Short-term GDP forecasting with a mixed frequency dynamic factor model with stochastic volatility*

Massimiliano Marcellino
European University Institute and Bocconi University
massimiliano.marcellino@EUI.eu

Mario Porqueddu
Banca d'Italia
mario.porqueddu@bancaditalia.it

Fabrizio Venditti
Banca d'Italia
fabrizio.venditti@bancaditalia.it

April 11, 2012

JEL Classification: E32, C22, E27

Abstract

In this paper we develop a mixed frequency dynamic factor model featuring stochastic shifts in the volatility of both the latent common factor and the idiosyncratic components. We take a Bayesian perspective and derive a Gibbs sampler to obtain the posterior density of the model parameters. This new tool is then used to investigate business cycle dynamics and for forecasting GDP growth at short-term horizons in the euro area. We discuss three sets of empirical results. First we use the model to evaluate the impact of macroeconomic releases on point and density forecast accuracy and on the width of forecast intervals. Second, we show how our setup allows to make a probabilistic assessment of the contribution of releases to forecast revisions. Third we design a pseudo out of sample forecasting exercise and examine point and density forecast accuracy. In line with findings in the Bayesian Vector Autoregressions (BVAR) literature we find that stochastic volatility contributes to an improvement in density forecast accuracy.

*We would like to thank participants at seminars at Bank of Italy, the 2011 RCEA Bayesian Workshop, the 2011 EABCN Workshop and the 2011 CFE Conference, for useful comments on a previous draft. The views expressed in the articles are those of the authors and do not involve the responsibility of the Bank of Italy.

1 Introduction

The conduct of monetary and fiscal policy relies on the timely assessment of current and future economic conditions. The task of providing an accurate picture of the current cyclical position is significantly plagued by the delay with which crucial economic indicators are released. GDP data, for example, are usually published with a 45 days delay both in the US and in the euro area. Important quantitative monthly indicators, like industrial production indexes, suffer more or less from the same publication delay. Survey data, on the other hand, provide very timely information as they are published roughly at the end of the reference month. Unfortunately, forecasts based on qualitative data only are known to be much less reliable than predictions based on quantitative information, see Banbura and Runstler (2007). The econometric literature has progressed significantly in the field of short-term forecasting in the past decade, and a number of tools have been developed, capable of dealing with the asynchronous timing of data releases, integrating data at different frequencies and dissecting the information content of monthly releases for tracking quarterly variables. Small and large scale factor models, in particular, have become the workhorse for short term forecasting. On the small scale side, building on the Stock and Watson (1989) coincident indicator, Mariano and Murasawa (2003), henceforth MM03, have proposed a unified framework for modeling quarterly GDP together with monthly indicators. The approach has recently been extended by Camacho and Perez-Quiros (2010) to accommodate real time issues and different GDP releases. On the large data side, research by Angelini et al. (2008) and Banbura and Modugno (2010) has documented the predictive content of a large number of indicators for GDP growth and also introduced new tools to link monthly data releases to GDP forecast revisions. These models are nowadays used on a regular basis to inform decision makers both at Central Banks as well as in private institutions.[1]

Although the literature has moved very rapidly, there are still some gaps between the demands posed by policy makers and the answers that the models discussed above can provide. In particular, policy makers have become more and more interested in having not only point forecasts, but also a model based assessment of the uncertainty surrounding the outlook. This is testified by the number of Central Banks that have started publishing fan charts and confidence bands around their medium/long term forecasts (Bank of England, Bank of Canada, Norges Bank, South Africa Reserve, the Sveriges Riksbank, the Bank of Italy and the US Fed). Despite the growing preference for a probabilistic assessment of economic projections, however, the focus of short-term forecasting models is still on point forecasts.

[1] An alternative approach to short-term forecasting with mixed frequency data is based on the MIDAS regressions introduced by Ghysels et al. (2004), see e.g. Clements and Galvão (2008), Foroni and Marcellino (2012) and Marcellino and Schumacher (2008) for macroeconomic applications. Mixed frequency VARs provide a third option, see e.g. Kuzin et al. (2011).

Another open issue in the field of short-term forecasting relates to parameter instability. As economic systems evolve and are hit by large shocks, the link between different indicators is likely to change over time, requiring some flexibility in the models parameters. The issue of forecast failure in the presence of structural breaks, which has been explored extensively in the case of points forecasts, has been recently extended to density forecasting. In particular Jore et al. (2010) show that changes in the underlying data generating process can severely hinder the accuracy of density forecasts produced with time invariant models. If structural breaks, however, are sufficiently large and distant in the past they do not pose serious problems. They can be easily detected with standard statistical tests and parameter instability can then be either incorporated in the model or simply bypassed by splitting the sample or adopting a rolling estimation scheme. These strategies, however, are not viable if breaks are, rather than large and discrete, small and continuous, a form of parameter time variation that has received a lot of attention in the macro empirical literature in the past decade. Indeed a number of papers, mainly concerned with structural monetary analysis, have documented the inadequacy of constant parameter models for describing macroeconomic data and have called for some form of slow, continuous, time variation in model parameters, see Benati and Surico (2008), Cogley and Sargent (2005) and Primiceri (2005). The message coming from this literature has raised little attention in the density forecast literature until the recent paper by Clark (2009), who shows that allowing for stochastic shifts in the volatility of the shocks in a BVAR significantly increases density forecast accuracy for a number of US variables.

In this paper we take stock of these issues and develop a mixed frequency small scale factor model that is suitable for producing density forecasts and that allows for time variation in some of the parameters. We start off with the basic setup of MM03 and twist it in two directions. Our first step consists of casting the model in a bayesian estimation framework. By treating the model parameters as random variables we can draw a distribution of forecasts from the predictive density making the model easily suitable for producing density forecasts. Secondly, following Clark (2009), we extend the model to allow for random shifts in the volatility of the underlying shocks. To clarify the expected gains of this modeling choice let us briefly describe the setup. In our factor model each variable is treated as the sum of two components, a latent factor which is common to all the variables and governs the amount of correlation across components, and an idiosyncratic component which is uncorrelated across variables. Dynamics are introduced by letting the common and the idiosyncratic components follow an autoregressive process subject to random shocks. Our innovation consists of allowing the variances of these shocks to vary continuously over time rather than being constant as in MM03.[2] We expect

[2]See Baumeister et al. (2010), Del Negro and Otrok (2008) and Korobilis (2009) for examples of bayesian dynamic factor models with stochastic volatility. In their models they do not handle mixed frequency data.

the model to boost the variance of the shocks in more turbulent times, hence producing wider confidence bands and providing a more accurate assessment of the uncertainty surrounding the median forecast. On the other hand, the richer structure of our model implies a much heavier parametrization, posing a trade-off between model flexibility and model parsimony, an issue that can be particularly important in relatively short samples and that can only be evaluated empirically.

After showing how to estimate the model we turn to an empirical application in which we use a small number of monthly indicators to predict quarterly GDP growth in the euro area. We present three sets of results. First we extend the approach proposed by Giannone et al. (2008), henceforth GRS, and show how successive macroeconomic releases not only improve point forecast accuracy but also increase the precision of density forecasts and reduce the width of forecast intervals. Second, we illustrate how, in a given quarter, our new tool can be applied not only to interpret the *news* content of monthly releases, like in Banbura and Modugno (2010), but also to assess how much 'confidence' the model places on the revisions implied by the release of monthly indicators. Third, we design a (pseudo) real time out of sample forecasting exercise and evaluate both the point and density forecasts produced by the model. In line with Clark (2009) we find that the introduction of stochastic volatility leads to an improvement of both point and density forecast accuracy.

The paper is structured as follows. In section 2 we describe the model. In section 3 we discuss the main steps of the Gibbs sampler used for simulating the posterior distribution of the parameters. Section 4 presents the empirical application and section 5 concludes.

2 The model

In this section we spell out the details of our model. Let $Y_{q,t}$ be a quarterly series, which can be seen as a monthly variable with its value associated to the third month of the quarter and missing observations in the first two months, and $Y_{m,t}$ a vector of k monthly series $Y_{mj,t}$, for $j = 1, 2, \ldots, k$ (from this point onwards we use the convention that whenever we write mj we mean the j_{th} element in the vector of monthly variables, for $j = 1, 2, \ldots, k$).[3] Now let $Y_{q,t}$ be the geometric mean of a latent random variable $Y^{\star}_{qt}$ such that:

$$lnY_{q,t} = \frac{1}{3}(lnY^{\star}_{q,t} + lnY^{\star}_{q,t-1} + lnY^{\star}_{q,t-2}) \tag{1}$$

[3] Although the setup can easily accommodate multiple quarterly variables we confine the model description to the single quarterly variable case to keep the notation as simple as possible.

Filtering both sides with the filter $(1-L^3)$, after some simple manipulation yields:

$$y_{q,t} = \frac{1}{3}y^\star_{q,t} + \frac{2}{3}y^\star_{q,t-1} + y^\star_{q,t-2} + \frac{2}{3}y^\star_{q,t-3} + \frac{1}{3}y^\star_{q,t-4} \tag{2}$$

where small case letters indicate growth rates over the previous three months: $y_{q,t} = \Delta_3 lnY_{q,t}$. We assume a dynamic (single) factor model for the latent process $y^\star_{q,t}$ and the monthly observed variables $y_{m,t}$.[4]

The system of measurement equations is:

$$\begin{pmatrix} y^\star_{q,t} \\ y_{m,t} \end{pmatrix} = \begin{pmatrix} \alpha^\star_1 \\ \alpha_2 \end{pmatrix} + \begin{pmatrix} \beta_q f_t \\ \beta_m f_t \end{pmatrix} + \begin{pmatrix} u_{q,t} \\ u_{m,t} \end{pmatrix} \tag{3}$$

The law of motions of the factor and of the the idiosyncratic disturbances of the quarterly and monthly variables are described by the following:

$$\Phi_f(L)f_t = v_t e^{\lambda_{f,t}/2} \tag{4}$$

$$\Phi_q(L)u_{q,t} = \epsilon_{q,t}\sigma_q e^{\lambda_{q,t}/2} \tag{5}$$

$$\Phi_{mj}(L)u_{mj,t} = \epsilon_{mj,t}\sigma_{mj} e^{\lambda_{mj,t}/2} \qquad j = 1, \ldots, k \tag{6}$$

where v_t, $\epsilon_{q,t}$ and $\epsilon_{mj,t}$ are uncorrelated N(0,1) and the $\Phi_i(L)$ polynomials are lag polynomials of order p_i:

$$\Phi_i(L) \quad = \quad 1 - \phi^i_1 \mathrm{L} - \phi^i_2 \mathrm{L}^2 - \cdots - \phi^i_{p_i} \mathrm{L}^{p_i} \tag{7}$$

for $i = f, q, mj$. The log-volatilities $\lambda_{i,t}$ follow a driftless random walk:

$$\lambda_{i,t} = \lambda_{i,t-1} + \theta_{i,t}\sigma_{\lambda,i} \qquad \theta_{i,t} \sim N(0,1) \tag{8}$$

for $i = f, q, mj$, and are assumed to be independent across equations. The hypothesis that the innovations to the idiosyncratic components of the factor and of the observable are cross-sectionally uncorrelated is an important identifying assumption that forces all the comovement in the panel to occur through the common factor.

Since the variable $y^\star_{q,t}$ is not observed, the measurement equations can be rewritten in terms

[4]The use of more than one common factor does not pose any additional technical difficulty as it would simply result in an enlargement of the state vector in the State Space representation of the model.

of the observable variable $y_{q,t}$ using the identity (2):

$$\begin{pmatrix} y_{q,t} \\ y_{m,t} \end{pmatrix} = \begin{pmatrix} \alpha_1 \\ \alpha_2 \end{pmatrix} + \begin{pmatrix} \beta_1(\frac{1}{3}f_t + \frac{2}{3}f_{t-1} + f_{t-2} + \frac{2}{3}f_{t-3} + \frac{1}{3}f_{t-4}) \\ \beta_2 f_t \end{pmatrix} + \begin{pmatrix} \frac{1}{3}u_{q,t} + \frac{2}{3}u_{q,t-1} + u_{q,t-2} + \frac{2}{3}u_{q,t-3} + \frac{1}{3}u_{q,t-4} \\ u_{m,t} \end{pmatrix} \quad (9)$$

where $\alpha_1 = 3\alpha_1^\star$. A more compact state space representation of the model is the following:

$$\boldsymbol{y}_t = F\boldsymbol{\mu}_t \quad (10)$$

$$\boldsymbol{\mu}_t = H\boldsymbol{\mu}_{t-1} + \eta_t \qquad \eta_t \sim N(0, Q_t) \quad (11)$$

$$\Lambda_t = \Lambda_{t-1} + \zeta_t \qquad \zeta_t \sim N(0, \Xi) \quad (12)$$

where $\boldsymbol{y}_t$ collects both quarterly and monthly variables, the state vector $\boldsymbol{\mu}_t$ includes the unobserved factor f_t and the idiosyncratic components ($u_{q,t}$ and $u_{m,t}$), the matrix F collects the factor loadings, H collects the autoregressive parameters of the laws of motion of the unobserved factors and of the idiosyncratic components, the time varying variance matrix Q_t is a diagonal matrix whose diagonal elements are the variances $e^{\lambda_{f,t}}, \sigma_q^2 e^{\lambda_{q,t}}, \sigma_m^2 e^{\lambda_{mj,t}}$, Λ_t is the vector of drifting volatilities and Ξ is a diagonal matrix collecting the variances that determine the amount of time variation of the log-volatilities ($\sigma^2_{\lambda,i}, i = f, q, mj$).

This model nests the one proposed by MM03, which can be easily recovered from our more general setup by shutting off the drifting volatilities, that is by setting $\Lambda_0 = 0$ and $\Xi = 0$. In this case the matrix Q_t is replaced by its constant counterpart Q.

To identify the model parameters some restrictions need to be placed. First, the scale of the factor loadings and of the factors cannot be separately identified, so we restrict the variance of the errors of the common factors to be 1 (see equation 4). Second, we need to fix the scale of the stochastic volatilities, which enter as a multiplicative element of the constant variances (see equations 4 to 6) and therefore can not be separately identified. We follow Del Negro and Otrok (2008) and identify the stochastic volatilities by setting to zero their initial condition $\lambda_i = 0$.

3 Model Estimation

The model is estimated with Bayesian methods using a Metropolis within Gibbs sampling procedure. The Gibbs sampler simplifies the daunting task of obtaining draws of the model parameters from the joint posterior distribution to a number of more manageable problems, which involve sampling from the distribution of a subset of parameters conditional on the remaining ones and on the data, see Kim and Nelson (1999). The sampling algorithm consists of six blocks, which we briefly describe in what follows. More precise details on the Gibbs sampler can be found in Appendix A.

3.1 Steps 1 and 2: drawing F and the time constant elements of Q_t

Since the model disturbances are uncorrelated, elements of the F matrix can be drawn row by row (equation by equation). Take the i^{th} measurement equation:

$$y_{i,t} = F(i)\boldsymbol{\mu}_t = \beta(i, L)f_t + \Phi_i(L)^{-1}\epsilon_{i,t}\sigma_i e^{\lambda_{i,t}/2} \tag{13}$$

Conditioning on f_t, $\Phi_i(L)$ and $\lambda_{i,t}$, this is a standard regression with autocorrelated and heteroscedastic disturbances. Pre-multiplying by $\Phi_i(L)$ and dividing by $e^{\lambda_{i,t}/2}$ one obtains a standard regression model with homoscedastic, uncorrelated residuals. Positing a Normal-gamma conjugate prior, the conditional posterior for $\beta(i, L)$ and σ_i is also Normal-gamma.

3.2 Step 3: drawing H

The transition matrix H can also be drawn row by row. Take the i^{th} transition equation:

$$\mu_{i,t} = \sum_{j=1}^{p_i} \phi_j \mu_{i,t-j} + \eta_{i,t} \tag{14}$$

Conditioning on $\mu_{i,t}$ and on the ith element of the Q_t matrix ($q_{i,t}$), this is a regression with heteroscedastic residuals. The residuals can be whitened by dividing by $q_{i,t}$. Positing a Normal prior for the regression coefficients the conditional posterior is also Normal.[5]

3.3 Step 4 and 5: drawing the stochastic volatilities

There are a number of methods for drawing the stochastic volatilities $\lambda_{i,t}$ and the variance $\sigma_{\lambda,i}$. We employ the Jacquier et al. (1994) algorithm, which involves drawing from a log-normal density and a Metropolis acceptance step. Details on the algorithm can be found in Cogley and Sargent (2005), Appendix B.2.5.

[5] When drawing from the conditional posterior we discard explosive roots.

3.4 Step 6: drawing μ_t

Conditioning on all the other parameters and on the data, draws of the state vector can obtained via a state vector simulation smoother as in Carter and Kohn (1994) or with the disturbance smoother proposed by Koopman and Durbin (2003). We resort to the latter, which turns out to be slightly more efficient from a computational point of view.

4 Empirical application: short-term forecasts of euro area GDP

In our empirical application we apply the model to the problem of forecasting euro area GDP growth at short horizons. Our information set consists of nine indicators, namely our target variable, which is the rate of growth of quarterly GDP, two Industrial Production indicators (the total index and the index for the Pulp and Paper sector), four surveys (the Germany IFO Business Climate Index, the Composite Purchasing Manager Index for the euro area, the Michigan Consumer Sentiment for the US, the euro area Economic Sentiment Indicator), the bilateral US dollar euro exchange rate and a the difference between the 3 months and the 10 years spread on US Government Bonds. These indicators, listed in Table 1 were selected from a large pool of candidate series adapting the algorithm used by Camacho and Perez-Quiros (2010) to our Bayesian setting. More details can be found in Appendix B. The empirical specification of the model also follows closely the one proposed by Camacho and Perez-Quiros (2010). In this setting the Industrial Production indexes and the interest rate spread load on the common factor contemporaneously. Survey data, on the other hand, are treated as if they were in phase with the year-on-year growth rate of the Industrial Production index, therefore loading a 11 terms moving average of the common factor. We also let the bilateral exchange rate enter the model in year-on-year percentage growth, the rationale being that pricing to market is likely to buffer temporary exchange rate short term movements with a variation in profit margins so that only more persistent fluctuations impact on economic growth. The exact specification of the state space matrices can be found in Appendix C.

Our empirical analysis proceeds as follows. After a brief discussion of how the priors are set (4.1), we estimate the model on the full sample to gauge the relative contributions of the various indicators to the common factor and also to evaluate if the model actually captures any significant shifts in the variance of the common and idiosyncratic errors (4.2). We then turn to three empirical exercises. The first one regards the typical situation of a forecaster that is required to update her forecasts at each release of new data. In this context we set off by replicating with our model the analysis of *news* performed by GRS and evaluate how point forecast accuracy is affected by data releases. We take advantage of the Bayesian nature of our model and extend GRS results to examine how new data affects density forecast accuracy and the width of forecast intervals (4.3). We then turn to a different concept of *news*, introduced in the literature by Banbura and Modugno (2010). We show how our set up adds a new dimension to their tool, as one can use draws from the posterior to derive a measure of uncertainty around the *news* content of each data (or block of data) release (4.4). Our third set of results concerns a fully fledged out-of-ample forecast exercise in which we assess the point and density forecasting performance of our model (4.5).

4.1 Priors

To set the prior hyperparameters we follow Primiceri (2005) and retain a three years training sample. Since the model features an unobserved component that is common to the indicators, we start by getting an initial estimate of the common factor f_t as the *cross-sectional* average of the monthly indicators f_t^{start} over this training sample. Conditioning on this we then get an estimate of the factor loadings with an OLS regression of the indicators on f_t^{start}. The prior distributions of the factor loadings are then centered around this $\hat{\beta}_{OLS}$. The prior is flat as we set the variance to $10^3 V(\beta_{OLS})$. By regressing the residuals of these OLS regressions $u_{OLS,t}$ on their first two lags we also obtain an OLS estimate of the autoregressive parameters of the idiosyncratic shocks to the observable indicators. The prior distributions of the ϕ_s are then centered on this estimate and their variance is set to $10^3 V(\phi_{i,OLS})$, where $i = q, h, s$. Similarly, we use this training sample estimate of the factor f_t^{start} to set the prior mean and variance of $\phi_{f,1}, \phi_{f,2}$. Finally, we need to set the degrees of freedom and the scales of the prior inverse-Gamma distributions for the variances of the idiosyncratic shocks. In all the IG distributions we set the degrees of freedom to 1, the minimum required to make the prior distributions proper while keeping the weight of the prior as low as possible. The prior scale parameters are set at the sum of square residuals of the OLS estimates obtained on the training sample data. The Gibbs sampler is initialized at the prior means.

4.2 Full sample results: loadings and volatilities

In this section we report the estimation results of our model for the entire sample which starts in January 1991 and ends in May 2011. Appendix D assesses the convergence of the Markov Chain to the ergodic distribution based on the inefficiency factors..

A first evaluation of the relative importance of the indicators that are included in the model is given by the posterior estimates of the factor loadings (β), which are shown in Table 2. The highest posterior median weight (0.49) is given to the Industrial production index, followed by GDP (0.38) and by the Industrial production index in the Pulp and Paper sector. Survey data receive roughly the same weight (around 0.1), with a slight prevalence given to the PMI and the weakest contribution coming from the Michigan US Consumer Survey. The annual rate of change of the euro-dollar exchange rate and the US spread have a counter-cyclical effect on GDP. The sign of these two parameters is easily rationalized by considering that these indicators typically *lead* the business cycle, so that their correlation with current cyclical conditions (measured by the common factor) is negative.

An alternative way to look at the relative importance of the different indicators in the estimation of the business cycle is given by the forecast weights used by Banbura and Runstler (2007). Differently from factor loadings, Kalman filter based forecast weights take into account the different timeliness of the indicators and are time varying. To see how these weights are computed consider the update equation of the unobserved states that is used in the Kalman filter to update the estimate of the unobserved states upon the arrival of new information at time t:

$$\boldsymbol{\mu}_{t\backslash t} = \boldsymbol{\mu}_{t\backslash t-1} + K_t v_t \tag{15}$$

where $\boldsymbol{\mu}_{t\backslash t-1}$ is the forecast of the states based on the information set available at time $t-1$, K_t is the Kalman gain and v_t is a forecast error defined below. Now, from the measurement and transition equations the following can be derived:

$$\boldsymbol{\mu}_{t\backslash t-1} = H\boldsymbol{\mu}_{t-1\backslash t-1} \tag{16}$$

$$v_t = \boldsymbol{y}_t - FH\boldsymbol{\mu}_{t-1\backslash t-1} \tag{17}$$

Plugging (16) and (17) in (15) an autoregressive representation of the filtered states can be obtained:

$$\begin{aligned} \boldsymbol{\mu}_{t-1\backslash t-1} &= H\boldsymbol{\mu}_{t-1\backslash t-1} + K_t(y_t - FH\boldsymbol{\mu}_{t-1\backslash t-1}) \\ &= (I - K_t F)H\boldsymbol{\mu}_{t-1\backslash t-1} + K_t\boldsymbol{y}_t \end{aligned} \tag{18}$$

Inverting equation (18) one obtains the moving average representation of the unobserved states as a function of the observed variables. Time variation in these weights stems from the time varying nature of the Kalman gain K_t. Given the ragged edged nature of data releases, when timely (survey) data are published they will receive relatively more weight than lagged (hard) indicators.

Table 3 shows these weights estimated using the last available vintage, that is May 2011.[6] Given publication lags, GDP growth is known only for the first quarter of 2011 and available Industrial Production data refer to March 2011. The following results emerge:

1. In the last month of each quarter, when GDP is observed, the Kalman filter matches GDP forecast with the actual figure, so that all the weight is given to GDP.
2. In the first and second month of the quarter, when GDP is unobserved but all the monthly indicators are known (the dataset is "balanced"), over half of the estimate of (unobserved) real activity growth depends on the two Industrial Production indexes. Soft indicators play a minor role, with a relatively stronger contribution coming from the Economic Sentiment Indicator
3. In the months when neither GDP nor Industrial Production data are available, the strongest contribution to GDP forecast comes from the Economic Sentiment Indicator, followed by the PMI.

The factor model delivers a smoothed monthly estimate of GDP growth, which we show in Figure 1 together with the €-Coin indicator of Altissimo et al. (2010), which provides an estimate of the monthly growth of euro area GDP after the removal of measurement errors, seasonal and other short-run fluctuations.[7] The estimated monthly GDP tracks very closely €-Coin in periods of relative stability (from 2000 to 2008). During the recent crisis the discrepancy between the two indicators widens considerably, owing to the fact that €-Coin tracks medium-run GDP growth, while our model targets actual GDP growth and is therefore affected by higher frequency fluctuations.

To see whether the model picks up any significant time variation in the variances of the common and idiosyncratic errors we plot the posterior median of selected members of Q_t together with their 68% confidence bands (Figure 2). Starting from the common factor (which can be seen as a measure of the underlying business cycle) the model identifies two shifts in volatility over the past twenty years. The former is a temporary increase at the beginning of the past decade, roughly around the brief recession experienced by the world economy in 2001. The latter, more persistent, starts between 2007/2008, and peaks in 2008, during the recent Great Recession.

We next look at the hard indicators that receive the largest weights in the estimation of the common factor (GDP and IP). Visual inspection of the variances of the idiosyncratic shocks to these two indicators reveals that volatility has been rather stable over most of the sample, with

[6]The weights are scaled as contributions to the forecasts of GDP.

[7]Details on €-Coin can be found in Altissimo et al. (2010). See also http://eurocoin.bancaditalia.it/

the exception of the latest recession, when it surged significantly until 2008 to fall thereafter. Finally, the variance of the US spread shows a slight upward trend during the Nineties and a much more persistent increase during the 2007/2009 recession, consistently with the financial origins of the recent economic downturn.

4.3 News and forecasts 1

Given the mixed-frequency nature of our model, GDP forecasts are continuously updated as new monthly data become available. The impact of data releases on forecast revisions can be assessed using the methodology developed by GRS. To clarify the spirit of the exercise, the concept of vintage needs to be formally introduced. The Ω_{v_j} vintage is defined as:

$$\Omega_{v_j} = \{X_{it/v_j}; t = 1, \ldots, T_{iv_j}, i = 1, \ldots, n\} \tag{19}$$

that is the information set Ω_{v_j} is composed of n indicators available from month 1 to month T_{iv_j}, where the date for which the last observation is available varies across indicators. Within our model, a GDP forecast is obtained as an expectation of future GDP conditional on this information set.

Now consider a new vintage $\Omega_{v_{j+1}}$, which differs from the previous one for the release of a new observation of the i^{th} indicator:

$$\Omega_{v_{j+1}} - \Omega_{v_j} = X_{it/v_{j+1}} \tag{20}$$

The updated information entails a change in the conditioning set and, consequently, a forecast revision. Notice that we work with final data vintages in a pseudo real time context, that is, we do not consider data revisions but only new end of sample releases. This means that, starting from a given point in time, we let the information set gradually *expand*, one indicator at the time.

Data releases can occur at different intervals within the month but, for simplicity, following GRS we set up a stylized calendar in which the order of release of the various indicators is kept fixed within the month. The stylized calendar is shown in Table 4. From a given point in time we start enlarging our dataset by including new data on Industrial Production, typically published around the middle of each month. In the second month of each quarter, right after Industrial Production data are made available, GDP data are included in the information set. From the third week onwards survey data start being published by various sources. Surveys cannot be clearly ranked in terms of timeliness, since their release dates sometimes cross each other. We use the convention to place the IFO index release first, followed by the PMI, the Economic

Sentiment Indicator and the US Michigan Consumer index. Finally we include exchange and interest rates, which enter the model as monthly averages of daily data.[8]

GRS evaluate how efficiently their large factor model incorporates data *news* in terms of Mean Squared Errors (MSE) reduction. In fact, since successive vintages carry more information, one can reasonably expect to see a systematic fall in the forecast error variance as indicators are updated. Exploiting the Bayesian nature of our model we add two dimensions to this metric. First, we look at the width of the forecast distribution at different horizons and investigate whether it shrinks as the information set expands. In a way this gives us some indication to whether the model forecast gains *confidence* as new information accrues and the forecast horizon decreases. Notice that this notion of forecast *confidence* is not strictly related to forecast accuracy, since the model might produce narrower confidence bands around its central forecast but the forecast distribution might well be moving further away rather than closer to the target as the information set expands. Second, we move beyond point forecast accuracy and evaluate the evolution of density forecast accuracy. To this end we use the log-score, that is the logarithm of the predictive density generated by the model evaluated at the outturn of the series. Since the log-score measures the probability that the model assigns to the actual value prior to its realization, we expect to see higher log-scores as the information set expands.

We view these three tools (Mean Squared Errors, confidence interval width and log-scores) as strongly complementary in the evaluation of the impact of news on GDP forecasts. Provided that point forecast accuracy increases with the arrival of more information (i.e. provided that the MSE falls), it is desirable to have less uncertainty around the forecast (i.e. it is desirable to have narrower confidence intervals) but this decrease in uncertainty must not come at the cost of lower density forecast accuracy (i.e. the log-score must rise).

GRS provide evidence that the precision of the signal increases within the month as new data are released in both an in-sample and an out of sample exercise. Due to computational constraints we provide evidence only on the in-sample effect of *news* in a sample period spanning the 2004-2001 period. Running a pseudo out-of-sample evaluation would, in fact, require us to run nine (the number of indicators updated each month) MCMC for each of the months considered in the empirical analysis, which would be computationally quite demanding. We are, however, reassured by two considerations. First, the differences between the in-sample and the out-of-sample analysis presented by GRS are minimal, so that the marginal impact of data releases seems to depend on the timeliness of the releases and on the forecast horizon much more than on parameter uncertainty (inherently higher in an out-of-sample setting). Second, in section 4.5 we provide genuine out-of-sample evidence that successive data releases increase

[8]This timing convention, which is the same used by GRS, somewhat penalizes financial variables as daily information on the dollar-euro and on the spread are disregarded.

point forecast accuracy, although in a simplified setting in which, instead of considering each series separately, we consider only two partitions of the information set, i.e. hard and soft data.

The timing of the analysis is the following. We consider releases from January 2004 to May 2011 and, in line with GRS, we forecast each quarter from the first month of the quarter to the first month of the subsequent one, that is we compute three nowcasts and one backast. For each month we update the vintages sequentially according to our stylized calendar, sample 1000 draws from the posterior, run the Kalman filter and smoother and, for each posterior draw, produce nine GDP estimates, corresponding to the release of each of the nine indicators, and consequently nine forecast errors. We compare our model forecasts with those of a naive constant growth model.

In Figure 3 we show the evolution of the Mean Squared Errors within the month, relative to the MSE obtained with the naive model. Three comments are in order. First, in the first month the MSE falls monotonically within the month, albeit at a very slow rate. Going from the first to the second month there is a discrete jump corresponding to the publication of the Industrial production index. Within the second month a large fall in the MSE occurs at the publication of the GDP for the previous quarter. When this information is released, the model can exploit the serial correlation of the target variable and provide a more precise estimate of current GDP growth. In this month the release of survey data and financial indicators offers some marginal improvement in the estimate of current GDP. From the third month onwards the contribution of soft data weakens further and only Industrial production provides some further refinement of the GDP estimate. Also notice that throughout the forecast cycle the MSE ratio remains below one, reflecting the valuable content of conjunctural indicators. All in all, these results are consistent with findings in the literature that stress the declining value of soft indicators and the increasing importance of hard data as the forecast horizon progressively shortens.

We next assess the evolution of forecast *confidence* over the forecasting cycle. Since for each of the nine data releases we have an entire distribution of GDP forecasts we can gauge forecast *confidence* by measuring the dispersion of these forecasts. As a measure of statistical dispersion we use the standardized interquartile range, that is the difference between the 75^{th} and the 25^{th} percentiles standardized by the median. We choose the interquartile range since it has some desirable statistical properties, in particular it is a robust statistics (i.e. it is not affected by outliers) and in a symmetric distribution it equals the median absolute deviation. The evolution of the interquartile range over the forecast cycle is depicted in Figure 4. Two results are worth stressing. First, the chart reveals a clear downward tendency in the dispersion of GDP estimates, indicating that the *confidence* that the model places on its GDP forecasts increases as conjunctural information accrues. Second, soft data play an important role in driving the reduction in forecast dispersion, especially at the very beginning of the forecast cycle when a strong fall in forecast uncertainty occurs as the first surveys become available.

Finally, in Figure 5 we show the evolution of the log-score (crossed line) together with the log-score obtained with the constant growth model (dotted line). Consistently with point forecast accuracy results, density forecast accuracy monotonically increases at the release of each new indicator, suggesting that as the forecast horizon shortens the model assigns (ex ante) a progressively higher probability to the actual GDP releases.

4.4 News and forecasts 2

In a recent paper Banbura and Modugno (2010) propose and derive an alternative way to map directly *news* into forecast revisions. They motivate this alternative measure of *news* by noticing that in factor models the forecast of the unobserved factors is a weighted average of present and past observable indicators, with weights endogenously assigned by the Kalman smoother. When the information set is enriched by a new release, the Kalman smoother incorporates the new information by revising the weights assigned to *all the available indicators* making it impossible to discern whether an improvement in forecast accuracy is due to the new release or to a revision of the weights assigned to other indicators. They therefore devise a way to dissect more precisely the contribution of each release to forecast revisions. Their method, whose technical details are described in Appendix E, is of particular interest in cases when, instead of considering the release of a single indicator, a whole block of data is released and the contribution of the *news* content of each single indicator needs to be assessed.

Our setup, by providing a quantification of the the uncertainty surrounding the news content of a new data (or block of data) release, provides a more complete picture of the forecast revision implied by the intra-monthly information flow.

We illustrate this point using as a case study the GDP forecast of the second quarter of 2010. We start nowcasting this GDP release in the first half of April, when the February Industrial Production numbers become available. We update our forecasts twice a month until the first half of August, right before the first GDP estimate is published. The first by-monthly update coincides with the release of a string of hard data, the second with the publication of survey and of the monthly averages of financial indicators. The resulting forecast updates are shown in Figure 6. The bars below the dotted line depict the contribution of the release (the *news*) of each new indicator computed according to Banbura and Modugno methodology.[9]

[9] In order to evaluate the direct effect of news it is assumed that model parameters are unchanged if confronted with the previous vintage.

At the beginning of the forecast cycle (mid-April) the prediction of the model stands quite far from the final outcome, as the model envisages barely positive growth against a GDP growth outturn of around 1%. Between the end of April and the middle of May positive signals coming from the survey first, and from Industrial production and the release of GDP data for Q1 afterwards, push the forecast progressively upwards. In May a false signal sent by the release of survey data depress again GDP growth expectations. From June onwards, positive news from both soft and hard data set the model forecasts on the right track and GDP predictions start fluctuating more or less around 1%, not far from the actual figure.

To complement the analysis with a measure of uncertainty on both (1) the overall revision implied by the release of an entire data block and (2) the contribution of each indicators to such revision, at each by-monthly update of our information set we draw 1000 forecasts from the predictive density and map each of these forecasts onto the *news*.

In Figure 7 we report estimated kernel densities of the overall revision to the forecast due to the release of 'hard' (upper panel) and 'soft' (lower panel) data between April and July. To show what the individual contributions look like we report in Figure 8 similar densities for two selected indicators, namely Industrial production and the Economic Sentiment Indicator, which appear to be responsible for most of the revisions over the forecast cycle.

From the comparison of these distributions with the information provided in Figure 6, the importance of having a tool to identify the credibility of forecast updates emerges quite clearly. In the second half of April and May, for example, the model picks up first a strong upward, then a strong downward revision due to the release of survey data, which can be largely be attributed to *news* in the Economic Sentiment Indicator. These releases contribute to a revision of GDP forecast by around three decimal points on the way up, a little less on the way down. However, results in Figures 7 and 8 show that in both months the overall revisions and the contribution of the Economic Sentiment Indicator to such revisions are measured with considerable uncertainty, calling for some caution in the interpretation of these forecast updates. In June and in July, on the other hand, as monthly information accumulates and the forecast horizon shortens, the dispersion of estimated revisions and contributions shrinks considerably.

4.5 Out of sample forecasting performance

The last empirical analysis we conduct is a pseudo out of sample forecast exercise. The design of the exercise is similar in spirit to the sequence of forecasts updates discussed in the previous section. In particular, for each quarterly GDP release we provide eight forecasts, starting from six months before the end the quarter of interest to one month afterwards (backcast). Taking as a target, for example, the third quarter of each year, we produce the first forecast in March and the last one in October. We update each of these projections twice a month, when, respectively, hard and soft data are released. The forecast exercise runs from the first quarter of 2006 to the last quarter of 2010. We contrast the forecasting performance of our model with that of a baseline setup, obtained by shutting off time variation in the volatility of the idiosyncratic components and of the latent common factor.

Starting from point forecast evaluation, the evolution of the Root Mean Squared Forecast Errors over the forecast cycle of the two competing models are shown in figures 9. RMSFE are reported as a ratio to the RMSFE attained by a trivial constant growth benchmark. Two observations are in order. First, time variation in the variances increases forecast accuracy, since the model with stochastic volatility has a lower relative RMSFE with respect to the baseline model over most of the forecast horizon, from the beginning to the first two nowcasts. From the end-month update of the second nowcast to the backcast, when more recent industrial production figures are unveiled, the two models deliver instead broadly similar results. Second, the RMSFEs of both models decline as the flow of information accumulates, yet the rate of their decline is quite different. Compared to the linear model, the model with stochastic volatility exploits more efficiently early data releases and starts outperforming the naive benchmark (i.e. its RMSFE falls below 1) earlier in the forecast cycle. The gap between the two RMSFE shrinks considerably when the previous quarter GDP is released (mid-month update of the second nowcast) and the baseline model can use the first order GDP autocorrelation to adjust the nowcast.

Turning to density forecast evaluation, we look first at coverage rates, that is the frequency with which the actual outcome falls within a given confidence interval. Given the popularity in Central Banks and among forecasters of confidence intervals and fan charts, this seems a natural starting point. If the model produces a density forecast which matches well the underlying unknown density function that has generated the data, one can expect that the actual coverage rate equals the nominal one. For example one should find that in our out-of-sample exercise GDP growth fell 10% of the times within our 10% confidence interval, 20% of the times within our 20% confidence interval and so forth. To gauge uncertainty we run a t-test on the null hypothesis that the actual coverage equals the nominal one.[10] In Table 5 we report the coverage rate for the baseline model (without stochastic volatility). We look at backcast (projections one month after the end of the quarter), nowcast (projections during the quarter), and 1 step ahead forecast (projections for the next quarter).

It is clear that the baseline model produces far too wide confidence intervals for the backcast and for the one quarter ahead forecast. In three and four cases out of ten, respectively, according

[10] As emphasized by Clark (2009) this test is slightly imprecise as it abstracts from parameter uncertainty

to the p-values these differences are statistically significant at the 10% confidence level. In the case of the nowcast the situation improves, with the tests rejecting only once.

Table 6 shows that adding stochastic volatility to the model brings sizeable gains in density forecast accuracy. In the case of the backcast, when the information set on the quarter of interest is almost complete, confidence intervals are extremely accurate, with actual coverage rates only once significantly different from the nominal ones, and never when the model is used for nowcasting. In the case of one step ahead forecasts the model with stochastic volatility records only two rejections.

Finally, we look at the normalized probability integral transforms (PITS) of the forecast errors, another popular tool for evaluating density forecasts. According to the testing framework developed by Berkowitz (2001), if the model forecast density matches the density that generated the data, the PITS should be independent standard normal. We follow Clark (2009) and test these conditions (zero mean, unit variance and no serial correlation) separately and jointly, with the Berkowitz (2001) likelihood ratio test for the joint null of zero mean, unity variance, and no serial correlation of order 1.

The p-values of these tests are presented in Table 7 for the baseline model without stochastic volatility and in Table 8 for the model with stochastic volatility.[11] For each month we consider a mid-month and an end-month update, so that we have six different results for the one step ahead forecast and for the nowcast and two results for the backast. Using a 10% significance level, for both models the tests generally fail to reject the null hypotheses of zero mean, unit variance and no serial correlation, with 11 rejections out of 48 cases in the baseline model and a couple more in the model with stochastic volatility. The joint Normality/Independence tests show p-values larger than 0.1 in all the cases considered. While no conclusive ranking emerges from this exercise, evidence on the PITS confirms that our model provides accurate density forecasts.

5 Conclusions

This paper introduces a mixed frequency factor model with stochastic volatility, and develops a Bayesian procedure for its estimation. The model deals with all the challenges faced by a forecaster that needs to produce updated quarterly GDP forecasts at each relevant data release, like data sampled at different frequencies and ragged-edge data. Differently from existing linear models, our setup allows for continuous shifts in the volatility of the errors of both the common

[11] Notice that the testing framework developed by Berkowitz (2001) for the PITS applies to one step ahead forecasting. As such it does not immediately apply to the context of nowcasting, where the impact of different data vintages and the presence of ragged edged data have to be taken into account.

factor and of the idiosyncratic errors, a feature that in the macro forecasting literature has been shown to improve both point and density forecast accuracy.

This measurement tool is applied to the problem of forecasting euro area GDP at short horizons. When estimated over the whole sample, the model picks up significant shifts in the volatility of the errors, with two peaks coinciding with the major recessionary episodes of the past twenty years.

We further illustrate how, in a given quarter, the factor model can be used to assess the uncertainty around the news content of monthly releases of hard, soft and financial indicators. Consistently with findings in the literature, we find that forecast accuracy improves significantly in connection with the release of monthly data as the forecast horizon decreases. Also, forecast uncertainty (measured by the width of the forecast distribution) progressively decreases as more information on the quarter of interest becomes available.

Finally, we design a (pseudo) real time out of sample forecasting exercise and evaluate out of sample point and density forecasts accuracy. In line with Clark (2009) we find that the introduction of stochastic volatility significantly contributes to an improvement in density forecast accuracy.

References

Altissimo, Filippo, Riccardo Cristadoro, Mario Forni, Marco Lippi, and Giovanni Veronese (2010), "New eurocoin: Tracking economic growth in real time." *The Review of Economic and Statistics*, 92, 1024–1034.

Angelini, Elena, Gonzalo Camba-Méndez, Domenico Giannone, Gerhard Rünstler, and Lucrezia Reichlin (2008), "Short-term forecasts of euro area gdp growth." Working Paper Series 949, European Central Bank.

Bai, Jushan and Serena Ng (2008), "Forecasting economic time series using targeted predictors." *Journal of Econometrics*, 146, 304–317.

Banbura, Marta and Michele Modugno (2010), "Maximum likelihood estimation of factor models on data sets with arbitrary pattern of missing data." Working Paper Series 1189, European Central Bank.

Banbura, Marta and Gerhard Runstler (2007), "A look into the factor model black box - publication lags and the role of hard and soft data in forecasting gdp." Working Paper Series 751, European Central Bank.

Baumeister, Christiane, Philip Liu, and Haroon Mumtaz (2010), "Changes in the transmission of monetary policy: evidence from a time-varying factor-augmented var." Bank of England working papers 401, Bank of England.

Benati, Luca and Paolo Surico (2008), "Evolving u.s. monetary policy and the decline of inflation predictability." *Journal of the European Economic Association*, 6, 634–646.

Berkowitz, Jeremy (2001), "Testing density forecasts, with applications to risk management." *Journal of Business and Economic Statistics*, 465–474.

Boivin, Jean and Serena Ng (2006), "Are more data always better for factor analysis?" *Journal of Econometrics*, 132, 169–194.

Camacho, Maximo and Gabriel Perez-Quiros (2010), "Introducing the euro-sting: Short-term indicator of euro area growth." *Journal of Applied Econometrics*, 25, 663–694.

Carter, Kevin and Robert Kohn (1994), "On gibbs sampling for state space models." *Biometrika*, 81, 541–553.

Clark, Todd E. (2009), "Real-time density forecasts from vars with stochastic volatility." Technical report.

Clements, Michael P and Ana Beatriz Galvão (2008), "Macroeconomic forecasting with mixed-frequency data." *Journal of Business & Economic Statistics*, 26, 546–554.

Cogley, Timothy and Thomas J. Sargent (2005), "Drift and volatilities: Monetary policies and outcomes in the post wwii u.s." *Review of Economic Dynamics*, 8, 262–302.

Del Negro, Marco and Christopher Otrok (2008), "Dynamic factor models with time-varying parameters: measuring changes in international business cycles." Technical report.

Foroni, Claudia and Massimiliano Marcellino (2012), "A comparison of mixed frequency approaches for modelling euro area macroeconomic variables." mimeo, European University Institute.

Geweke, John (1991), "Evaluating the accuracy of sampling-based approaches to the calculation of posterior moments." Technical report.

Ghysels, Eric, Pedro Santa-Clara, and Rossen Valkanov (2004), "The midas touch: Mixed data sampling regression models." University of California at Los Angeles, Anderson Graduate School of Management 4852, Anderson Graduate School of Management, UCLA.

Giannone, Domenico, Lucrezia Reichlin, and David Small (2008), "Nowcasting: The real-time informational content of macroeconomic data." *Journal of Monetary Economics*, 55, 665–676.

Jacquier, Eric, Nicholas G. Polson, and Peter E. Rossi (1994), "Bayesian analysis of stochastic volatility models." *Journal of Business & Economic Statistics*, 12, 371–89.

Jore, Anne Sofie, James Mitchell, and Shaun P. Vahey (2010), "Combining forecast densities from vars with uncertain instabilities." *Journal of Applied Econometrics*, 25, 621–634.

Kim, C. and R. Nelson (1999), *State Space Models with Regime Switching*. MIT Press, Cambridge, Massachussets.

Koopman, S. J. and J. Durbin (2003), "Filtering and smoothing of state vector for diffuse state-space models." *Journal of Time Series Analysis*, 24, 85–98.

Korobilis, Dimitris (2009), "Assessing the transmission of monetary policy shocks using dynamic factor models." Working Papers 0914, University of Strathclyde Business School, Department of Economics.

Kuzin, Vladimir, Massimiliano Marcellino, and Christian Schumacher (2011), "Midas vs. mixed-frequency var: Nowcasting gdp in the euro area." *International Journal of Forecasting*, 27, 529–542.

Marcellino, Massimiliano and Christian Schumacher (2008), "Factor-midas for now- and forecasting with ragged-edge data: A model comparison for german gdp." CEPR Discussion Papers 6708, C.E.P.R. Discussion Papers.

Mariano, Roberto S. and Yasutomo Murasawa (2003), "A new coincident index of business cycles based on monthly and quarterly series." *Journal of Applied Econometrics*, 18, 427–443.

Primiceri, Giorgio (2005), "Time varying structural vector autoregressions and monetary policy." *Review of Economic Studies*, 72.

Stock, J.H and M.W. Watson (1989), "New indexes of coincident and leading economic indicators." In *NBER Macroeconomics Annual*, 351–393, Blanchard, O. and S. Fischer (eds). MIT Press, Cambridge, MA.

A Details of the Gibbs sampler

We describe in more details the six blocks that compose our Gibbs sampler procedure:

A.1 Block 1: drawing the factor loadings $\beta_q, \beta_h, \beta_s$

In the first block of the Gibbs sampler we draw the slopes, conditional on all the other parameters of the model. To see how this is done let us start from the measurement equation of the hard indicator:

$$y_{h,t} = \beta_h f_t + u_{h,t} \tag{21}$$

where the law of motion of the idiosyncratic shock is $u_{h,t} = \phi_{h,1}u_{h,t-1} + \phi_{h,2}u_{h,t-2} + \epsilon_{h,t}e^{\lambda_{h,t}/2}$ and $\epsilon_{h,t} \sim N(0, \sigma_h)$. Since we are conditioning on all the parameters, on the factor f_t and on the stochastic volatilities $\lambda_{h,t}$ we can treat this equation as a simple regression with autocorrelated and heteroscedastic residuals. To whiten the residuals, we quasi-difference the equation by filtering both sides with the filter $1 - \phi_{h,1}L - \phi_{h,1}L^2$ and dividing each observation by $e^{\lambda_{h,t}/2}$:

$$y^\star_{h,t} = \beta_h x^\star_t + \epsilon_{h,t} \tag{22}$$

where $x^\star_t = (1 - \phi_{h,1}L - \phi_{h,1}L^2)f_t/e^{\lambda_{h,t}/2}$. Then positing a Normal prior: $p(\beta_h) \sim N(\underline{\beta_h}, \underline{\sigma_{\beta,h}})$ the conditional posterior is also normal $N \sim (\overline{\beta_h}, \overline{\sigma_{\beta,h}})$ where

$$\overline{\sigma_{\beta,h}} = (\underline{\sigma_{\beta,h}}^{-1} + \sigma_h^{-2}\sum_{t=1}^{T} x^{\star 2}_t) \tag{23}$$

$$\overline{\beta_h} = \overline{\sigma_{\beta,h}}(\underline{\sigma_{\beta,h}}^{-1}\underline{\beta_h} + \sigma_h^{-2}\sum_{t=1}^{T} x^\star_t y^\star_t) \tag{24}$$

The case of survey variables can be treated accordingly after noticing that $x^\star_t = (1-\phi_{h,1}L - \phi_{h,1}L^2)\sum_{j=0}^{11} f_{t-j}/e^{\lambda_{s,t}/2}$. In the case of quarterly variables two adjustments are needed. First, since the variable is observed only every three months only these observations can be used for estimating the factor loading. Second, in the measurement equation an MA(4) regression error appears:

$$y_{q,t} = \beta_q w(L) f_t + w(L)u_{q,t} \tag{25}$$

where $w(L) = \frac{1}{3} + \frac{2}{3}L + L^2 + \frac{2}{3}L^3 + \frac{1}{3}L^4$. Furthermore the error term u_t is an AR(2) process $u_{q,t} = \phi_{q,1}u_{q,t-1} + \phi_{q,2}u_{q,t-2} + \epsilon_{q,t}e^{\lambda_{q,t}/2}$. Our estimation strategy consists of working out the variance covariance matrix of the error terms of equation (25), $\Phi(\phi_{q,1}, \phi_{q,2}, \sigma^2_q)$, which we can treat at this step of the sampler as if it were known. Then it suffices to divide each observation by $e^{\lambda_{q,t}/2}$ and premultiply both sides of the equation by $\Phi^{-\frac{1}{2}}$ to obtain a standard regression with uncorrelated residuals. We are now in the familiar setting in which we can posit a normal prior and draw β_q from a normal posterior.

A.2 Block 2: drawing $\phi_{f,1}, \phi_{f,2}, \phi_{q,1}, \phi_{q,2}, \phi_{h,1}, \phi_{h,2}, \phi_{s,1}, \phi_{s,2}$

To draw the parameters that govern the autocorrelation of the idiosyncratic shocks first notice that since we are conditioning on the state vector μ_t, we can treat the common factor f_t and the residuals $u_{q,t}, u_{h,t}, u_{s,t}$ as known. The transition equations become standard regression problems which can be analyzed separately (again after pre-whitening to take into account the stochastic volatility components). We employ normal priors $p([\phi_{j,1}\phi_{j,2}]') \sim N(\underline{\phi_j}, \underline{\Sigma_{\phi,j}})$, where $j = f, q, h, s$, and for each equation we draw from the respective normal conjugate posteriors. We rule out explosive roots by drawing from the untruncated Normal posterior and discarding draws if the roots of $\phi_j(L) = 0$ lie outside the unit circle.

A.3 Block 3: drawing the innovation variances $\sigma_f^2, \sigma_q^2, \sigma_h^2, \sigma_s^2$

The variances of the innovations to the idiosyncratic shocks can also be easily drawn once we condition on the state vector μ_t, on the ϕ_s and on the stochastic volatilities. We again proceed by treating the transition equations one at the time. Let us consider a generic element of the state vector $\mu_{i,t}$. Its law of motion is:

$$\mu_{i,t} = \phi_{i,1}\mu_{i,t-1} + \phi_{i,2}\mu_{i,t-2} + \eta_{i,t} \quad \eta_{i,t} \sim N(0, \sigma_i^2 e^{\lambda_{i,t}}) \tag{26}$$

For the innovation variance σ_i^2 we posit an inverse-Gamma prior $p(\sigma_i^2) = IG(n_i, s_i^2)$. Since the prior is conjugate it can be interpreted as adding n_i artificial observations to the state variable $\mu_{i,t}$. The prior embodies the belief that the sum of squared residuals of these artificial observations equals s_i^2:

$$s_i^2 = \frac{1}{n_i} \sum_{t=1}^{n_i} (\mu_{t,i}^{\star} - \phi_{i,1}\mu_{t-1,i}^{\star} - \phi_{i,2}\mu_{t-2,i}^{\star})^2 \tag{27}$$

Given our assumption that the idiosyncratic shocks are normal the posterior is also an inverse-Gamma, $IG(T + n_i, \frac{n_i s_i^2 + T d_i^2}{T + n_i})$ where:

$$d_i^2 = \frac{1}{n_i} \sum_{t=1}^{n_i} (\mu_{t,i} - \phi_{i,1}\mu_{t-1,i} - \phi_{i,2}\mu_{t-2,i})^2 \tag{28}$$

The weight of the prior is therefore proportional to the prior degree of freedom parameter n_i.

A.4 Block 4: drawing the state vector μ_t

Since the model can be cast in state space draws of the state vector can obtained via a state vector simulation smoother as in Carter and Kohn (1994) or with the disturbance smoother proposed by Koopman and Durbin (2003). We resort to the latter, which turns out to be slightly more efficient from a computational point of view.[12]

A.5 Block 5: drawing $\lambda_{i,t}$

To sample the stochastic volatilities $\lambda_{i,t}$ notice that conditional on all parameters and on the states μ_t the orthogonal innovations $\eta_{i,t}/\sigma_{h,i}$ are observable. The $\lambda_{i,t}$ can then be sampled adopting the date-by-date blocking scheme developed by Jacquier et al. (1994).[13].

A.6 Block 6: drawing $\sigma^2_{h,i}$

The final block of the sampler involves drawing the variances of the log-volatilities. Conditioning on the log-volatilities and postulating an inverse-Gamma prior distribution, the $\sigma^2_{h,i}$ can also be drawn from an inverse Gamma posterior.

B The selection of the monthly indicators

Small scale models have their own “curse of dimensionality": since they rely on a small set of indicators, they are prone to the criticism of potentially leaving out relevant information compared to factor models that use hundreds of time series. In the literature, however, the initial enthusiasm for the use of very large sets of data has started waning when some authors have pointed out that models that use a smaller set of accurately targeted predictors might deliver more accurate forecasts. Bai and Ng (2008) and Boivin and Ng (2006), for example, question the usefulness of 'too much information' for forecasting purposes. The former, in particular, shows that a number of variable selection techniques (already widely used in biomedical statistics where the number of covariates is typically very large) give encouraging results when applied to economic time series.

[12]We did not find sizeable gains in using the univariate version of the Durbin and Koopman smoother.

[13]Details on the algorithm, which involves a Metropolis Hastings step within the Gibbs sampler, can be found in Cogley and Sargent (2005), Appendix $B.2.5$

To make the choice of the indicators to be included in our model as objective as possible we proceed as follows. We start by considering a dataset of more than a hundred variables for the period 1987-2011[14], and select a subset of 39 indicators similar to those employed in Angelini et al. (2008) and in Camacho and Perez-Quiros (2010). We then set a priori four *core* variables that we decide to include in the model, which are Industrial Production for the euro area (IP), the composite Purchasing Manager Index (PMI), the European Commission Economic Sentiment Indicator (ESI) and the Germany IFO Business Climate Index. To select the remaining variables, we calculate as a benchmark the percentage of GDP variance explained by the factor computed from the core variables only, as in Camacho and Perez-Quiros (2010), and design an algorithm for the selection of a set of additional indicators which maximize this statistic.

1. We evaluate datasets with all *core* variables and one other variable at a time in order to calculate the explained variance, and the probability that it is higher than in the dataset with *core* variables only. In this way we obtain a ranking of the other series.

2. We add a variable at a time, starting with the ones with an higher probability to increase the explained variance with respect to the benchmark; we keep the variable only if this probability increases. We end up with the small set of 8 variables described in the main text.

C The state space specification in the empirical application

The specification we adopt follows Camacho and Perez-Quiros (2010) where surveys are modeled as a 12 terms moving average of the unobserved factor, while hard variables load the factor contemporaneously. This amounts to imposing that surveys are in phase with the year on year growth rate of Industrial Production (and of the other hard indicators). To get an idea of the state representation of the model while keeping notation to a minimum we present the case of a toy model with one quarterly variable, one hard indicator and one soft indicator in which all the idiosyncratic shocks follow an AR(2) process. The more general case can be easily derived from this example. The loading matrix F in the measurement equation (10) can be written as:

$$F = \begin{pmatrix} \beta_q \frac{1}{3} & \beta_q \frac{2}{3} & \beta_q & \beta_q \frac{2}{3} & \beta_q \frac{1}{3} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \frac{1}{3} & \frac{2}{3} & 1 & \frac{2}{3} & \frac{1}{3} & 0 & 0 & 0 & 0 \\ \beta_h & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ \beta_s & \beta_s & \beta_s & \beta_s & \beta_s & \beta_s & \beta_s & \beta_s & \beta_s & \beta_s & \beta_s & \beta_s & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \end{pmatrix} \quad (29)$$

[14]The series are those used to compile €-Coin (see Altissimo et al. (2010)). For a description of the dataset see `http://eurocoin.bancaditalia.it/`

where β_q, β_h and β_s are the loadings of, respectively, the quarterly variable, the hard and the soft indicators. The state vector is:

$$\mu_t = \begin{pmatrix} f_t & f_{t-1} & \dots & f_{t-11} & u_{q,t} & \dots & u_{q,t-4} & u_{h,t} & u_{h,t-1} & u_{s,t} & u_{s,t-1} \end{pmatrix}' \tag{30}$$

The transition matrix is:

$$H = \begin{pmatrix} \phi_{f,1} & \phi_{f,2} & 0 & 0 & 0 & \dots & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & \dots & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & \dots & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & 0 & \dots & \phi_{q,1} & \phi_{q,2} & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & \dots & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & \dots & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & 0 & \dots & 0 & 0 & 0 & 0 & 0 & \phi_{h,1} & \phi_{h,2} & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & \dots & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & \dots & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \phi_{s,1} & \phi_{s,2} \\ 0 & 0 & 0 & 0 & 0 & \dots & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \end{pmatrix} \tag{31}$$

Since the idiosyncratic shocks are collected in the state vector the matrix R_t is a (k+2) dimension zero matrix while the matrix Q_t is a diagonal matrix which collects all the variances:

$$Q_t = diag\begin{pmatrix} 1 & 0 & 0 & 0 & \dots & \sigma_q^2 e^{\lambda_{q,t}} & 0 & 0 & 0 & 0 & \sigma_h^2 e^{\lambda_{h,t}} & 0 & \sigma_s^2 e^{\lambda_{s,t}} & 0 \end{pmatrix} \tag{32}$$

D Assessing the convergence of the Markov chain to the ergodic distribution

We assess the convergence of the Markov chain to the ergodic distribution by looking at the autocorrelation properties of the draws across sets of parameters. In the full sample estimate of the models we run 30000 replications and retain the last 5000 draws. As a measure of convergence of the Markov Chain we consider the inefficiency factors (henceforth, IFs) of the draws, which are defined as the inverse of the relative numerical efficiency measure (RNE) of Geweke (1991). The RNE is computed considering one parameter at the time and using the sequence of draws as time dimension. Specifically the RNE is defined as:

$$RNE = (2\pi)^{-1} \frac{1}{S(0)} \int_{-\pi}^{\pi} S(\omega) d\omega \tag{33}$$

where $S(\omega)$ is the spectral density of the draws of a given parameter at the frequency ω. The denominator $S(0)$ is the spectral density at the zero frequency, a measure of the long run variance of the draws. The spectral densities are estimated by the smoothed periodogram using a 32 points Bartlett triangular window which weighs less more distant autocorrelations. In Figures 10 and 11 we present the IFs. As the figures show the autocorrelation of the draws is very low, with values of the IFs overall below two, that is ten times lower than the threshold (twenty) which can be considered as satisfactory, as stressed by Primiceri (2005).

E News and forecast revisions

In their paper Banbura and Modugno (2010) derive a way to decompose a forecast revision as a linear function of news.

They denote as Ω_v a vintage of data corresponding to a statistical data release v, which as an example can be mid-month for industrial production and end of month for surveys, in order to define *news* as:

$$I_{v+1,j} = y_{i_j,t_j} - E[y_{i_j,t_j} | \Omega_v] \tag{34}$$

the surprise incorporated in a new data with respect to what was expected given information Ω_v. A forecast revision is defined as:

$$E[y_{k,t_k} | I_{v+1}] = E[y_{k,t_k} | \Omega_{v+1}] - E[y_{k,t_k} | \Omega_v] \tag{35}$$

and can be expressed as weighted average of news:

$$E[y_{k,t_k} | I_{v+1}] = B_{v+1} I_{v+1} = E[y_{k,t_k} I'_{v+1}] E[I_{v+1} I'_{v+1}]^{-1} I_{v+1} \tag{36}$$

where:

$$E[y_{k,t_k} I_{v+1,j}] = H_k E[(\mu_{t_k} - E(\mu_{t_k} | \Omega_v))(\mu_{t_j} - E(\mu_{t_j} | \Omega_v)')] H'_{i_j} \tag{37}$$

$$E[I_{v+1,j} I_{v+1,l}] = H_{i_j} E[(\mu_{t_j} - E(\mu_{t_j} | \Omega_v))(\mu_{t_l} - E(\mu_{t_l} | \Omega_v)')] H'_{i_l} \tag{38}$$

where $E[(\mu_{t_j} - E(\mu_{t_j}|\Omega_v))(\mu_{t_l} - E(\mu_{t_l}|\Omega_v)')]$ is the state vector covariance matrix obtained as a by-product of the Kalman Smoother.

Table 1: Variable selection summary

Indicator	Country
Quarterly series	
GDP	Euro Area
Monthly series	
Industrial Production	Euro Area
Industrial Production - Pulp/paper	Euro Area
Business Climate - IFO	Germany
Economic Sentiment Indicator	Euro Area
PMI composite	Euro Area
dollar-euro	US-Euro
10y-3m spread	US
Michigan Consumer Sentiment	US

Table 2: Factor Loadings - posterior estimates

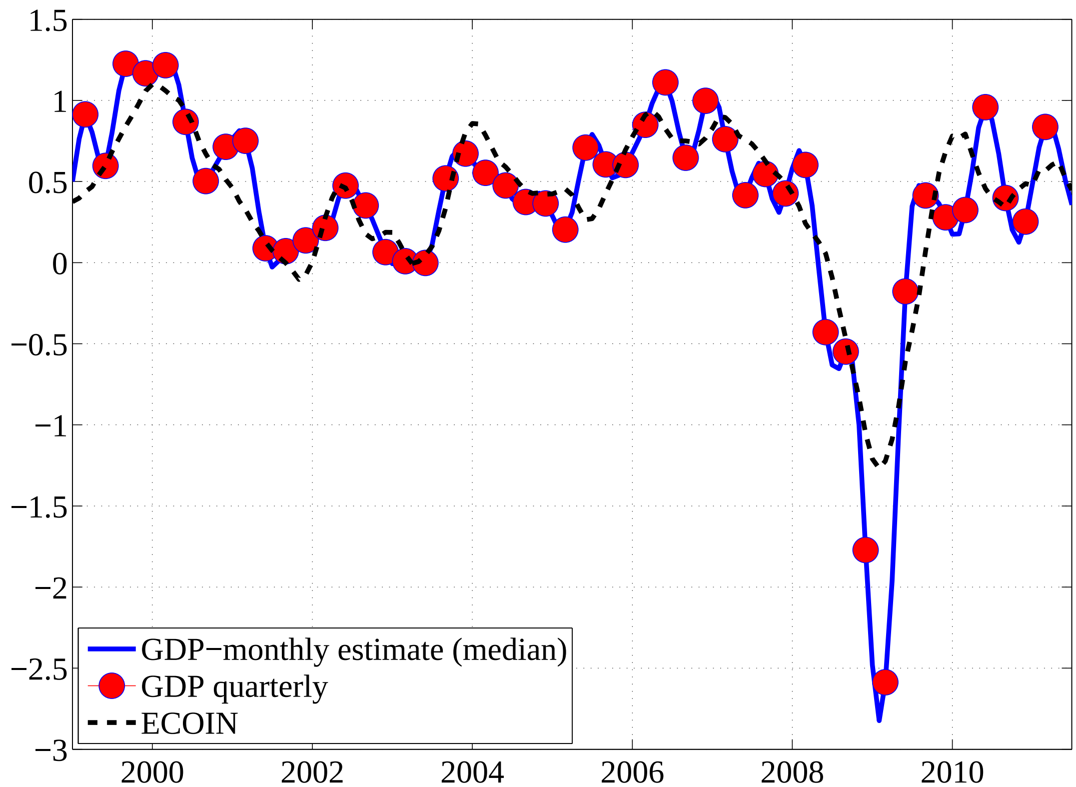
Percentiles	25th	50th	75th
GDP	0.27	0.38	0.54
IP	0.40	0.49	0.60
IP-PULP	0.23	0.29	0.36
IFO	0.10	0.12	0.13
ESI	0.10	0.12	0.14
PMI	0.12	0.13	0.15
dollar-euro	-0.08	-0.05	-0.02
US-spread	-0.06	-0.04	-0.02
Michigan Consumer	0.04	0.06	0.08

Table 3: Forecast weights

	GDP	IP	IP-PULP	IFO	ESI	PMI	dollar-euro	US-spread	Michigan
Oct10	0.00	0.43	0.17	0.07	0.23	0.11	-0.01	-0.01	0.01
Nov10	0.00	0.44	0.16	0.07	0.23	0.10	-0.01	-0.01	0.01
Dec10	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Jan11	0.00	0.43	0.17	0.07	0.23	0.11	-0.01	-0.01	0.01
Feb11	0.00	0.44	0.16	0.07	0.23	0.10	-0.01	-0.01	0.01
Mar11	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Apr11	0.00	0.00	0.00	0.17	0.58	0.27	-0.02	-0.03	0.03
May11	0.00	0.00	0.00	0.17	0.58	0.26	-0.02	-0.03	0.03

Note to Table 3. Percentage contribution of the indicators in forecasting GDP computed as in Banbura and Runstler (2007).

Figure 1: GDP: Median monthly estimate and €-Coin



Note to Figure 1: the blue continuous line is the monthly GDP growth estimate produced by the factor model with stochastic volatility. The dotted line is the €-Coin indicator. €-Coin is an indicator of the medium-term component of quarter on quarter euro area GDP growth published each month by the Bank of Italy. For details see `http://eurocoin.cepr.org/`.

Figure 2: Stochastic volatility for the common factor and for selected variables

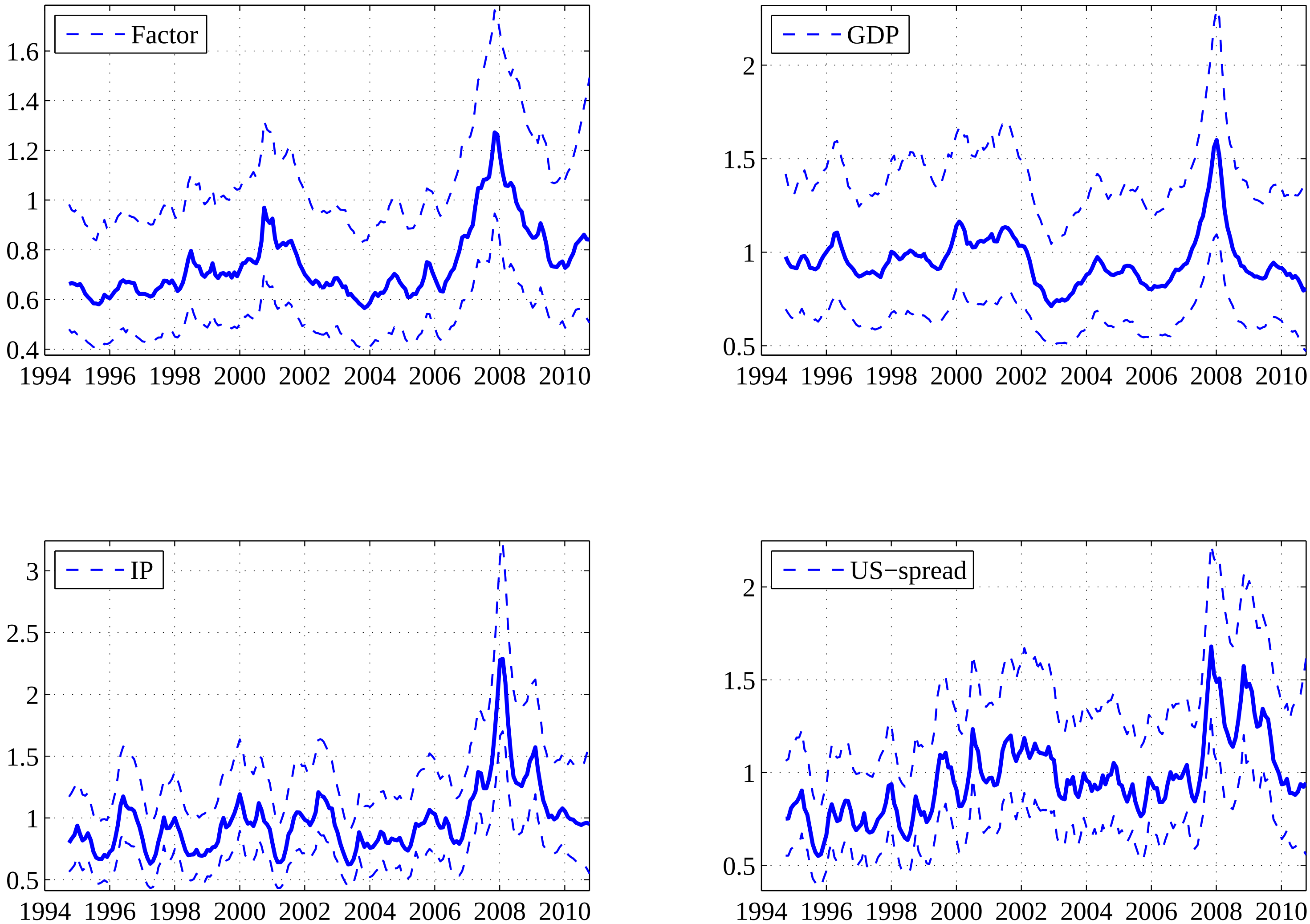
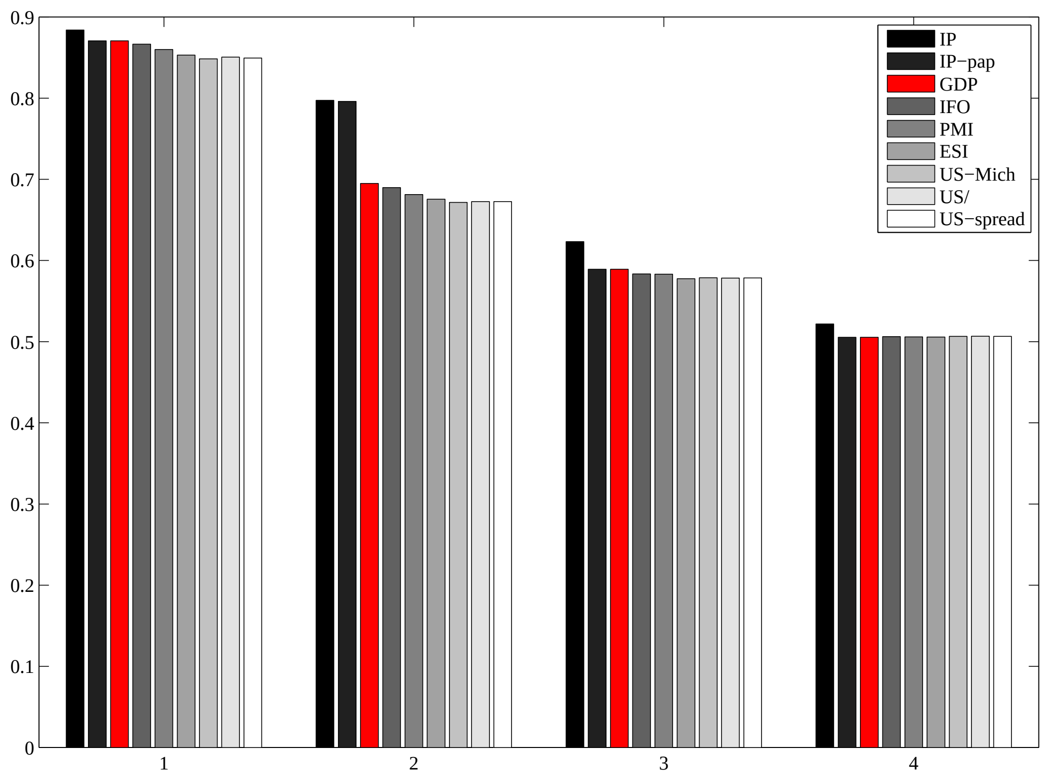


Table 4: Stylized data release calendar

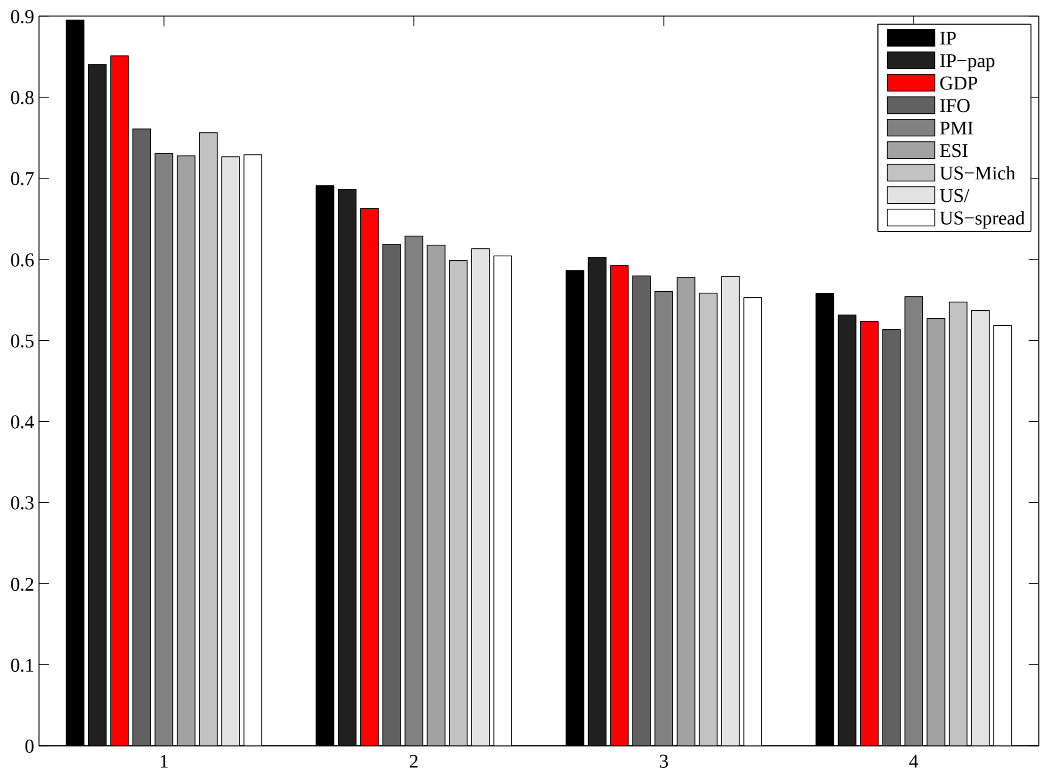
Indicator	Timing	Publication lag	Frequency
IP	$11^{th} - 15^{th}$ of month	2	Monthly
IP-PULP	$11^{th} - 15^{th}$ of month	2	Monthly
GDP	1 day after IP	2	Quarterly
IFO	$20^{th} - 30^{th}$ of month	0	Monthly
PMI	$20^{th} - 30^{th}$ of month	0	Monthly
ESI	$20^{th} - 30^{th}$ of month	0	Monthly
Michigan Consumer	Last Friday of the month	0	Monthly
dollar-euro	Last day of month(Monthly ave.)	0	Monthly
US-spread	Last day of month(Monthly ave.)	0	Monthly

Figure 3: RMSE at different releases



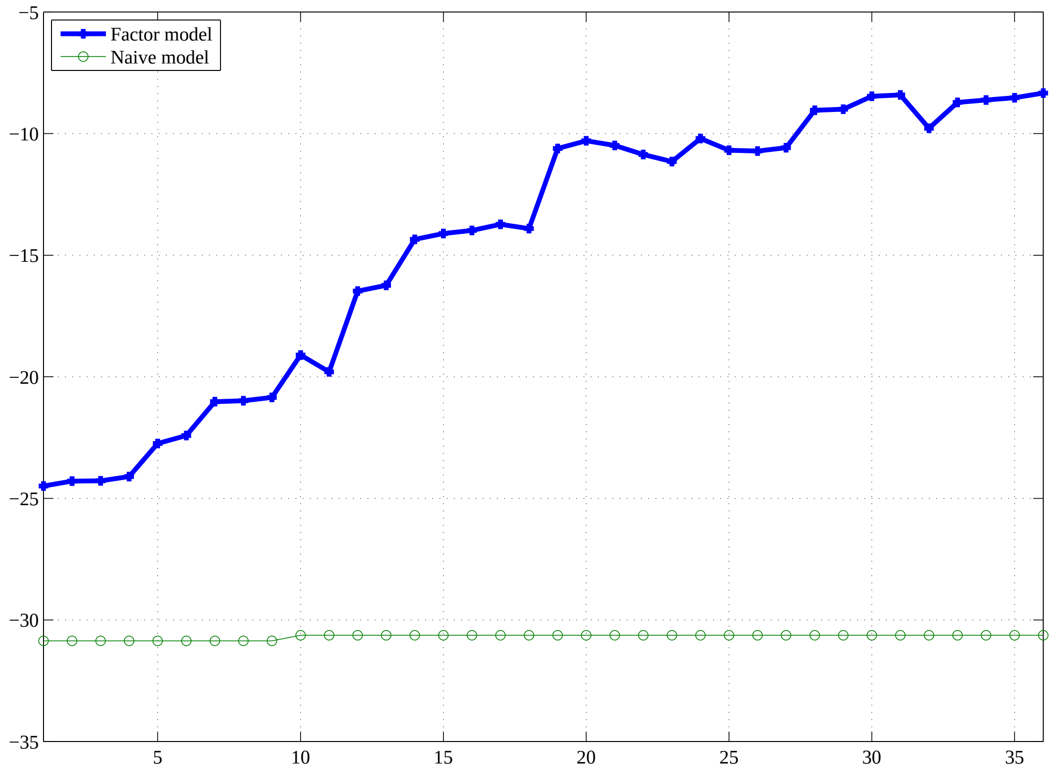
Note to Figure 3: the Figure shows the ratio of the RMSE of the factor model with stochastic volatility to that of a naive constant growth model for each of the indicated data release. Data releases follow the stylized calendar 4.

Figure 4: Forecast dispersion at different releases



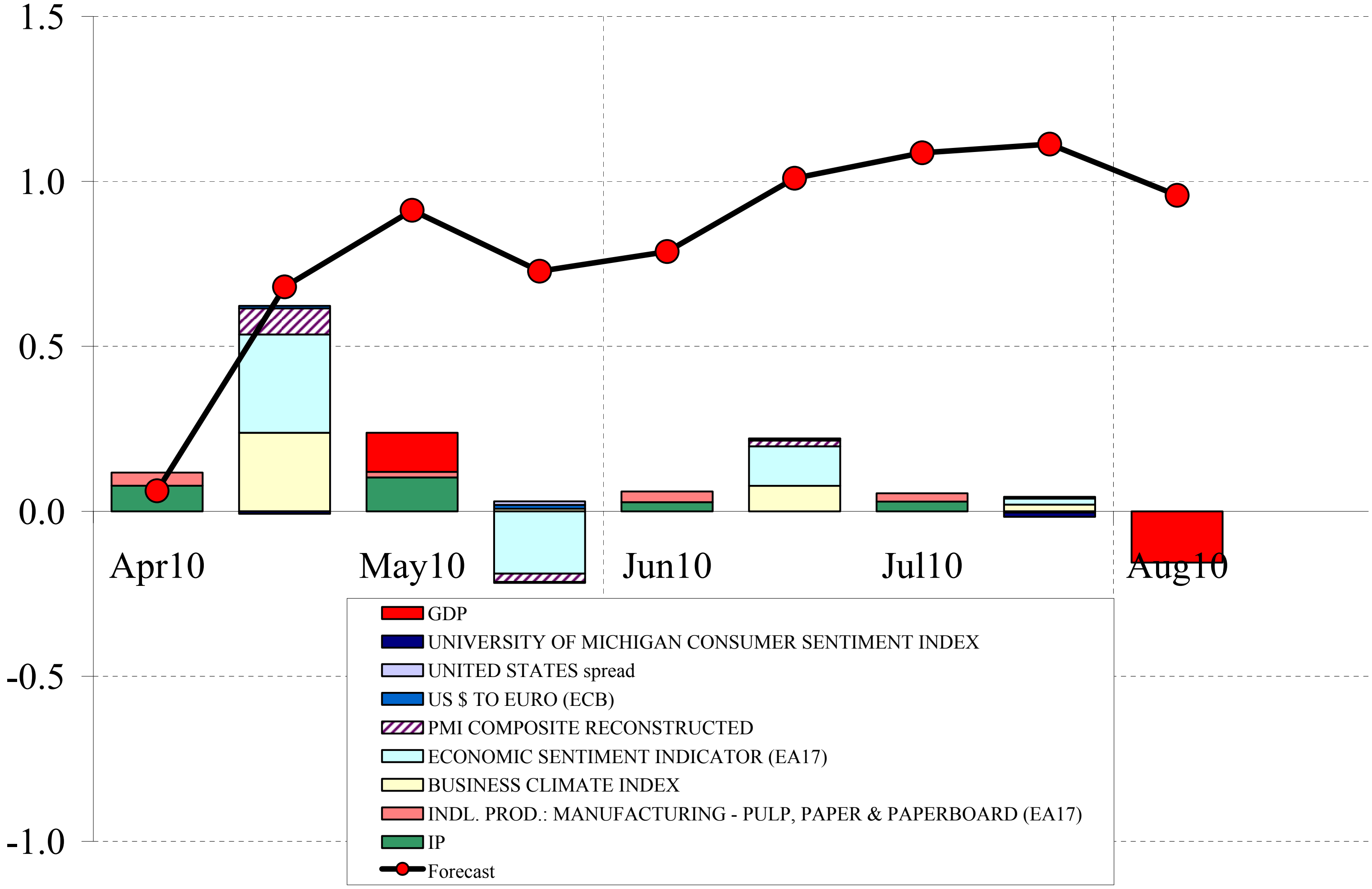
Note to Figure 4: the Figure shows the difference between the 75 and the 25 percentiles (both scaled by the median) of the forecast distribution obtained with the factor model with stochastic volatility updated at each data release. Data releases follow the stylized calendar 4.

Figure 5: Log-predictive score at different releases



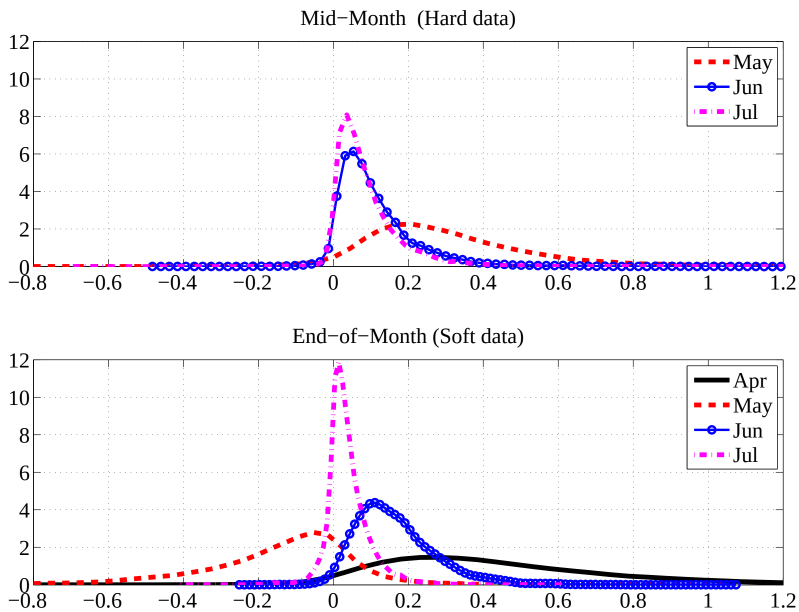
Note to Figure 5: the Figure shows the log-predictive score of the factor model with stochastic volatility updated at each data release and of the naive constant growth model. Data releases follow the stylized calendar 4.

Figure 6: Forecast revisions 2010Q2



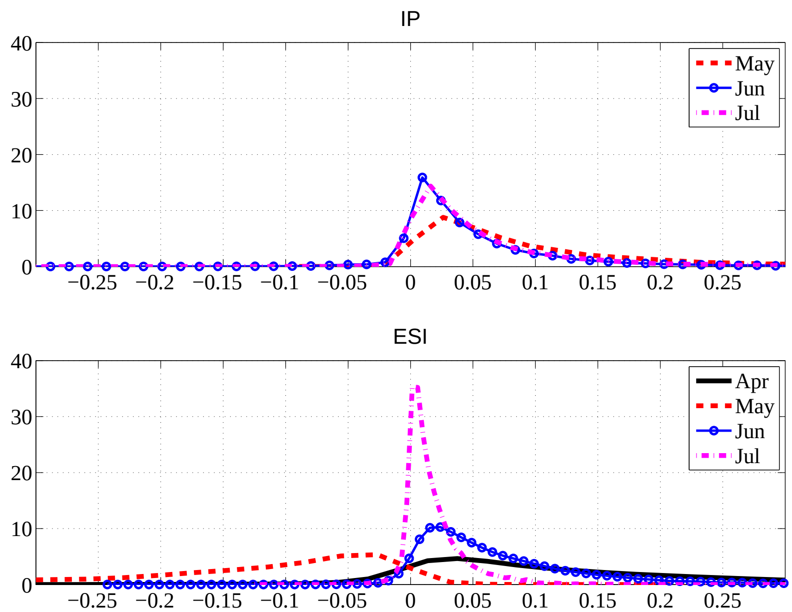
Note to Figure 6: the Figure shows the by-monthly GDP forecasts revisions relative to the second quarter of 2010 and the contributions of the new releases. The firs forecast update is at the end of April, the last update in the middle of August.

Figure 7: Revisions Density evolution 2010Q2



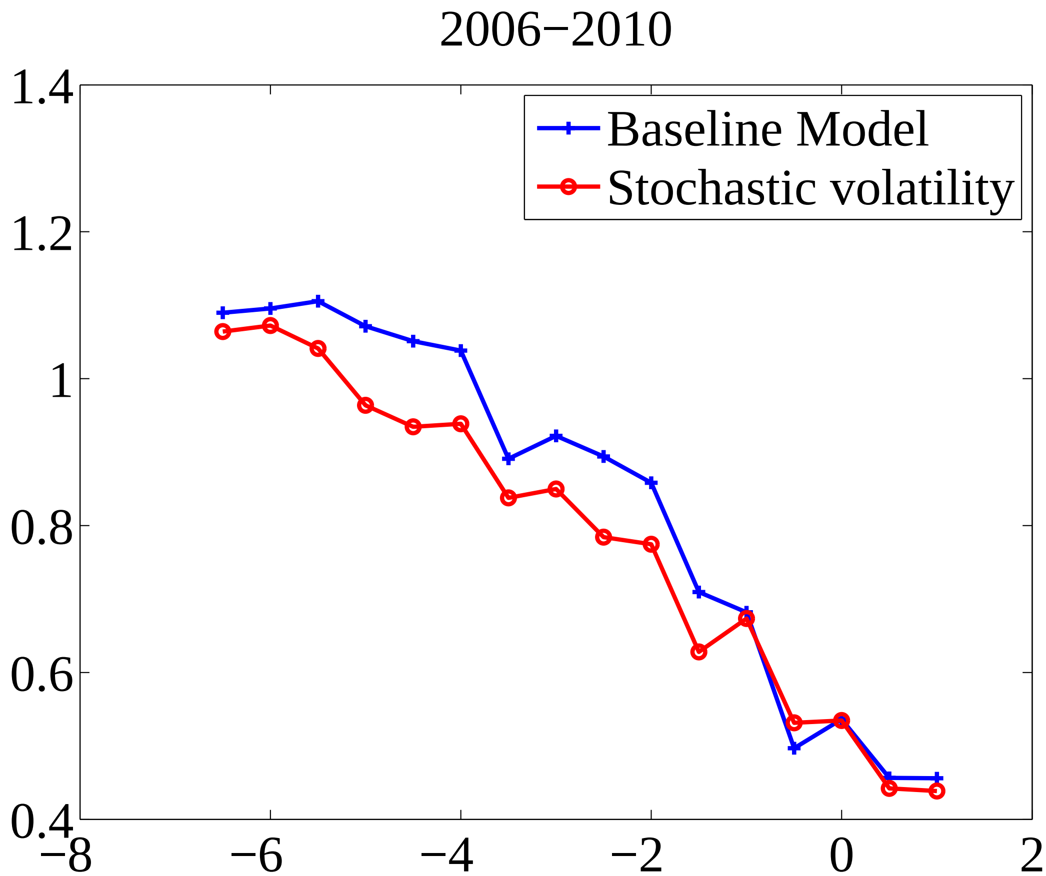
Note to Figure 7: the Figure shows the distributions of the by-monthly GDP forecasts revisions relative to the second quarter of 2010. The density estimation is based on a normal kernel function, using an optimal window parameter function of number of data points. The distribution is based on 1000 draws from the predictive density.

Figure 8: Revisions Density evolution 2010Q2: selected indicators



Note to Figure 8: the Figure shows the distributions of the contributions of Industrial Production and of the Economic Sentiment Indicator to the by-monthly GDP forecasts revisions relative to the second quarter of 2010. The density estimation is based on a normal kernel function, using an optimal window parameter function of number of data points. The distribution is based on 1000 draws from the predictive density.

Figure 9: RMSE



Note to Figure 9: the Figure shows the RMSFE of the factor model with stochastic volatility and of a baseline factor model without stochastic volatility between the first quarter of 2006 to the last quarter of 2010. The forecast horizon goes from six months ahead to one month after the end of the quarter of interest (backast). Therefore the first forecast is produced with the information set available in the middle of September 2005, the last one with data released at the end of January 2011.

Table 5: Coverage Rates - Baseline Model

Nom Cov	Backcast		Nowcast		1 step ahead	
	Coverage	P-value	Coverage	P-value	Coverage	P-value
0.1	0.14	0.63	0.15	0.25	0.17	0.15
0.2	0.32	0.26	0.23	0.60	0.26	0.29
0.3	0.50	**0.08**	0.41	**0.08**	0.42	**0.05**
0.4	0.59	**0.09**	0.50	0.11	0.55	**0.02**
0.5	0.59	0.41	0.56	0.33	0.58	0.22
0.6	0.64	0.73	0.67	0.26	0.59	0.88
0.7	0.77	0.44	0.73	0.62	0.61	0.13
0.8	0.86	0.41	0.79	0.81	0.65	**0.01**
0.9	1.00	**0.09**	0.88	0.60	0.74	**0.01**

Note to Table 5. Nominal and estimated Coverage Probabilites and p-values for the hypothesis that they are equal. Significant differences a the 10% level are in boldface.

Table 6: Coverage Rates - Model with Stochastic Volatility

Nom Cov	Backcast		Nowcast		1 step ahead	
	Coverage	P-value	Coverage	P-value	Coverage	P-value
0.1	0.09	0.89	0.14	0.40	0.05	**0.04**
0.2	0.18	0.83	0.26	0.29	0.23	0.60
0.3	0.32	0.86	0.32	0.75	0.30	0.96
0.4	0.45	0.62	0.41	0.88	0.44	0.52
0.5	0.59	0.41	0.47	0.63	0.48	0.81
0.6	0.68	0.43	0.61	0.92	0.58	0.69
0.7	0.86	**0.04**	0.73	0.62	0.71	0.83
0.8	0.91	0.10	0.82	0.71	0.73	0.19
0.9	0.95	0.24	0.89	0.87	0.77	**0.02**

Note to Table 6. Nominal and estimated Coverage Probabilites and p-values for the hypothesis that they are equal. Significant differences a the 10% level are in boldface.

Table 7: Density forecast evaluation using the *pits*: Baseline Model

	Backcast		Nowcast						One step ahead					
	1	2	1	2	3	4	5	6	1	2	3	4	5	6
Mean	0.53	0.54	0.52	0.43	0.32	0.39	0.62	0.41	0.45	0.34	0.17	0.18	0.33	0.43
Variance	**0.00**	0.33	0.05	0.01	0.34	0.35	0.22	0.92	0.03	0.09	0.90	0.96	0.99	0.13
AR(1)	0.14	0.07	0.72	0.21	0.14	0.20	0.08	0.03	0.27	0.20	0.17	0.07	0.05	0.03
Joint	0.20	0.57	0.73	0.50	0.44	0.59	0.63	0.47	0.49	0.39	0.54	0.49	0.48	0.37

Note to Table 7. P-values for the null hypotheses of zero mean, unit variance, no serial correlation and joint Normality/Indipendence of forecast errors at different horizons. Backcast refers to two weeks (1) and one month (2) after the end of the quarter of interest. Nowcast refers to the first two weeks (1), the first month (2)and so on of the quarter of interest. One step ahead to the next quarter in the same periods as in Nowcast.

Table 8: Density forecast evaluation using the *pits*: Model with Stochastic Volatility

	Backcast		Nowcast						One step ahead					
	1	2	1	2	3	4	5	6	1	2	3	4	5	6
Mean	0.13	0.00	0.16	0.12	0.08	0.08	0.01	0.13	0.22	0.13	0.05	0.04	0.00	0.22
Variance	**0.05**	**0.00**	0.84	0.38	0.63	0.98	0.24	0.83	0.06	0.20	0.20	0.01	0.01	0.78
AR(1)	0.53	0.63	0.46	0.73	0.02	0.02	0.22	0.67	0.53	0.34	0.22	**0.03**	0.91	0.20
Joint	0.50	0.37	0.71	0.68	0.11	0.24	0.51	0.84	0.53	0.33	0.43	0.22	0.15	0.61

Note to Table 8. P-values for the null hypotheses of zero mean, unit variance, no serial correlation and joint Normality/Indipendence of forecast errors at different horizons. Backcast refers to two weeks (1) and one month (2) after the end of the quarter of interest. Nowcast refers to the first two weeks (1), the first month (2)and so on of the quarter of interest. One step ahead to the next quarter in the same periods as in Nowcast.

Figure 10: Inefficiency factors for the draws from the ergodic distribution for the time constant hyperparameters and for the states

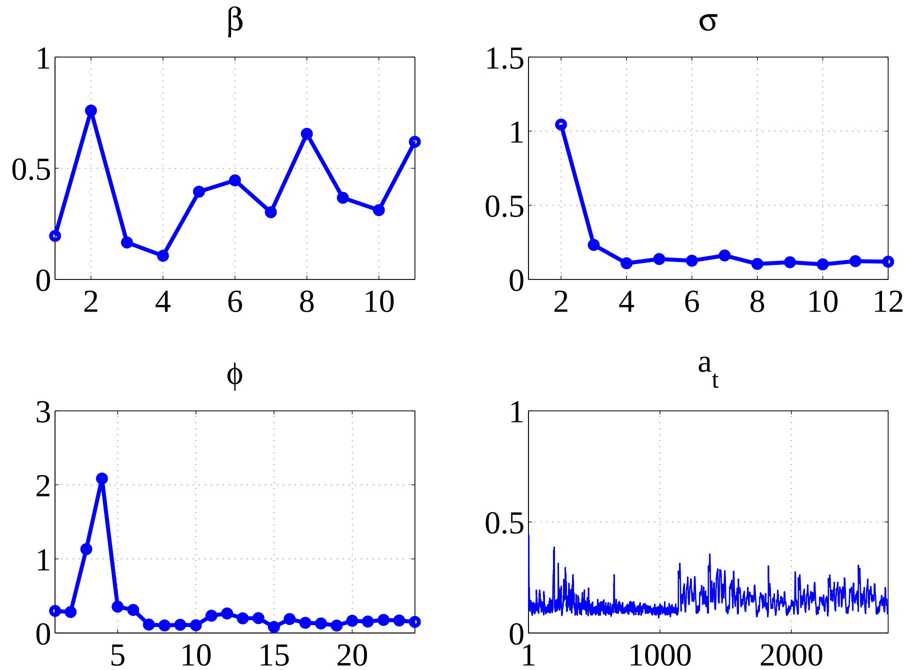


Figure 11: Inefficiency factors for the draws from the ergodic distribution for the stochastic volatilities