW model*. This time-domain method was introduced in Stock and M. W. Watson, 2002a, Stock and M. W. Watson, 2002b. The factors are estimated by computing static principal components of the variables in the dataset. Let y_{it} be the variable of the dataset to be forecasted at time t, its h-step-ahead prediction equation (also called *Diffusion Forecast Index*) is obtained by regressing y_{it+h} on the factors and on y_{it} itself. Lags of the factors and of y_{it} may be added.

(ii) *FHLR model*. This frequency-domain method was proposed in Forni, Hallin, Lippi, and Reichlin, 2000, Forni, Hallin, Lippi, and Reichlin, 2005 and requires the computation of two steps. In a first step, the common component χ_t, the idiosyncratic component ξ_t and their covariances are estimated using a frequency-domain method introduced in Forni, Hallin, Lippi, and Reichlin, 2000 named *Dynamic Principal Component*. In the second step, the factors are estimated by computing Generalized Principal Components.

(iii) *FHLZ model*. This frequency-domain method was proposed in Forni, Hallin, Lippi, and Zaffaroni, 2015, Forni, Hallin, Lippi, and Zaffaroni, 2016. Here, the underlying assumption in (i) and (ii) that the common components span a finite-dimensional space as n tends to infinity is relaxed. The estimation of the parameters is much more complex though.

There exists some literature comparing the forecasting performances of SW and FHLR, but universal consensus still does not seem to have been reached. Theoretically, time-domain methods consider only relations among the variables at the same time, whereas frequency-domain methods exploit leaded and lagged relations among the variables. However time-domain methods require less parameters to be calibrated. Hence they are more robust to misspecification than frequency-domain methods. Empirically, in Boivin and Ng, 2005 Boivin and Ng found that SW generally outperforms FHLR on US macroeconomic data, whereas D'agostino and Giannone in D'Agostino and Giannone, 2012 found no relevant differences in the performance of both methods on the whole sample (even though heterogeneity is found in subsamples). In Schumacher, 2007 Schumacher found that FHLR outperforms SW on the prediction of the German GDP. The same conclusions are drawn in Reijer et al., 2005 over the forecasting of Dutch GDP. So far, a systematic comparison of the forecasting performances of SW, FHLR and FHLZ can be found only in Forni, Giovannelli, et al., 2016. Here, Forni et al. conducted a forecasting exercise on an US macroeconomic dataset, where they took an autoregressive process of order 4 as a benchmark. They showed that FHLZ oupterforms SW, FHLR and the benchmark both for Industrial Production and Inflation during the Great Moderation. In the Great Recession, the forecasting performances of the Industrial Production change dramatically: all factor models are outperformed by the benchmark and SW and FHLR outperform FHLZ. Hence, Forni et al. concluded that, due to its more dynamical structure, FHLZ tends to be the best performing method in "stationary period", but it loses ground during regime changes. Also, they showed that FHLZ tends to be outperforming on nominal variables and FHLR on real variables.

In this chapter, a large macroeconomic dataset consisting of 176 EU macroeconomic and financial time series observed at monthly frequency over the period from February 1986 to November 2015 is used to analyse the forecasting performance of these methods. To achieve stationarity, the series are deseasonalized and transformed. No treatment for outliers is applied.

As in Forni, Giovannelli, et al., 2016, the EU dataset is split into two subsamples. The first, from February 1986 to December 2000, will be used to calibrate the models, i.e. to produce in-sample forecasts of the variables of the EU dataset for several specifications of SW, FHLR and FHLZ. Then, for each class of models, we selected the specification which shows the minimum mean squared prediction error. These models are then run and compared in the remaining subsample, from January 2001 to November 2015.

The chapter is structured as follows. In Section 2, the factor models are discussed in detail. In Section 3, the main features of the dataset are illustrated. In Section 4, the calibration process of the models is described. In Section 5, results are discussed and Section 6 concludes.

1.2 An overview on Dynamic Factor Models

'

Let $\{\mathbf{x_t} = (x_{1t} \dots x_{nt})' | t = 1, \dots, T\}$ be a n-dimensional vector of time series, which will be denoted as $\mathbf{x_t}$ for simplicity. $\mathbf{x_t}$ will be assumed to be weakly-stationary, purely non deterministic with zero mean and unit variance. Let us assume that the following decomposition holds:

$$\mathbf{x_t} = \chi_t + \xi_t. \tag{1.1}$$

The process $\chi_t = (\chi_{1t} \dots \chi_{nt})'$ is called *common component*. It will be assumed that χ_{it} is stationary and that χ_{it} is costationary with χ_{jt} for all $i, j = 1, \dots, n$ such that $i \neq j$. The process ξ_t is called *idiosyncratic component*. It will be assumed that ξ_{it} is stationary and that ξ_{it} is costationary with ξ_{jt} for all $i, j = 1, \dots, n$ such that

$i \neq j$. A distinction between *static* and *dynamic* factor models can be made according to the functional form selected for the common component. In static factor models, the common component can be modeled as $\chi_t = \mathbf{\Lambda F}_t$ where $\mathbf{\Lambda} \in \mathbb{R}^{n \times r}$ is called the factor-loading matrix and $\mathbf{F}_t \in \mathbb{R}^r$ is the vector of the static factors at time t, with $r << n$. In dynamic factor models, the common component takes into account also the lags of the factors. Hence, the common component can be modeled as $\chi_t = \mathbf{\Lambda}(L)\mathbf{u}_t$. Here L is the lag operator, $\mathbf{\Lambda}(L) \in \mathbb{R}^{n \times q}$ is called the factor-loading matrix and $\mathbf{u}_t \in \mathbb{R}^q$ is the vector of the dynamic factors at time t, with $q << n$. In Forni and Lippi, 2001, Hallin and Lippi, 2013, it is shown that $\mathbf{u}_t$ is a orthonormal white noise process. Moreover, fixed a maximum lag order s for the matrix $\mathbf{\Lambda}(L)$, a dynamic factor model can be rewritten as a static factor model by staking all of the factors with their lags in a single vector, i.e. by imposing $\mathbf{F}_t = (\mathbf{u}_t \ldots \mathbf{u}_{t-s})$. More details can be found in Forni, Giannone, et al., 2009. This way, it holds that $r << q << n$. The three different dynamic factor models for estimating the factors are discussed in the following subsections.

1.2.1 The SW model

In Stock and M. W. Watson, 2002a, Stock and M. W. Watson, 2002b, Stock and Watson proposed a static factor model whose components are estimated by means of static principal component. Let $\hat{\mathbf{\Gamma}} = T^{-1} \sum_{t=1}^{T} \mathbf{x}_t \mathbf{x}_t'$ be the sample covariance matrix of $\mathbf{x}_t$. By computing the eigenvalues of $\hat{\mathbf{\Gamma}}$ and stacking them into the matrix $\mathbf{P} = (\mathbf{P}_1 \ldots \mathbf{P}_r)'$, with $\mathbf{P}_i$ the eigenvector corresponding to the i-th largest eigenvalue, we can compute the factors $\hat{\mathbf{F}}_t = \mathbf{P}'\mathbf{x}_t$ The h-step ahead SW forecasting equation, also called *Diffusion Forecast Index*, can be computed by regressing x_{it+h} on the factors $\hat{\mathbf{F}}_t$ and x_{it}. Lags of both $\hat{\mathbf{F}}_t$ and x_{it} may be added. Hence, the Diffusion Forecast Index can be modeled as

$$x_{it+h|t} = \mathbf{a_i}(L)\hat{\mathbf{F}}_t + b_i(L)x_{it}, \tag{1.2}$$

where $\mathbf{a_i}(L)$ is a $n \times r$ vector of polynomials of degree α_i and $b_i(L)$ is a scalar polynomial of degree β_i.

1.2.2 The FHLR model

This model was proposed by Forni et al. in Forni, Hallin, Lippi, and Reichlin, 2000, Forni, Hallin, Lippi, and Reichlin, 2005. It is articulated in two steps:

(i) *Estimation of the common and idiosyncratic component*: let $\hat{\mathbf{\Gamma}}(k) = T^{-1} \sum_{t=1}^{T} \mathbf{x}_t \mathbf{x}_{t-k}'$ be the sample autocovariance of $\mathbf{x}_t$ at lag k. In order to consistently estimate the spectral density matrix of $\mathbf{x}_t$, we can compute this estimator $\hat{\mathbf{\Sigma}}(\theta) = \sum_{k=-d}^{d} w_k \hat{\mathbf{\Gamma}}(k) e^{-ik\theta}$ with w_k being the weights of a kernel function. By computing the spectral decomposition of $\hat{\mathbf{\Sigma}}(\theta)$ for all θ, the spectral density matrix of the common component can be reconstructed by computing $\hat{\mathbf{\Sigma}}_\chi(\theta) = \hat{\mathbf{P}}(\theta)\mathbf{\Lambda}(\theta)\hat{\mathbf{P}}^*(\theta)$, where $\mathbf{\Lambda}(\theta)$ denotes the diagonal matrix whose entries are the q largest eigenvalues of $\hat{\mathbf{\Sigma}}(\theta)$, $\hat{\mathbf{P}}(\theta) = (\hat{\mathbf{P}}_1(\theta) \ldots \hat{\mathbf{P}}_q(\theta))$ the matrix whose columns are the corresponding eigenvectors and $\hat{\mathbf{P}}^*(\theta)$ indicates the hermitian conjugate of $\hat{\mathbf{P}}(\theta)$. The spectral density matrix of the idiosyncratic component can be reconstructed by differencing $\hat{\mathbf{\Sigma}}_\xi(\theta) = \hat{\mathbf{\Sigma}}(\theta) - \hat{\mathbf{\Sigma}}_\chi(\theta)$ Autocovariances at lag k of the common and the idiosyncratic component can be obtained by computing the inverse Fourier transform of their estimated spectral density matrix.

(ii) *Estimation of the factors*: now, the estimated covariance matrix of the common component and of the idiosyncratic component are used to solve the generalized principal components problem:

$$\hat{\mathbf{\Sigma}}_{\xi}^{0}\mathbf{P} = \hat{\mathbf{\Sigma}}_{\chi}^{0}\mathbf{PD}, \tag{1.3}$$

s.t. $\mathbf{P}'\hat{\mathbf{\Sigma}}_{\chi}{}^{0}\mathbf{P} = \mathbf{I}$, where $\mathbf{D}$ is a diagonal matrix whose entries are the r largest eigenvalues of the pair $(\hat{\mathbf{\Sigma}}_{\xi}^{0} and \hat{\mathbf{\Sigma}}_{\chi}^{0})$ and $\mathbf{P}$ is the matrix containing the corresponding eigenvectors. The first r factors are defined as $\hat{\mathbf{F}}_t = \mathbf{P}'\mathbf{x}_t$

By means of the projections:

$$\begin{aligned} \hat{\chi}_{it+h|t} &= Proj[\chi_{it+h}|\hat{\mathbf{F}}_t] \\ \hat{\xi}_{it+h|t} &= Proj[\xi_{it+h}|\mathbf{x}_{it} \dots \mathbf{x}_{it-p}], \end{aligned} \tag{1.4}$$

the h-step ahead FHLR forecasting equation can be finally derived as

$$\chi_{it+h|t} = \hat{\chi}_{it+h|t} + \hat{\xi}_{it+h|t}. \tag{1.5}$$

1.2.3 The FHLZ model

This model was proposed by Forni et al. in Forni, Hallin, Lippi, and Zaffaroni, 2015, Forni, Hallin, Lippi, and Zaffaroni, 2016. Differently from the previous models, here the assumption that the common component spans a finite-dimensional space is relaxed. This model is articulated in the steps listed below:

(i) *Estimation of the spectral density matrix* $\hat{\mathbf{\Sigma}}^{\mathbf{x}}(\theta)$ *of* $\mathbf{x}_t$: the spectral density matrix $\hat{\mathbf{\Sigma}}^{\mathbf{x}}(\theta)$ can be estimated as $\hat{\mathbf{\Sigma}}^{\mathbf{x}}(\theta) = \frac{1}{2\pi}\sum_{k=-d}^{d}\omega_k\hat{\mathbf{\Gamma}}(k)e^{-ik\theta}$ with ω_k representing the weights of a kernel function.

(ii) *Estimation of the spectral density matrix* $\hat{\mathbf{\Sigma}}^{\chi}(\theta)$ *of* χ_t: it is obtained by computing the dynamic principal components of $\hat{\mathbf{\Sigma}}^{\mathbf{x}}(\theta)$ and then by selecting its q principal components which are associated to the largest eigenvalues. For more details, see Forni, Hallin, Lippi, and Reichlin, 2000, Forni, Hallin, Lippi, and Zaffaroni, 2015.

(iii) *Estimation of the autocovariance matrices* $\hat{\mathbf{\Gamma}}_k^{\chi}$ *of* χ_t: The autocovariances of χ_t are estimated by means of the Wiener-Khinchin-Einstein theorem.

(iv) *Estimation of the VAR matrices* $\hat{\mathbf{A}}^k(L)$: under general assumptions, the common component admits a unique blockwise autoregressive representation of the form:

$$\begin{bmatrix} \mathbf{A}^1(L) & 0 & \dots & 0 & \dots \\ 0 & \mathbf{A}^2(L) & \dots & 0 & \dots \\ \vdots & \vdots & \ddots & & \\ 0 & 0 & \dots & \mathbf{A}^k(L) & \dots \\ \vdots & & & & \ddots \end{bmatrix} \chi_t = \begin{bmatrix} \mathbf{R}^1 \\ \mathbf{R}^2 \\ \vdots \\ \mathbf{R}^k \\ \vdots \end{bmatrix} u_t, \tag{1.6}$$

where $\mathbf{A}^k(L) \in \mathbb{R}^{(q+1)\times(q+1)}$ is a polynomial matrix of finite degree ml and $\mathbf{R}^k \in \mathbb{R}^{(q+1)\times q}$. To estimate the VAR matrices $\mathbf{A}^k(L)$, the covariances $\hat{\mathbf{\Sigma}}^{\chi}(\theta)$ are employed.

(v) *Estimation of the matrices* $\mathbf{R}^k$ *and the shock* $\mathbf{u}_t$: these estimates can be recovered by applying standard principal components to the process $\mathbf{A}(L)\mathbf{x}_t$.

By inverting equation (6) it follows that: $\chi_t = [\mathbf{A}(L)^{-1}]\mathbf{R}\mathbf{u}_t = \mathbf{C}(L)\mathbf{u}_t = \mathbf{C}_0\mathbf{u}_t + \mathbf{C}_1\mathbf{u}_{t-1} + ...$ where $\hat{\chi}_t \in \mathbb{R}^n$ and $\hat{\mathbf{A}}(L), \hat{\mathbf{R}}, \hat{\mathbf{W}}(L) \in \mathbb{R}^{n\times n}$.

The h-step ahead FHLZ forecasting equation is reported below:

$$\chi_{it+h|t} = \mathbf{C}_h\mathbf{u}_t + \mathbf{C}_{h+1}\mathbf{u}_{t-1} + ... \,. \tag{1.7}$$

1.3 Description of the dataset

In this empirical application, a large macroeconomic dataset consisting of 176 EU macroeconomic and financial time series observed at monthly frequency is employed. This dataset contains real variables (import/export price indexes, employment, Industrial Production) and nominal variables (money aggregates, consumer price indexes, wages), asset prices (stock prices and exchange rates) and surveys. Further details can be found in Appendix A. To achieve stationarity, several series are deseasonalized and transformed. No treatment for outliers is applied. In addition to SW, FHLR, FHLZ, the forecasts of an autoregressive process (AR) of order 4 are computed as a benchmark. We remark that such a choice for the benchmark is made in Forni, Giovannelli, et al., 2016. The dataset is divided in two parts: a *calibration sample*, ranging from February 1986 to December 2000, which will be employed to select the most performing specification of each model, and a *proper sample*, ranging from January 2001 to November 2015, which will be employed to compare the selected specifications of each model. As in Stock and M. W. Watson, 2002b, D'Agostino and Giannone, 2012, to assess the forecasting performances, the variables which are taken into account are the level of the logarithm of the Industrial Production (IP) and the yearly change of the logarithm of the Consumer Price Index (CPI). Forecasts are computed h-months ahead, with $h \in \{1, 3, 6, 12, 24\}$. For each methods, we employed a rolling-window scheme $[t-l, t]$, whose size l will be determined in the calibration sample. To assess the forecasting performance of each model, the mean-squared forecasting error (MSFE) is employed as a metric:

$$MSFE^h(i) = \frac{1}{(T_{end}-h) - T_{begin} + 1} \sum_{k=T_{begin}}^{T_{end}-h} SFE^h(i), \tag{1.8}$$

where T_{begin}, T_{end} stand, respectively, for the first and the last date in the dataset and $i \in \{SW, FHLR, FHLZ, AR\}$. SFE^h stands for h-step ahead squared forecasting error and is defined as $SFE^h(i) = (y^i_{t+h|T} - y_{t+h})^2$, where $y^i_{t+h|T}$ is the forecasted value at horizon h of the variable y_t by the method i and y_{t+h} is its real value.

1.4 Calibration

The calibration procedure is basically the same as in Forni, Giovannelli, et al., 2016, but is more robust since the size of the rolling window is taken into account as a parameter to be tuned. The calibration sample, ranging from February 1986 to December 2000, will be used to calibrate the methods SW, FHLR, FHLZ. Namely, this portion of the dataset will be used to select the best performing specifications of each class of models. To compare the performances of two different methods, say α, β, at a certain horizon h, the ratio mean-squared forecasting error (rMSFE) will be computed. Such metric is defined as

$$rMSFE^h(\alpha,\beta) = \frac{MSFE^h(\alpha)}{MSFE^h(\beta)} \,. \tag{1.9}$$

It should be clear that the following cases can arise:

(i) $rMSFE^h(\alpha,\beta) < 1$: then the method α shows a lower MSFE than β at forecasting horizon h on the calibration sample;

(ii) $rMSFE^h(\alpha,\beta) > 1$: then the method α shows a higher MSFE than β at forecasting horizon h on the calibration sample;

(iii) $rMSFE^h(\alpha,\beta) = 1$: then the method α shows the same MSFE than β at forecasting horizon h on the calibration sample.

As in Forni, Giovannelli, et al., 2016, for each method we will select as the most performing configuration the values of the parameters which guarantee the lowest mean $rMSFE^h$ over all $h \in \{1, 3, 6, 12, 24\}$.

1.4.1 Calibration of SW

To produce forecastings by means of equation (1.2), the following parameters must be calibrated:

(i) *the number of static factors* r: ranging from 1 to 10. Also, a comparison with Bai & Ng criterium (BN) with maximum 12 factors has been made.

(ii) *the degree* α *of* $\mathbf{a}(L)$: ranging from 1 to 10.

(iii) *the degree* β *of* $b(L)$: ranging from 0 to 10.

(iv) *the size* l *of the rolling window*: ranging from 5 to 12 years.

By selecting the values of the parameters which guarantee the lowest mean rMSFE, the chosen configuration for the IP is the following:

$$(r, \alpha, \beta, l) = (4, 1, 0, 7). \tag{1.10}$$

Instead, the chosen configuration for the CPI is the following:

$$(r, \alpha, \beta, l) = (4, 1, 9, 12). \tag{1.11}$$

1.4.2 Calibration of FHLR

To produce forecastings by means of equation (1.5), the following parameters must be calibrated:

(i) *the number of static factors* r: ranging from 1 to 10. Also, a comparison with Bai & Ng criterium (BN) with maximum 12 factors has been carried out.

(ii) *the number of dynamic factors* q: ranging from 0 to 10. Also, a comparison with Hallin-Liska criterium (HL) with maximum 12 factors has been carried out.

(iii) *the type of kernel k*: ranging in the set {Triangular, Rectangular, Parzen, Gaussian, Exponential, Cosine, Tukey, Hann}.

(iv) *The lag window d for spectral density estimation*: ranging in the set {25, 35, 40}.

(v) *the size l of the rolling window*: ranging from 5 to 12 years.

By selecting the values of the parameters which guarantee the lowest mean rMSFE, the chosen configuration for the IP is the following:

$$(r, q, k, d, l) = (4, 3, Tukey, 25, 12). \tag{1.12}$$

Instead, the chosen configuration for the CPI is the following:

$$(r, q, k, d, l) = (1, 3, Cosine, 25, 7). \tag{1.13}$$

1.4.3 Calibration of FHLZ

To produce forecastings by means of equation (1.7), the following parameters must be calibrated:

(i) *the number of dynamic factors q*: ranging from 1 to 5. Also, a comparison with Hallin-Liska criterium has been carried out.

(ii) *the type of kernel k*: ranging in the set {Triangular, Rectangular, Parzen, Gaussian, Exponential, Cosine, Tukey, Hann}.

(iii) *the lag window d for spectral density estimation*: ranging in the set {25, 35, 40}.

(iv) *the maximum lag ml for the matrix $\mathbf{A}^k(L)$*: ranging from 1 to 5.

(v) *the size l of the rolling window*: ranging from 5 to 12 years.

By selecting the values of the parameters which guarantee the lowest mean rMSFE, the chosen configuration for the IP is the following:

$$(q, k, d, ml, l) = (4, Parzen, 25, 4, 12). \tag{1.14}$$

Instead, the chosen configuration for the CPI is the following:

$$(q, k, d, ml, l) = (2, Rectangular, 25, 1, 7). \tag{1.15}$$

1.5 Results on the proper sample

1.5.1 Forecasting of the Industrial Production and the Inflation

Now, the forecasting performances of the three dynamic factor models over the IP and CPI are compared on the proper sample, which starts on January 2001 and ends on November 2015. To forecast the IP, we changed the size of the rolling window employed in SW from 7 to 12 years. Instead, to forecast the CPI, we changed the size of the rolling window employed in SW from 12 to 7 years. Hence, to forecast the IP and the CPI, all dynamic factor models employ a rolling window of the same length. The common benchmark for the factor models is the autoregressive process (AR) of order four. In table 1.1, the average rMSFEs (relative to the AR) of the selected dynamic factor models are reported for the IP and CPI on the whole proper sample. However, as reported by CEPR, during the proper sample, the european economy faces two crisis periods: the first starts on May 2008 and ends on January 2009. The second starts on September 2011 and ends on March 2013. Hence, it is reasonable to assess whether the relative forecasting performances of the three dynamic factor models present a relevant change during the crisis periods. In table 1.2, the average rMSFEs (relative to the AR) of the three dynamic factor models are reported for the IP and the CPI from January 2001 to April 2008 (i.e. before the first crisis on the proper sample). In table 1.3, the average rMSFEs (relative to the AR) of the three dynamic factor models are reported for the IP and the CPI from January 2001 to August 2011 (i.e. before the second crisis on the evaluation sample). One, two or three asterisks indicate that the null of equal performance of the three factor models relative to AR is rejected at, respectively, the 1%, 5%, 10% significance by the Diebold-Mariano Test. One, two or three daggers indicate that the null of equal performance of FHLR, FHLZ relative to SW is rejected at, respectively, the 1%, 5%, 10% significance by the Giacomini-White Test. For further details about Giacomini-White Test, see Giacomini and White, 2006.

TABLE 1.1: rMSFEs on the whole sample: IP on the left, CPI on the right.

	IP			CPI		
$h/model$	$FHLZ$	$FHLR$	SW	$FHLZ$	$FHLR$	SW
1	0.97	0.95	0.92	0.95	1.01	1.17
3	0.84	0.82	0.82	0.88^{***}	0.95	1.08
6	0.66	0.66	0.68	0.85	0.94	1.04
12	0.80	0.81	0.79	0.85	0.87	1.18
24	0.94	0.94	0.92	1.02	1.02	1.47

TABLE 1.2: rMSFEs from January 2001 to April 2008: IP on the left, CPI on the right.

	IP			CPI		
$h/model$	$FHLZ$	$FHLR$	SW	$FHLZ$	$FHLR$	SW
1	1.15	1.20	1.24	0.96	0.98	0.99^{***}
3	0.80	0.82	0.79	1.05	1.07	1.14
6	0.92	1.03	1.15	$1.14^{**\dagger}$	$1.17^{**\dagger\dagger}$	1.32^{*}
12	0.91	1.03	1.03	$1.00^{***\dagger}$	$1.04^{**\dagger\dagger}$	1.32^{*}
24	0.85	0.84^{*}	0.86^{*}	0.79	0.82	1.09

TABLE 1.3: rMSFEs January 2001 to August 2011: IP on the left, CPI on the right.

	IP			CPI		
$h/model$	$FHLZ$	$FHLR$	SW	$FHLZ$	$FHLR$	SW
1	0.95	0.87	0.82	0.97	1.06	1.24
3	0.84	0.77	0.78	0.90	0.97	1.08
6	0.65	0.65	0.66	0.85	0.92	0.97
12	0.79	0.80	0.76	0.85	0.87	1.11
24	0.98	0.98	0.93	1.24	1.23	1.82

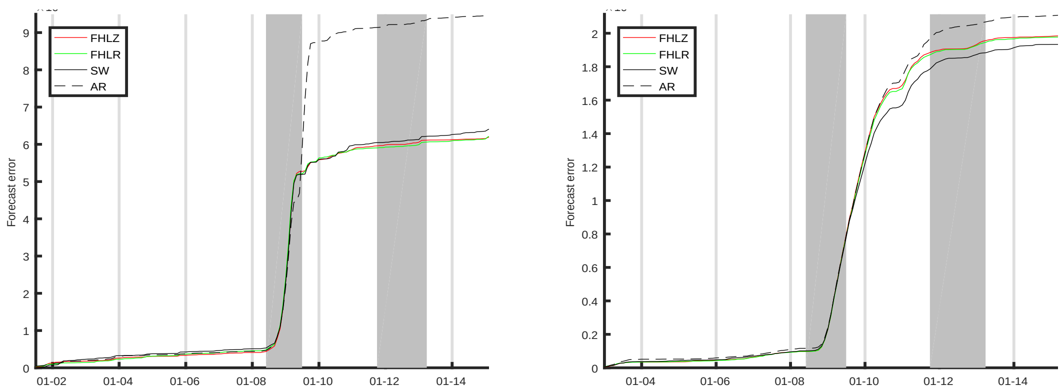

FIGURE 1.1: Cumulative sum of the squared forecasting error at $h = 6$ for the IP (on the left). Cumulative sum of the squared forecasting error at $h = 24$ for the IP (on the right).

The relative performances of all methods tend to improve especially after the first crisis at horizons $h \in \{1, 3, 6, 12\}$. This holds both for IP and CPI. This is also illustrated in figure 1.1 (on the left), in which the graphs of the cumulated sums of the squared forecasting error at $h = 6$ for the IP are reported. The corresponding plots are similar for different values of $h \neq 24$ and are not reported here. The shaded areas correspond to the two crisis periods in the proper sample, according to CEPR. A relevant gradient can be observed during the first crisis for all methods, including the benchmark, followed by a period of consecutive flatness. However, after the first crisis, the sum of the cumulative forecasting errors of AR increases substantially in comparison with those of the dynamic factor models (e.g., on October 2009, the sum of the cumulative forecasting errors of AR increases, on average, more than 50% in comparison with those of the dynamic factor models). Similar results are obtained for the CPI and are not reported here.

Instead, at horizon $h = 24$, the relative performances of all methods tend to worsen after the first crisis. This is also illustrated in figure 1.1 (on the right), in which the graph of the cumulated sums of the squared forecasting error for the IP are reported. In this plot, a gradient during the first crisis can be observed whose slope is less remarked than in the left-side plot. This behaviour seems to persist until the beginning of the second crisis, after which a period of consecutive flatness arises. However, no substantial increase in the sum of the cumulative forecasting errors of AR appears in comparison with the dynamic factor models. Similar results are obtained for the CPI and are not reported here.

As in Forni, Giovannelli, et al., 2016, to assess the forecasting performance of each couple of methods locally, each time series of the dataset is smoothed by a centered moving average of length $m = 61$ (with coefficients equal to $1/m$) and then the Fluctuation test is run, at 5% significance level. To give some insights about the Fluctiation Test, we should remind the reader that it relies on the following statistic:

$$F_{t,m} = \hat{\sigma}^{-1} m^{-1/2} \sum_{j=t-m/2}^{t+m/2+1} \Delta L_j(\alpha_{j-h}, \beta_{j-h}) \tag{1.16}$$

$t = T_{begin} + h + m/2, ..., t_{end} - m/2 + 1$, where $\Delta L_j(\alpha_{j-h}, \beta_{j-h})$ identifies a loss function for methods α, β evaluated at time $j - h$ (in our setting, the loss function is defined as the difference between the MSFE of methods α, β), $\hat{\sigma}$ is a HAC estimator of

$$\sigma = \lim_{R \to +\infty} E(R^{-1/2} \sum_{j=T_{begin}+h}^{T_{end}} \Delta L_j(\alpha_{j-h}, \beta_{j-h}))^2 \tag{1.17}$$

with R identifying the size of the proper sample. Let $\tau \in [0, 1]$. It can be proved that, under general hypotheses, for the null hypothesis $H_0 : E[\Delta L_j(\alpha_{j-h}, \beta_{j-h})] = 0$ for all $t = T_{begin} + h, ..., T_{end}$ we have that:

$$F_{t,m} \sim [\mathbf{B}(\tau + \mu/2) - \mathbf{B}(\tau - \mu/2)]/\sqrt(\mu) \tag{1.18}$$

where $t = [\tau P], m = [\mu R]$ and $\mathbf{B}(.)$ is a standard univariate Brownian motion. the critical valus for significance level α are $\pm k_\alpha$, where k_α solves the following equation:

$$Prob\{sup_\tau |[\mathbf{B}(\tau + \mu/2) - \mathbf{B}(\tau - \mu/2)]/\sqrt(\mu)| > k_\alpha\} = \alpha. \tag{1.19}$$

Simulated values of (α, k_α) and further details can be found in Giacomini and Rossi, 2010.

The results for the IP at horizons $h \in \{6, 12, 24\}$ are reported in Figure 1.2. All methods outperform AR significatively from the first crisis on. At horizon $h = 6$, FHLR and FHLZ outperforms SW on average on the whole sample, except during the first crisis. Instead, at horizons $h \in \{12, 24\}$, SW outperforms FHLR and FHLZ from the first crisis on. FHLR tends to outperform FHLZ in the two crisis periods. Outside the crisis periods, FHLZ tends, on average, to outperfom FHLR at horizons $h \in \{6, 12\}$. Instead, at horizon $h = 24$, FHLR outperforms FHLZ from the first crisis on.

The results for the CPI at horizons $h \in \{6, 12, 24\}$ are reported in Figure 1.3. FHLR and FHLZ tend to outperfom SW at all horizons, except FHLR at horizon $h = 6$ during the first crisis.

FHLR and FHLZ tend to outperfom SW at all horizons, except FHLR at horizon $h = 6$ during the first crisis. FHLR and FHLZ outperform AR at horizons $h \in \{6, 12\}$. At horizon $h = 24$, AR outperforms FHLR and FHLZ from the first crisis on. SW outperforms AR during the two crisis periods at horizons $h \in \{6, 12\}$. At horizons $h \in \{12, 24\}$, AR outperforms SW from the second crisis on. At all horizons, FHLZ outperforms FHLR during the first crisis. At horizons $h \in \{6, 12\}$, this behaviour seems to be persistent. Instead, at horizon $h = 24$, FHLR outperforms FHLZ from the second crisis on.

1.5.2 Forecasting of the dataset

As in Forni, Giovannelli, et al., 2016, this exercise has been extended to the other variables in the dataset. In table 1.4, we report the mean rMSFEs of each class of real (the first three) and nominal (the others) time series, taking AR as a benchmark. Similar results are obtained for the median and not reported here. The best performances are given in bold. We have excluded the categories "Demand", "Money" and "Exchange Rates', since AR outperformed all factor models at any horizon.

AR outperforms all methods at horizon $h = 1$ for all cathegories, except for Interest Rates. Also, AR generally outperforms the dynamic factor models at shorter horizons (within 6 months) and it is the best methods also at $h = 12$ for Unemployment and Wages. As to the factor models, FHLZ seem to be the most accurate methods for real variables. Instead, no method seems to be sistematically the most performing in forecasting the nominal variables. In any case, dynamic methods seem to produce more precise forecasts than SW both on nominal and real variables. In table 1.5, the

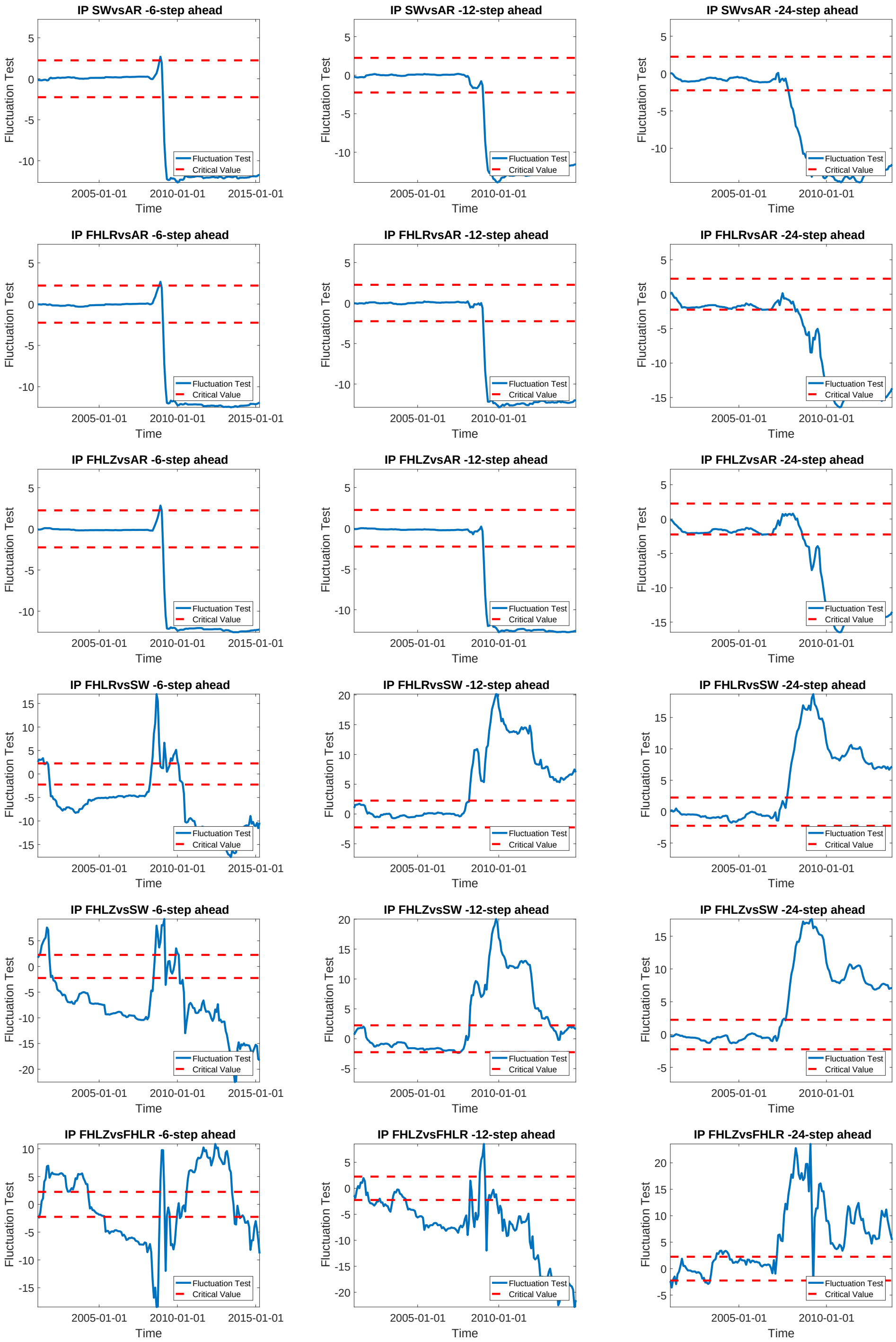

FIGURE 1.2: Fluctuation test for the IP.

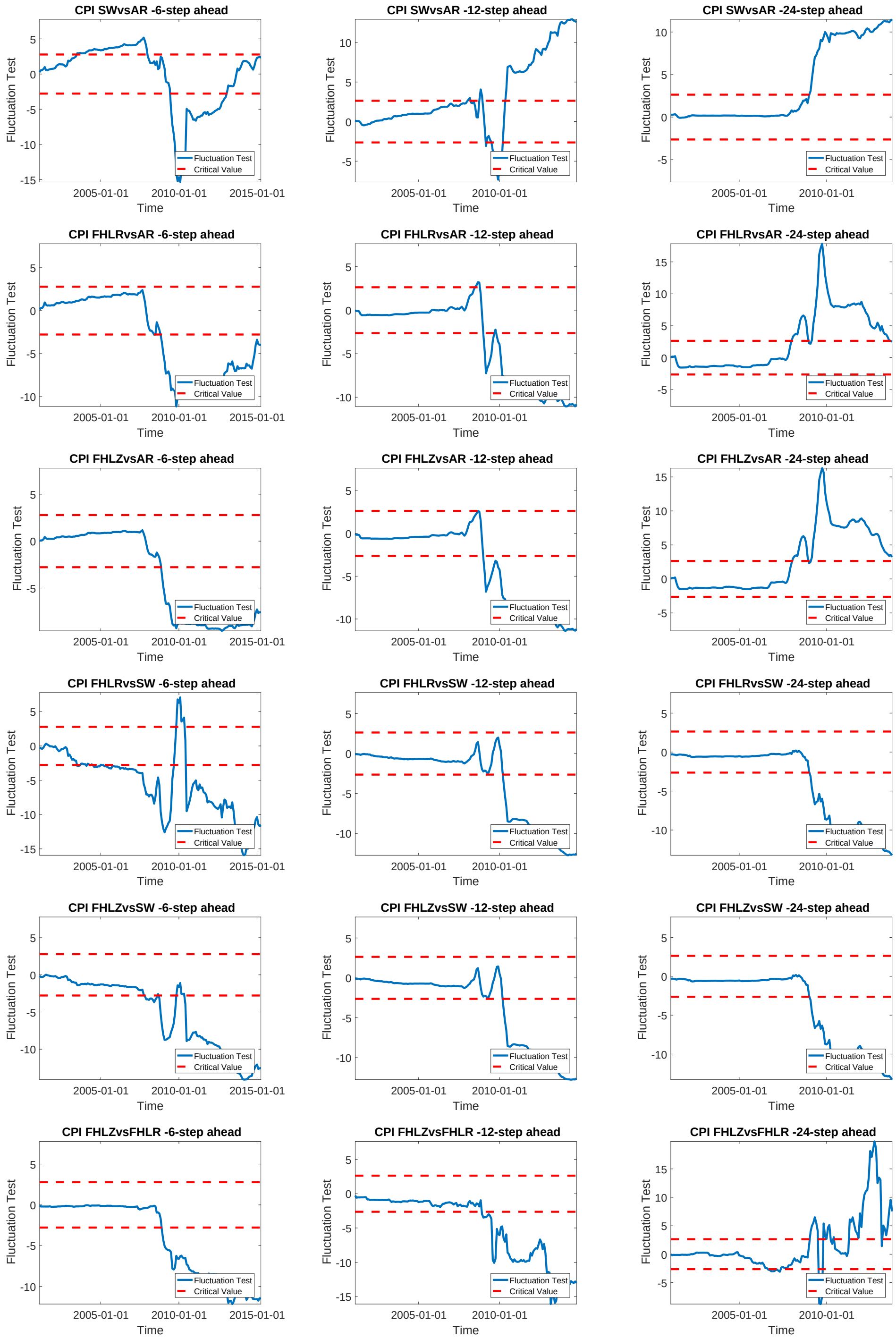

FIGURE 1.3: Fluctuation test for the CPI.

TABLE 1.4: Mean rMSFE for category.

SW					
	$h = 1$	$h = 3$	$h = 6$	$h = 12$	$h = 24$
Import-Export	1.20	0.94	1.00	0.99	1.04
Unemployment	1.28	1.17	1.16	1.10	1.00
Industrial Production	1.12	1.01	0.93	**0.92**	0.96
Prices	1.01	**0.97**	**0.96**	1.01	1.16
Wages	1.19	1.28	1.39	1.33	1.31
Surveys	1.12	1.09	1.06	**0.94**	**0.94**
Interest Rates	1.06	1.11	1.08	1.06	1.00
Stock Prices	1.12	1.10	1.07	1.01	1.03

FHLR					
	$h = 1$	$h = 3$	$h = 6$	$h = 12$	$h = 24$
Import-Export	1.18	**0.93**	0.97	0.99	1.00
Unemployment	1.29	1.17	1.15	1.06	0.99
Industrial Production	1.14	1.00	0.93	0.94	0.98
Prices	1.19	1.17	1.11	1.04	1.11
Wages	1.07	1.09	1.14	1.02	**0.98**
Surveys	1.14	1.12	1.04	1.01	1.09
Interest Rates	0.94	0.93	0.91	**0.89**	**0.84**
Stock Prices	1.01	1.01	**0.99**	1.01	**0.99**

FHLZ					
	$h = 1$	$h = 3$	$h = 6$	$h = 12$	$h = 24$
Import-Export	1.16	0.97	**0.96**	**0.97**	1.00
Unemployment	1.22	1.09	1.07	1.02	**0.97**
Industrial Production	1.11	**0.99**	**0.91**	**0.92**	0.98
Prices	1.11	1.11	1.10	1.10	**0.92**
Wages	1.36	1.49	1.55	1.48	1.57
Surveys	1.06	1.09	1.06	1.09	0.99
Interest Rates	**0.78**	**0.73**	**0.80**	0.91	1.08
Stock Prices	1.13	1.14	1.13	1.22	1.29

distributions of the rMSFEs of the dynamic models for each cathegories are reported. The configuration of the parameters is the one chosen in the calibration process for the forecasting of the IP. Similar results are obtained for the configurations adopted for the CPI and are not reported here.

TABLE 1.5: Distributions of the rMSFEs.

SW					
Percentile	0.05	0.25	0.50	0.75	0.95
$h = 1$	0.89	1.02	1.08	1.17	1.53
$h = 3$	0.83	0.97	1.05	1.11	1.41
$h = 6$	0.81	0.96	1.03	1.13	1.39
$h = 12$	0.80	0.92	0.99	1.10	1.38
$h = 24$	0.82	0.94	1.03	1.14	1.39

FHLR					
Percentile	0.05	0.25	0.50	0.75	0.95
$h = 1$	0.91	0.98	1.04	1.14	1.48
$h = 3$	0.83	0.94	1.02	1.07	1.30
$h = 6$	0.82	0.95	1.01	1.07	1.30
$h = 12$	0.85	0.96	1.01	1.05	1.25
$h = 24$	0.88	0.96	1.00	1.05	1.19

FHLZ					
Percentile	0.05	0.25	0.50	0.75	0.95
$h = 1$	0.91	0.98	1.02	1.10	1.43
$h = 3$	0.85	0.95	0.99	1.07	1.28
$h = 6$	0.80	0.94	0.97	1.03	1.23
$h = 12$	0.83	0.93	0.97	1.01	1.17
$h = 24$	0.86	0.95	0.98	1.04	1.16

FHLZ is the method which shows better performances for at least half of the series. More precisely, FHLZ is as accurate as AR till half of the series at horizon $h = 1$. At the other horizons, it is as accurate as AR till the 75-th percentile. SW is outperformed by frequency domain methods at all horizons and at almost all percentiles. Within frequency domain methods, FHLZ performs the best at 5-th and 95-th percentile.

1.6 Conclusions

In this chapter, the forecasting performances of SW, FHLR and FHLZ are compared on a EU macroeconomic dataset which spans from February 1986 to November 2015. The EU dataset is split into two parts: the first is called *calibration set* and is used to tune the parameters of the dynamic factor models. The second is called *proper sample* and is used to compare the forecasting performances of the three dynamic factor models. In the proper sample, all methods outperform AR after the first crisis at horizons $h \neq 24$. Instead, at horizon $h = 24$, all methods tend to lose ground against AR after the first crisis. Moreover, FHLZ generally outperforms all methods in forecasting of the CPI. Instead, no method seems to outperform the others in forecasting the IP, but all dynamic factor models outperform the benchmark AR. In the forecasting the whole dataset exercise, FHLZ is the most performing method in predicting real variables, whereas no significant evidences appear on the prediction of nominal variables. However, apart from Interest Rates, all methods seem to perform poorly in comparison with AR.

Chapter 2

A non-standard approach to the calibration of selected dynamic factor models in macroeconomic forecasting

2.1 Introduction

In this chapter, a comparative analysis of the forecasting performance of three Large-Dimensional Dynamic Factor Models is presented. As a key feature, Dynamic Factor Models represent each variable in a dataset as the sum of two orthogonal terms: a *common component* χ_t, driven by a reduced (as compared to the number of series in the dataset) number of common factors, and an *idiosyncratic component* ξ_t, which represents measurement errors or local features. Among the different versions of the Dynamic Factor Models we selected the same ones that we have described in Chpater 1. Below, some insights of these methods are reportd:

(i) *SW model*. This time-domain method was introduced in Stock and M. W. Watson, 2002a, Stock and M. W. Watson, 2002b. The factors are estimated by computing static principal components of the variables in the dataset. Let y_{it} be the variable of the dataset to be forecasted at time t, its h-step-ahead prediction equation (also called *Diffusion Forecast Index*) is obtained by regressing y_{it+h} on the factors and on y_{it} itself. Lags of the factors and of y_{it} may be added.

(ii) *FHLR model*. This frequency-domain method was proposed in Forni, Hallin, Lippi, and Reichlin, 2000, Forni, Hallin, Lippi, and Reichlin, 2005 and requires the computation of two steps. In a first step, the common component χ_t, the idiosyncratic component ξ_t and their covariances are estimated using a frequency-domain method introduced in Forni, Hallin, Lippi, and Reichlin, 2000 named *Dynamic Principal Component*. In the second step, the factors are estimated by computing Generalized Principal Components.

(iii) *FHLZ model*. This frequency-domain method was proposed in Forni, Hallin, Lippi, and Zaffaroni, 2015, Forni, Hallin, Lippi, and Zaffaroni, 2016. Here, the underlying assumption in (i) and (ii) that the common components span a finite-dimensional space as n tends to infinity is relaxed.

There exists some literature comparing the forecasting performances of SW and FHLR, but universal consensus still does not seem to have been reached. Theoretically, time-domain methods consider only relations among the variables at the same time, whereas frequency-domain methods exploit leaded and lagged relations among the variables. However time-domain methods require less parameters to be calibrated. Hence they are more robust to misspecification than frequency-domain methods. Instead, a systematic comparison of the forecasting performances of SW, FHLR and FHLZ can be found only in Forni, Giovannelli, et al., 2016, Della Marra, 2017. Forni, Giovannelli, et al., 2016 conducted a forecasting exercise on a US macroeconomic dataset, taking an autoregressive process of order 4 as a benchmark. They showed that FHLZ oupterforms SW, FHLR and the benchmark both for Industrial Production and Inflation during the Great Moderation. In the Great Recession, the forecasting performances of the Industrial Production change dramatically: all factor models are outperformed by the benchmark. SW and FHLR outperform FHLZ. Hence, Forni et al. concluded that, due to its more dynamical structure, FHLZ tends to be the best performing method in "stationary periods", but it loses ground during regime changes. Also, they showed that FHLZ tends to be outperforming on nominal variables and FHLR on real variables. Della Marra, 2017 conducted a forecasting exercise on an EU macroeconomic dataset. The global settings of his exercise are basically the same as in Forni, Giovannelli, et al., 2016, but also the length of the rolling window is suboptimally selected during the calibration process. He found that, on the proper sample, FHLZ is the most performing for the Inflation. However, mixed evidences appear over the proper sample for the Industrial Production.

In this chapter, the EU dataset is the same employed in Della Marra, 2017. This dataset is split into two subsamples. The first, from February 1986 to December 2000, is used to calibrate the models, i.e. to produce in-sample forecasts of the variables of the EU dataset for several specifications of SW, FHLR and FHLZ. Since each method is characterized by several parameters, an exhaustive exploration of the parameter space would be computationally infeasbile. Hence, the innovative aspect of this paper is that we employ genetic algorithms to explore more efficiently the parameter space in the calibration sample of each dataset. Then, for each class of models, we select the specification which shows the minimum mean squared prediction error. These models are then run and compared in the remaining subsample, from January 2001 to November 2015. Instead, the US dataset employed in our exercise is accurately described in Forni, Giovannelli, et al., 2016. This dataset is split into two subsamples. The first, from March 1960 to December 1984, is used to calibrate the models. Then, for each class of models, we select the specification which shows the minimum mean squared prediction error. These models are then run and compared in the remaining subsample, from January 1985 to October 2014.

The chapter is structured as follows. In Section 2, the main features of genetic algorithms are presented. In Section 3, the calibration process of the models is described. In Section 4, the results achieved on the EU dataset are discussed. In Section 5, the results achieved on the US dataset are discussed. Section 6 concludes.

2.2 An overview on genetic algorithms

The motivating idea of genetic algorithms is to start with a population of binary strings, randomply sampled, which are morphed and recombined to produce new strings with "better" characteristics, i.e. lower values for a preselected information criterion. We start with an initial population represented by a $N \times m$ matrix made up of 0s and 1s. N represents the length of each string and m is the chosen size of the population. Following Kapetanios, 2007 and Kapetanios, Marcellino, and Papailias, 2014, we denote this population matrix by $\mathbf{P_0}$. The genetic algorithm involves defining a transition from $\mathbf{P_i}$ to $\mathbf{P_{i+1}}$. The algorithm could be described in the following steps:

1. For $\mathbf{P_i}$ create a $m \times 1$ "fitness" vector, $\mathbf{p_i}$, by calculating for each column of $\mathbf{P_i}$ its "fitness". The choice of the "fitness" function is completely open and depends on the problem. For our purposes it is the opposite of the information criterion. Normalise $\mathbf{p_i}$, such that its elements lie in $(0, 1)$ and add up to 1. Denote this vector by $\mathbf{p_i^\star}$. Treat $\mathbf{p_i^\star}$ as a vector of probabilities and resample m times out of $\mathbf{P_i}$ with replacement, using the vector $\mathbf{p_i^\star}$ as the probabilities with which each string will be sampled. So "fit" strings are more likely to be chosen. Denote the resampled population matrix by $\mathbf{P^1_{i+1}}$.

2. Perform cross over on $\mathbf{P^1_{i+1}}$. For cross over we do the following: Arrange all strings in $\mathbf{P^1_{i+1}}$, in pairs (assume that m is even) where the pairings are randomly drawn. Denote a generic pair by $(a_1^\alpha, a_2^\alpha, ..., a_N^\alpha)(a_1^\beta, a_2^\beta, ..., a_N^\beta)$. Choose a random integer between 2 and $N - 1$. Denote this by j. Replace the pair by the following pair: $(a_1^\alpha, a_2^\alpha, ..., a_j^\alpha, a_{j+1}^\beta, ..., a_N^\alpha)(a_1^\beta, a_2^\beta, ..., a_j^\beta, a_{j+1}^\alpha, ..., a_N^\beta)$. Perform cross over on each pair with probability p_c. Denote the new population by $\mathbf{P^2_{i+1}}$. Usually p_c is set to some number around $0.5 - 0.6$.

3. Perform mutation on $\mathbf{P^2_{i+1}}$. This amounts to flipping the bits (0 or 1) of $\mathbf{P^2_{i+1}}$ with probability $\mathbf{p_m}$. $\mathbf{p_m}$ is usually set to a small number, say 0.01. After mutation the resulting population is $\mathbf{P_{i+1}}$.

These steps are repeated a prespecified number of times. Each set of steps is referred to as generation in the genetic literature. If a string is to be chosen this is the one with maximum fitness. For every generation we store the identity of the string with maximum "fitness". Further, this string is allowed to remain intact for that generation. So it gets chosen with probability one in step 1 of the algorithm and does not undergo neither cross-over nor mutation. At the end of the algorithm the string with the lowest preselectd information criterion value over all members of the populations and all generations is chosen. There has been considerable work on the theoretical properties of genetic algorithms. Hartl and Belew, 1990 have shown that with probability approaching one, the population at the $n - th$ generation will contain the global maximum as $n \to +\infty$. Perhaps the most relevant result from that work is Theorem 4.1 of Hartl and Belew, 1990. This theorem states that as long as:

- the sequence of the maximum fitnesses in the population across generations is monotonically increasing;

- any point in the model space is reachable from any other point by means of mutation and cross-over in a finite number of steps;

then the global maximum will be attained as $n \to +\infty$. Both these conditions hold for the algorithm described above. The first condition holds by the requirement that the string with the maximum fitness is always kept intact in the population. The second condition holds since any string of finite length can be obtained from another by cross-over and mutation with non-zero probability in a finite number of steps. For more details on the theory of genetic algorithms see also Del Moral, Kallel, and Rowe, 2001 and in Morinaka et al., 2001.

2.3 Description of the two datasets and of the calibration process

Both dataset contain real variables (import/export price indexes, employment, Industrial Production) and nominal variables (money aggregates, consumer price indexes, wages), asset prices (stock prices and exchange rates) and surveys. To achieve stationarity, several series are deseasonalized and transformed. No treatment for outliers is applied. In addition to SW, FHLR, FHLZ, the forecasts of an autoregressive process (AR) are computed. The order p of the AR process is determined in the calibration process. As in Stock and M. W. Watson, 2002b, D'Agostino and Giannone, 2012, to assess the forecasting performances, the variables which are taken into account are the level of the logarithm of the Industrial Production (IP) and the yearly change of the logarithm of the Consumer Price Index (CPI). Forecasts are computed h-months ahead, with $h \in \{1, 3, 6, 12, 24\}$. For each methods, we employ a rolling-window scheme $[t-l, t]$, whose size l is determined in the calibration sample. To assess the forecasting performance of each model, the mean-squared forecasting error (MSFE) is employed as a metric.

Since each method is characterized by several parameters, an exhaustive exploration of the parameter space would be computationally infeasbile. Hence, we employ genetic algorithms to explore more efficiently the parameter space in the calibration sample of each dataset. At each epoque, the population of the genetic algorithm is a subset of the strings containing all the possibile configurations of the parameters. We set the fitness as the inverse of its MSFE. For each method, we iterate the genetic algorithm ten times on the calibration sample of the two datasets. The fitness of each individual is stored in a data structure. Eventually, for each method we select as the most performing configuration the one endowed with the greatest fitness. More precisely, we select the configuration with the highest fitness that has been assessed during each of the ten runs of the genetic algorithms, independently from the final solutions obtained at each run. The parameters of each run of the genetic algorithm are the following:

(i) *Population size of the genetic algorithm at each generation= 100;*

(ii) *Crossover fraction = 0.6;*

(iii) *Number of individuals who passes to the next generation = 25;*

(iv) *Mutation = Gaussian model (adds a random number chosen from a Gaussian distribution, to each entry of the parent vector).*

The stopping criteria of each run of the genetic algorithm are the following:

(i) *Maximum number of generations = 1000;*

(ii) *Maximum number of generations in which the difference between the average fitness is less than the threshold* 10^{-7} *= 5;*

(iii) *Fitness of an individual in the last generation tending to* ∞*.*

The same procedure is described in Kapetanios, 2007 and in Kapetanios, Marcellino, and Papailias, 2014 , but the main purpose of these articles is to achieve suboptimal variable selection in a regressive setting. The convergence of each run of the genetic algorithms is graphically shown by plotting the boxplot of the results.

2.4 Calibration

In the calibration procedure, the metrics employed to select the most performing configuration of the parameters for each method is the same as in the section 1.4 (i.e. the mean $rMSFE^h$ over all $h \in \{1, 3, 6, 12, 24\}$). In the EU dataset, the calibration sample, ranging from February 1986 to December 2000, will be used to calibrate the methods SW, FHLR, FHLZ and the benchmark. Namely, this portion of the dataset will be used to select the best performing specifications of each class of models. Instead, in the US dataset, the calibration sample will range from March 1960 to December 1984.

2.5 Results on the EU dataset

2.5.1 Calibration of SW

To produce forecastings by means of equation (1.2), the following parameters must be calibrated:

(i) *the number of static factors* r: ranging from 1 to 10. Also, a comparison with Bai & Ng criterium (BN) with maximum 12 factors has been made.

(ii) *the degree* α *of* $\mathbf{a}(L)$: ranging from 1 to 10.

(iii) *the degree* β *of* $b(L)$: ranging from 0 to 10.

(iv) *the size* l *of the rolling window*: ranging from 5 to 12 years.

After the ten runs of the genetic algorithms in the calibration process, the individual granted with the maximum fitness for the IP is the following:

$$(r, \alpha, \beta, l) = (5, 1, 0, 11). \tag{2.1}$$

Instead, the individual granted with the maximum fitness for the CPI is the following::

$$(r, \alpha, \beta, l) = (1, 0, 1, 7). \tag{2.2}$$

2.5.2 Calibration of FHLR

To produce forecastings by means of equation (1.5), the following parameters must be calibrated:

(i) *the number of static factors* r: ranging from 1 to 10. Also, a comparison with Bai & Ng criterium (BN) with maximum 12 factors has been carried out.

(ii) *the number of dynamic factors* q: ranging from 0 to 10. Also, a comparison with Hallin-Liska criterium (HL) with maximum 12 factors has been carried out.

(iii) *the type of kernel* k: ranging in the set {Triangular, Rectangular, Parzen, Gaussian, Exponential, Cosine, Tukey, Hann}.

(iv) *The lag window* d *for spectral density estimation*: ranging in the set {25, 35, 40}.

(v) *the size* l *of the rolling window*: ranging from 5 to 12 years.

After the ten runs of the genetic algorithms in the calibration process, the individual granted with the maximum fitness for the IP is the following:

$$(r, q, k, d, l) = (10, 3, Cosine, 35, 11). \tag{2.3}$$

Instead, the individual granted with the maximum fitness for the CPI is the following:

$$(r, q, k, d, l) = (8, 4, Hann, 25, 7). \tag{2.4}$$

2.5.3 Calibration of FHLZ

To produce forecastings by means of equation (1.7), the following parameters must be calibrated:

(i) *the number of dynamic factors* q: ranging from 1 to 5. Also, a comparison with Hallin-Liska criterium has been carried out.

(ii) *the type of kernel* k: ranging in the set {Triangular, Rectangular, Parzen, Gaussian, Exponential, Cosine, Tukey, Hann}.

(iii) *the lag window* d *for spectral density estimation*: ranging in the set {25, 35, 40}.

(iv) *the maximum lag* ml *for the matrix* $\mathbf{A}^k(L)$: ranging from 1 to 5.

(v) *the size* l *of the rolling window*: ranging from 5 to 12 years.

After the ten runs of the genetic algorithms in the calibration process, the individual granted with the maximum fitness for the IP is the following:

$$(q, k, d, ml, l) = (4, Parzen, 25, 4, 11). \tag{2.5}$$

Instead, the individual granted with the maximum fitness for the CPI is the following:

$$(q, k, d, ml, l) = (2, Parzen, 35, 1, 7). \tag{2.6}$$

2.5.4 Calibration of the benchmark

To calibrate the benchmark $AR(p)$, the only parameter that needs to be fixed is the order p. In our exercise, we let p range from 1 to 13. By selecting the values of the parameter p which guarantee the lowest mean rMSFE, the chosen configuration for the IP is the following:

$$p = 2. \tag{2.7}$$

Instead, the chosen configuration for the CPI is the following:

$$p = 1. \tag{2.8}$$

2.5.5 Empirical proof of the convergence of the runs of the genetic algorithm

To give an empirical proof of the convergence of the genetic algorithm, in figure 2.1 the boxplots of the results of the ten runs of each selected dynamic factor models for the IP are reported.

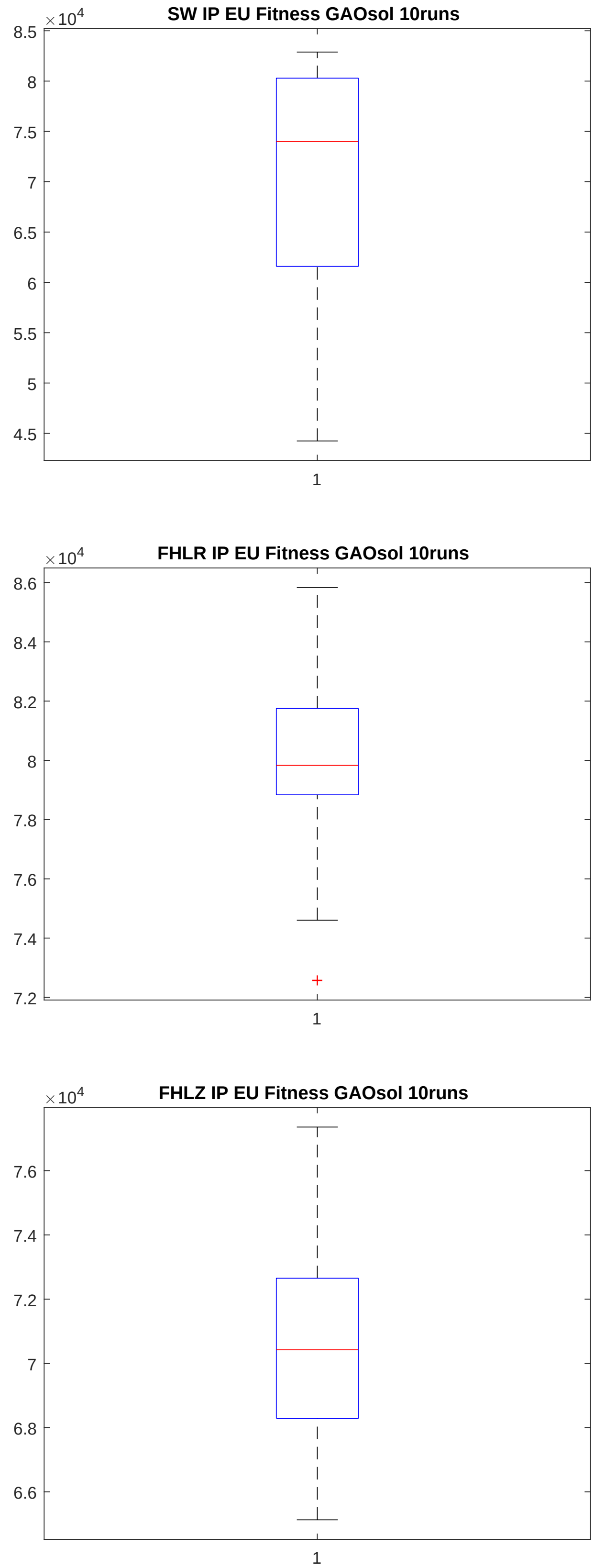

FIGURE 2.1: Boxplot of the results of the ten runs of the genetic algorithm for SW, FHLR and FHLZ over the IP.

Since the results achieved for all dynamic factor models span a narrow region, we can conclude that the ten runs of the genetic algorithms parametrized as in subsection 2.3 for all methods over IP have reached convergence. In addition, we can see that dynamic methods show better results since the ten runs span a narrower region than SW. In figure 2.2 the boxplots of the results of the ten runs of each selected dynamic factor models for the CPI are reported.

Since the results achieved for all dynamic factor models span a narrow region, we can conclude that the ten runs of the genetic algorithms parametrized as in subsection 2.3 for all methods over CPI have reached convergence. We can see that FHLZ show better results since its ten runs of the genetic algorithm span a narrower region than the other methods. In addition, the boxplot of FHLR covers a smaller region than SW.

2.5.6 Forecasting of the Industrial Production and the Inflation

Now, the forecasting performances of the three dynamic factor models over the IP and CPI are compared on the proper sample, which starts on January 2001 and ends on November 2015. The common benchmark for the factor models is the autoregressive process (AR) of order $p = 2$ for the IP and $p = 1$ for the CPI. In table 2.1, the average rMSFEs (relative to the AR) of the selected dynamic factor models are reported for the IP and CPI on the whole proper sample. However, as reported by CEPR, during the proper sample, the european economy faces two crisis periods: the first starts on May 2008 and ends on January 2009. The second starts on September 2011 and ends on March 2013. Hence, it is reasonable to assess whether the relative forecasting performances of the three dynamic factor models present a relevant change during the crisis periods. In table 2.2, the average rMSFEs (relative to the AR) of the three dynamic factor models are reported for the IP and the CPI from January 2001 to April 2008 (i.e. before the first crisis on the proper sample). In table 2.3, the average rMSFEs (relative to the AR) of the three dynamic factor models are reported for the IP and the CPI from January 2001 to August 2011 (i.e. before the second crisis on the evaluation sample). One, two or three asterisks indicate that the null of equal performance of the three factor models relative to AR is rejected at, respectively, the 1%, 5%, 10% significance by the Diebold-Mariano Test. One, two or three daggers indicate that the null of equal performance of FHLR, FHLZ relative to SW is rejected at, respectively, the 1%, 5%, 10% significance by the Giacomini-White Test. For further details about Giacomini-White Test, see Giacomini and White, 2006.

TABLE 2.1: rMSFEs on the whole sample: IP on the left, CPI on the right.

	IP			CPI		
$h/model$	$FHLZ$	$FHLR$	SW	$FHLZ$	$FHLR$	SW
1	0.89^{***}	0.83^{**}	0.85	0.97	0.94	1.03
3	0.96	0.88^{**}	0.92	0.93^{***}	1.01	1.00
6	0.88	0.87	0.93	0.92^{***}	0.98	1.01
12	0.89	0.90	0.91	0.95^{***}	1.01	0.97
24	0.96	0.88^{***}	0.97	0.95	0.91	0.98

The relative performances of all methods tend to improve especially after the first crisis at horizons $h \in \{1, 3, 6, 12\}$. In addition, after the first crisis such results tend to lose statistical significativity (in a slightly less severe way than in Della Marra,

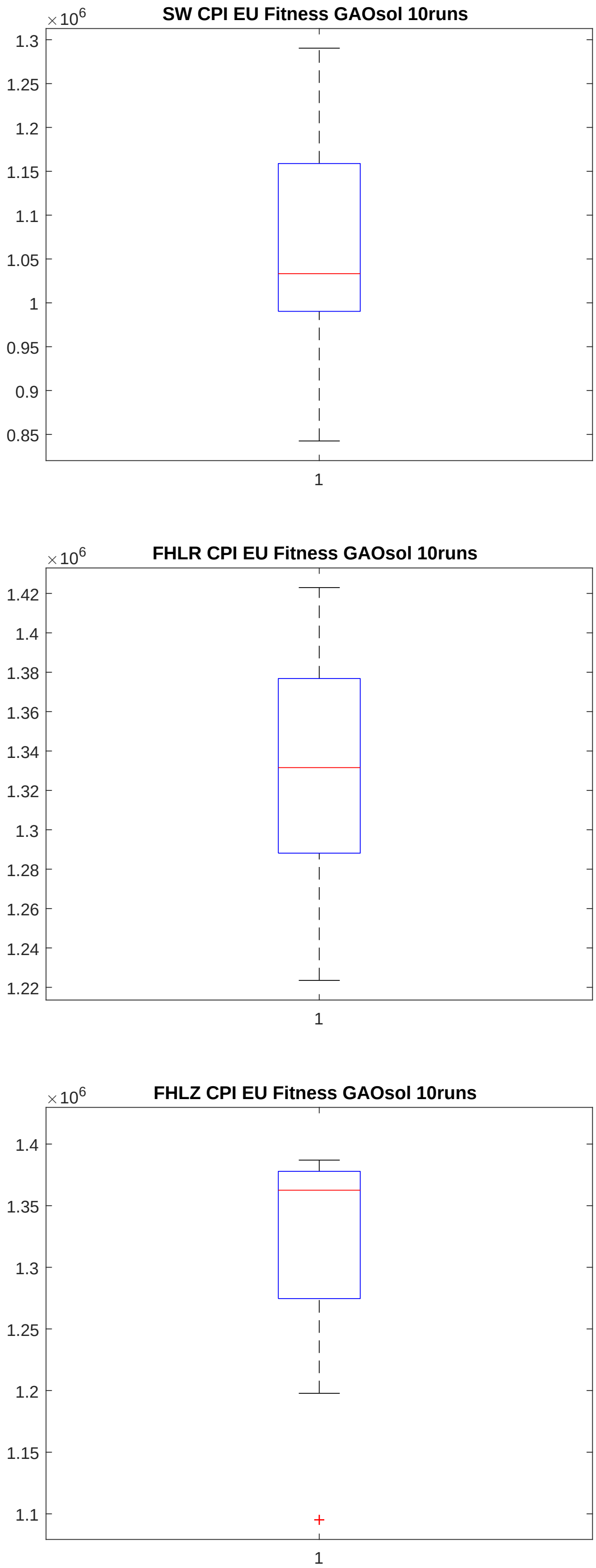

FIGURE 2.2: Boxplot of the results of the ten runs of the genetic algorithm for SW, FHLR and FHLZ over the CPI.

TABLE 2.2: rMSFEs from January 2001 to April 2008: IP on the left, CPI on the right.

	IP			CPI		
$h/model$	$FHLZ$	$FHLR$	SW	$FHLZ$	$FHLR$	SW
1	1.12	1.08	$1.25^{\dagger\dagger}$	$0.96^{\dagger\dagger}$	0.97	1.10
3	$0.81^{**\dagger\dagger}$	0.85	0.87	0.97	0.95	1.05
6	$0.83^{*\dagger\dagger\dagger}$	0.94	1.11	1.04	$1.09^{\dagger}$	1.07
12	$0.88^{\dagger\dagger}$	1.02	1.04	1.02	$1.09^{\dagger}$	1.08
24	0.91^{***}	0.84	1.02	$1.00^{\dagger}$	$0.91^{\dagger\dagger}$	1.04

TABLE 2.3: rMSFEs January 2001 to August 2011: IP on the left, CPI on the right.

	IP			CPI		
$h/model$	$FHLZ$	$FHLR$	SW	$FHLZ$	$FHLR$	SW
1	0.89	0.80^{**}	0.78^{***}	$0.93^{\dagger\dagger}$	0.89^{**}	1.03
3	0.98	$0.87^{**\dagger\dagger\dagger}$	0.88^{**}	0.92	0.97	0.97
6	0.89	0.86	0.91	0.91	0.93	0.98
12	0.89	0.88	0.87	0.94	0.96	0.92
24	0.97	0.89	0.96	1.17	1.03	1.10

2017 though). This holds both for IP and CPI. For the latter variable, the loss in statistical significativity is more pronounced. This is also illustrated in figure 2.3, in which the graphs of the cumulated sums of the squared forecasting error for the IP are reported. The shaded areas correspond to the two crisis periods in the proper sample, according to CEPR. A relevant gradient can be observed during the first crisis for all methods, including the benchmark, followed by a period of consecutive flatness. However, after the first crisis, the sum of the cumulative forecasting errors of AR increases substantially in comparison with those of the dynamic factor models at horizons $h \in (6, 12)$. Hence, this may explain why the performance of the benchmark tend to worsen in comparison of those of the selected dynamic factor models between the two crises. Instead, at horizon $h = 24$, the sum of the cumulative forecasts errors of AR tend to increase analogously in comparison of those of the factor models. This may be the reason why, for $h = 24$, the forecasting performances of dynamic factor models tend to slightly lose ground in comparison of that of the banchmark between the two crises. Similar results for the CPI are reported in figure 2.4.

As in Forni, Giovannelli, et al., 2016 and in Della Marra, 2017, to assess the forecasting performance of each couple of methods locally, each time series of the dataset is smoothed by a centered moving average of length $m = 61$ (with coefficients equal to $1/m$) and then the Fluctuation test is run, at 5% significance level. Further details about this test can be found in subsection 1.5.1 and in Giacomini and Rossi, 2010. The results for the IP at horizons $h \in \{6, 12, 24\}$ are reported in Figure 2.5. All factor models outeperform significantly the benchmark from the first crisis on at all horizons. SW tend to outperform the dynamic methods between the two crises. Instead, outside the period between the two crises, the dynamic methods show significantly better performances than SW. As to the performance of the dynamic methods, FHLR outperforms FHLZ between the two crises. To sum up, FHLR tend to outperform the other methods. But, this does not hold true in the period between the two crises, in which SW seems to be the most performing method. These results are similar

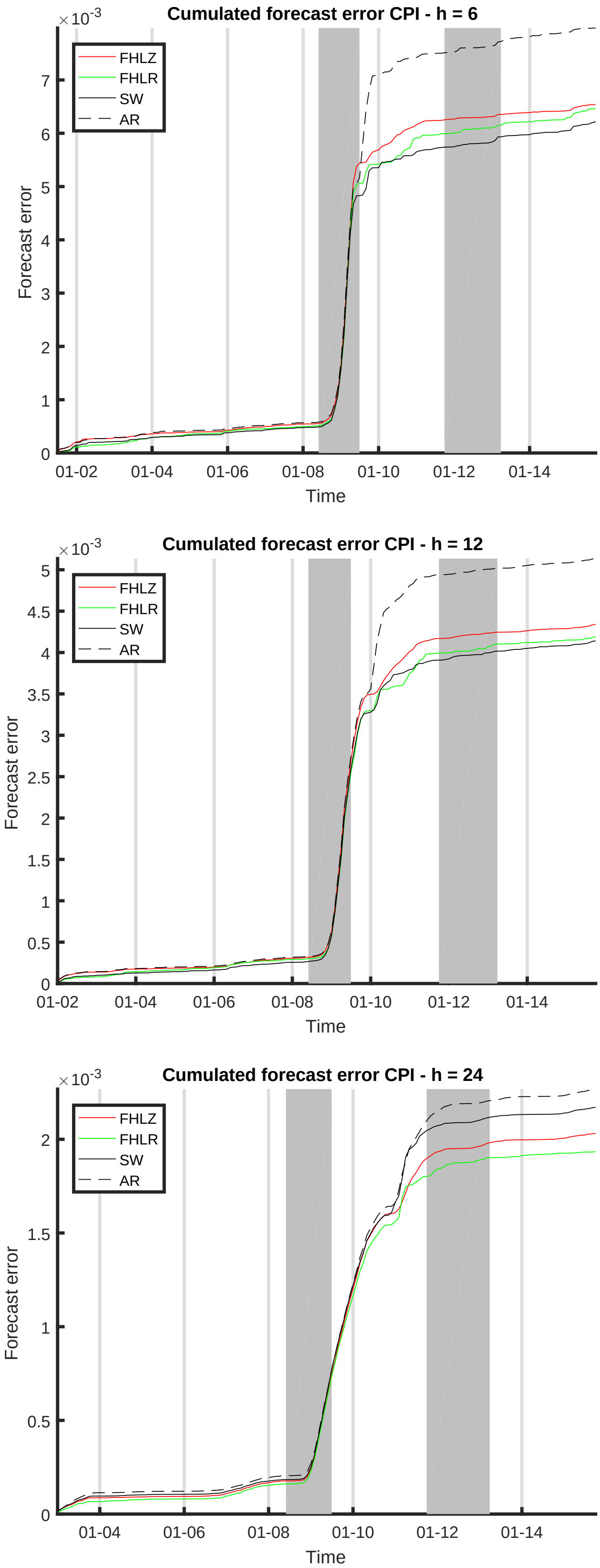

FIGURE 2.3: Cumulative sum of the squared forecasting error for the IP at horizons $h = 6$, $h = 12$ and $h = 24$.

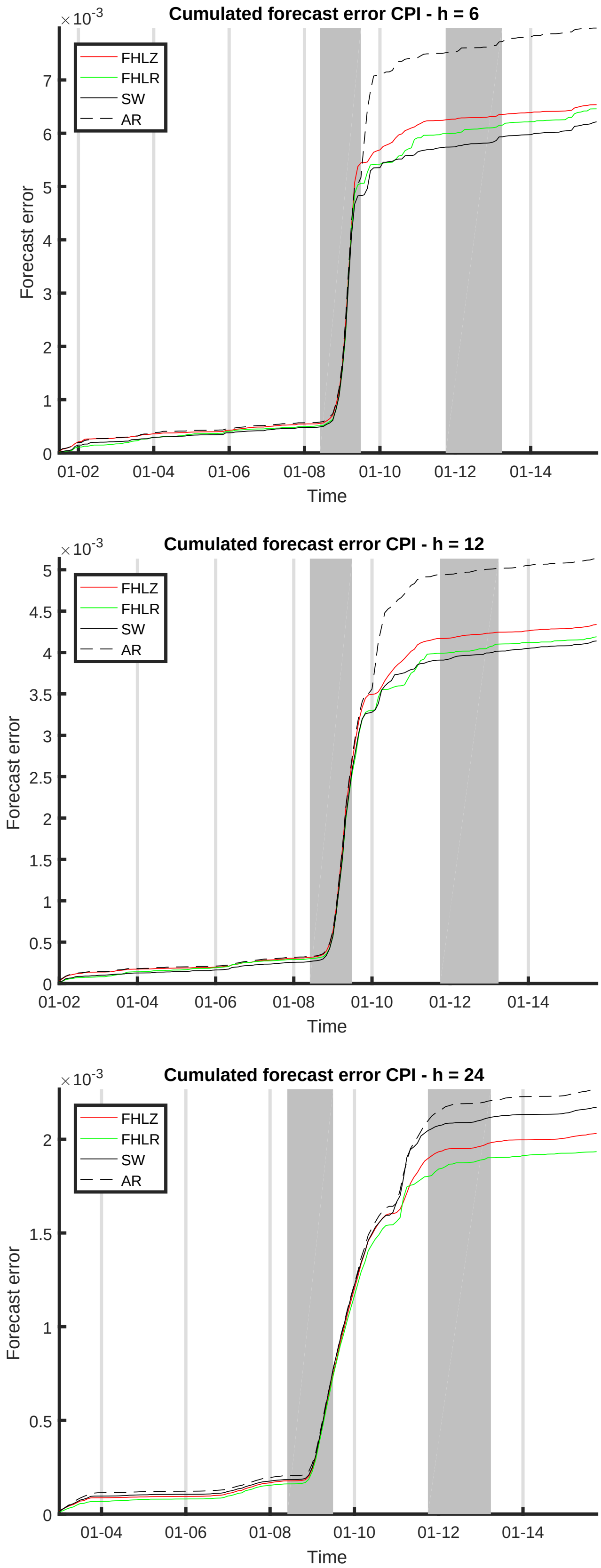

FIGURE 2.4: Cumulative sum of the squared forecasting error for the CPI at horizons $h = 6$, $h = 12$ and $h = 24$.

to those obtained in Della Marra, 2017, but in our exercise the relative perfomance of FHLR in comparison with SW are neater. The results for the CPI at horizons $h \in \{6, 12, 24\}$ are reported in Figure 2.6. All methods outperform AR significatively from the first crisis on. At horizon $h \in 12, 24$, FHLR and FHLZ outperform SW on average on the whole sample, except between the two crises. As to the comparison of dynamic methods,at horizons $h \in 6, 12$ FHLZ globally outperforms FHLR from the first crisis on. Instead, at horizon $h = 6$, FHLR globally outperforms FHLZ from the first crisis on. In comparison with Della Marra, 2017, FHLR shows slightly better forecasting performance in comparison with other methods. In addition, SW seems to be the most performing method between the two crises.

FHLR and FHLZ tend to outperfom SW at all horizons, except FHLR at horizon $h = 6$ during the first crisis. FHLR and FHLZ outperform AR at horizons $h \in \{6, 12\}$. At horizon $h = 24$, AR outperforms FHLR and FHLZ from the first crisis on. SW outperforms AR during the two crisis periods at horizons $h \in \{6, 12\}$. At horizons $h \in \{12, 24\}$, AR outperforms SW from the second crisis on. At all horizons, FHLZ outperforms FHLR during the first crisis. At horizons $h \in \{6, 12\}$, this behaviour seems to be persistent. Instead, at horizon $h = 24$, FHLR outperforms FHLZ from the second crisis on.

2.5.7 Forecasting of the dataset

As in Forni, Giovannelli, et al., 2016 and in Della Marra, 2017, this exercise has been extended to the other variables in the dataset. In table 2.4, we report the median rMSFEs of each class of real (the first three) and nominal (the others) time series, taking AR as a benchmark. Similar results are obtained for the mean and not reported here. The best performances are given in bold. We have excluded the categories ”Demand” and ”Money” since AR outperformed all factor models at any horizon.

As to the real variables, FHLR seems to be globally the most performing method. As in Della Marra, 2017,no methods outperforms the benchmark at horizon $h = 1$ or on the category ”Demand”. But, differently from Della Marra, 2017, FHLZ does not outperform FHLR. Instead, as to the nominal variables, FHLZ seems to be the most performing method (especially for the categories ”Interest Rates” and ”Stock Prices”), followed by FHLR (which is the most performing method for the category ”Survey”). To conclude, SW does not show outstanding performance in any category. In table 2.5, the distributions of the rMSFEs of the dynamic models for each cathegories are reported. The configuration of the parameters is the one chosen in the calibration process for the forecasting of the IP.

By analysing the distributions of the rMSFEs, it comes out that dynamic methods tend to outperform SW on real variables. FHLR outperforms FHLZ up to the 50-th percentile. In comparison with Della Marra, 2017, the forecasting performance of FHLR has improved remarkably. The results obtained for the configurations adopted for the CPI are reported in table 2.6.

By analysing the distributions of the rMSFEs, it comes out that dynamic methods tend to outperform SW on real variables. FHLZ globally outperforms FHLR, especially at long horizons ($h \geq 6$) . In comparison with Della Marra, 2017, the forecasting performance of FHLZ has improved remarkably.

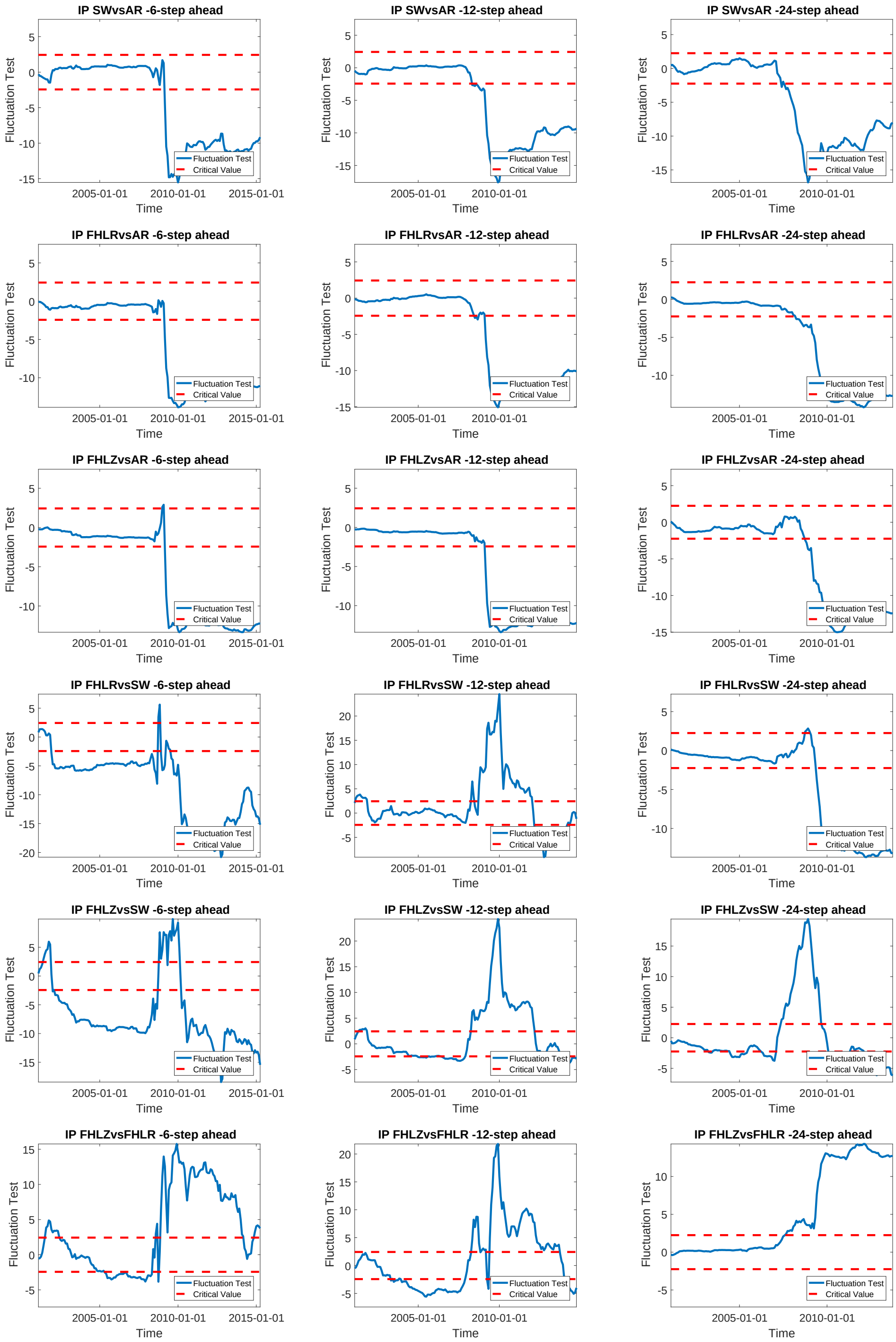

FIGURE 2.5: Fluctuation test for the IP.

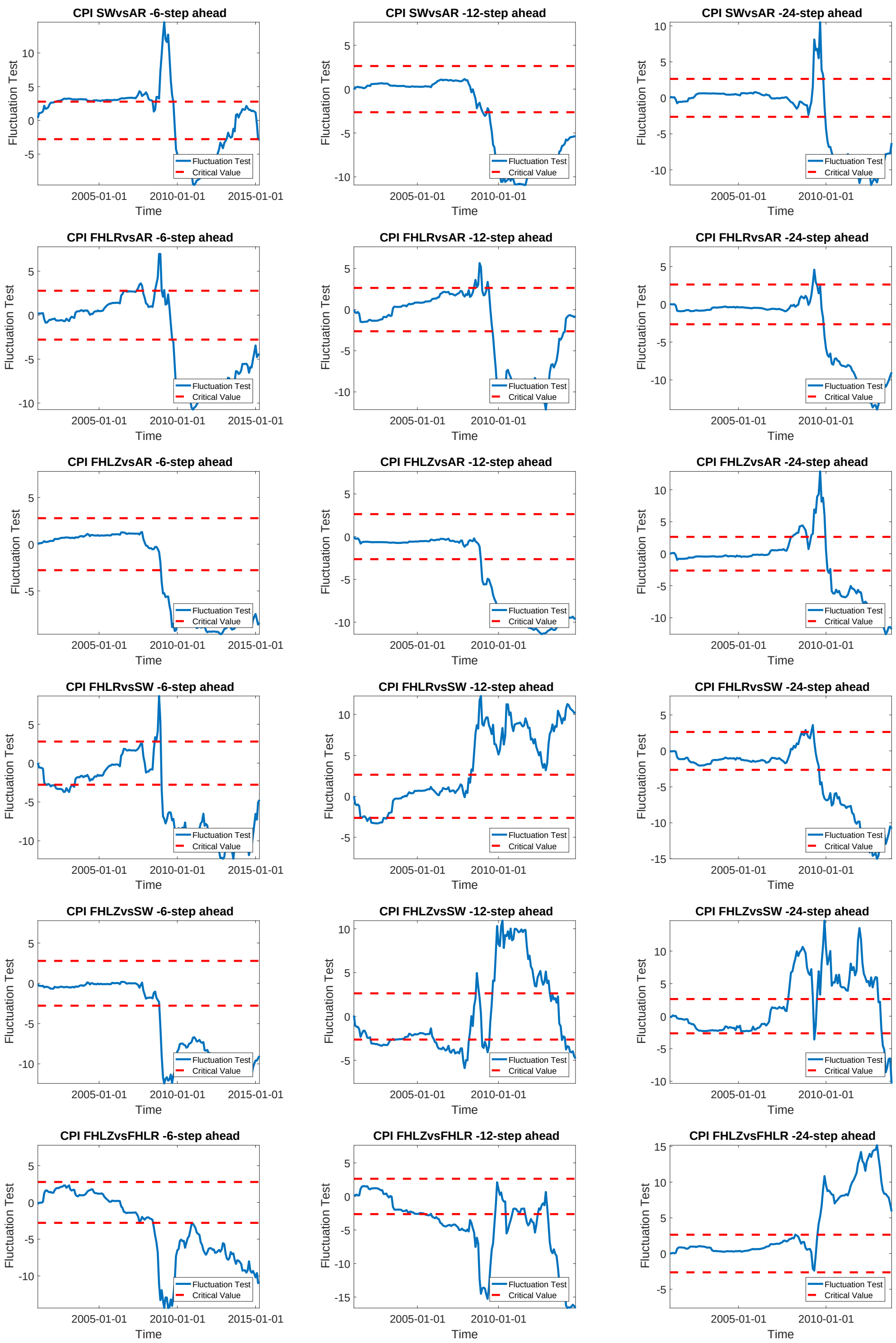

FIGURE 2.6: Fluctuation test for the CPI.

TABLE 2.4: Median rMSFE for category.

SW					
	$h = 1$	$h = 3$	$h = 6$	$h = 12$	$h = 24$
Import-Export	1.10	0.97	1.10	1.03	1.09
Unemployment	1.14	1.07	1.07	1.06	1.01
Industrial Production	1.05	1.03	1.02	1.00	1.02
Exchange Rates	1.13	1.08	1.06	1.05	1.21
Prices	1.03	1.01	1.02	0.97	1.04
Wages	1.03	1.08	1.01	1.06	1.08
Surveys	0.99	1.00	1.04	1.00	1.02
Interest Rates	1.02	1.08	1.04	0.98	0.98
Stock Prices	0.99	1.02	1.05	1.02	1.10

FHLR					
	$h = 1$	$h = 3$	$h = 6$	$h = 12$	$h = 24$
Import-Export	1.07	**0.92**	0.98	1.03	**0.95**
Unemployment	1.08	**0.98**	**0.94**	**0.99**	0.99
Industrial Production	**1.00**	**0.93**	**0.95**	**0.98**	**0.94**
Exchange Rates	1.04	1.14	1.20	1.13	1.06
Prices	1.01	1.03	**0.99**	0.99	**0.91**
Wages	1.09	1.13	1.16	1.10	1.03
Surveys	**0.97**	**0.94**	**0.97**	**0.94**	**0.91**
Interest Rates	1.02	1.07	1.03	1.00	0.93
Stock Prices	0.99	1.05	1.10	1.01	0.98

FHLZ					
	$h = 1$	$h = 3$	$h = 6$	$h = 12$	$h = 24$
Import-Export	1.08	0.95	**0.95**	**0.97**	1.00
Unemployment	1.09	1.00	1.00	1.00	**0.96**
Industrial Production	1.03	0.96	0.96	0.99	0.99
Exchange Rates	1.03	1.01	1.01	**0.98**	**0.96**
Prices	1.05	1.02	0.99	**0.96**	0.96
Wages	1.05	1.02	1.01	1.00	**1.00**
Surveys	0.99	1.02	1.00	0.99	0.98
Interest Rates	**0.98**	**0.97**	**0.96**	**0.96**	**0.93**
Stock Prices	**0.98**	**0.99**	**0.98**	**0.98**	**0.98**

TABLE 2.5: Distributions of the rMSFEs: configuration for the IP.

SW					
Percentile	0.05	0.25	0.50	0.75	0.95
$h = 1$	0.91	1.02	1.11	1.18	1.49
$h = 3$	0.86	1.00	1.07	1.15	1.43
$h = 6$	0.83	0.99	1.07	1.13	1.41
$h = 12$	0.84	0.97	1.05	1.12	1.35
$h = 24$	0.91	1.00	1.08	1.16	1.39

FHLR					
Percentile	0.05	0.25	0.50	0.75	0.95
$h = 1$	0.84	0.96	1.03	1.09	1.42
$h = 3$	0.82	0.93	0.99	1.08	1.32
$h = 6$	0.81	0.92	0.98	1.09	1.32
$h = 12$	0.85	0.93	1.00	1.10	1.34
$h = 24$	0.78	0.90	0.96	1.05	1.20

FHLZ					
Percentile	0.05	0.25	0.50	0.75	0.95
$h = 1$	0.90	0.99	1.02	1.10	1.40
$h = 3$	0.90	0.97	1.00	1.06	1.25
$h = 6$	0.88	0.96	0.98	1.03	1.16
$h = 12$	0.89	0.96	0.98	1.01	1.12
$h = 24$	0.92	0.97	0.99	1.02	1.13

TABLE 2.6: Distributions of the rMSFEs: configuration for the CPI.

SW					
Percentile	0.05	0.25	0.50	0.75	0.95
$h = 1$	0.89	0.98	1.03	1.11	1.23
$h = 3$	0.83	0.98	1.02	1.08	1.17
$h = 6$	0.89	0.97	1.03	1.07	1.14
$h = 12$	0.87	0.96	1.01	1.04	1.12
$h = 24$	0.95	1.00	1.05	1.12	1.25

FHLR					
Percentile	0.05	0.25	0.50	0.75	0.95
$h = 1$	0.84	0.97	1.03	1.10	1.31
$h = 3$	0.86	0.98	1.05	1.11	1.26
$h = 6$	0.86	0.96	1.04	1.12	1.24
$h = 12$	0.88	0.94	1.01	1.06	1.18
$h = 24$	0.81	0.92	0.96	1.01	1.10

FHLZ					
Percentile	0.05	0.25	0.50	0.75	0.95
$h = 1$	0.93	0.98	1.02	1.13	1.36
$h = 3$	0.91	0.97	1.00	1.05	1.23
$h = 6$	0.92	0.97	0.99	1.02	1.15
$h = 12$	0.01	0.96	0.98	1.00	1.08
$h = 24$	0.90	0.96	0.98	1.00	1.08

2.6 Results on the US dataset

2.6.1 Calibration of SW

To produce forecastings by means of equation (1.2), the following parameters must be calibrated:

(i) *the number of static factors r*: ranging from 1 to 10. Also, a comparison with Bai & Ng criterium (BN) with maximum 12 factors has been made.

(ii) *the degree α of* $\mathbf{a}(L)$: ranging from 1 to 10.

(iii) *the degree β of* $b(L)$: ranging from 0 to 10.

(iv) *the size l of the rolling window*: ranging from 5 to 12 years.

By selecting the values of the parameters which guarantee the lowest mean rMSFE, the chosen configuration for the IP is the following:

$$(r, \alpha, \beta, l) = (BN, 0, 0, 12). \tag{2.9}$$

Instead, the chosen configuration for the CPI is the following:

$$(r, \alpha, \beta, l) = (3, 1, 10, 15). \tag{2.10}$$

2.6.2 Calibration of FHLR

To produce forecastings by means of equation (1.5), the following parameters must be calibrated:

(i) *the number of static factors r*: ranging from 1 to 10. Also, a comparison with Bai & Ng criterium (BN) with maximum 12 factors has been carried out.

(ii) *the number of dynamic factors q*: ranging from 0 to 10. Also, a comparison with Hallin-Liska criterium (HL) with maximum 12 factors has been carried out.

(iii) *the type of kernel k*: ranging in the set {Triangular, Rectangular, Parzen, Gaussian, Exponential, Cosine, Tukey, Hann}.

(iv) *The lag window d for spectral density estimation*: ranging in the set {25, 35, 40}.

(v) *the size l of the rolling window*: ranging from 5 to 12 years.

By selecting the values of the parameters which guarantee the lowest mean rMSFE, the chosen configuration for the IP is the following:

$$(r, q, k, d, l) = (9, 2, Exponential, 40, 12). \tag{2.11}$$

Instead, the chosen configuration for the CPI is the following:

$$(r, q, k, d, l) = (6, 1, Hann, 25, 15). \tag{2.12}$$

2.6.3 Calibration of FHLZ

To produce forecastings by means of equation (1.7), the following parameters must be calibrated:

(i) *the number of dynamic factors q*: ranging from 1 to 5. Also, a comparison with Hallin-Liska criterium has been carried out.

(ii) *the type of kernel k*: ranging in the set {Triangular, Rectangular, Parzen, Gaussian, Exponential, Cosine, Tukey, Hann}.

(iii) *the lag window d for spectral density estimation*: ranging in the set {25, 35, 40}.

(iv) *the maximum lag ml for the matrix $\mathbf{A}^k(L)$*: ranging from 1 to 5.

(v) *the size l of the rolling window*: ranging from 5 to 12 years.

By selecting the values of the parameters which guarantee the lowest mean rMSFE, the chosen configuration for the IP is the following:

$$(q, k, d, ml, l) = (5, Triangular, 40, 2, 12). \tag{2.13}$$

Instead, the chosen configuration for the CPI is the following:

$$(q, k, d, ml, l) = (5, Triangular, 25, 5, 15). \tag{2.14}$$

2.6.4 Calibration of the benchmark

To calibrate the benchmark $AR(p)$, the only parameter that needs to be fixed is the order p. In our exercise, we let p range from 1 to 13. By selecting the values of the parameter p which guarantee the lowest mean rMSFE, the chosen configuration for the IP is the following:

$$p = 2 \tag{2.15}$$

Instead, the chosen configuration for the CPI is the following:

$$p = 9 \tag{2.16}$$

2.6.5 Empirical proof of the convergence of the runs of the genetic algorithm

To give an empirical proof of the convergence of the genetic algorithm, in figure 2.7 the boxplots of the results of the ten runs of each selected dynamic factor models for the IP are reported.

Since the results achieved for all dynamic factor models span a narrow region, we can conclude that the ten runs of the genetic algorithms parametrized as in subsection 2.4 for all methods over IP have reached convergence. We can see that FHLR show better results since its ten runs of the genetic algorithm span a narrower region than the other methods. In addition, the boxplot of FHLZ covers a slightly smaller region than SW. In figure 2.8 the boxplots of the results of the ten runs of each selected dynamic factor models for the CPI are reported.

Since the results achieved for all dynamic factor models span a narrow region, we can conclude that the ten runs of the genetic algorithms parametrized as in subsection 2.4 for all methods over CPI have reached convergence. We can see that

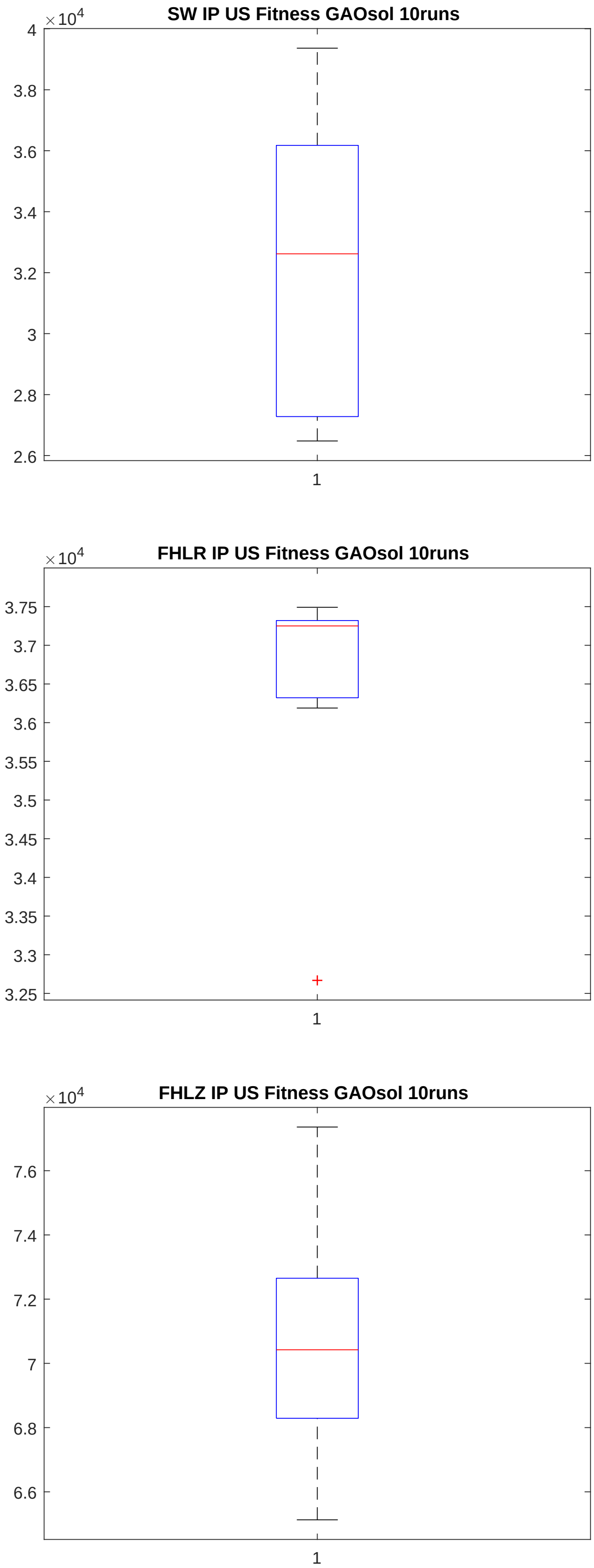

FIGURE 2.7: Boxplot of the results of the ten runs of the genetic algorithm for SW, FHLR and FHLZ over the IP.

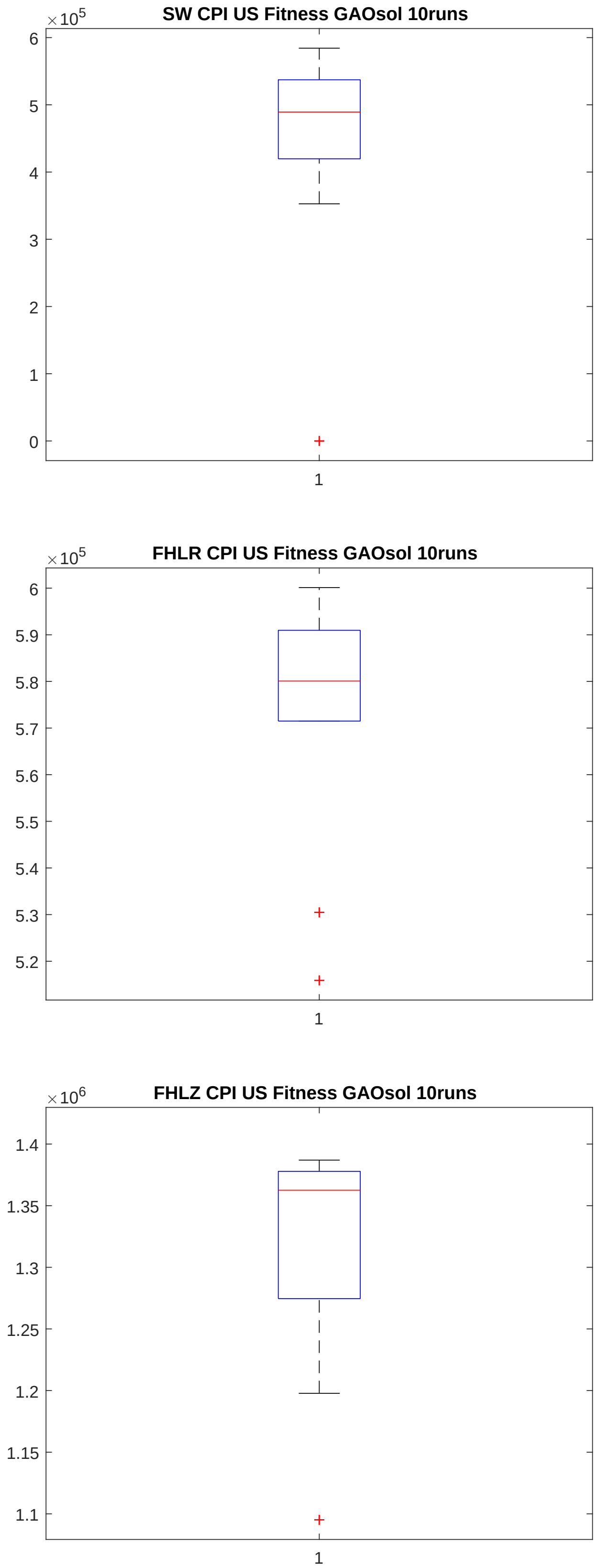

FIGURE 2.8: Boxplot of the results of the ten runs of the genetic algorithm for SW, FHLR and FHLZ over the CPI.

FHLZ show better results since its ten runs of the genetic algorithm span a narrower region than the other methods. In addition, the boxplot of FHLR covers a smaller region than SW.

2.6.6 Forecasting of the Industrial Production and the Inflation

Now, the forecasting performances of the three dynamic factor models over the IP and CPI are compared on the proper sample, which starts on March 1960 and ends on October 2014. The common benchmark for the factor models is the autoregressive process (AR) of order $p = 2$ for the IP and $p = 9$ for the CPI. In table 2.7, the average rMSFEs (relative to the AR) of the selected dynamic factor models are reported for the IP and CPI on the whole proper sample. However, as reported by NBER, during the proper sample, the american economy faces a crisis period which starts on December 2007 and ends on June 2009. Hence, it is reasonable to assess whether the relative forecasting performances of the three dynamic factor models present a relevant change during the crisis period. In table 2.8, the average rMSFEs (relative to the AR) of the three dynamic factor models are reported for the IP and the CPI from March 1960 to Nove,ber 2007 (i.e. before the crisis on the proper sample). One, two or three asterisks indicate that the null of equal performance of the three factor models relative to AR is rejected at, respectively, the 1%, 5%, 10% significance by the Diebold-Mariano Test. One, two or three daggers indicate that the null of equal performance of FHLR, FHLZ relative to SW is rejected at, respectively, the 1%, 5%, 10% significance by the Giacomini-White Test. For further details about Giacomini-White Test, see Giacomini and White, 2006.

TABLE 2.7: rMSFEs on the whole sample: IP on the left, CPI on the right.

	IP			CPI		
$h/model$	$FHLZ$	$FHLR$	SW	$FHLZ$	$FHLR$	SW
1	0.90^{**}	0.90^{*}	0.88^{**}	$0.91^{*\dagger\dagger\dagger}$	1.00	1.05
3	1.07	1.00	0.90	0.83^{**}	0.88	1.00
6	1.11	1.11	1.09	0.94	0.98	1.17
12	1.02	1.13	1.08	0.98	1.00	1.13
24	0.99	$1.21^{\dagger\dagger\dagger}$	$1.18^{\dagger\dagger\dagger}$	1.09	1.00	1.38

TABLE 2.8: rMSFEs from January 2001 to April 2008: IP on the left, CPI on the right.

	IP			CPI		
$h/model$	$FHLZ$	$FHLR$	SW	$FHLZ$	$FHLR$	SW
1	0.93^{**}	$0.88?* \dagger \dagger$	0.93^{***}	$0.90^{*\dagger}$	1.06	0.96
3	0.94	0.94	1.05	$0.94^{\dagger\dagger\dagger}$	0.95	1.11
6	0.95	1.05	1.24	0.89^{***}	0.90	1.09
12	0.97	1.02	$1.23^{\dagger\dagger\dagger}$	$0.79^{**\dagger\dagger}$	0.77^{***}	$1.01^{\dagger\dagger\dagger}$
24	0.99	1.01	$1.22^{\dagger\dagger}$	0.87^{***}	0.69^{**}	0.82

The relative performances of the dynamic methods tend to improve especially after the first crisis at all horizons both or the IP and the CPI. However, this does not hold for SW in the IP scenario. In addition, as for the IP, the loss of statistical significativity is less pronounced than in Forni, Giovannelli, et al., 2016. Instead, as for the CPI, the loss of statistical significativity is more pronounced than in Forni, Giovannelli, et al., 2016. This is also illustrated in figure 2.9, in which the graphs of the cumulated sums of the squared forecasting error for the IP are reported. The shaded area corresponds to the crisis period in the proper sample, according to NBER. A relevant gradient can be observed from the Great Recession on for all methods, including the benchmark, followed by a period of consecutive flatness. However, from the Great Recession on, the sum of the cumulative forecasting errors of AR increases substantially in comparison with those of the dynamic factor models at horizons $h \in (6, 12)$. Instead, at $h = 24$, this does not hold only for FHLZ. This may be the reason why, for $h = 24$, the forecasting performances of dynamic factor models tend to slightly lose ground in comparison of that of the banchmark between the two crises. Similar results for the CPI are reported in figure 2.10.

As in Forni, Giovannelli, et al., 2016 and in Della Marra, 2017, to assess the forecasting performance of each couple of methods locally, each time series of the dataset is smoothed by a centered moving average of length $m = 61$ (with coefficients equal to $1/m$) and then the Fluctuation test is run, at 5% significance level. Further details about this test can be found in subsection 1.5.1 and in Giacomini and Rossi, 2010 The results for the IP at horizons $h \in \{6, 12, 24\}$ are reported in Figure 2.11. The benchmark globally shows significantly better results than the factor models from the Great Recession on. However, this does not hold for SW at horizons $h \in 6, 12$ and for FHLZ at horizon $h = 24$ during the Great Recession, as can be seen in figure 2.9. SW tends to outperform the dynamic methods from the Great Recession on, apart from FHLZ at horizon $h = 24$. As in Forni, Giovannelli, et al., 2016, FHLR outeperforms FHLZ from the Great Recession on. The results for the CPI at horizons $h \in \{6, 12, 24\}$ are reported in Figure 2.12. No factor model seems to outperform the benchmark from the Great Recession on. Instead, before the Great Recession, the contrary seems to hold. Dynamic methods show significant better performances than SW at all horizons, except for $h = 12$. FHLZ outperforms FHLR outside the Great Recession. Apart from this period, mixed evidences appear as far as the comparison between dynamic methods is concerned. Hence, as to the perfomances of dynamic methods, we can draw less clear conclusions than in Forni, Giovannelli, et al., 2016.

2.6.7 Forecasting of the dataset

As in Forni, Giovannelli, et al., 2016 and in Della Marra, 2017, this exercise has been extended to the other variables in the dataset. In table 2.9, we report the median rMSFEs of each class of real (the first three) and nominal (the others) time series, taking AR as a benchmark. Similar results are obtained for the mean and not reported here. The best performances are given in bold.

As to the real variables, FHLR seems to be globally the most performing method, as in Forni, Giovannelli, et al., 2016. We also notice that FHLR shows better performance over the category ”Housing’ (differently from Forni, Giovannelli, et al., 2016, in which FHLR shows better performances over the categories ”Unemployment” and ”Inventories”). Instead, as to the nominal variables, FHLZ seems to be the most performing method (especially for the categories ”Interest Rates” and ”Stock Prices”, as in Forni, Giovannelli, et al., 2016), followed by FHLR (which is the most performing method for the category ”Money”). To conclude, SW does not show outstanding performance in any category, as in Forni, Giovannelli, et al., 2016. It is interesting to notice that, differently from Forni, Giovannelli, et al., 2016, no categories are excluded from our exercise since genetically calibrated factor models outperform the

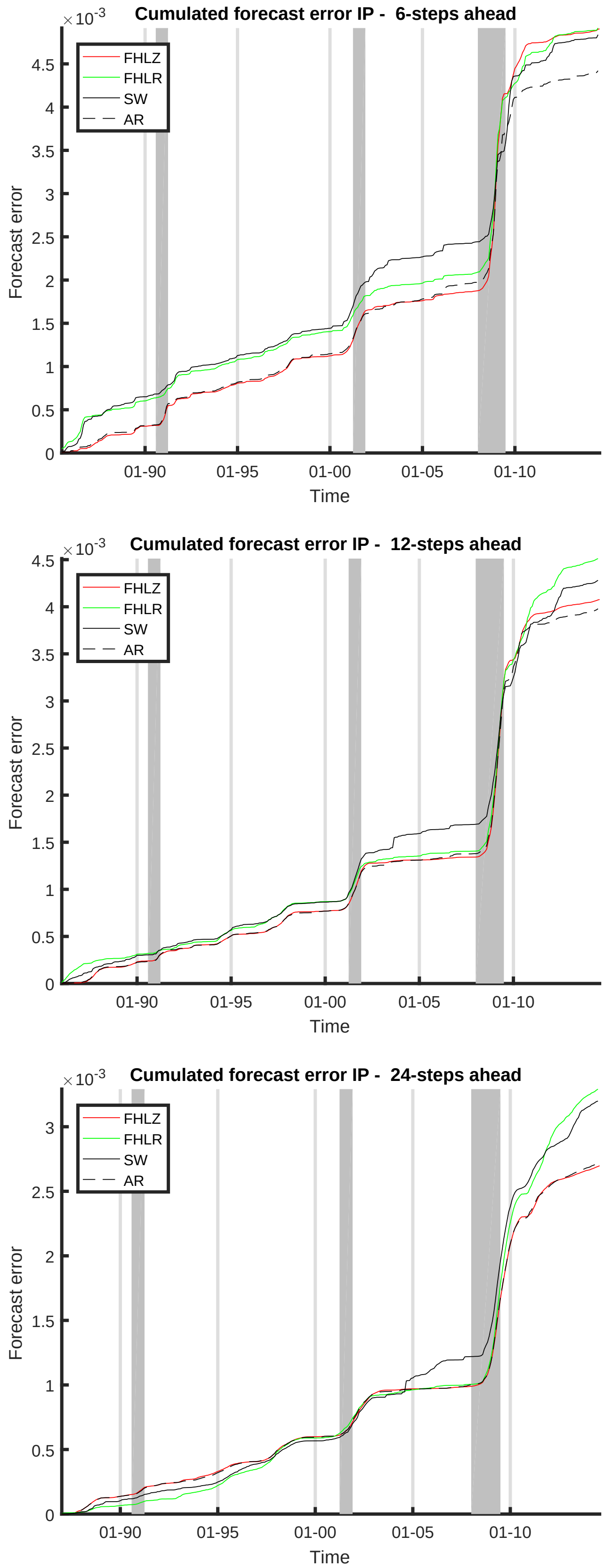

FIGURE 2.9: Cumulative sum of the squared forecasting error for the IP at horizons $h = 6$, $h = 12$ and $h = 24$.

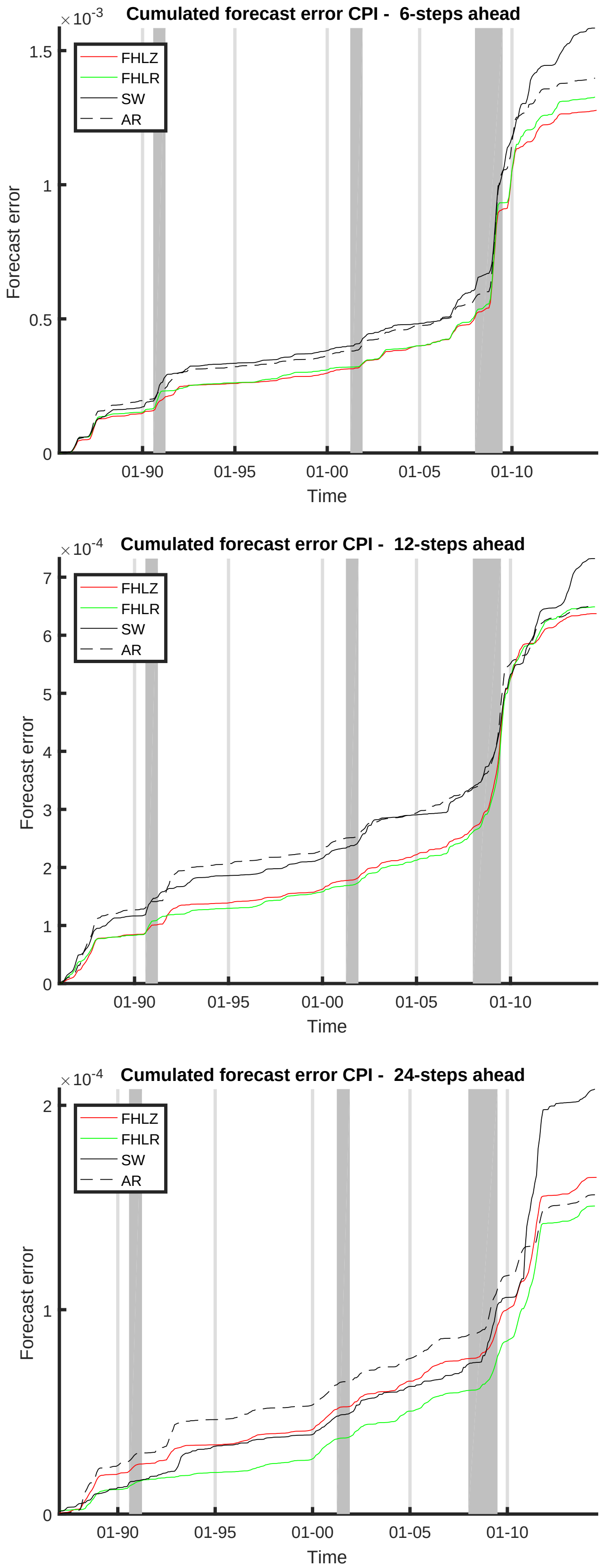

FIGURE 2.10: Cumulative sum of the squared forecasting error for the CPI at horizons $h = 6$, $h = 12$ and $h = 24$.

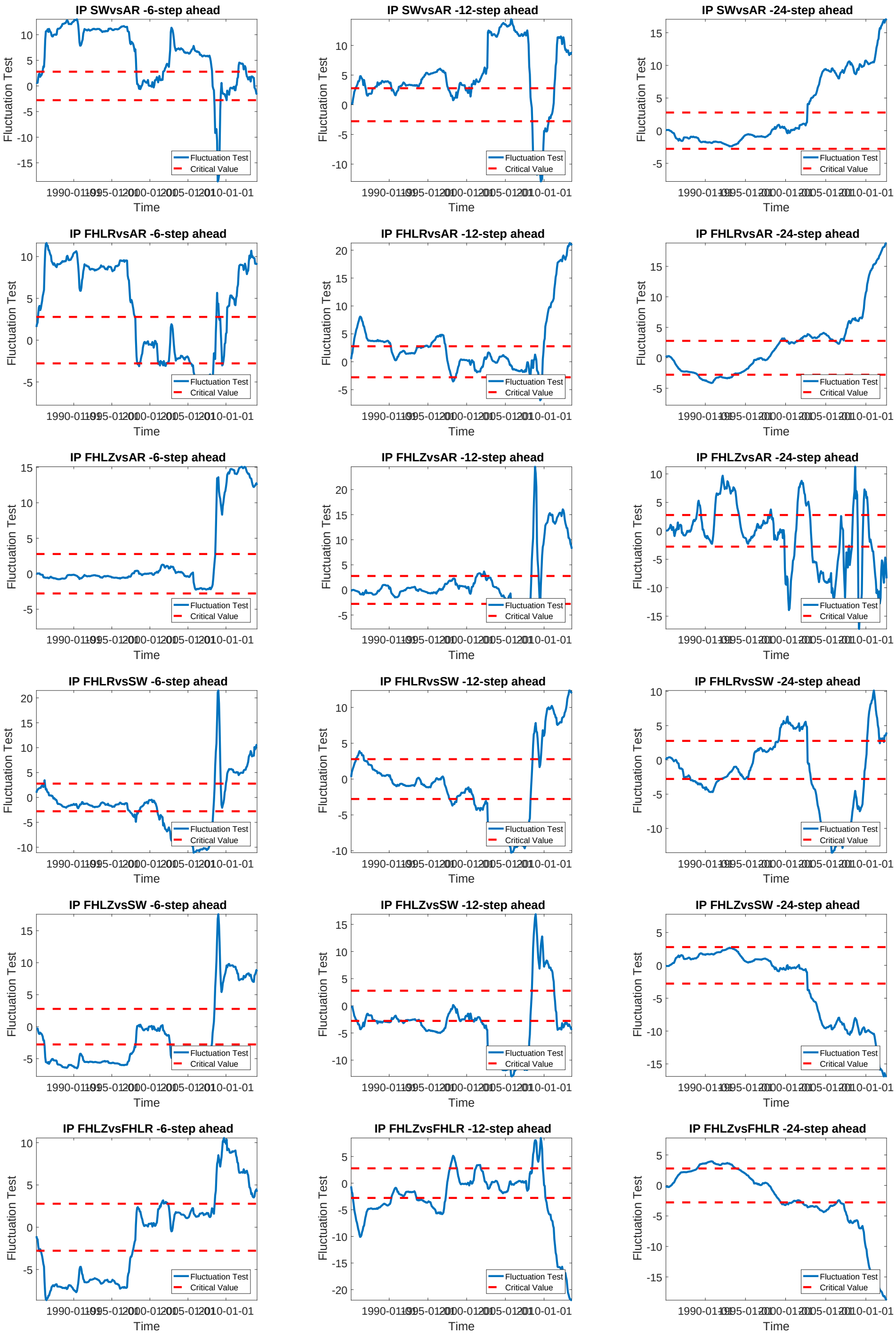

FIGURE 2.11: Fluctuation test for the IP.

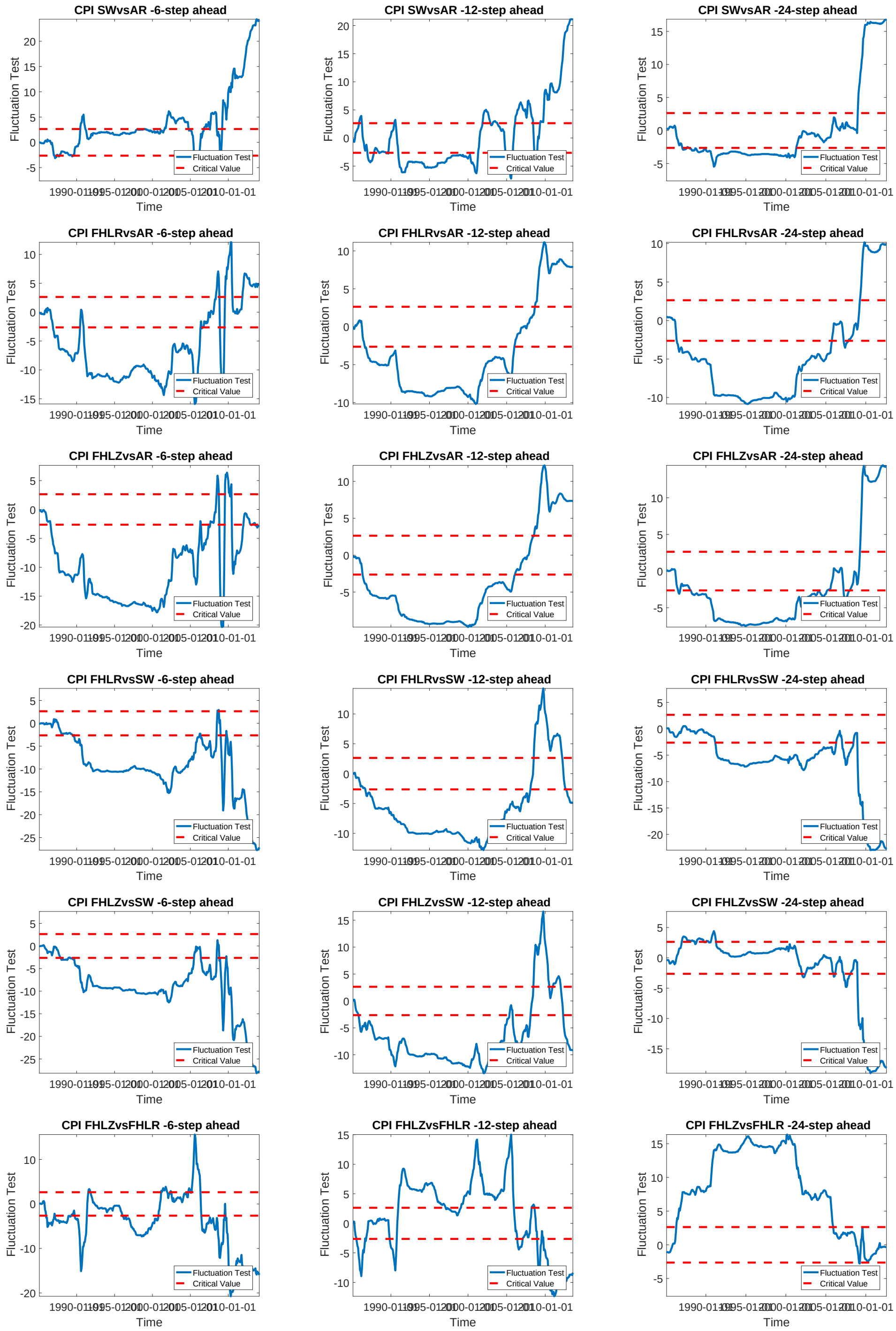

FIGURE 2.12: Fluctuation test for the CPI.

TABLE 2.9: Median rMSFE for category.

SW					
	$h = 1$	$h = 3$	$h = 6$	$h = 12$	$h = 24$
Industrial Production	1.01	**0.97**	1.08	1.09	1.18
Employment	**0.99**	**0.98**	1.01	1.08	1.10
Unemployment Rate	0.94	**0.76**	**0.67**	**0.79**	1.09
Housing	1.01	1.02	1.02	1.02	1.02
Inventories	1.67	1.19	1.11	1.09	1.23
Prices	1.11	1.09	1.18	1.12	1.18
Wages	1.16	1.19	1.26	1.08	1.11
Interest Rates	1.44	1.13	1.21	1.30	1.39
Money	1.21	1.13	1.21	1.22	1.10
Exchange Rates	1.15	1.16	1.17	1.05	1.12
Stock Prices	1.14	1.21	1.54	1.68	1.42

FHLR					
	$h = 1$	$h = 3$	$h = 6$	$h = 12$	$h = 24$
Industrial Production	**0.99**	1.00	1.00	1.12	1.20
Employment	1.02	1.12	1.15	1.16	1.12
Unemployment Rate	**0.90**	0.77	0.70	0.82	1.08
Housing	**0.01**	**0.02**	**0.01**	**0.02**	**0.03**
Inventories	1.23	1.09	1.04	1.10	1.13
Prices	0.94	0.94	1.00	1.01	1.05
Wages	1.02	1.00	1.02	1.06	1.02
Interest Rates	1.15	0.96	1.00	0.95	0.90
Money	**0.97**	**0.86**	**0.94**	**0.87**	0.94
Exchange Rates	1.04	1.01	0.99	0.93	**0.89**
Stock Prices	0.95	1.05	1.03	1.14	1.10

FHLZ					
	$h = 1$	$h = 3$	$h = 6$	$h = 12$	$h = 24$
Industrial Production	1.00	1.00	**0.99**	1.00	**0.98**
Employment	1.08	1.28	1.30	1.21	1.05
Unemployment Rate	0.95	0.88	0.95	0.97	**0.98**
Housing	0.03	0.03	0.04	0.05	0.05
Inventories	1.05	1.03	1.02	1.00	1.02
Prices	**0.91**	**0.93**	**0.98**	1.04	1.09
Wages	**0.99**	**1.00**	1.00	1.00	1.00
Interest Rates	**0.96**	**0.95**	**0.98**	**0.94**	**0.90**
Money	0.98	0.89	0.98	0.93	**0.93**
Exchange Rates	1.03	**0.99**	**0.98**	**0.93**	0.92
Stock Prices	**0.93**	**0.87**	**0.83**	**0.81**	**0.83**

benchmark for every category at least on one horizon. In table 2.10, the distributions of the rMSFEs of the dynamic models for each cathegories are reported. The configuration of the parameters is the one chosen in the calibration process for the forecasting of the IP. By analysing the distributions of the rMSFEs for the IP, it comes out that dynamic methods tend to outperform SW on real variables. FHLZ outperforms the benchmark up to the 50-th percentile and the other factor models up to the 75-th percentile.

TABLE 2.10: Distributions of the rMSFEs: configuration for the IP.

SW					
Percentile	0.05	0.25	0.50	0.75	0.95
$h = 1$	0.89	0.99	1.06	1.16	5.34
$h = 3$	0.77	0.98	1.07	1.18	1.98
$h = 6$	0.69	1.02	1.13	1.25	1.48
$h = 12$	0.76	1.02	1.09	1.24	1.63
$h = 24$	0.94	1.05	1.15	1.27	1.73

FHLR					
Percentile	0.05	0.25	0.50	0.75	0.95
$h = 1$	0.10	0.96	1.03	1.13	3.81
$h = 3$	0.10	0.94	1.02	1.11	1.60
$h = 6$	0.15	0.98	1.04	1.14	1.46
$h = 12$	0.18	1.02	1.10	1.18	1.34
$h = 24$	0.19	1.02	1.11	1.20	1.35

FHLZ					
Percentile	0.05	0.25	0.50	0.75	0.95
$h = 1$	0.10	0.96	1.01	1.10	3.90
$h = 3$	0.10	0.973	0.98	1.12	1.76
$h = 6$	0.12	0.96	1.00	1.07	1.49
$h = 12$	0.15	0.96	1.00	1.07	1.27
$h = 24$	0.19	0.96	0.99	1.05	1.14

The results obtained for the configurations adopted for the CPI are reported in table 2.11. By analysing the distributions of the rMSFEs for the CPI, it comes out that dynamic methods tend to outperform SW. FHLZ globally outperforms FHLR on the whole distribution.

2.7 Conclusions

In this chapter, we have shown that FHLR tends globally to outperform the other methods on the real variables and that FHLZ tends globally to outperform the other methods on the nominal variables. As to EU dataset, Della Marra, 2017 found similar results for the CPI, but mixed evidences appeared for the IP. As to the US dataset,

TABLE 2.11: Distributions of the rMSFEs: configuration for the CPI.

SW					
Percentile	0.05	0.25	0.50	0.75	0.95
$h = 1$	0.93	1.00	1.12	1.23	1.87
$h = 3$	0.85	1.00	1.09	1.20	1.39
$h = 6$	0.88	1.10	1.17	1.28	1.53
$h = 12$	0.95	1.04	1.16	1.30	1.68
$h = 24$	0.88	1.05	1.20	1.32	1.61

FHLR					
Percentile	0.05	0.25	0.50	0.75	0.95
$h = 1$	0.06	0.93	1.00	1.15	5.35
$h = 3$	0.05	0.89	0.96	1.10	1.91
$h = 6$	0.07	0.91	1.00	1.10	1.62
$h = 12$	0.12	0.89	0.97	1.07	1.28
$h = 24$	0.11	0.88	0.94	1.05	1.17

FHLZ					
Percentile	0.05	0.25	0.50	0.75	0.95
$h = 1$	0.05	0.91	0.97	1.11	4.56
$h = 3$	0.04	0.87	0.95	1.09	1.94
$h = 6$	0.06	0.91	0.98	1.05	1.56
$h = 12$	0.11	0.87	0.94	1.05	1.24
$h = 24$	0.11	0.87	0.93	1.05	1.22

Forni, Giovannelli, et al., 2016 found similar but less significative results, especially on the forecasting of the whole dataset. Hence, we have empirically shown that the calibration process plays a crucial role in these applications, since a more efficient exploration of the paramter space allowed us to empirically prove the superiority of frequence-domain dynamic factor models against time-domain factor models in a macroeconomic forecasting setting.

Chapter 3

Linear or non-linear estimates of the factors? This is the dilemma

3.1 Introduction

In the last decade, extending forecasting models beyond linear models has been considered one of the golden directions in which theoretical and empirical econometrics would move, especially in the light of the latest crisis periods which took place in the last decade over the global economy. As exposed in Calhoun and Elliott, 2012, it is reasonable to assume that nonlinear models may show better forecasting results than the linear ones. In fact, almost all economic models of the macroeconomy show some sources of nonlinearity. For instance, dynamic stochastic general equilibrium (DSGE) models of the economy predict complicated nonlinear relationships between past and future variables. Indeed it would be highly surprising if the relationships between current and future outcomes of variables were simply linear. In spite of these considerations, this future of nonlinear models used to predict outcomes has not materialized. In forecasting macroeconomic data, a number of studies over the last decade or so have found a very small role for parametric nonlinear models. Nonparametric models have also only been used for small set of variables, typically not for macroeconomic forecasting applications. Recent surveys of forecasting (e.g., Stock and M. Watson, 1998) show little influence of nonlinear models in macroeconomic forecasting. As to our knowledge, only in Marcellino, 2004 there are some evidences that a particular class of nonlinear models (neural networks) may produce better macroeconomic forecasting than several classes of linear models. Moreover, they also showed that there is no relationship between the forecast horizon and the performance of nonlinear models.

In this chapter, we extend a previous work from Banbura and Modugno, 2014, in which they modify the Expecetd - Maximization algorithm (EM) to estimate the parameter of a dynamic factor model on a dataset with data sampled at different frequencies and with an arbitrary pattern of missing data. Their approach relies on a dynamic factor model which is formulated, under general assumptions, in a state-space form whose parameters are estimated by means of the EM algorithm. To estimate the conditional moments of the latent factors, they employ the Kalman filter under the assumptions of linearity and gaussianity of the data in the model. In our work, we estimate the conditional moments of the latent factor relaxing the assumption of linearity and gaussianity of the underlying DGP. To produce such estimates, we employ the particle filter, which has become pretty popular in the last fifteen years and whose application in the field of engineering and economics are rapidly growing (see Gustaffson, 2010). However, in our application we will not deal the case of missing data, since an extension of the Particle Filter for arbitrary pattern of missing data is still debated (for some references, see Zhang et al., 2014, Gopaluni,

B. T. Schön, and A. G. Wills, 2009). Moreover, we will only work on the "Small" dataset described in Banbura and Modugno, 2014, to avoid the curse of dimensionality which characterises the Particle Filter. To justify our choice, we will report that, as in Banbura and Modugno, 2014, the forecasting performance on the "Medium" and on the "Large" dataset does not seem to improve from the Small case. An extensive description of this dataset can be found in the Appendix C. To conclude, we empirically show that the number of particles employed in the Particle Filter do not seem to change dramatically its macroeconomic forecasting performance. This chapter is organised as follows. Section 2 presents some theoretical issues about filtering in state-space model. In this section, the Kalman Filter and the Particle Filter are briefly described. Section 3 describes the empirical macroeconomic framework of our application. Section 4 shows the main forecasting results achieved both by means of the Kalman Filter and of the Particle Filter. Section 5 concludes.

3.2 An overview on filtering

In this paragraph, we will briefly describe the filtering problem following the same lines as Arulampalam et al., 2002. Given a state-space model defined by equations (3.1):

$$\begin{cases} x_t = g_t(x_{t-1}, v_{t-1}) \\ y_t = h_t(x_t, w_t), \end{cases} \tag{3.1}$$

where $x_t, t \in \mathbb{N}$ is a unobservable state variable, $y_t, t \in \mathbb{N}$ is a measurable variable, $v_t \sim \mathcal{N}(0, Q_{t-1}), w_t \sim \mathcal{N}(0, R_t), t \in \mathbb{N}$ i.i.d. two gaussian noise processes with zero mean and covariance matrices, respectively, Q_{t-1}, R_t, $g_t \in \mathbb{R}^{n_x} \times \mathbb{R}^{n_v} \to \mathbb{R}^{n_x}$ and $h_t \in \mathbb{R}^{n_x} \times \mathbb{R}^{n_w} \to \mathbb{R}^{n_y}$ two (possibly) nonlinear functions of the state x_t, n_x is the dimension of the state space x_t, n_y is the dimension of the space formed by y_t and n_v, n_w are the dimensions of, respectively, the noise processes v_t and w_t. Under this setting, we are concerned with finding the best (in the sense of minimum mean squared error) linear estimates of $x_t, t \in \mathbb{N}$ in terms of all avalible observations $y_1, y_2, ..., y_t$. This is the so called "filtering problem".
From a Bayesian perspective, solving this problem requires to recursively calculate some degree of belief in the state x_t at time t, given the different values of $y_1, y_2, ..., y_t$ up to time t. Hence, we shall compute the probability density function $p(x_t|y_1, y_2, ..., y_t)$. By assumption, the prior $p(x_0|y_0) := p(x_0|\emptyset) = p(x_0)$ is assumed to be known.
The probability density function $p(x_t|y_1, y_2, ..., t_t)$ may be obtained, recursively, from the two following stages:

(i) *Prediction:* let $p(x_{t-1}|y_1, y_2, ..., y_{t-1})$ be known at time $t-1$. The prediction stage involves using the model (3.1) to obtain the prior probability density function at time t by the following formula:

$$p(x_{t-1}|y_1, y_2, ..., y_{t-1}) = \int p(x_t|x_{t-1})p(x_{t-1}|y_1, ..., y_{t-1})dx_{t-1}. \tag{3.2}$$

In formula (3.2), we basically employed Chapman - Kolmogorov equation under the assumption that x_t describes a Markov process of order one ($p(x_t|x_{t-1}, y_1, y_2, ..., y_{t-1}) = p(x_t|x_{t-1}) = g_t(x_{t-1}, v_{t-1})$, where the latest passage holds due to (3.1));

(ii) *Update:* at time t, as a new measurement y_t becomes available, this can be used to update the prior by means of the Bayes rule

$$p(x_t|y_1, y_2, ..., y_t) = \frac{p(y_t|x_t)p(x_t|y_1, y_2, ..., y_{t-1})}{p(y_t|y_1, y_2, ..., y_{t-1})}, \tag{3.3}$$

with

$$p(y_t|y_1, y_2, ..., y_{t-1}) = \int p(y_t|x_t)p(x_t|y_1, ..., y_{t-1})dx_t. \tag{3.4}$$

In equation (3.4), for the likelihood function holds that $p(y_t|x_t) = h_t(x_t, w_t)$, due to (3.1). In the update stage, the measurement y_t is used to modify the prior density to obtain the required posterior density of the current state.

The recurrence relations (3.3) and (3.4) form the basis for the optimal Bayesian solution. This recursive propagation of the posterior density cannot be determined analytically in general. Solutions do exist in a restrictive set of cases, including the Kalman filter and other methods that we will not discuss here. We also describe how, when the analytic solution is intractable, particle filter may approximate the optimal Bayesian solution. Other methods (extended Kalman Filter, etc.) will not be discussed here. For further details, see Arulampalam et al., 2002.

3.2.1 The Kalman Filter

The Kalman filter assumes that the posterior density $p(x_{t-1}|y_1, y_2, ..., y_{t-1})$ at every time step t follows a Gaussian distribution. Under this assumption, it can be proved that also $p(x_t|y_1, y_2, ..., y_{t-1})$ is Gaussian if the following conditions hold:

(i) v_{t-1}, w_t follow a Gaussian distribution whose parameters are known;

(ii)

$$\begin{cases} x_t = g_t(x_{t-1}, v_{t-1}) := F_t x_{t-1} + v_{t-1} \\ y_t = h_t(x_t, w_t) := H_t x_t + w_t, \end{cases} \tag{3.5}$$

where F_t, H_t are known matrices for all t.

The Kalman filter algorithm, which was derived using (3.3) and (3.4), can then be viewed as the following recursive relationship:

$$\begin{cases} x_{t-1}|y_1, y_2, ..., y_{t-1} \sim \mathcal{N}(x_{t-1}; m_{t-1|t-1}, P_{t-1|t-1}) \\ x_t|y_1, y_2, ..., y_{t-1} \sim \mathcal{N}(x_t; m_{t|t-1}, P_{t|t-1}) \\ x_t|y_1, y_2, ..., y_t \sim \mathcal{N}(x_t; m_{t|t}, P_{t|t}), \end{cases} \tag{3.6}$$

where

$$\begin{cases} m_{t|t-1} = F_t m_{t-1|t-1} \\ P_{t|t-1} = Q_{t-1} + F_t P_{t-1|t-1} F_t^T \\ m_{t|t} = m_{t|t-1} + K_t(y_t - H_t m_{t|t-1}) \\ P_{t|t} = P_{t|t-1} - K_t H_t P_{t|t-1}, \end{cases} \tag{3.7}$$

and where $\mathcal{N}(\lambda; \mu, \Sigma)$ is a Gaussian density with argument λ, mean μ and covariance matrix Σ. The following formulas also hold:

$$\begin{cases} S_t = H_t P_{t|t-1} H_t^T + R_t \\ K_t = P_{t|t-1} H_t^T S_t^{-1}. \end{cases} \tag{3.8}$$

S_t is the covariance matrix of the innovation term $y_t - H_t m_{t|t-1}$ and K_t is called the "Kalman gain". In the above equations, the transpose of a matrix Γ is denoted by Γ^T. This is the optimal solution to equations (3.2) - (3.3) if the (highly restrictive) assumptions hold. The implication is that no algorithm can ever do better than a Kalman filter in this linear Gaussian setting. It should be noted that it is possible to derive the same results using an ordinary least squares approach (OLS).

3.2.2 The Particle Filter

When the prediction and update steps of the optimal filtering are not analytically tractable, one has to resort to approximate methods, such as Monte Carlo sampling. Sequential importance sampling (SIS) is the most basic Monte Carlo method used for this purpose, which is also the basic model for most sequential Monte Carlo filters developed over the last years. In deriving the SIS algorithm, it is useful to consider the full posterior distribution at time $t-1$, $p(x_0, ..., x_{t-1}|y_1, ..., y_{t-1})$, rather than the filtering distribution, $p(x_{t-1}|y_1, ..., y_{t-1})$, which is just the marginal of the full posterior distribution with respect to x_{t-1}. The idea in SIS is to approximate the posterior distribution at time $t-1$, $p(x_0, ..., x_{t-1}|y_1, ..., y_{t-1})$, with a weighted set of samples $\{x_0^i, ..., x_{t-1}, w_{t-1}\}_{i=1}^N$, also called particles, and recursively update these particles to obtain an approximation to the posterior distribution at the next time step: $p(x_0, ..., x_t|y_1, ..., y_t)$.
SIS is based on importance sampling. In importance sampling, one approximates a target distribution $p(x)$, using samples drawn from a proposal distribution $q(x)$. Importance sampling is generally used when it is difficult to sample directly from the target distribution itself, but much easier to sample from the proposal distribution. To compensate for the discrepancy between the target and proposal distributions, one has to weight each sample x^i by $w^i \propto \pi(x^i)/q(x^i)$ where $\pi(x)$ is a function that is proportional to $p(x)$ and that we know how to evaluate. Applied to our posterior distribution at time $t-1$, importance sampling yields:

$$p(x_0, ..., x_{t-1}|y_1, ..., y_{t-1}) \approx \sum_{i=1}^{N} w_{t-1}^i \delta_{x_0^i, ..., x_{t-1}^i}. \tag{3.9}$$

The crucial idea in SIS is to update the particles $x_0^i, ..., x_{t-1}^i$ and their weights w_{t-1}^i such that they would approximate the posterior distribution at the next time step: $p(x_0, ..., x_t|y_1, ..., y_t)$. To do this, we first assume that the proposal distribution at time t can be factorized as follows:

$$q(x_0, ..., x_t|y_1, ..., y_t) = q(x_t|x_0, ..., x_{t-1}, y_1, ..., y_t) q(x_0, ..., x_{t-1}|y_1, ..., y_{t-1}), \tag{3.10}$$

so that we can simply augment each particle $x_0^i, ..., x_{t-1}^i$ at time $t-1$ with a new state x_t^i at time t sampled from $q(x_t|x_0, ..., x_{t-1}, y_1, ..., y_t)$. To update the weights, w_{t-1}^i, we note that according importance sampling, the weights of the particles at time k should be as follows:

$$w_t^i \propto \frac{p(x_0^i, ..., x_t^i|y_1, ..., y_t)}{q(x_0^i, ..., x_t^i|y_1, ..., y_t)}. \tag{3.11}$$

We will not give the full derivation here, but it turns out that it is possible to express (3.11) recursively in terms of w_{k-1}^i. It can be shown that the final result is the following:

$$w_t^i \propto w_{t-1}^i \frac{p(y_t|x_t^i)p(x_t^i|x_{t-1}^i)}{q(x_t^i|x_0^i, ..., x_{t-1}^i, y_1, ..., y_t)}. \tag{3.12}$$

The weights $\{w_k^i\}_{i=1}^N$ are then normalized to sum to 1. If we also assume that $q(x_t|x_0, ..., x_{t-1}, y_1, ..., y_t) = q(x_t|x_{t-1}, y_t)$ so that the proposal at the next time step only depends on the most recent state and the most recent measurement, then we only need to store x_{t-1}^i and generate the particle at the next time step from the distribution $q(x_t|x_{t-1}^i, y_t)$. Thus, in this case, the update equations simplify to:

$$\begin{cases} x_t^i \sim q(x_t|x_{t-1}^i, y_t) \\ w_t^i \propto w_{t-1}^i \frac{p(y_t|x_t^i)p(x_t^i|x_{t-1}^i)}{q(x_t^i|x_{t-1}^i, y_t)}. \end{cases} \tag{3.13}$$

3.3 Econometric Framework

In this section, we describe the framework adopted in this work, which is basically the same as in Banbura and Modugno, 2014. Let $y_t = [y_{1,t}, y_{2,t}, ..., y_{n,t}]', t = 1, ..., T$, denote a stationary standardized n-dimensional vector process. We assume that y_t admits the following factor model representation:

$$y_t = \Lambda f_t + \xi_t, \tag{3.14}$$

where $f_t \in \mathbb{R}^{r\times 1}$ is the vector of unobserved common factors and $\xi_t = [\xi_{1,t}, \xi_{2,t}, ..., \xi_{n,t}]'$ is the idiosyncratic component, uncorrelated with f_t at all leads and lags.
$\Lambda \in \mathbb{R}^{n\times r}$ is the matrix of the factor loadings. $\chi_t = \Lambda f_t$ is referred to as the common component. It is assumed that $\xi_t \sim i.i.d.\mathcal{N}(0, R)$ with R a diagonal matrix, i.e. y_t follows an exact factor model. The assumption of no cross-correlation in the idiosyncratic component could be too restrictive, in particular in the case of large cross-sections. However, Doz, Giannone, and Reichlin, 2012 show that when idiosyncratic components are weakly cross-correlated the factors can be consistently estimated, as $n, T \to \infty$, by quasi maximum likelihood, where the misspecified model is the exact factor model. Consequently, the estimators considered above are asymptotically valid also in the case of the approximate factor model.Stock and M. W. Watson, 2002a prove a similar result for factor estimators based on principal components. Differently from Banbura and Modugno, 2014, in our exercise we will only consider the case in which ξ_t follows an AR(1) process. To achieve a state space representation of the stochastic process y_t, it is assumed that the common factors f_t follow a stationary VAR(p) process of order:

$$f_t = \sum_{t=1}^{p} A_i f_{t-p} + u_t, \tag{3.15}$$

where $A_1, ..., A_p \in \mathbb{R}^{r\times r}$ are matrices of autoregressive coefficients (which we will collect in the vector $A = [A_1, ..., A_p]$) and $u_t \sim i.i.d.\mathcal{N}(0, Q)$, with Q a diagonal matrix. In our exercise, we will assume that $p \in \{1, 2\}$ out of simplicity. As in Banbura and Modugno, 2014, to model the serial correlation in the idiosyncratic component ξ_t, we assume that it can be decomposed as

$$\begin{cases} \xi_{i,t} = \tilde{\xi}_{i,t} + \epsilon_{i,t} \\ \epsilon_{i,t} = \alpha_i \tilde{\epsilon}_{i,t-1} + e_{i,t}, \end{cases} \tag{3.16}$$

where $\epsilon_{i,t} \sim i.i.d.\mathcal{N}(0,k), e_{i,t} \sim i.i.d.\mathcal{N}(0,\sigma_i^2)$. $\epsilon_t = [\epsilon_{1,t},...,\epsilon_{n,t}]', \xi_t = [\xi_{1,t},...,\xi_{n,t}]$ are cross-sectionally uncorrelated and k is a small number. Combining equations from (3.14) to (3.16) we get the following state-space representation:

$$\begin{cases} y_t = \tilde{\Lambda}\tilde{f}_t + \xi_t \\ \tilde{f}_t = \sum_{t=1}^{p} \tilde{A}_i \tilde{f}_{t-p} + \tilde{u}_t, \end{cases} \tag{3.17}$$

where $\epsilon_t \sim i.i.d.\mathcal{N}(0,\tilde{R}), \tilde{u}_t \sim i.i.d.\mathcal{N}(0,\tilde{Q})$, $e_t = [e_{1,t},...,e_{n,t}]', \tilde{f}_t = \begin{bmatrix} f_t \\ \tilde{\xi}_t \end{bmatrix}$, $\tilde{u}_t = \begin{bmatrix} u_t \\ e_t \end{bmatrix}, \tilde{\Lambda} = [\Lambda, I], \tilde{A} = \begin{bmatrix} A & 0 \\ 0 & diag(\alpha_1,...,\alpha_n) \end{bmatrix}, \tilde{Q} = \begin{bmatrix} Q & 0 \\ 0 & diag(\sigma_1,...,\sigma_n) \end{bmatrix}$, $\tilde{R} = diag(k,...,k)$.

Since f_t are unobserved, the maximum likelihood estimators of the parameters of the state-space model in equation (3.17), which we collect in $\theta = \{\Lambda, A, Q, \alpha_i, \sigma_i^2\}$, are in general not available in closed form. On the other hand, a direct numerical maximization of the likelihood is computationally demanding, in particular for large n owing to the large number of parameters. As in Banbura and Modugno, 2014, in this chapter we adopt an approach based on the EM algorithm, which was proposed by Dempster, Laird, and Rubin, 1977 as a general solution to problems for which incomplete or latent data make the likelihood intractable or difficult to deal with. The essential idea of the algorithm is to write the likelihood as if the data were complete and to iterate between two steps: in the expectation step we "fill in" the missing data in the likelihood, while in the maximization step we re-optimize this expectation. Under some regularity conditions, the EM algorithm converges towards a local maximum of the likelihood function. Let us denote the joint log-likelihood of y_t and $f_t, t = 1,...,T$ by $l(Y,F;\theta)$, where $Y = [y_1,...,y_T]$ and $F = [f_1,...,f_T]$. Given the available data $\Omega_T \subseteq Y$ for the state-space model given by equations $(3.14) - (3.15)$ the EM algorithm proceeds in a sequence of two alternating steps:

(i) *E-step:* the expectation of the log-likelihood conditional on the data is calculated using the estimates from the previous iteration, $\theta(j)$:

$$L(\theta,\theta(j)) = \mathbb{E}_{\theta(j)}[l(Y,F;\theta)|\Omega_T]; \tag{3.18}$$

(ii) *M-step:* the parameters are re-estimated through the maximization of the expected log-likelihood with respect to θ:

$$\theta(j+1) = arg\max_{\theta} L(\theta,\theta(j)). \tag{3.19}$$

By solving the optimization problem in equation (3.19), we can derive the following iterative formulas:

$$A(j+1) = (\sum_{t=1}^{T} \mathbb{E}_{\theta(j)}[f_t f_{t-1}^{'}|\Omega_T])(\sum_{t=1}^{T} \mathbb{E}_{\theta(j)}[f_{t-1} f_{t-1}^{'}|\Omega_T])^{-1}, \tag{3.20}$$

$$\alpha_i(j+1) = (\sum_{t=1}^{T} \mathbb{E}_{\theta(j)}[\tilde{\xi}_{i,t}\tilde{\xi}_{i,t-1}|\Omega_T])(\sum_{t=1}^{T} \mathbb{E}_{\theta(j)}[\tilde{\xi}^2_{i,t-1}|\Omega_T])^{-1}, \tag{3.21}$$

$$\sigma_i^2(j+1) = \frac{1}{T}(\sum_{t=1}^{T} \mathbb{E}_{\theta(j)}[\tilde{\xi}^2_{i,t}|\Omega_T] - \alpha_i(j+1)\sum_{t=1}^{T} \mathbb{E}_{\theta(j)}[\tilde{\xi}_{i,t}\tilde{\xi}_{i,t-1}|\Omega_T]). \tag{3.22}$$

As in Banbura and Modugno, 2014, in this econometric framework it also possible to impose resctrictions on the parameters of the model. Since we are modelling both monthly and quarterly variables in unified framework, we adopt the restrictions on the factor loadings which are described in Mariano and Murasawa, 2003, which can be posed in the form $H_\Lambda ved(\Lambda) = k_\Lambda$ (see Bork, 2009). Hence, under such assumptions, the matrix Λ can be estimated by the following formula:

$$vec(\Lambda_r(j+1)) = vec(\Lambda_u(j+1)) + (\sum_{t=1}^{T} \mathbb{E}_{\theta(j)}[f_t f_t'|\Omega_T] \otimes \tilde{R}_t)H_\Lambda^{`}\cdot$$
$$\cdot (H_\Lambda(\sum_{t=1}^{T} \mathbb{E}_{\theta(j)}[f_t f_t'|\Omega_T] \otimes \tilde{R}_t)H_\Lambda')^{-1}(k_\Lambda - H_\Lambda vec(\Lambda_u(j+1))), \tag{3.23}$$

where $\Lambda_u(j+1)$ is given by the following formula:

$$vec(\Lambda_u(j+1)) = (\sum_{t=1}^{T} \mathbb{E}_{\theta(j)}[f_t f_t'|\Omega_T] \otimes I_t)^{-1} vec(\sum_{t=1}^{T} I_t y_t \mathbb{E}_{\theta(j)}[f_t'|\Omega_T] +$$
$$- I_t y_t \mathbb{E}_{\theta(j)}[\xi_t f_t'|\Omega_T]), \tag{3.24}$$

with I_t being the identity matrix of size n and $\otimes$ the Kronecker product. As in Banbura and Modugno, 2014, the conditional moments of the latent factors $\mathbb{E}_{\theta(j)}[f_t|\Omega_T]$, $\mathbb{E}_{\theta(j)}[f_t f_t'|\Omega_T]$, $\mathbb{E}_{\theta(j)}[f_{t-1} f_{t-1}'|\Omega_T]$, $\mathbb{E}_{\theta(j)}[f_t f_{t-1}'|\Omega_T]$ can be obtained through the Kalman filter, described in subsection 3.2.1, for which the assumptions described in equations (3.16) - (3.17) must hold. In this chapter, we will also estimate the conditional moments of the latent factors $\mathbb{E}_{\theta(j)}[f_t|\Omega_T]$, $\mathbb{E}_{\theta(j)}[f_t f_t'|\Omega_T]$, $\mathbb{E}_{\theta(j)}[f_{t-1} f_{t-1}'|\Omega_T]$, $\mathbb{E}_{\theta(j)}[f_t f_{t-1}'|\Omega_T]$ by relaxing the assumptions of linearity and gaussianity of the model expressed in equations (3.16) - (3.17). To produce such estimates, we will employ the Particle Filter which has been described in subsection 3.2.2.

3.4 Forecasting the EU GDP

3.4.1 Description of the application

In this section, analogously to Banbura and Modugno, 2014, we employ their methodology, described in subsection 3.3, to forecast the EU Gross Domestic Product (GDP). We evaluate the methodology only on the ”Small” dataset from Banbura and Modugno, 2014. Such a choice was also made in order to cope with the curse of dimensionality which tipically occurs when applying the Particle Filter algorithm to high-dimensional problems (for further details, see Gustaffson, 2010). Below, some further insights about the ”Small” dataset are given:

(i) *Small:* contains the main indicators of real activity on the total economy, such as industrial production, orders, retail sales, unemployment rate, European Commission Economic Sentiment Indicator, Purchasing Manager Index, GDP or employment (14 series in total). It also contains financial series such as stock prices index or prices of raw materials. For further details, see Appendix C.

This dataset contains both monthly and quarterly variables. We combine the information from such time series following Mariano and Murasawa, 2003 and assume that the frequency of the model is monthly, and for each quarterly variable we construct a partially observed monthly counterpart (see the Appendix in Banbura and Modugno, 2014 for more details). The series observed on a daily basis are converted to monthly frequency by taking monthly averages. We evaluate the average precision of the forecasts for the "Small" dataset in an out-of-sample exercise. For each reference quarter, we produce a sequence of projections, starting with the forecast based on the information available in the first month of the preceding quarter, seven months ahead of the GDP flash release. The second forecast is produced with the information that would be available one month later and the last forecast is based on the information available in the first month of the following quarter, one month before the flash release. We denote projections based on the information in the preceding, current and following quarter (with respect to the forecast reference quarter) as Q(-1), Q(0) and Q(+1) respectively. Forecasts made in the first, second and third months of the quarter are referred to as M1, M2 and M3 respectively. For the measure of prediction accuracy we choose the root mean squared forecast error (RMSFE), which is defined as the square root of the right-hand member of equation (1.8) (not to be confused with the rMSFE defined in equation (1.9)). The evaluation sample ranges from 2000:Q1 to 2007:Q4. The estimation sample starts in January 1993 (hence, we employ 84 observations to compute the first forecastings). We opted for a recursive-window scheme, which means that the sample length increases each time that more information becomes available. We run the out-of-sample forecast evaluation for specifications, including one to five factors ($r \in \{1, 2, 3, 4, 5\}$) and one or two lags in the VAR ($p \in \{1, 2\}$) and we take averages over the specifications. We evaluate the forecasts for the "Small" dataset under the assumption of an AR(1) idiosyncratic component.

3.4.2 Results

Table 3.1 presents the results for the different forecast horizons, from the first month of the preceding quarter, Q(-1)M1, to the first month of the following quarter, Q(+1)M1. "Average" gives the average forecast error for all considered horizons. The column "KF" illustrates the results obtained by estimating the conditional moments of the factors by means of the Kalman Filter. The column "PF_30", instead, illustrates the results obtained by estimating the conditional moments of the factors by means of the Particle Filter with 30 particles.
As in Banbura and Modugno, 2014, we selected the following univariate benchmarks:

(i)*AR(p):* an autoregressive process whose order p is determined by means of Akaike information criterion (AIC);

(ii)*Mean:* the sample mean;

and the following multivariate benchmark:

(i)*BR:* model described in Banbura and Rünstler, 2011. In their application, they used such a model to forecast the euro area GDP growth from a dataset of 76 monthly series, comprising real activity measures, financial data and surveys.

TABLE 3.1: RMSFE for the EU GDP:2000:Q1 - 2007:Q4

	Factor models		Benchmarks		
	KF	PF_30	AR	$Mean$	BR
$Q(-1)M1$	0.27	0.32	0.33	0.32	0.26
$Q(-1)M2$	0.24	0.32	0.32	0.32	0.24
$Q(-1)M3$	0.24	0.32	0.32	0.32	0.21
$Q(0)M1$	0.23	0.32	0.32	0.32	0.21
$Q(0)M2$	0.21	0.31	0.27	0.31	0.22
$Q(-1)M3$	0.19	0.31	0.27	0.31	0.21
$Q(+1)M1$	0.18	0.31	0.27	0.31	0.18
$Average$	0.22	0.32	0.30	0.31	0.22

We can see that all the Kalman Filter performs much better than the univariate benchmarks and than the Particle Filter, with largest improvements for shortest forecast horizons. This confirms that the importance of the sources of nonlinearity in the DGP and of non-Gaussianity in the distribution of the error terms in the factor model is negligible in macroeconomic forecasting. Since the underlying assumptions of the Kalman Filter seems to hold, it should provide the exact solution to equations (3.3) - (3.4), in contrast to an approximate one computed by means of the Particle Filter. Hence, this might explain why the Kalman Filter outperforms the Particle Filter. Moreover, the Particle Filter is outperformed also by the model described by Banbura and Rünstler, 2011 (whose RMSEs are reported in table 3.1 under the column "BR") . This might be explained by the fact that this model relies on principal component and also on the Kalman Filter. The Particle Filter is also outperformed by the AR benchmark and shows similar performances of the sample mean benchmark. To test the importance of the number of particles in the Particle Filter, we have repeated the same exercise for different configurations of this algorithm. To be more precise, for the number of particles $npar$, we tested $npar \in \{10, 50, 100\}$. The results achieved for the different forecast horizons are reported in table 3.2, where "PF_10" stands for Particle Filter with 10 particles, "PF_50" stands for Particle Filter with 50 particles and "PF_100" stands for Particle Filter with 100 particles.

TABLE 3.2: RMSFE of different configurations of the PF for the EU GDP:2000:Q1 - 2007:Q4

	PF_10	PF_50	PF_100
$Q(-1)M1$	0.32	0.32	0.32
$Q(-1)M2$	0.24	0.32	0.31
$Q(-1)M3$	0.32	0.32	0.32
$Q(0)M1$	0.32	0.32	0.31
$Q(0)M2$	0.31	0.31	0.31
$Q(-1)M3$	0.31	0.31	0.31
$Q(+1)M1$	0.31	0.31	0.30
$Average$	0.32	0.32	0.31

We can see that the forecasting performance of the Particle Filter does not seem to change by augmenting the number of particles. To be more precise, there is a slight improvement in the performance of the Particle Filter with 100 particles but, due to the curse of dimensionality which affects the Particle Filter, a greater number of particles implies a larger waste of computational time. Hence, we can conclude that the number of particles does not seem to be a critical parameter for the macroeconomic forecasting performance of the Particle Filter.

3.5 Conclusions

In this chapter, we have seen that the assumptions of linearity of the DGP and of a gaussian distribution for the error terms of the dynamic factor model seems to hold for macroeconomic forecasting. Hence, consistently with Stock and M. Watson, 1998 among others, we can conclude that a linear estimate of the conditional moments of the factors is the right answer to the initial dilemma of this chapter. However, it should be pointed out that, in this study, we have focused on the estimation of the factors, without considering the effects of the sources of nonlinearities in the DGP and of the non-Gaussian distributions of the error terms in the estimation of the other parameters of the model (3.13) - (3.15), contained in the vector $\theta = \{\Lambda, A, Q, \alpha_i, \sigma_i^2\}$. To conduct a more extensive analysis, we should modify the EM algorithm so to relax the aforementioned assumptions of linearity and gaussianity. Up to our knowledge, the only extension of the EM in this direction can be found in T. B. Schön, A. Wills, and Ninness, 2011. This extension to our econometric framework is left for future research.

Appendix A

The EU dataset and transformations used in Chapters 1 and 2

In table A.1, the series which compose the EU dataset are reported. *Tcode* identifies the tranformation (further details are given below) and *Des* identifies the deseasonalized flag (1 if the series is deseasonalized, 0 otherwise).

TABLE A1: List of the EU time series.

	Number	Long Desc.	Tcode	Des
1	BDM1....A	BD MONEY SUPPLY - M1 - CURA	6	1
2	BDM2C...B	BD MONEY SUPPLY - M2 CURA	6	0
3	BDM3C...B	MONEY SUPPLY - M3 - CURA	6	0
4	FRM1....A	FR MONEY SUPPLY - M1 - CURN	6	1
5	FRM2....A	FR MONEY SUPPLY - M2 - CURN	6	1
6	FRM3....A	FR MONEY SUPPLY - M3 - CURN	6	1
7	ITM1....A	IT MONEY SUPPLY: M1 - CURN	6	1
8	ITM2....A	IT MONEY SUPPLY: M2 - CURN	6	1
9	ITM3....A	MONEY SUPPLY: M3 - CURN	6	1
10	NLM1....A	NL MONEY SUPPLY - M1 - CURN	6	1
11	NLM2....A	NL MONEY SUPPLY - M2 - CURN	6	1
12	NLM3....A	NL MONEY SUPPLY - M3 - CURN	6	1
13	EMECBM1.B	EM MONEY SUPPLY: M1 - CURA	6	0
14	EMM2....B	EM MONEY SUPPLY: M2 - CURA	6	0
15	EMECBM3.B	EM MONEY SUPPLY: M3 - CURA	6	0
16	NLIMPGDSA	NL IMPORTS - CIF - CURN	5	1
17	NLEXPGDSA	NL EXPORTS - FOB - CURN	5	1
18	FRIMPGDSB	FR IMPORTS FOB - CURA	5	1
19	FREXPGDSB	FR EXPORTS FOB - CURA	5	1
20	ESOXT003b	ES ITS EXPORTS F.O.B. TOTAL - CURA	5	1
21	ESOXT009b	ES ITS IMPORTS C.I.F. TOTAL - CURA	5	1
22	ESEXPGDSD	ES EXPORTS - CONA	5	1
23	ESIMPGDSD	ES IMPORTS - CONA	5	1
24	ESEXPPRCF	ES EXPORT UNIT VALUE INDEX - NADJ	5	1
25	ESIMPPRCF	ES IMPORT UNIT VALUE INDEX - NADJ	5	1
26	BDEXPGDSB	BD EXPORTS OF GOODS (FOB) - CURA	5	1
27	BDIMPGDSB	BD IMPORTS OF GOODS (CIF) - CURA	5	1
28	BDEXPPRCF	BD EXPORT PRICE INDEX - NADJ	7	1
29	BDIMPPRCF	BD IMPORT PRICE INDEX - NADJ	7	1
30	ITEXPPRCF	IT EXPORT UNIT VALUE INDEX - NADJ	7	1

31	BDOCC011	BD REAL EFFECTIVE EXCHANGE RATES - CPI BASED - NADJ	5	0
32	BGOCC011	BG REAL EFFECTIVE EXCHANGE RATES - CPI BASED - NADJ	5	0
33	ESOCC011	ES REAL EFFECTIVE EXCHANGE RATES - CPI BASED - NADJ	5	0
34	FNOCC011	FN REAL EFFECTIVE EXCHANGE RATES - CPI BASED - NADJ	5	0
35	FROCC011	FR REAL EFFECTIVE EXCHANGE RATES - CPI BASED -NADJ	5	0
36	GROCC011	GR REAL EFFECTIVE EXCHANGE RATES - CPI BASED - NADJ	5	0
37	IROCC011	IR REAL EFFECTIVE EXCHANGE RATES - CPI BASED - NADJ	5	0
38	ITOCC011	IT REAL EFFECTIVE EXCHANGE RATES - CPI BASED - NADJ	5	0
39	NLOCC011	NL REAL EFFECTIVE EXCHANGE RATES - CPI BASED - NADJ	5	0
40	OEOCC011	OE REAL EFFECTIVE EXCHANGE RATES - CPI BASED - NADJ	5	0
41	PTOCC011	PT REAL EFFECTIVE EXCHANGE RATES - CPI BASED - NADJ	5	0
42	BDESPPINF	BD PPI: MIG - NON-DURABLE CONSUMER GOODS - NADJ	7	0
43	BDPROPRCF	BD PPI: INDL. PRODUCTS, TOTAL, SOLD ON THE DOMESTIC MARKET -NADJ	7	0
44	BDESPPIEF	BD PPI: MIG - ENERGY - NADJ	7	0
45	FRESPPITF	FR PPI: MIG - INTERMEDIATE GOODS - NADJ	7	0
46	ITESPPINF	IT PPI: MIG - NON-DURABLE CONSUMER GOODS -NADJ	7	0
47	ITESPPIEF	IT PPI: MIG - ENERGY - NADJ	7	0
48	ESESPPITF	ES PPI: MIG - INTERMEDIATE GOODS - NADJ	7	0
49	ESESPPINF	ES PPI: MIG - NON-DURABLE CONSUMER GOODS - NADJ	7	0
50	ESPPDCNSF	ES PPI - CONSUMER GOODS, DURABLES - NADJ	7	0
51	ESPPINVSF	ES PPI - CAPITAL GOODS - NADJ	7	0
52	ESESPPIEF	ES PPI: MIG - ENERGY - NADJ	7	0
53	ESPROPRCF	ES PPI -NADJ	7	1
54	BGESPPITF	BG PPI: MIG - INTERMEDIATE GOODS - NADJ	7	0
55	BGESPPINF	BG PPI: MIG - NON-DURABLE CONSUMER GOODS - NADJ	7	1
56	BGESPPIIF	BG PPI: INDUSTRY - NADJ	7	0
57	NLESPPITF	NL PPI: MIG - INTERMEDIATE GOODS - NADJ	7	0
58	EKPROPRCF	EK PPI: INDUSTRY - NADJ	7	0
59	ITCPWORKF	IT CPI EXCLUDING TOBACCO (FOI) - NADJ	7	0
60	ITCP7500F	IT CPI (1975=100) - NADJ	7	0
61	ITRAWPRCF	IT RAW MATERIALS PRICE INDEX - NADJ	7	0
62	ITPROPRCF	IT PPI - NADJ	7	0
63	FRCONPRAF	FR CPI (LINKED & REBASED) - NADJ	7	0
64	FRAGPRC.F	FR AGRICULTURAL PRICE INDEX - NADJ	7	0
65	FRAGIIGSF	FR AGRICULTURAL INPUT PRICES - INVESTMENT GOODS & SERVICES - NADJ	7	1
66	BDCP7500F	BD CPI (1975=100) - NADJ	7	1
67	ESCONPRCF	ES CPI - NADJ	7	0
68	NLCONPRCF	NL CPI - NADJ	7	0
69	EMCONPRCF	EM CPI - NADJ	7	0
70	BDI..RELF	BD REAL EFFECTIVE FX RATE (REER) BASED ON UNIT LABOUR COSTS - NADJ	5	0
71	BDMWAGINF	BD WAGE&SALARY LEVEL,MTHLY BASIS - PRDG.SECT.(PAN BD M0191) NADJ	5	1
72	ESWAGES.F	ES WAGES: INCOME INDICATOR - VOLN	5	1
73	ESWAGES%F	ES WAGES: INCOME INDICATOR (%YOY) - VOLN	2	0
74	FRI..RELF	FR REAL EFFECTIVE FX RATE (REER) BASED ON UNIT LABOUR COSTS - NADJ	5	0
75	ITI..RELF	IT REAL EFFECTIVE FX RATE (REER) BASED ON UNIT LABOUR COSTS - NADJ	5	1
76	ITWAGES.F	IT CONTRACTUAL HOURLY WAGE: ALL WORKERS - NADJ	5	1
77	ITOLC007H	IT HOURLY WAGE RATE: INDUSTRY INCL. CONSTRUCTION - PROXY NADJ	5	1
78	ITWAGES%F	IT CONTRACTUAL HOURLY WAGE: ALL WORKERS (%YOY) - NADJ	5	0
79	NLI..RELF	NL REAL EFFECTIVE FX RATE (REER) BASED ON UNIT LABOUR COSTS - NADJ	5	0
80	NLOLC007H	NL HOURLY WAGE RATE: MFG - PROXY NADJ	5	1

81	BDIPTOT.G	BD INDUSTRIAL PRODUCTION INCLUDING CONSTRUCTION (CAL ADJ) - VOLA	5	0
82	BDESPISDH	BD IPI: MIG - DURABLE CONSUMER GOODS, VOLUME IOP (WDA) - VOLN	5	1
83	BDESPIESH	BD IPI: MIG-CAPITAL GOODS, VOLUME INDEX OF PRODUCTION (WDA) - VOLN	5	1
84	BDESPISNH	BD IPI: MIG - NON-DURABLE CONSUMER GOODS, VOLUME IOP (WDA) - VOLN	5	1
85	ESIPINTGH	ES INDUSTRIAL PRODUCTION - INTERMEDIATE GOODS - VOLN	5	1
86	ESIPINVSH	ES INDUSTRIAL PRODUCTION - CAPITAL GOODS - VOLN	5	1
87	ESESIBASG	ES IPI: MANUFACTURE OF BASIC METALS, VOLUME IOP (WDA) - VOLA	5	0
88	ESIPOMNPH	ES INDUSTRIAL PRODUCTION - OTHER NON-METAL MINERAL PRODUCTS - VOLN	5	1
89	ESIPTOT.G	ES INDUSTRIAL PRODUCTION (WDA) - VOLA	5	0
90	ITIPTOT.G	IT INDUSTRIAL PRODUCTION - VOLA	5	0
91	NLIPTOT.G	NL INDUSTRIAL PRODUCTION EXCLUDING CONSTRUCTION - VOLA	5	0
92	FRIPTOT.G	FR INDUSTRIAL PRODUCTION - VOLA	5	0
93	EU18	EK PRODUCTION - TOTAL INDUSTRY EXCL. CONSTRUCTION - VOLA	5	0
94	BDNEWORDE	BD MANUFACTURING ORDERS - SADJ	5	0
95	BDRVNCARP	BD NEW PASSENGER CAR REGISTRATIONS - VOLN	5	1
96	BGACECARP	BG NEW PASSENGER CAR REGISTRATIONS - VOLN	5	1
97	ESCAR...O	ES REGISTRATIONS: PASSENGER CAR - VOLA	5	0
98	FRCARREGO	FR NEW CAR REGISTRATIONS (CAL ADJ) -VOLA	5	0
99	FRHCONMFD	FR HOUSEHOLD CONSUMPTION - MANUFACTURED GOODS - CONA	5	0
100	FRHCONDGD	FR HOUSEHOLD CONSUMPTION - DURABLE GOODS - CONA	5	0
101	ITNEWORDF	IT NEW ORDERS - NADJ	5	1
102	ITRETTOTF	IT RETAIL SALES - NADJ	5	1
103	BDCNFCONQ	BD CONSUMER CONFIDENCE INDICATOR - GERMANY - SADJ	2	0
104	BGCNFCONQ	BG BNB CONS. SVY.: CONSUMER CONFIDENCE INDICATOR (EP) - SADJ	2	0
105	BGCNFBUSQ	BG BUSINESS INDICATOR SURVEY - ECONOMY - SADJ	2	0
106	BGEUSIOBQ	BG IND.: OVERALL - ORD BOOKS - SADJ	2	0
107	BG000183Q	BG BNB BUS. SVY. - MANUFACTURING - NOT SMOOTHED - SADJ	2	0
108	BG000186Q	BG BNB BUS. SVY. - BUILDING - NOT SMOOTHED - SADJ	2	0
109	BG000189Q	BG BNB BUS. SVY. - TRADE - NOT SMOOTHED - SADJ	2	0
110	BGSURECSQ	BG BNB CONS.SVY.: ECON.SITUATION- FCST. OVER NEXT 12 MONTHS - SADJ	2	0
111	BGSURPUHQ	BG BNB CONS.SVY.: MAJOR HH.PURCH-FCST.OVER NEXT 12 MONTHS(EP)	2	0
112	ESINT384R	ES PRODUCTION LEVEL - INDUSTRY - NADJ	2	0
113	FRINDSYNQ	FR SURVEY: MANUFACTURING - SYNTHETIC BUSINESS INDICATOR - SADJ	2	0
114	FRSURPMPQ	FR SURVEY: MANUFACTURING OUTPUT - RECENT OUTPUT TREND - SADJ	2	0
115	FRSURGMPQ	FR SURVEY: MANUFACTURING OUTPUT - ORDER BOOK & DEMAND - SADJ	2	0
116	FRSURGPDQ	FR SURVEY: MANUFACTURING OUTPUT LEVEL - GENERAL OUTLOOK - SADJ	2	0
117	FRSURTMPQ	FR SURVEY: MANUFACTURING OUTPUT - PERSONAL OUTLOOK - SADJ	2	0
118	ITHHFECSR	IT HOUSEHOLD CONFIDENCE SURVEY: FUTURE FINANCIAL POSITION - NADJ	2	0
119	ITCNFCONQ	IT HOUSEHOLD CONFIDENCE INDEX - SADJ	5	0
120	NLCNFBUSQ	NL CBS MFG. SVY.: PRODUCER CONFIDENCE INDEX - SADJ	2	0
121	NLEUSCPCR	NL CONSUMER SURVEY: MAJOR PURCH.OVER NEXT 12 MONTHS-NETHERLANDS	2	0
122	EKCNFBUSQ	EK INDUSTRIAL CONFIDENCE INDICATOR - EA - SADJ	2	0
123	EMEUSCCIQ	EM CONSUMER CONFIDENCE INDICATOR - EA - SADJ	2	0
124	EKEUSIPAQ	EK INDUSTRY SURVEY: PRODUCTION EXPECTATIONS (EA) - SADJ	2	0
125	EKEUBCI.R	EK BUSINESS CLIMATE INDICATOR-COMMON FACTOR IN IND. (EA) - NADJ	2	0
126	EKEUSESIG	EK ECONOMIC SENTIMENT INDICATOR (EA18) - VOLA	5	0
127	EMGBOND.	EM GOVERNMENT BOND YIELD - 10 YEAR	2	0
128	EMECB2Y.	EM GOVERNMENT BOND YIELD - 2 YEAR	2	0
129	EMECB3Y.	EM GOVERNMENT BOND YIELD - 3 YEAR	2	0
130	EMECB5Y.	EM GOVERNMENT BOND YIELD - 5 YEAR	2	0

131	EMECB7Y.	EM GOVERNMENT BOND YIELD - 7 YEAR	2	0
132	BDESSFUB	BD HARMONISED GOVERNMENT 10-YEAR BOND YIELD	2	0
133	FRESSFUB	FR HARMONISED GOVERNMENT 10-YEAR BOND YIELD	2	0
134	ESESSFUB	ES HARMONISED GOVERNMENT 10-YEAR BOND YIELD	2	0
135	BGESSFUB	BG HARMONISED GOVERNMENT 10-YEAR BOND YIELD	2	0
136	ITESSFUB	IT HARMONISED GOVERNMENT 10-YEAR BOND YIELD	2	0
137	ITINTER3	IT INTERBANK DEPOSIT RATE-AVERAGE ON 3-MONTHS DEPOSITS	2	0
138	MSEROP$ E	MSCI EUROPE U$ - PRICE INDEX	5	0
139	INDGSIT& E	ITALY-DS Inds Gds & Svs - PRICE INDEX	5	0
140	INDGSBD& E	GERMANY-DS Inds Gds & Svs - PRICE INDEX	5	0
141	INDGSFR& E	FRANCE-DS Inds Gds & Svs - PRICE INDEX	5	0
142	INDUSBD E	GERMANY-DS Industrials - PRICE INDEX	5	0
143	INDUSFR E	FRANCE-DS Industrials - PRICE INDEX	5	0
144	INDUSIT E	ITALY-DS Industrials - PRICE INDEX	5	0
145	FINANFR E	FRANCE-DS Financials - PRICE INDEX	5	0
146	FINANBD E	GERMANY-DS Financials - PRICE INDEX	5	0
147	FINANIT E	ITALY-DS Financials - PRICE INDEX	5	0
148	CNSMGFR E	FRANCE-DS Consumer Gds - PRICE INDEX	5	0
149	CNSMGBD E	GERMANY-DS Consumer Gds - PRICE INDEX	5	0
150	CNSMGIT E	ITALY-DS Consumer Gds - PRICE INDEX	5	0
151	OILGSEM E	EMU-DS Oil & Gas - PRICE INDEX	5	0
152	BMATREM E	EMU-DS Basic Mats - PRICE INDEX	5	0
153	INDUSEM E	EMU-DS Industrials - PRICE INDEX	5	0
154	RITDVEM E	EMU-DS Divers. REITs - PRICE INDEX	5	0
155	CNSMGEM E	EMU-DS Consumer Gds - PRICE INDEX	5	0
156	HLTHCEM E	EMU-DS Health Care - PRICE INDEX	5	0
157	TELCMEM E	EMU-DS Telecom - PRICE INDEX	5	0
158	UTILSEM E	EMU-DS Utilities - PRICE INDEX	5	0
159	FINANEM E	EMU-DS Financials - PRICE INDEX	5	0
160	CNSMSEM E	EMU-DS Consumer Svs - PRICE INDEX	5	0
161	TECNOEM E	EMU-DS Technology - PRICE INDEX	5	0
162	EMSHRPRCF	EM DATASTREAM EURO SHARE PRICE INDEX (MONTHLY AVERAGE) - NADJ	5	0
163	BDMLM006Q	BD REGISTERED UNEMPLOYMENT: RATE (ALL PERSONS) - SADJ	2	0
164	BDMLM005O	BD REGISTERED UNEMPLOYMENT: LEVEL (ALL PERSONS) - VOLA	5	0
165	ITMLFT15O	IT HARMONISED UNEMPLOYMENT: LEVEL, ALL PERSONS (ALL AGES) - VOLA	5	0
166	ITMLRT16Q	IT HARMONISED UNEMPLOYMENT: RATE, ALL PERSONS (ALL AGES) - SADJ	2	0
167	ITMLRT14Q	IT HARMONISED UNEMPLOYMENT: RATE, ALL PERSONS (AGES 15-24) - SADJ	2	0
168	ITMLRF16Q	IT HARMONISED UNEMPLOYMENT: RATE, FEMMES (ALL AGES) - SADJ	2	0
169	ITMLRM16Q	IT HARMONISED UNEMPLOYMENT: RATE, HOMMES (ALL AGES) - SADJ	2	0
170	FRESTUNPO	FR UNEMPLOYMENT: TOTAL - TOTAL - VOLA	5	0
171	FRMLRT14Q	FR HARMONISED UNEMPLOYMENT: RATE, ALL PERSONS (AGES 15-24) - SADJ	2	0
172	FRMLRT15Q	FR HARMONISED UNEMPLMT.: RATE,ALL PERSONS(AGES 25 AND OVER) - SADJ	2	0
173	FRMLRT16Q	FR HARMONISED UNEMPLOYMENT: RATE, ALL PERSONS (ALL AGES) -SADJ	2	0
174	FRMLRF16Q	FR HARMONISED UNEMPLOYMENT: RATE, FEMMES (ALL AGES) - SADJ	2	0
175	FRMLRm16Q	FR HARMONISED UNEMPLOYMENT: RATE, HOMMES (ALL AGES) - SADJ	2	0
176	ESMLM005O	ES HARMONISED UNEMPLOYMENT: LEVEL, ALL PERSONS (ALL AGES) - VOLA	5	0

The Transformation Codes, *Tcode*, in table A.1, are defined as reported below:

$$
\begin{aligned}
&z_{it} = x_{it}\delta(T_{code} - 1) + [(1 - L)x_{it}]\delta(T_{code} - 2) \\
&+[(1 - L)(1 - L^{12})x_{it}]\delta(T_{code} - 3) + \ln(x_{it})\delta(T_{code} - 4) \\
&+(1 - L)\ln(x_{it})\delta(T_{code} - 5) + (1 - L)(1 - L^{12})\ln(x_{it})\delta(T_{code} - 6) \\
&+(1 - L)^2\ln(x_{it})\delta(T_{code} - 7),
\end{aligned}
\tag{A.1}
$$

where x_{it} represents the raw time series x_i at time t and $\delta(.)$ the Dirac delta function.

Appendix B

The US dataset and transformations used in Chapter 2

In table B.1, the series which compose the US dataset are reported. *Tcode* identifies the tranformation (further details are given below). All time series are deseasonalized.

TABLE B1: List of the US time series.

	Mnemonic	Description	Tcode
1	USIPTOT.G	US INDUSTRIAL PRODUCTION - TOTAL INDEX VOLA	5
2	USIPMPROH	US INDL PROD - PRDS, TOTAL VOLN	5
3	USIPMFING	US INDL PROD - FINAL PRODUCTS, TOTAL VOLA	5
4	USIPMCOGG	US INDL PROD - CONSUMER GOODS VOLA	5
5	USIPMDUCG	US INDL PROD - DURABLE CONSUMER GOODS VOLA	5
6	USIPMNOCG	US INDL PROD - NONDURABLE CONSUMER GOODS VOLA	5
7	USIPMBUQG	US INDL PROD - BUSINESS EQUIPMENT VOLA	5
8	USIPMMATG	US INDL PROD - MATERIALS, TOTAL VOLA	5
9	USIPMDUMH	US INDL PROD - DURABLE GOODS MATERIALS VOLN	5
10	USIPMNDMG	US INDL PROD - NONDURB GOODS MATERIALS VOLA	5
11	USIPMFGSG	US INDUSTRIAL PRODUCTION - MANUFACTURING (SIC) VOLA	5
12	USIP512UG	US INDL PROD - RESIDENTIAL UTILITIES VOLA	5
13	USIP512FG	US INDL PROD - FUELS VOLA	5
14	NAPM	ISM Manufacturing: PMI Composite Index	1
15	MCUMFN	Capacity Utilization: Manufacturing (NAICS)	1
16	AHE: goods	AVG HRLY EARNINGS, PROD WRKRS, NONFARM - GOODS-PRODUCING	6
17	AHE: const	AVG HRLY EARNINGS, PROD WRKRS, NONFARM - CONSTRUCTION	6
18	AHE: mfg	AVG HRLY EARNINGS, PROD WRKRS, NONFARM - MFG	6
19	Real AHE: goods	REAL AVG HRLY EARNINGS, PROD WRKRS, NONFARM - GOODS-PRODUCING	5
20	Real AHE: const	REAL AVG HRLY EARNINGS, PROD WRKRS, NONFARM - CONSTRUCTION	5
21	Real AHE: mfg	REAL AVG HRLY EARNINGS, PROD WRKRS, NONFARM - MFG	5
22	USEMIP..O	US EMPLOYED - TOTAL PRIVATE VOLA	5
23	USEMPG..O	US EMPLOYED - GOODS-PRODUCING VOLA	5
24	CES1021000001	EMPLOYEES, NONFARM - MINING	5
25	USEM23..O	US EMPLOYED - CONSTRUCTION VOLA	5
26	USEMPMANO	US EMPLOYED - MANUFACTURING VOLA	5
27	USEMIMD.O	US EMPLOYED - DURABLE GOODS VOLA	5
28	USEMIMN.O	US EMPLOYED - NONDURABLE GOODS VOLA	5
29	USEMPS..O	US EMPLOYED - SERVICE-PROVIDING VOLA	5
30	USEMIT..O	US EMPLOYED - TRADE, TRANSPORTATION, & UTILITIES VOLA	5
31	USEM42..O	US EMPLOYED - WHOLESALE TRADE VOLA	5
32	USEMIR..O	US EMPLOYED - RETAIL TRADE VOLA	5

33	USEMIF..O	US EMPLOYED - FINANCIAL ACTIVITIES VOLA	5
34	USEMIG..O	US EMPLOYED - GOVERNMENT VOLA	5
35	USEMPTOTO	US TOTAL CIVILIAN EMPLOYMENT VOLA	5
36	USEMPALLO	US EMPLOYED - NONFARM INDUSTRIES TOTAL (PAYROLL SURVEY) VOLA	5
37	UNRATE	Civilian Unemployment Rate	2
38	UEMPMEAN	Average (Mean) Duration of Unemployment	2
39	USUNWK5.O	US UNEMPLOYED FOR LESS THAN 5 WEEKS VOLA	5
40	USUNWK14O	US UNEMPLOYED FOR 5 TO 14 WEEKS VOLA	5
41	USUNPLNGE	US UNEMPLOYED FOR 15 WEEKS OR MORE VOLA	5
42	USUNWK26O	US UNEMPLOYED FOR 15 TO 26 WEEKS VOLA	5
43	USUNWK27O	US UNEMPLOYED FOR 27 WEEKS & OVER VOLA	5
44	USHKPG..P	US AVG HOURS PROD WRKRS-GOODS-PRODUCING VOLN	1
45	USHXPMANO	US AVG OVERTIME HOURS - MANUFACTURING VOLA	2
46	USHOUSATE	US NEW PRIVATE HOUSING UNITS AUTHORIZED BY BLDG.PERMIT (AR) VOLA	4
47	USHOUSE.O	US NEW PRIVATE HOUSING UNITS STARTED (AR) VOLA	4
48	HOUSTNE	Housing Starts in Northeast Census Region	4
49	HOUSTMW	Housing Starts in Midwest Census Region	4
50	HOUSTS	Housing Starts in South Census Region	4
51	HOUSTW	Housing Starts in West Census Region	4
52	FEDFUNDS	Effective Federal Funds Rate	2
53	TB3MS	3-Month Treasury Bill: Secondary Market Rate	2
54	TB6MS	6-Month Treasury Bill: Secondary Market Rate	2
55	GS1	1-Year Treasury Constant Maturity Rate	2
56	GS5	5-Year Treasury Constant Maturity Rate	2
57	GS10	10-Year Treasury Constant Maturity Rate	2
58	AAA	Moody's Seasoned Aaa Corporate Bond Yield	2
59	BAA	Moody's Seasoned Baa Corporate Bond Yield	2
60	Sfygm6	fygm6-fygm3	1
61	Sfygt1	fygt1-fygm3	1
62	Sfygt10	fygt10-fygm3	1
63	sFYAAAC	FYAAAC-Fygt10	1
64	sFYBAAC	FYBAAC-Fygt10	1
65	M1SL	M1 Money Stock	6
66	MZMSL	MZM Money Stock	6
67	M2SL	M2 Money Stock	6
68	BOGAMBSL	Board of Governors Monetary Base, Adj. for Changes in Res. Requirements	6
69	USTOTRSAB	US TOTAL RESERVES OF DEPOSITORY INSTITUTIONS CURA	6
70	USNBRRSAB	US NONBORROWED RESERVES OF DEPOSITORY INSTITUTIONS CURA	3
71	BUSLOANS	Commercial and Industrial Loans at All Commercial Banks	6
72	NONREVSL	Total Nonrevolving Credit Owned and Securitized, Outstanding	6
73	USCP...CE	US CHAIN-TYPE PRICE INDEX FOR PERSONAL CONSMPTN.EXPENDITURE SADJ	6
74	USCONDUCE	US CHAIN-TYPE PRICE INDEX FOR PCE - DURABLES SADJ	6
75	USCONNDCE	US CHAIN-TYPE PRICE INDEX FOR PCE - NONDURABLE GOODS SADJ	6
76	USCONSRCE	US CHAIN-TYPE PRICE INDEX FOR PCE - SERVICES SADJ	6
77	CPIAUCSL	CPI for All Urban Consumers: All Items	6
78	CPILFESL	CPI for All Urban Consumers: All Items Less Food & Energy	6
79	PCEPILFE	PCE Excluding Food and Energy (Chain-Type Price Index)	6
80	PPIFGS	PPI: Finished Goods	6
81	PPIFCG	PPI: Finished Consumer Goods	6
82	PPIITM	PPI: Intermediate Materials: Supplies & Components	6

83	PPICRM	PPI: Crude Materials for Further Processing	6
84	PWCMSAR	Real PPI:CRUDE MATERIALS (82=100,SA) (PWSMSA/PCEPILFE)	5
85	CRBSPOT	CRB BLS Spot Index (1967=100) - PRICE INDEX	6
86	PSCCOMR	Real SPOT MARKET PRICE INDEX (PSCCOM/PCEPILFE)	5
87	USPCIPCOF	US PPI - CRUDE PETROLEUM NADJ	6
88	PW561R	PPI Crude (Relative to Core PCE) (pw561/PCEPiLFE)	5
89	NAPMPRI	ISM Manufacturing: Prices Index	1
90	EXRUS	UNITED STATES;EFFECTIVE EXCHANGE RATE(MERM)(INDEX NO.)	5
91	EXSZUS	Switzerland / U.S. Foreign Exchange Rate	5
92	EXJPUS	Japan / U.S. Foreign Exchange Rate	5
93	EXUSUK	U.S. / U.K. Foreign Exchange Rate	5
94	EXCAUS	Canada / U.S. Foreign Exchange Rate	5
95	US500STK	US STANDARD & POOR'S INDEX OF 500 COMMON STOCKS(MONTHLY AVE NADJ	5
96	USS&PIND	US STANDARD & POORS' SHARE PRICE INDEX - INDUSTRIALS (EP)	5
97	USSPDIVY	US STANDARD AND POORS' 500 COMPOSITE - DIVIDEND YLD	2
98	USSPRPER	US STANDARD AND POORS' 500 COMPOSITE - REAL P/E RATIO	2
99	USSHRPRCF	US DOW JONES INDUSTRIALS SHARE PRICE INDEX (EP) NADJ	5
100	USUMCONEH	US UNIV OF MICHIGAN CONSUMER SENTIMENT - EXPECTATIONS VOLN	2
101	NAPM	ISM Manufacturing: PMI Composite Index	1
102	NAPMNOI	ISM Manufacturing: New Orders Index	1
103	NAPMSDI	ISM Manufacturing: Supplier Deliveries Index	1
104	NAPMII	ISM Manufacturing: Inventories Index	1
105	USIPNOMAD	US MANUFACTURERS NEW ORDERS, CONSUMER GOODS & MATERIALS CONA	5
106	USNONDCGD	US MANUFACTURERS NEW ORDERS - NONDEFENSE CAPITAL GOODS CONA	5
107	CPIULFSL	CPI for All Urban Consumers: All Items Less Food	6
108	CUSR0000SA0L5	CPI for All Urban Consumers: All items less medical care	6
109	CUSR0000SA0L2	CPI for All Urban Consumers: All items less shelter	6
110	CPIAPPSL	CPI for All Urban Consumers: Apparel	6
111	CUSR0000SAC	CPI for All Urban Consumers: Commodities	6
112	CUSR0000SAD	CPI for All Urban Consumers: Durables	6
113	CPIMEDSL	CPI for All Urban Consumers: Medical Care	6
114	CUSR0000SAS	CPI for All Urban Consumers: Services	6
115	CPITRNSL	CPI for All Urban Consumers: Transportation	6

The Transformation Codes, *Tcode*, in table B.1, are defined as reported below:

$$\begin{aligned} &z_{it} = x_{it}\delta(T_{code}-1) + [(1-L)x_{it}]\delta(T_{code}-2) \\ &+[(1-L)(1-L^{12})x_{it}]\delta(T_{code}-3) + \ln(x_{it})\delta(T_{code}-4) \\ &+(1-L)\ln(x_{it})\delta(T_{code}-5) + (1-L)(1-L^{12})\ln(x_{it})\delta(T_{code}-6), \end{aligned} \tag{B.1}$$

where x_{it} represents the raw time series x_i at time t and $\delta(.)$ the Dirac delta function.

Appendix C

The EU/US short dataset and transformations used in Chapter 3

In table C.1, the series which compose the short EU/US dataset are reported. *Tcode* identifies the tranformation (further details are given below).

TABLE C1: List of the EU/US time series.

	Group	EU/US	Long Description	Tcode
1	Real	EU	IP total excluded constructions	4
2	Real	EU	New passenger cars	4
3	Real	EU	New orders	4
4	Real	EU	Retail turnover (deflated)	4
5	Survey	EU	ECS Economic Sentiment indicator	3
6	Survey	EU	PMS Manufacturing: purchasing manager index	3
7	Real	EU	Unemployment Rate	3
8	Real	EU	Extra EA trade, export value	4
9	Financial	EU	Dow Jones Euro Stoxx Price Index	4
10	Financial	EU	World market price of raw materials. Index total. Euro	4
11	Real	EU	UE GDP	4
12	Real	EU	Employment	4
13	Survey	EU	ECS Capacity utilisation	3
14	Real	US	US GDP	4

The Transformation Codes, *Tcode*, in table C.1, are defined as reported below:

$$\begin{aligned} &z_{it} = x_{it}\delta(T_{code} - 1) + \ln(x_{it})\delta(T_{code} - 2) \\ &+(1 - L)x_{it}\delta(T_{code} - 3) + (1 - L)\ln(x_{it})\delta(T_{code} - 4), \end{aligned} \tag{C.1}$$

where x_{it} represents the raw time series x_i at time t and $\delta(.)$ the Dirac delta function.

Bibliography

Arulampalam, M. S. et al. (2002). "A tutorial on particle filters for online nonlinear/non-Gaussian Bayesian tracking." In: *IEEE Transactions on signal processing* 50.2, pp. 174–188.

Banbura, M. and M. Modugno (2014). "Maximum likelihood estimation of factor models on datasets with arbitrary pattern of missing data". In: *Journal of Applied Econometrics* 29.1, pp. 133–160.

Banbura, M. and G. Rünstler (2011). "A look into the factor model black box: publication lags and the role of hard and soft data in forecasting GDP." In: *International Journal of Forecasting* 27.2, pp. 333–346.

Boivin, J. and S. Ng (2005). "Understanding and comparing factor-based forecasts". In: *International Journal of Central Banking* 1.3, pp. 117–151.

Bork, L. (2009). *Estimating US monetary policy shocks using a factor-augmented vector autoregression: An EM algorithm approach.* Tech. rep. CRATES Research Paper.

Calhoun, G. and G. Elliott (2012). *Why Do Nonlinear Models Provide Poor Macroeconomic Forecasts?* Tech. rep. Seventh ECB Workshop on Forecasting Techniques-New Directions for Forecasting, Frankfurt am Main, Germany.

D'Agostino, A. and D. Giannone (2012). "Comparing Alternative Predictors Based on Large-Panel Factor Models". In: *Oxford bulletin of economics and statistics* 74.2, pp. 306–326.

Del Moral, P., L. Kallel, and J. Rowe (2001). "Modeling genetic algorithms with interacting particle systems". In: *Revista de Matematica. Teoria y aplicaciones.* 8.2.

Della Marra, F. (2017). "A forecasting performance comparison of dynamic factor models based on static and dynamic methods". In: *Communications in Applied and Industrial Mathematics* 8.1, pp. 44–66. ISSN: 2038-0909.

Dempster, A. P., N. M. Laird, and D. B. Rubin (1977). "Maximum likelihood from incomplete data via the EM algorithm". In: *ournal of the royal statistical society. Series B (methodological)*, pp. 1–38.

Doz, C., D. Giannone, and L. Reichlin (2012). "A quasimaximum likelihood approach for large, approximate dynamic factor models". In: *Review of economics and statistics* 94.4, pp. 1014–1024.

Forni, M., D. Giannone, et al. (2009). "Opening the black box: Structural factor models with large cross sections". In: *Econometric Theory* 25.05, pp. 1319–1347.

Forni, M., A. Giovannelli, et al. (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". In: *Review of Economics and Statistics* 82.4, pp. 540–554.

— (2005). "The generalized dynamic factor model: One-sided estimation and forecasting". In: *Journal of the American Statistical Association* 100, pp. 830–840.

Forni, M., M. Hallin, M. Lippi, and P. Zaffaroni (2015). "Dynamic factor model with infinite dimensional factor space: representation". In: *Journal of Econometrics* 185, pp. 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, Einaudi Institute for Economics and Finance.

Forni, M. and M. Lippi (2001). "The generalized dynamic factor model: representation theory". In: *Econometric theory* 17.06, pp. 1113–1141.

Giacomini, R. and B. Rossi (2010). "Forecast comparisons in unstable environments". In: *Journal of Applied Econometrics* 25.4, pp. 595–620.

Giacomini, R. and H. White (2006). "Tests of Conditional Predictive Ability". In: *Econometrica* 74.6, pp. 1545–1578.

Gopaluni, R. B., B. T. Schön, and A. G. Wills, eds. (2009). *Particle filter approach to nonlinear system identification under missing observations with a real application.* Vol. 42. 10.

Gustaffson, F. (2010). "Particle filter theory and practice with positioning applications." In: *IEEE Aerospace and Electronic Systems Magazine.* 25.7, pp. 53–82.

Habibnia, A. (2017). "Nonlinear Forecasting Using a Large Number of Predictors." PhD thesis. London School of Economics.

Hallin, M. and M. Lippi (2013). "Factor models in high-dimensional time seriesA time-domain approach". In: *Stochastic Processes and their Applications* 123.7, pp. 2678–2695.

Hartl, R. F. and R. K. Belew (1990). *A global convergence proof for a class of genetic algorithms.* Tech. rep. University of Technology, Vienna.

Kapetanios, G. (2007). "Variable selection in regression models using nonstandard optimisation of information criteria." In: *Computational Statistics & Data Analysis* 52.1, pp. 4–15.

Kapetanios, G., G. M. Marcellino, and F. Papailias (2014). *Variable selection for large unbalanced datasets using non-standard optimisation of information criteria and variable reduction methods*. Tech. rep. Quantf research.

Marcellino, G. M. (2004). "Forecast pooling for European macroeconomic variables." In: *Oxford Bulletin of Economics and Statistics* 66.1, pp. 91–112.

Mariano, R. S. and Y. Murasawa (2003). "A new coincident index of business cycles based on monthly and quarterly series." In: *Journal of Applied Econometrics* 18.4, pp. 427–443.

Morinaka, Y. et al. (2001). *The L-index: An indexing structure for efficient subsequence matching in time sequence databases.* Tech. rep. Proc. 5th Pacific Asia Conerence on Knowledge Discovery and Data Mining.

Reijer, A. H. J. den et al. (2005). *Forecasting Dutch GDP using large scale factor models*. Tech. rep. Netherlands Central Bank, Research Department.

Schön, T. B., A. Wills, and B. Ninness (2011). "System identification of nonlinear state-space models." In: *Automatica* 47.1, pp. 39–49.

Schumacher, C. (2007). "Forecasting German GDP using alternative factor models based on large datasets". In: *Journal of Forecasting* 26.4, pp. 271–302.

Stock, J. H. and M. Watson (1998). *A Comparison of Linear and Nonlinear Univariate Models for Forecasting Macroeconomic Time Series.* Tech. rep. National Bureau of Economic Research.

Stock, J. H. and M. W. Watson (2002a). "Forecasting using principal components from a large number of predictors". In: *Journal of the American Statistical Association* 97.460, pp. 1167–1179.

— (2002b). "Macroeconomic forecasting using diffusion indexes". In: *Journal of Business & Economic Statistics* 20.2, pp. 147–162.

Zhang, X. P. et al. (2014). *Convergence analysis of multiple imputations particle filters for dealing with missing data in nonlinear problems.* Tech. rep. Circuits and Systems (ISCAS), 2014 IEEE International Symposium on IEEE.