Combined Density Nowcasting in an Uncertain Economic Environment*

Knut Are Aastveit† Francesco Ravazzolo‡ Herman Van Dijk§

Preliminary version: October 2, 2014

Abstract

We introduce a Combined Density Nowcasting (CDN) approach to Dynamic Factor Models (DFM) that accounts for time-varying uncertainty of several model and data features in order to provide more accurate and complete density nowcasts. The combination weights depend on past nowcasting performance and other learning mechanisms that are incorporated in a Bayesian Sequential Monte Carlo method which re-balances the set of nowcasted densities in every period using the updated information on the time-varying weights. In this way, we are able to weight data uncertainty, parameter uncertainty, model uncertainty, including model incompleteness, and uncertainty in the combination of weights in a coherent way. Using experiments with simulated data we show that the CDN approach works particularly well in the presence of large data uncertainty and model incompleteness. For empirical analysis we use U.S. real-time data and obtain as results that our CDN approach gives more accurate density nowcasts of GDP growth than a model selection strategy and other combination strategies throughout the quarter. The relative gains are particularly large for the two first months of the quarter. CDN performs also well with respect to focusing on the tails and delivers probabilities of stagnation, measured as negative growth, that provide good signals for calling recessions in real time, and that are in line with forecasts from the Survey of Professional Forecasters.

JEL-codes: C11, C13, C32, C53, E37

Keywords: Density forecast combination; Survey forecast; Bayesian filtering; Sequential Monte Carlo Nowcasting; Real-time data

*We would like to thank Jonathan Wright, seminar and conference participants at Norges Bank, the Bank of Portugal, the Allied Social Science Association annual meeting in Philadelphia, the European Meeting of the Econometric Society in Toulouse, International Association for Applied Econometrics annual meeting in London and the International Symposium of Forecasting in Rotterdam. The views expressed in this paper are those of the authors and should not be attributed to Norges Bank.

†Norges Bank, Knut-Are.Aastveit@norges-bank.no

‡Norges Bank and BI Norwegian Business School, Francesco.Ravazzolo@norges-bank.no

§Econometric Institute, Erasmus University Rotterdam, Econometrics Department VU University Amsterdam and Tinbergen Institute, hkvandijk@ese.eur.nl

1 Introduction

Economic forecast and decision making in real time are, in recent years, made under a high degree of uncertainty. One prominent feature of this uncertainty is that many key statistics are released with a long delay, are subsequently revised and are available at different frequencies. Therefore, professional economists in business and government, whose job is to track the swings in the economy and to make forecasts that inform decision-makers in real time, prefer to examine a large number of potential relevant time series. In this context factor models provide a convenient and efficient tool to exploit information in a large panel of time series in a systematic way by allowing for information reduction in a parsimonious manner while retaining forecasting power. This is achieved by summarizing the information of the many data releases within a few common factors.

Several studies have found such factor models very useful for forecasting, see e.g., Stock and Watson (2002a,b, 2006), Forni et al. (2005) and Boivin and Ng (2005). A recent study by Giannone et al. (2008) shows that they are particularly suitable for *nowcasting*. The basic principle of nowcasting is the exploitation of the information which is published early and possibly at higher frequencies than the target variable of interest in order to obtain an "early estimate" before the official number becomes available, see Evans (2005) and Banbura et al. (2011). A key challenge is dealing with the differences in data release dates that cause the available information set to differ over points in time within the quarter. This is what Wallis (1986) coined the "ragged edge" of data. Giannone et al. (2008) evaluate point nowcasts from a dynamic factor model and highlight the importance of using non-synchronous data release. These authors show that the root mean square forecasting error decreases monotonically with each release.

The recent academic literature on factor models and nowcasting has focused on developing single models that increase forecast accuracy in terms of point nowcasts, see, among others, Banbura and Modugno (2014) and Banbura and Rünstler (2011). As there is considerable uncertainty regarding several features of the model specification, for example, choice of variables to include in the large data set, choice of number of factors, choice of lag length, etc., recent work by Clark and McCracken (2009, 2010) suggested to follow the idea of Bates and Granger (1969) and combine forecasts from a wide range of models with different features in order to reduce these problems.[1] Surprisingly however, few studies in the nowcasting literature focus on combining nowcasts from different models, Kuzin et al. (2013) and Aastveit et al. (2014) being notable exceptions. Furthermore, the research interest in forecast combination has more recently focused on the construction of combinations of predictive densities and not point forecasts, see e.g. Hall and Mitchell (2007) and Jore et al. (2010).[2] A recent extension to density forecasting is to allow for time varying model weights with learning and model set incompleteness, see Billio et al. (2013). Using a combination scheme that allows for model set incompleteness, seems particularly suitable for nowcasting, as economic decision makers produce their nowcasts based on both incomplete data information (ragged edge problem) and uncertainty about the true data generating process.

In this paper, we introduce a Combined Density Nowcasting (CDN) approach to Dynamic Factor Models (DFM) that accounts for time-varying uncertainty of several model and data features in order to provide more accurate and complete density nowcasts. The combination weights depend on past nowcasting performance and other learning mechanisms that are incorporated in a Bayesian Sequential Monte Carlo method which re-balances the set of nowcasted densities in every period using the updated information on the time-varying weights.[3] In this way, we are able to weight data uncertainty, parameter uncertainty, model uncertainty, including model incompleteness, and uncertainty in the combination of weights in a coherent way. We address the aforementioned sources of uncertainty using a large unbalanced real-time macroeconomic data set for the United States and combine predictive density nowcast from 4 different DFMs. The 4 DFMs varies in terms of the number of factors included.

In statistical terms, our CDN approach results in a convolution of a set of three probability density functions: the conditional density of the nowcasts of individual models, the conditional density of the latent weights of the combination scheme, and finally the density of the combination scheme. The integral of this product of three densities does not have a closed form solution and has to be evaluated numerically. The algorithm that we use to approximate the weight and combination scheme densities is an extension of Billio et al. (2013) to the case of

[1]The idea of combining forecasts from different models have been widely used for economic forecasting. Timmermann (2006) provides an extensive survey of different combination methods.

[2]See also Aastveit et al. (2014) for a nowcasting application.

[3]Note the analogy with dynamic portfolio management of a set of assets where periodically a rebalancing of the assets occurs depending on the dynamic pattern of the weights that incorporate past performance of the assets.

dynamic factor models with model incompleteness and data uncertainty. The application of a Sequential Monte Carlo method leads to good approximations of these two densities in the convolution. The procedure is computational intensive, when the number of models to combine increases. However, by making use of recent increases in computing power and recent advances in parallel programming technique it is feasible to apply the non-linear time-varying weights to the 4 factor models at different points in time during the quarter. In doing so, we apply the MATLAB package DeCo (Density Combination), developed by Casarin et al. (2013), which provide an efficient implementation of the algorithm in Billio et al. (2013) based on CPU and GPU parallel computing.

We first implement simulation experiments in order to understand the role of incompleteness for nowcasting . We distinguish between data incompleteness (ragged edge problem) and model set incompleteness (the true model is not a part of the forecasters' model space) and compare point and density nowcasting performance from our CDN approach with the performance of a Bayesian Model Averaging (BMA) approach and the ex post best individual model. The results illustrate that all three approaches provide accurate point and density nowcasts when there is no incompleteness. However, when data incompleteness and/or model set incompleteness is present, the point and density nowcasting performance from the CDN approach is superior to both the BMA approach and the ex post best individual model, providing considerably more accurate nowcasts.

Next, we show the usefulness of our CDN approach applied to 4 different DFMs for nowcasting GDP growth using U.S. real-time data. We divide data into different blocks, according to their release date within the quarter, and update the density nowcasts at three different points in time during each month of the quarter for the evaluation period 1990Q2-2010Q3. Our experiment refers to a professional economist who is interested in dealing with various forms of uncertainty in real-time, including model specifications.

We find that the CDN approach outperforms a BMA approach, a selection strategy and even the ex-post best individual model in terms of density nowcasting performance for all blocks. Interestingly, the relative gains in terms of improved density nowcasts are larger for the first blocks of the quarter than for the last blocks of the quarter. By studying the standard deviation of the combination residual, we show that this is higher for the earlier blocks in the quarter than for the later blocks in the quarter, indicating that incompleteness plays a larger role in the early part of the quarter. Thus, there are clear gains in terms of improved nowcasting performance from our CDN when incompleteness is present. Finally, the standard deviations of the combination residuals fluctuate over time and seem to increase during economic downturns. We document that the CDN approach also performs well with respect to focusing on the tails and delivers probabilities of stagnation, measured as the probability of negative growth, that provide timely warning signals for calling a recession and are in line with forecasts from the Survey of Professional Forecasters.

The structure of the paper is as follows. Section 2 introduces our CDN approach. Section 3 describes the data. Section 4 contains results using simulated data and Section 5 provides results of the application of the proposed method to U.S. nowcasting. Section 6 concludes. In the Appendix, we provide additional figures.

2 Combined Density Nowcasting to Dynamic Factor Models

There is considerable empirical evidence that Dynamic Factor Models (DFMs) provide accurate short-term forecasts, see e.g., Giannone et al. (2008) and Banbura and Modugno (2014). These models are particularly useful in a data rich environment, where common latent factors and shocks are assumed to drive the co-movements between aggregate and disaggregate variables and the real-time data flow is inherently high dimensional with data released at different frequencies. We build on this literature and propose a general model structure which can deal with both uncertainty related to data due to different sample frequencies and data releases, and uncertainty regarding model specification, such as selecting the number of factors and the information set.

We start to describe how individual factor models cope with data uncertainty. Next, we specify the convolution of the three probability density functions that involve a novel combination scheme that deals with model uncertainty including model incompleteness and we end with a brief description of the algorithms used to evaluate the convolution of densities.

2.1 Individual Factor Model

Assume we have a monthly (m) unbalanced dataset X_{t_m}, where the unbalancedness is due to data being released at different points in time (ragged edge). Let $X_{t_m} = (x_{1,t_m}, \ldots, x_{N,t_m})'$ be a vector of observable and stationary monthly variables which have been standardized to have mean equal to zero and variance equal to one. A dynamic factor model is then given by the following observation equation:

$$X_{t_m} = \chi_{t_m} + \epsilon_{t_m} = \Lambda F_{t_m} + \epsilon_{t_m} \tag{1}$$

where Λ is a $(n \times r)$ matrix of factor loadings, $F_m = \begin{pmatrix} f_{1t_m}, & \dots, & f_{rt_m} \end{pmatrix}'$ is the static common factors and $\epsilon_{t_m} = \begin{pmatrix} \epsilon_{1t_m}, & \dots, & \epsilon_{nt_m} \end{pmatrix}'$ is an idiosyncratic component with zero expectation and $\Psi_{t_m} = E\left[\epsilon_{t_m}\epsilon'_{t_m}\right]$ as covariance matrix.

The dynamics of the common factors follows a VAR process:

$$F_{t_m} = AF_{t_m-1} + Bu_{t_m} \tag{2}$$

where $u_m \sim WN\left(0, I_s\right)$, B is a $(r \times s)$ matrix of full rank s, A is a $(r \times r)$ matrix where all roots of $\det(I_r - Az)$ lie outside the unit circle. The idiosyncratic and VAR residuals are assumed to be independent:

$$\begin{bmatrix} \epsilon_{t_m} \\ u_{t_m} \end{bmatrix} \sim i.i.d.N\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} R & 0 \\ 0 & Q \end{bmatrix} \right) \tag{3}$$

with R set to be diagonal.[4]

Lastly, predictions of quarterly GDP growth, y_{t_q}, are obtained by using a bridge equation where nowcast of quarterly GDP growth (y_{t_q}) are expressed as a linear function of the expected common factors:

$$y_{t_q} = \alpha + \beta' F_{t_q} + \varsigma_{t_q} \tag{4}$$

The monthly factors F_{t_m} given k initial conditions, are first forecasted over the remainder of the quarter using equation (2) to produce the quarterly aggregate density $p(F_{t_q+h}|k)$. To obtain quarterly aggregates of the monthly factors, $(F_{t_q} = F^{(3)}_{t_m})$, we use the same approach as Giannone et al. (2008) and Aastveit et al. (2014). Prior to estimating equation (1) and (2), we transform

[4] The estimates are robust to violations of this assumption, see e.g. Banbura et al. (2012)

each monthly variable to correspond to a quarterly quantity when observed at the end of the quarter. Quarterly differences are therefore calculated as $x_{t_q} = x_{t_m}^{(3)} = (1-L_m^3)(1+L_m+L_m^2)Z_{t_m}$, where L_m is the monthly lag operator and Z_{t_m} is the raw data. Likewise quarterly growth rates are calculated as $x_{t_q} = x_{t_m}^{(3)} = (1 - L_m^3)(1 + L_m + L_m^2)logZ_{t_m}$.

In order to estimate equations (1), (2) and (4) one can make use of Bayesian approaches based on Monte Carlo or frequentist estimation principles. In our case we take a pragmatic approach and make use of standard frequentist approaches based on bootstrapping in order to estimate equations (1), (2) and (4), and then compute $p(F_{t_q+h}|k)$ and $p(\widetilde{y}_{t_q+h}|F_{t_q+h}, k)$. Here we apply the boostrapping approach developed in Aastveit et al. (2014) and refer to that paper for more details. Thus, motivated by Fernandez et al. (2001) and Sala-I-Martin et al. (2004), we make use of Bayesian averaging of frequentist estimates, extending their Bayesian averaging approach to account for time-varying weights and model set incompleteness.[5]

2.2 A convolution of combination, weight and predictive densities

While the dynamic factor model can cope with unbalanced data and provide forecasts of quarterly GDP growth using monthly information, there is considerable uncertainty regarding model specification, such as selecting the number of factors (r) and other components of the information set (X). This can potentially result in, say K, different DFM specifications. Selection criteria and various testing procedure have been proposed in order to address such problems, see e.g. Bai and Ng (2006).

Instead, we propose to follow the approach by Strachan and Dijk (2013), and rely on Bayesian combination of several model features. We extend their approach of using fixed model weights to the situation where we combine a set of predictive densities of model and data features using time-varying weights as well as allowing for model incompleteness, meaning that the true model is not necessarily included in the model set. Given that we obtain a combined predictive density of quarterly growth, we can report tail probabilities of such features as high, low an even negative growth.

The combined density is a convolution of the conditional predictive densities of the the different models, the conditional densities of the latent weights and the density of the combination scheme. Assume that there are K specifications of different models, then we propose to compute

[5] We leave for further research to make an efficient Bayesian estimation procedure for the DFM that we use.

the combined density nowcast of GDP growth $p(y_{t_q+h}|y_{1:t_q})$ as:

$$p(y_{t_q+h}|y_{1:t_q}) = \int_{\widetilde{y}_{t_q+h}} \left(\int_{w_{t_q+h}} p(y_{t_q+h}|\widetilde{y}_{t_q+h}, w_{t_q+h}, F_{t_q+h}, K) p(w_{t_q+h}|w_{t_q}) dw_{t_q+h} \right) p(\widetilde{y}_{t_q+h}|F_{t_q+h}, K) p(F_{t_q+h}|K) d\widetilde{y}_{t_q+h} \tag{5}$$

where $p(y_{t_q+h}|\widetilde{y}_{t_q+h}, w_{t_q+h}, F_{t_q+h}, K)$ is the combination scheme for the K different predictive densities with a first-order Markov combination weights distributed as $p(w_{t_q+h}|w_{t_q})$; $p(\widetilde{y}_{t_q+h}|F_{t_q+h}, K)$ is a vector of K predictive densities for the variable y_{t_q+h} following equation (4) with K different initial conditions; and $p(F_{t_q+h}|K)$ is a $K-$vector of the predictive densities for the factors given by equation (2) with K different initial conditions. Notice that the combined density $p(y_{t_q+h}|y_{1:t_q})$ is computed in a recursive way depending on past data. In the previous section, we described how we estimated the set of predictive densities $p(\widetilde{y}_{t_q+h}|F_{t_q+h}, K)$; hereby we discuss how to compute the other densities.

We make use of a Gaussian combination scheme which allows for model incompleteness. This is done via the following specification:

$$p(y_{t_q+h}|\widetilde{y}_{t_q+h}, w_{t_q+h}, F_{t_q+h}, K) \propto \exp\left\{ -\frac{1}{2} \left(y_{t_q+h} - w_{t_q+h}\tilde{y}_{t_q+h} \right)' \sigma^{-1} \left(y_{t_q+h} - w_{t_q+h}\tilde{y}_{t_q+h} \right) \right\} \tag{6}$$

where $w_{t_q+h} = (w_{1,t_q+h}, ..., w_{K,t_q+h})$ is a $(1 \times K)$ vector containing the K densities for the combination weights and $\tilde{y}_{t_q+h}$ is a $(K \times 1)$ vector containing the K predictive densities $p(\widetilde{y}_{t_q+h}|F_{t_q+h}, k)$, $k = 1, ..., K$.

In our modeling strategy, combination residuals are estimated and their distribution follows a Gaussian process with mean zero and standard deviation σ, providing a probabilistic measure of the incompleteness of the model set. In other words, the model that is specified in equation (6) can be written as:

$$y_{t_{q+h}} = w_{t_q+h}\tilde{y}_{t_q+h} + \zeta_{t_q+h} \tag{7}$$

with $\zeta_{t_q+h} \sim \mathcal{N}(0, \sigma^2)$.

Thirdly, the combination weights w_{t_q+h} have a probabilistic distribution in the unit interval and we model them as logistic transforms, given as

$$w_{k,t_q+h} = \frac{\exp\{z_{k,t_q+h}\}}{\sum_{j=1}^{M} \exp\{z_{j,t_q+h}\}}, \; k = 1, ..., K \tag{8}$$

and

$$p(z_{t_q+h}|z_{t_q}, \tilde{y}_{t_q-\tau:t_q}) \propto \exp\left\{-\frac{1}{2}\left(\Delta z_{t_q+h} - \Delta e_{t_q+h}\right)' \Lambda^{-1} \left(\Delta z_{t_q+h} - \Delta e_{t_q+h}\right)\right\} \tag{9}$$

with $\Delta z_{t_q+h} = z_{t_q+h} - z_{t_q}$, $z_{t_q+h} = (z_{1,t_q+h}, ..., z_{K,t_q+h})$ and $\Delta e_{t_q+h} = e_{t_q+h} - e_{t_q}$ where $e_{t_q+h} = (e_{1,t_q+h}, ..., e_{K,t_q+h})$ is a learning function based on past predictive performances. Therefore, z_{t_q+h} is a latent process evolving over time with dynamics following a first-order Markov specification depending on past performances which describes the contribution of each model in the combination. The logistic transformation restricts weights to be in the unit interval. Furthermore, we define

$$e_{k,t_q+h} = (1-\lambda) \sum_{i=\tau}^{t_q} \lambda^{i-1} e_{k,i}, \; k = 1, .., K$$

where λ is a discount factor, and $(t_q - \tau + 1)$ is the length of the learning parameter. In the empirical application we set $\lambda = 0.95$ and $\tau = 1$. Following the discussion in Gneiting (2011), we propose that different scoring rules should be applied depending on the user preference. Therefore, a user interested in point forecasting could focus on mean square prediction errors; a user with a more general loss function should focus on scores that are based on density forecasting, such as the log score, see section 2.4. A user just interested to standard Bayesian updating and no learning based on past performance scores can set $\Delta e_{t_q+h} = 0$ and weights will be driven by a process equal to the previous values plus a news component normally distributed with zero mean and Λ covariance matrix.

If the three densities in equation (5) did all belong to the normal family with no dynamics, the integral in (5) could be solved analytically or by simple numerical methods like direct Monte Carlo simulation. The integral in our case does not have a closed form solution. However, there is a perfect analogy between the set up of the equations in our CDN approach and the model specification in nonlinear State Space literature. Therefore we can apply a Sequential Monte Carlo method by interpreting the combination scheme in a state space formulation. Equation (6) is the measurement or observable equation; equations (8) and (9) are nonlinear transition equations. Equations (1), (2) and (4) can be interpreted as being equivalent to the parameter equations in the nonlinear State Space but they can also be presented as being part of a general State Space model where the nonlinear filtering methods are to be used to approximate these densities. Note that equation (5) accounts for several sources of uncertainty, including different sample frequencies, different data releases, different information sets and model specifications.

The convolution has such useful properties like commutative, associative and distributive laws that enable is to be flexible in the order of integration and other properties. As mentioned, we use sequential Monte Carlo integration to solve part of the integral in (5) by using the regularized version of the Liu and West (2001) procedure and we make use of draws from the K individual predictive densities. Therefore $\tilde{y}^{j}_{t_q+h,k}$ is a draw from $p(\widetilde{y}_{t_q+h}|F_{t_q+h},1)$ and $w^{d-j}_{t_q+h,1}$ is a draw from $w_{t_q+h,k}$ with $k = 1, ..., K$.

Our methodology is very general and allows to convolute predictive densities provided by various methods (parametric Bayesian and Frequentist models as well as nonparametric methods), given the condition that $p(\widetilde{y}_{t_q+h}|F_{t_q+h},K)$ represents densities. We repeat that in the empirical applications in section 5 we construct predictive densities using frequentist bootstrapping methods and combine these predictive densities using Bayesian inference. The algorithm is explained in detail in the next section.

Our approach accounts for various sources of uncertainty, such as data uncertainty, parameter uncertainty, model uncertainty; and it estimates a time-varying weight w_{k,t_q+h} based on past predictive density performance for each of these components. The resulting predictive density will integrate out the aforementioned sources of uncertainty while allowing for model incompleteness. We label this as a Combined Density Nowcasting approach applied to Dynamic Factor Models.

2.3 Algorithm and parallelization

The main steps to estimate our general model are:

Step 1: Estimate K DFM models and generate draws for $\widetilde{F}_{k,t_m+h}$, $k = 1, ..., K$.

Step 2: Conditional on $\widetilde{F}_{k,t_m+h}$ generate draws of $\widetilde{y}_{t_q+h}$, $k = 1, ..., K$

Step 3: Combine the predictions from the K models, accounting for uncertainty on the number of factors (r) and information set (X), using the convolution mechanism.

We elaborate on each step briefly.

Step 1: The following bootstrap procedure is used to construct simulated forecasts. Let $\hat{A}_0 = [\hat{A}_1, \ldots, \hat{A}_p]$, $\hat{B}_0$, $\hat{u}_{0,t_m^x}$, $\hat{\xi}_{0,t_m^x}$, $\hat{\Lambda}_0$, $\hat{\alpha}_0$, $\hat{\beta}_0$, and $\hat{e}_{0,t_m+h_m}$ denote the initial point estimates. Then, for $d = 1, ..., 2000$:

1. Simulate monthly $\tilde{F}_{t_m^x} = \sum_{i=1}^{p} \hat{A}_i \tilde{F}_{t_m^x - i} + \hat{B}_0 u^*_{t_m^x}$, where $u^*_{t_m^x}$ is re-sampled from $\hat{u}_{0,t_m^x}$.

2. Simulate $\tilde{X}_{t_m^x} = \hat{\Lambda}_0 \tilde{F}_{t_m^x} + \xi^*_{t_m^x}$, where $\xi^*_{t_m^x}$ is re-sampled from $\hat{\xi}_{0,t_m^x}$.

3. Based on $\tilde{X}_{t_m^x}$, re-estimate the model to get a new set of parameter and factor estimates. Use these to generate factor forecasts according to equation (2), where shock uncertainty is included by re-sampling from $\hat{u}_{0,t_m^x}$.

Step 2: Estimate equation (4) based on the monthly factor estimated in the previous step and converted to quarterly as described in the previous section, and construct forecasts for $\widetilde{y}_{t_q+h}$ where shock uncertainty is included by re-sampling from $\hat{e}_{0,t_m+h_m}$.

Step 3: Apply an extension of the parallelized version of the sequential Monte Carlo algorithm of Billio et al. (2013) and Casarin et al. (2013) to the case of Dynamic Factor Models. For a technical description of this algorithm, we refer the reader to Casarin et al. (2013). Here, we provide some details on the prior. The combination weights are [0,1]-valued processes and one can interpret them a sequence of prior probabilities over the set of models. In our framework, the prior probability on the set of models is random, as opposite to the standard model selection or BMA frameworks, where the model prior is fixed. The likelihood, given by the combination scheme, allows us to compute the posterior distribution on the model set. In this sense the proposed combination scheme shares some similarities with the dilution and hierarchical model set prior distributions for BMA, proposed in George (2010) and Ley and Steel (2009) respectively. The learning strategy also plays a crucial role and we propose to use scores depending on the loss function of interest. In the next section we describe our scores for forecast evaluation and for each metric we apply the corresponding score in the learning mechanism in (9). For all the cases, we also consider standard Bayesian updating.

We repeat steps 1-3 recursively for every block in each quarter vintage. The exercise is very time consuming and requires parallelization to be implemented. We parallelize the code in two directions. First, step 1 and step 2 are parallelized across models, vintages and blocks. Then, step 3 is parallelized across draws using the MATLAB toolbox DeCo described in Casarin et al. (2013).[6]

2.4 Forecast evaluation

The aim of this paper is to provide an efficient methodology which deals with various sources of uncertainty in order to improve nowcast accuracy. As most other papers focusing on nowcasting do, we provide first some results on point forecasts. However, as these forecasts are only optimal for a small and restricted group of loss functions, our main focus is on density forecasting. When evaluating the predictive nowcasts, we evaluate both the full distribution as well as their tails. For notational simplicity, we define $t = t_q$ in the remaining part of the paper.

To shed light on the predictive ability of our methodology, we consider several evaluation statistics for point and density forecasts previously proposed in the literature. Suppose we have $k = 1, ..., K$ different approaches to nowcast GDP. We compare point forecasts in terms of Root Mean Square Prediction Errors (RMSPE)

$$RMSPE_k = \sqrt{\frac{1}{t^*}\sum_{t=\underline{t}}^{\bar{t}} e_{k,t+h}}$$

where $t^* = \bar{t} - \underline{t} + h$, $\bar{t}$ and $\underline{t}$ denote the beginning and end of the evaluation period, and $e_{k,t+h}$ is the h-step ahead square prediction error of model k.

The complete predictive densities are evaluated using the Kullback Leibler Information Criterion (KLIC) based measure, utilizing the expected difference in the Logarithmic Scores of the candidate forecast densities; see, for example, Mitchell and Hall (2005), Hall and Mitchell (2007), Amisano and Giacomini (2007) and Kascha and Ravazzolo (2010). The KLIC chooses

[6] If the user was in the last vintage and block, parallelization across models in steps 1 and 2 and parallelization across predictive draws in step 3 are required to derive predictive densities for future values.

the model that on average gives the higher probability to events that actually occurred. Specifically, the KLIC distance between the true density $p(y_{t+h}|y_{1:t})$ of a random variable y_{t+h} and some candidate density $p(\tilde{y}_{k,t+h}|y_{1:t})$ obtained from model k is defined as

$$\begin{aligned} \mathrm{KLIC}_{k,t+h} &= \int p(y_{t+h}|y_{1:t}) \ln \frac{p(y_{t+h}|y_{1:t})}{p(\tilde{y}_{k,t+h}|y_{1:t})} dy_{t+h}, \\ &= \mathbb{E}_t[\ln p(y_{t+h}|y_{1:t}) - \ln p(\tilde{y}_{k,t+h}|y_{1:t}))]. \end{aligned} \tag{10}$$

where $\mathbb{E}_t(\cdot) = \mathbb{E}(\cdot|\mathcal{F}_t)$ is the conditional expectation given information set $\mathcal{F}_t$ at time t. An estimate can be obtained from the average of the sample information, $y_{\underline{t}+1}, \ldots, y_{\bar{t}+1}$, on $p(y_{t+h}|y_{1:t})$ and $p(\tilde{y}_{k,t+h}|y_{1:t})$:

$$\overline{KLIC}_k = \frac{1}{t^*} \sum_{t=\underline{t}}^{\bar{t}} [\ln p(y_{t+h}|y_{1:t}) - \ln p(\tilde{y}_{k,t+h}|y_{1:t})]. \tag{11}$$

Although we do not pursue the approach of finding the true density, we can still rank the different densities, $p(\tilde{y}_{k,t+h}|y_{1:t})$, $k = 1, \ldots, K$ by different criteria. For the comparison of two competing models, it is sufficient to consider the Logarithmic Score (LS), which corresponds to the latter term in the above sum,

$$LS_k = -\frac{1}{t^*} \sum_{t=\underline{t}}^{\bar{t}} \ln p(\tilde{y}_{k,t+h}|y_{1:t}), \tag{12}$$

for all k and to choose the model for which it is minimal, or, as we report in our tables, its opposite is maximal.

3 Data

We consider in total 120 monthly leading indicators to nowcast quarterly GDP growth in the United States. Our real-time dataset is similar to the one used in Aastveit et al. (2014).[7] As in that paper, we use the last available data vintage as real-time observations for consumer prices and survey data if the real-time data vintage is not available. For other series, such as disaggregated measures of industrial production, real-time vintage data exist only for parts

[7] The main source is the ALFRED (ArchivaL Federal Reserve Economic Data) database maintained by the Federal Reserve Bank of St. Louis. In addition some series are also collected from the Federal Reserve Bank of Philadelphia's Real-Time Data Set for Macroeconomists, see Croushore and Stark (2001).

of the evaluation period. For these variables, we use the first available real-time vintage and truncate these series backwards recursively. Finally, for financial data, we construct monthly averages of daily observations.

Following Banbura and Rünstler (2011) we divide the data into "soft data" and "hard data". The first set includes 38 surveys and financial indicators and reflects market expectations, as opposed to the latter set that includes 82 measures of GDP components (e.g. industrial production), the labor market and prices. The soft data are often timely available (i.e. early in the quarter), while real activity data are published with a significant delay but this latter category is considered to contain a more precise signal for GDP forecasting.

The full forecast evaluation period runs from 1990Q2 to 2010Q3. We use monthly real-time data with quarterly vintages from 1990Q3 to 2010Q4, i.e., we do not take account of data revisions in the monthly variables within a quarter.[8] The starting point of the estimation period is 1982M1. We study nowcasts at 9 different points in time during a quarter. They correspond to the beginning, middle and end of each month in the quarter. Since GDP measures are released approximately 20-25 days after the end of the quarter, our exercise also includes 2 backcasts, calculated at the beginning and the middle of the first month after the quarter of interest. See Table 1 for information on the final 11 blocks. When nowcasting GDP growth, the choice of a benchmark for the "actual" measure of GDP is not obvious (see Stark and Croushore (2002) for a discussion of alternative benchmarks). We follow Romer and Romer (2000) in using the second available estimate of GDP as the actual measure.

4 Simulation Exercise

In this section we implement several simulation exercises to understand what are the role of data incompleteness and model incompleteness for nowcasting. In practice, economic decision makers produce their nowcasts based on incomplete data information (ragged edge problem) and uncertainty about the true data generating process (DGP). In the simulation exercises below, we will therefore distinguish between different degrees of incompleteness. We refer to *weak incompleteness* in the case where the forecaster produce forecast based on missing observations of data (i.e. the ragged edge problem). The DGP is in this case assumed to be a part of the

[8]The quarterly vintages reflect information available just before the first release of the GDP estimate.

Table 1. Block information

Block	Time	Horizon
	Nowcasting	
1	Start of first month of quarter	2-step ahead
2	10th of first month of quarter (after inflation release)	2-step ahead
3	Around 20-25th of first month of quarter (after GDP release)	1-step ahead
4	Start of second month of quarter	1-step ahead
5	10th of second month of quarter (after inflation release)	1-step ahead
6	Around 20-25th of Second month of quarter	1-step ahead
7	Start of thirds month of quarter	1-step ahead
8	10th of Third month of quarter (after inflation release)	1-step ahead
9	Around 20-25th of third month of quarter	1-step ahead
	Backcasting	
10	Start of fourth month of quarter	1-step ahead
11	10th of fourth month of quarter (after inflation release)	1-step ahead

The table shows time in the quarter and forecast horizon for the 11 blocks.

forecasters' model space. On the contrary, when the DGP is not a part of the forecasts' model space, we refer to this as a case of *strong incompleteness*.

We run 4 simulation exercises, where in each exercise we produce recursive density nowcasts for 60 quarters. For the 3 first simulation exercises, we simulate y_t assuming that the DGP (DGP1) follows a dynamic factor model, described in section 2.1, with 2 factors extracted at the end of the sample (corresponding to the information set at Block 11). In the last simulation exercise, we assume that the DGP (DGP2) follows a VAR(4) in GDP growth, the unemployment rate, core PCE in inflation, and the federal funds rates. DGP2 is estimated from a balanced panel at the end of the sample. In each simulation exercise, we compare the performance of our CDN approach, both in terms of point nowcasts (MSPE) and density nowcasts (LS), with a Bayesian Model Averaging (BMA) approach as well as the best ex-post individual model.

In the first simulation exercise, (Sim1), we estimate (and combine) 4 individual DFMs with 1-4 factors extracted from a panel corresponding to the information at Block 11. Thus, in this exercise the DGP is a part of the model space and there is therefore no model set incompleteness and no data incompleteness. We introduce weak incompleteness in the second simulation exercise (Sim2). We estimate (and combine) the same individual DFMs with 1-4 factors. The only difference from Sim1 is that the models are now estimated with incomplete data information. More precisely, the models are estimated using data that corresponds to the information that is available when nowcasting at the middle of the quarter (i.e., Block 5).

Hence, there is data incompleteness, but no model incompleteness.

The last two simulation exercises focus on cases of strong incompleteness (cases where both data incompleteness and model incompleteness is present). In the third simulation exercise, (Sim3), we estimate (and combine) 4 individual DFMs. However, we assume that for some reason, the factors are only estimated based on the "hard data" variables in our data set (i.e. we assume that no survey data are available to the forecaster). Thus, there is weak model incompleteness, since the "true" model (which is a DFM with 2 factors extracted from the full data set) is within the model space, but all the models are misspecified in terms of using the wrong data set (i.e. using just a subset of all the "true" data series in order to extract the factors). In addition, we also assume that there is data incompleteness as in Sim2. In the final simulation exercise (Sim4), we also assume a different DGP. In this case, we assume that DGP follows a VAR(4) (DGP2) in GDP growth, the unemployment rate, inflation and the interest rate, while we again estimate and combine individual DFMs with 1-4 factors extracted from all the available data series (i.e. our estimated models are similar to the ones in the Sim2 exercise).

Table 2. Simulation results

	BMA	**Best model**	**CDN**
	Sim1: No incompleteness		
LS	-0.251	**0.224**	0.074
MSPE	0.028	0.025	**0.024**
	Sim2: Weak incompleteness		
LS	-3.882	-3.875	**-0.459**
MSPE	0.198	0.161	**0.147**
	Sim3: Strong incompleteness		
LS	-4.359	-4.328	**-0.457**
MSPE	0.241	0.240	**0.169**
	Sim4: Strong incompleteness		
LS	-0.567	-0.555	-0.325
MSPE	0.205	0.186	0.112

The table reports results from the 4 simulation exercises, showing the average log score (LS) and mean square prediction error (MSPE) for three different prediction methods: standard Bayesian model averaging based on predictive likelihood (BMA), the ex-post best performing model and our combined density nowcasting (CDN) approach applied to dynamic factor models. Bold numbers indicates the most accurate model for different statistics.

Table 2 reports results from the simulation exercises. When there is no model incompleteness the best individual model, CDN and BMA perform very similar in terms of point forecasts. There are some differences in terms on density forecasting performance, where the CDN approach clearly outperforms the BMA approach. As expected, the best individual model outperforms both BMA and CDN in terms of density nowcasting. Still, the results indicate that the CDN approach seems work well in the case where there is no data and model incompleteness. When introducing data and model incompleteness, there are clear gains from using our CDN approach relative to the other strategies. Starting with the case of weak incompleteness (i.e., Sim 2 where only data incompleteness is present), our CDN approach substantially improves upon the BMA approach, both in terms of point and density nowcasts performance. Interestingly, the CDN approach also outperforms the ex-post best individual model. This result is rather striking, as the only source of incompleteness, is missing data observations (ragged edge problem). Thus, this indicates that using a combination scheme that allows for model incompleteness is important in the case where data observations are missing. The relative improvements, compared to the other strategies, are even more evident in the cases of strong incompleteness (Sim3 and Sim4). Comparing the nowcasting performance from our CDN approach with the BMA approach and ex-post best individual model, indicates that there is a scope for substantial improvements in performance by using a combination scheme that allows for model incompleteness, when both data and model set incompleteness is present.[9]

5 Empirical Application

In this section, we analyze the performance of our CDN approach for nowcasting U.S. real GDP growth. The main goal of the exercise is to examine the nowcasting performance of our CDN approach and to study the role of model incompleteness for nowcasting.

5.1 Point and density nowcasts of GDP growth

We produce density nowcasts/backcasts for GDP growth at 11 different points in time, described in section 3, using 4 different DFMs. The models differ in terms of the numbers of factors included.[10] Our exercise refers to a researcher who construct nowcasts in real time accounting for

[9]Note that since DGP1 and DGP2 are rather different, it may be misleading to compare the absolute performance for each model from the two different simulation exercises (Sim3 and Sim4).

[10]We obtained very similar results when using 12 different DFMs: 4 models extracting factors from the hard data; 4 models using the soft data; and 4 models using all the data. For each group, we then considered 1 to 4 factors, resulting in 4 different DFM specifications for each data group. In general, the models using factors

various forms of uncertainty, including uncertainty related to model specification. We consider three different model specification strategies:

1. SEL: A selection strategy where we recursively pick the model with the highest realized cumulative log score at each point in time throughout the evaluation period.

2. BMA: A Bayesian model averaging approach based on predictive likelihood.

3. CDN: Our Combined Density Nowcasting approach, applied to the 4 DFMs

Table 3 reports results for the three different model specification strategies at the 11 different points in time (blocks) during the quarter. In addition, we also report results for the best performing ex-post individual model (labeled Ex-Post). The first column, shows the LS and MSPE for the BMA approach, while all other columns report measures relative to the BMA performance. The table reveals three interesting results. First, with the exception of the results for Block 1 and Block 2, the point nowcasting accuracy from the different models are very similar. Second, the CDN approach provides more accurate density nowcasts than BMA and SEL for all of the blocks. It also provides more accurate density nowcasts for all blocks than the ex-post best individual model, with the only exceptions being results for block 8 and block 10, where Ex Post performs slightly better than the CDN. Overall, this indicates that there are clear gains in terms of improved nowcastng performance from our CDN approach when we take into account the whole density shape of the nowcasts. Third, the relative gains in terms of improved density nowcasts, are larger for the first blocks of the quarter than for the last blocks of the quarter. This supports the findings from the simulation exercises in section 4, which showed that the gains from the CDN approach are larger when uncertainty is high, and thus the incompleteness is strong. In the early part of the quarter the data incompleteness (denoted as weak incompleteness) is larger than in the latter part of the quarter. In addition, when the data uncertainty is high, it is also more likely that it becomes harder to detect the "true" DGP than when the data uncertainty is low. That is, it is also more likely that model incompleteness is present when data uncertainty is high.

To illustrate the role of higher incompleteness, figure 1 shows the standard deviations of the combination residuals for the incomplete model sets, see equation 7, over time for Block

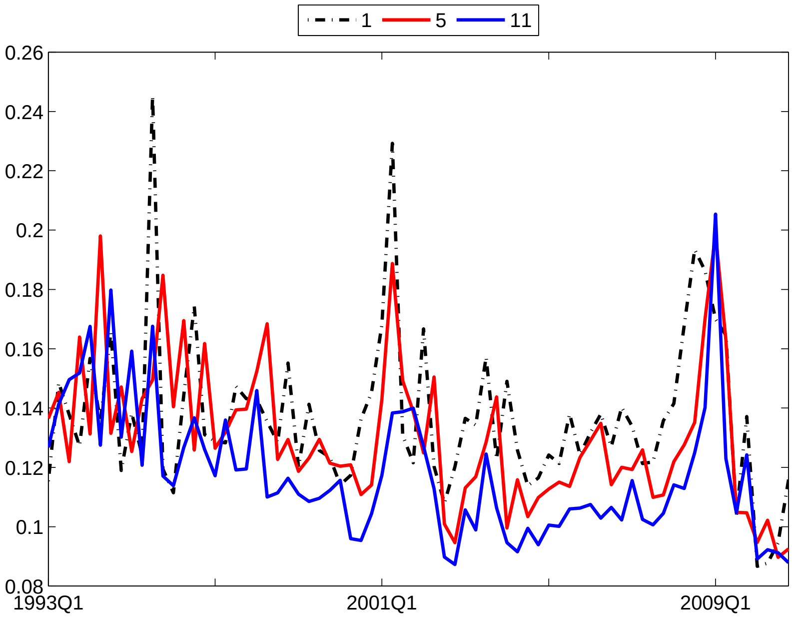
extracted form all the data series were superior to the models extracting factors from either hard or soft data. For brevity, and in order to save computational time, we therefore only report results when combining 4 different DFMs.

Table 3. Point and density forecasting

	BMA	SEL	Ex Post	CDN
		Block 1		
LS	-1.441	1.124	0.926	**0.590**
MSPE	0.583	0.988	**0.524**	0.542
		Block 2		
LS	-1.101	1.117	0.954	**0.715**
MSPE	0.317	1.032	0.959	**0.924**
		Block 3		
LS	-0.980	0.987	0.977	**0.814**
MSPE	0.289	0.989	**0.983**	1.025
		Block 4		
LS	-0.892	0.997	0.978	**0.862**
MSPE	0.275	0.991	**0.977**	1.007
		Block 5		
LS	-0.768	0.991	0.961	**0.897**
MSPE	0.241	0.990	**0.969**	1.002
		Block 6		
LS	-0.788	0.993	0.964	**0.882**
MSPE	0.247	0.989	**0.969**	0.984
		Block 7		
LS	-0.743	0.990	0.953	**0.911**
MSPE	0.242	0.991	**0.958**	0.969
		Block 8		
LS	-0.619	1.000	**0.968**	0.995
MSPE	0.203	0.995	**0.972**	1.024
		Block 9		
LS	-0.655	0.998	0.965	**0.949**
MSPE	0.218	1.002	0.979	**0.973**
		Block 10		
LS	-0.594	1.023	**0.951**	0.998
MSPE	0.189	1.011	**0.980**	1.031
		Block 11		
LS	-0.610	0.995	0.952	**0.931**
MSPE	0.187	0.991	**0.974**	0.989

The table shows average log score (LS) and mean square prediction error (MSPE) for four different prediction methods: standard Bayesian model averaging based on predictive likelihood (BMA), selecting the model with highest recursive score at each point in time (SEL), the ex-post best performing model and our combined density nowcasting (CDN) approach applied to dynamic factor models for different blocks. The results in the second, third and fourth column show LS and MSPE relative to the BMA measure. Bold numbers indicates the most accurate model for different statistics. See Table 1 for information on different blocks.

Figure 1. Standard deviation of the combination residuals



Standard deviation of the combination residuals for incomplete model sets from equation 7, for Block 1, Block 5 and Block 11.

1, Block 5 and Block 11. The figure reveals two interesting observations. First, for most of the time observations, the standard deviation of the combination residuals is higher for Block 1 than Block 5 and Block 11, and higher for Block 5 than Block 11. This observation therefore confirms that incompleteness is higher in the early part of the quarter than in the later part of the quarter. Second, the standard deviations of the combination residuals fluctuate over time. Interestingly, the standard deviation of the combination residual is high in 2001 and in the latter part of 2008 and the early part of 2009. This coincides with the U.S. economy being in a recession. The high standard deviation is evident for Block 1 and Block 5 for the 2001-recession, and even more pronounced for the Great Recession, increasing the standard deviation for the combination residual for all blocks. In section 5.2 we will more carefully study the performance of our CDN approach during economic downturns.

Figure 2 shows the weights associated with the 4 dynamic factor models for Block 1, Block 5 and Block 11. We notice the large uncertainty on the weights, with substantial variation over time. There is a clear indication that DFMs with either 1 or 2 factors obtain higher weights than DFMs with 3 and 4 factors. Moreover, the weights are also changing between the blocks. Finally, the red dotted line in each subfigure shows the corresponding weights obtained by the BMA approach. Comparing the CDN weights with the BMA weights, we see two interesting differences. First, the median of the CDN weights and BMA weights differ substantially, where there are much larger movements over time from the BMA weights. Second, the BMA approach selects much more extreme weights, attaching almost all the weights to one single model, consistent with findings in Amisano and Geweke (2013). The main difference between our CDN approach and the BMA approach, is that our weighting scheme allows for model incompleteness.[11]

Finally, figure 3 shows a full set of recursive real-time out-of-sample density nowcasts for U.S. GDP growth for the period 1990Q2-2010Q3 at three different blocks (Block 1, 5 and 11). The three panels illustrate how the precision of the predictive densities improves, i.e., being more narrow and centered around the actual GDP values, as more information becomes available.

5.2 Prediction of the business cycle phases

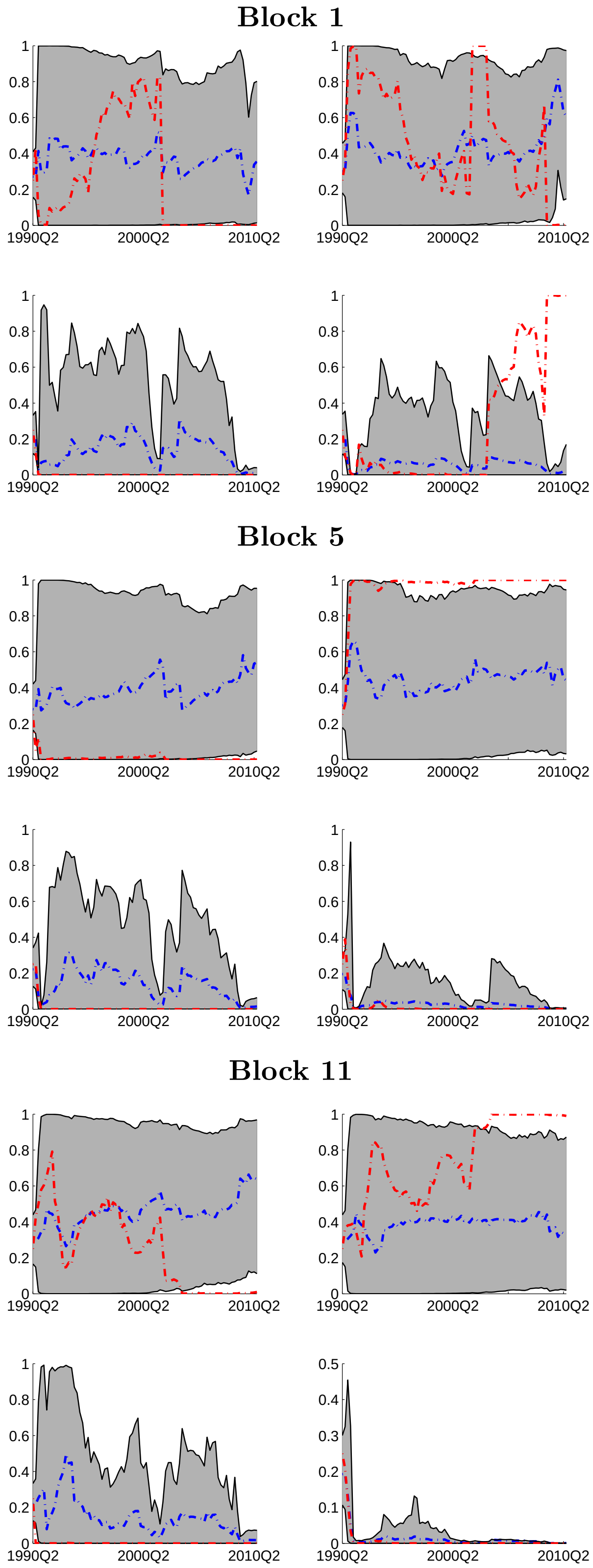
In the previous section we have shown that the CDN approach provides accurate nowcasts when focusing on the entire distribution of GDP growth. The full distribution of the CDN can also be used to compute probabilities to be in specific phases of the business cycle. There is a large literature on estimating and timely detect turning points and economic downturns, see e.g., Harding and Pagan (2002), Chauvet and Piger (2008), Hamilton (2011) and Stock and Watson (2014). The individual economists in the The Survey of Professional Forecasters (SPF) also report forecasts of the probability of a decline in the level of real GDP in the current quarter and the following four quarters. Motivated by this, we use the CDN to study the probability of negative growth in the current quarter (i.e., GDP growth nowcasts below 0).

Figure 4 compares the recursive probabilities of negative growth in the current quarter from our CDN approach with the mean responses for the probability of negative growth in the current quarter provided by the SPF. To ensure that the information set used to construct our CDN nowcasts are as similar as possible to the information available when the SPF forecasts where made, we report report CDN nowcasts for Block 5.[12] By comparing the the CDN and the SPF forecasts with actual GDP growth (shown by the bars), we find that both the CDN and SPF

[11] The BMA weights based on predictive likelihood will also take into account past predictive performance scores.

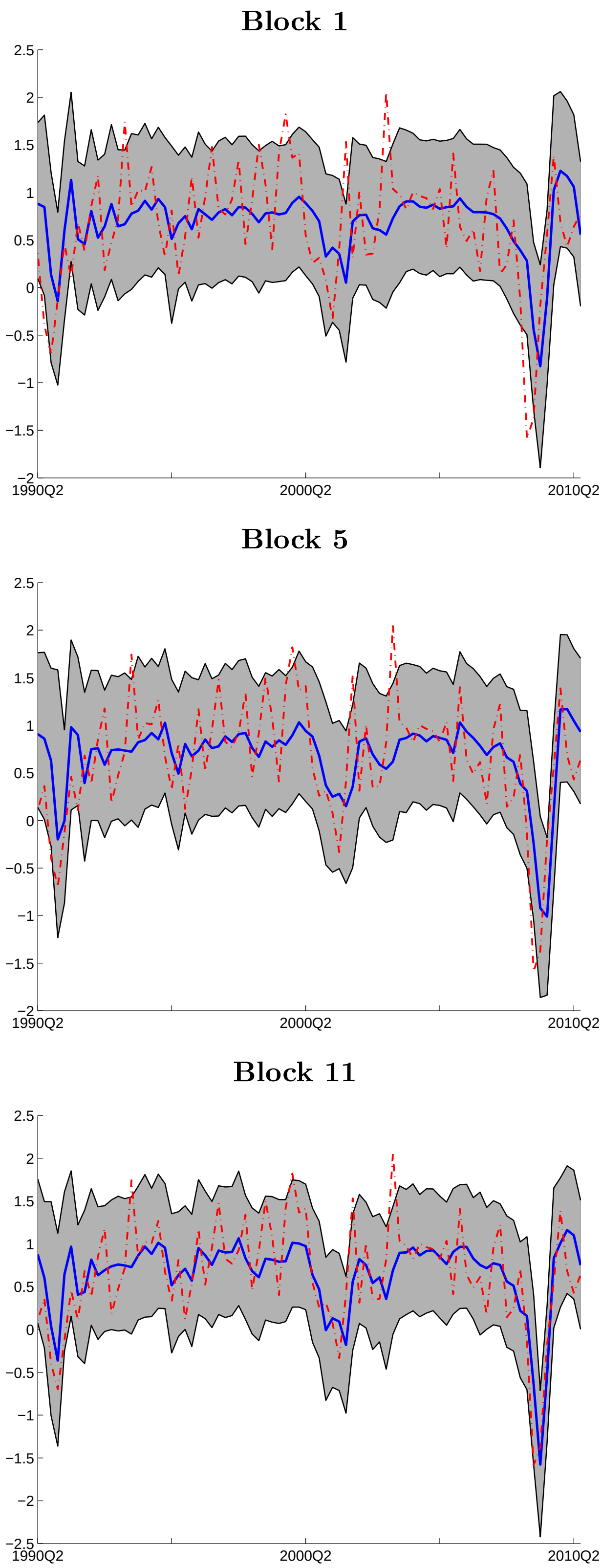
[12] Block 5 corresponds to the information set a few days prior to the release of the SPF forecasts.

Figure 2. Time-varying weights



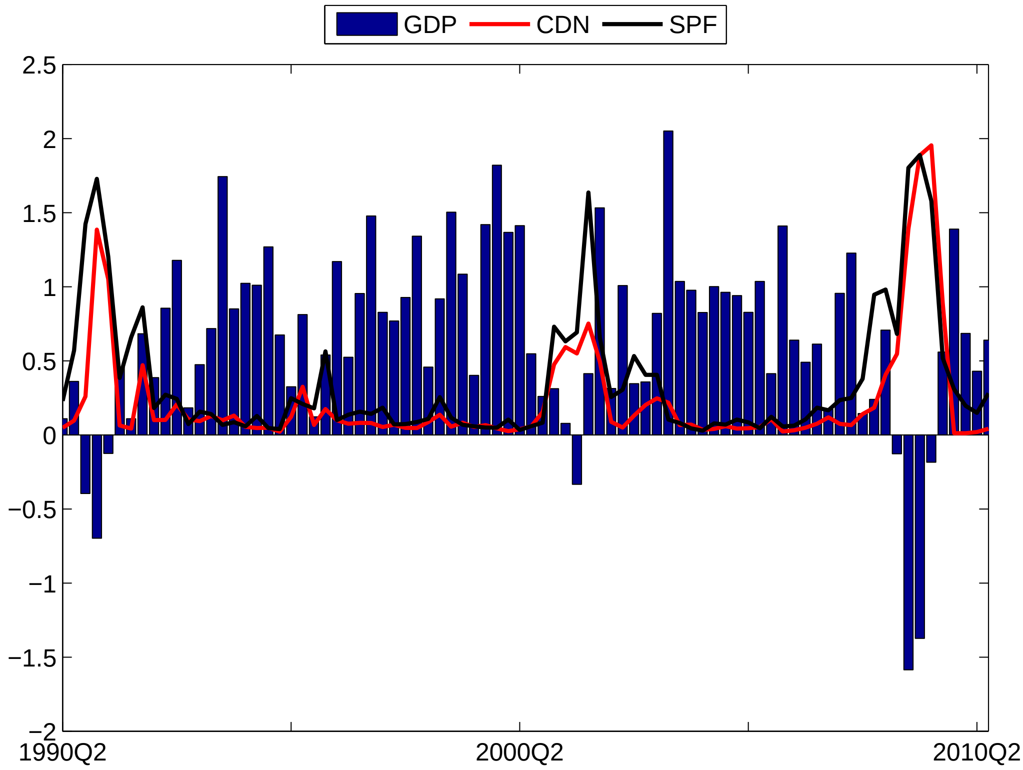
The figures plot the 90% credibility intervals of the model posterior weights and their medians (blue dotted lines) for Block 1, 5, and 11. The first row of each sub-figure shows weights for DFM models with 1 and 2 factors. The second row of each sub-figure shows weights for DFM models with 3 and 4 factors. The red dotted line shows the weights attached to each model using the BMA approach.

Figure 3. Recursive Nowcasts



The figures plot recursive nowcasts for Block 1, 5 and 11. The shaded areas show the 90% credibility intervals of the predictive densities and their medians (blue dotted lines). The red dotted line show actual GDP, measured as the 2nd release.

Figure 4. Probabilities of negative growth



Probabilities over time of negative quarterly growth given by the CDN approach and SPF. The red and black lines plot the probabilities scaled by 2 (therefore covering the interval [0,2]); the bars plot the realization.

forecats deliver timely and accurate forecasts of negative growth.

To provide insights about which method is more accurate, we can compute concordance statistics (CS). The concordance statistics counts the proportion of time during which the predicted and the actual GDP series are in the same state. We assume two states, either being in a state of negative growth or in a state of positive growth. We say that a model predicts negative growth for the current quarter if the probability of negative growth is 50% or larger. Comparing the CS for the CDN with the SPF, we find that they perform equally good with $CS = 0.963$.

6 Conclusion

In this paper, we have introduced a Combined Density Factor Model (CDFM) approach that accounts for time-varying uncertainty of several model and data features in order to provide more accurate and complete density nowcasts. The combination weights depend on past nowcasting performance and other learning mechanisms that are incorporated in a Bayesian Sequential Monte Carlo method which re-balances the set of nowcasted densities in every period using the updated information on the time varying weights. In this way, we are able to weight data uncertainty, parameter uncertainty, model uncertainty, including model incompleteness, and uncertainty in the combination of weights in a coherent way.

We first implement simulation experiments in order to understand the role of incompleteness for nowcasting, distinguishing between data incompleteness (ragged edge problem) and model set incompleteness (the true model is not a part of the forecasters' model space). By comparing point and density nowcasting performance from our CDN approach with the performance of a Bayesian Model Averaging (BMA) approach and the ex post best individual model, we find that the CDN approach provides superior nowcasts.

We then show the usefulness of our CDN approach applied to 4 different DFMs for nowcasting GDP growth using U.S. real-time data. The experiment refers to a professional economist who is interested in dealing with various forms of uncertainty in real-time. We therefore data into different blocks, according to their release date within the quarter, and update the density nowcasts at three different points in time during each month of the quarter for the evaluation period 1990Q2-2010Q3.

We find that the CDN approach outperforms a BMA approach, a selection strategy and even the ex-post best individual model in terms of density nowcasting performance for all blocks. Interestingly, the relative gains in terms of improved density nowcasts are larger for the first blocks of the quarter than for the last blocks of the quarter. By studying the standard deviation of the combination residual, we show that this is higher for the earlier blocks in the quarter than for the later blocks in the quarter, indicating that incompleteness plays a larger role in the early part of the quarter. Thus, there are clear gains in terms of improved nowcasting performance from our CDN when incompleteness is present. Finally, the standard deviations of the combination residuals fluctuate over time and seem to increase during economic downturns. We document that the CDN approach also performs well with respect to focusing on the tails and delivers probabilities of stagnation, measured as the probability of negative growth, that are timely and in line with forecasts from the Survey of Professional Forecasters.

References

Aastveit, K., K. Gerdrup, A. Jore, and L. Thorsrud (2014). Nowcasting GDP in real time: A density combination approach. *Journal of Business and Economic Statistics 32*(1), 48–68.

Amisano, G. and J. Geweke (2013). Prediction using several macroeconomic models. Working Paper Series, European Central Bank 1537, European Central Bank.

Amisano, G. and R. Giacomini (2007). Comparing density foreasts via weighted likelihood ratio tests. *25*(2), 177–190.

Bai, J. and S. Ng (2006). Confidence intervals for diffusion index forecasts and inference for factor-augmented regressions. *Econometrica 74*(4), 1133–1150.

Banbura, M., D. Giannone, M. Modugno, and L. Reichlin (2012). Now-casting and the real-time data flow. In G. Elliott and A. Timmermann (Eds.), *Handbook of Economic Forecasting Volume 2*. Elsevier North Holland.

Banbura, M., D. Giannone, and L. Reichlin (2011). Nowcasting. In M. P. Clements and D. F. Hendry (Eds.), *The Oxford Handbook of Economic Forecasting*. Oxford Handbooks in Economics.

Banbura, M. and M. Modugno (2014). Maximum likelihood estimation of factor models on data sets with arbitrary pattern of missing data. *Journal of Applied Econometrics 29*(1), 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. *International Journal of Forecasting 27*(2), 333–346.

Bates, J. and C. Granger (1969). The combination of forecasts. *Operations Research Quarterly 20*(4), 451–468.

Billio, M., R. Casarin, F. Ravazzolo, and H. K. van Dijk (2013). Time-varying combinations of predictive densities using nonlinear filtering. *Journal of Econometrics 177*(2), 213–232.

Boivin, J. and S. Ng (2005). Understanding and comparing factor-based forecasts. *International Journal of Central Banking 1*(3), 117–151.

Casarin, R., S. Grassi, F. Ravazzolo, and H. K. van Dijk (2013). Parallel sequential monte carlo for efficient density combination: The deco matlab toolbox. Technical report.

Chauvet, M. and J. Piger (2008). Comparison of the real-time performance of business cycle dating methods. *Journal of Business and Economic Statistics 26*, 42–49.

Clark, T. E. and M. W. McCracken (2009). Improving forecast accuracy by combining recursive and rolling forecasts. *International Economic Review 50*(2), 363–395.

Clark, T. E. and M. W. McCracken (2010). Averaging forecasts from vars with uncertain instabilities. *Journal of Applied Econometrics 25*(1), 5–29.

Croushore, D. and T. Stark (2001). A real-time data set for macroeconomists. *Journal of Econometrics 105*(1), 111–130.

Evans, M. D. (2005). Where are we now? real-time estimates of the macro economy. *International Journal of Central Banking 1*(2), 127–175.

Fernandez, C., E. Ley, and M. F. J. Steel (2001). Model uncertainty in cross-country growth regressions. *Journal of Applied Econometrics 16*(5), 563–76.

Forni, M., M. Hallin, M. Lippi, and L. Reichlin (2005). The generalized dynamic factor model: One-sided estimation and forecasting. *Journal of the American Statistical Association 100*, 830–840.

George, E. I. (2010). Dilution priors: Compensating for model space redundancy. In J. O. Berger, T. T. Cai, and I. M. Johnstone (Eds.), *Borrowing Strength: Theory Powering Applications - A Festschrift for Lawrence D. Brown*, pp. 158–165. Beachwood, Ohio, USA: Institute of Mathematical Statistics.

Giannone, D., L. Reichlin, and D. Small (2008). Nowcasting: The real-time informational content of macroeconomic data. *Journal of Monetary Economics 55*(4), 665–676.

Gneiting, T. (2011). Making and evaluating point forecasts. *Journal of the American Statistical Association 106*, 746–762.

Hall, S. G. and J. Mitchell (2007). Combining density forecasts. *International Journal of Forecasting 23*(1), 1–13.

Hamilton, J. D. (2011). Calling recessions in real time. *International Journal of Forecasting 27(4)*, 1006–1026.

Harding, D. and A. Pagan (2002). Dissecting the Cycle: A Methodological Investigation. *Journal of Monetary Economics 49*, 365–381.

Jore, A. S., J. Mitchell, and S. P. Vahey (2010). Combining forecast densities from vars with uncertain instabilities. *Journal of Applied Econometrics 25*(4), 621–634.

Kascha, C. and F. Ravazzolo (2010). Combining inflation density forecasts. *Journal of Forecasting 29*(1-2), 231–250.

Kuzin, V., M. Marcellino, and C. Schumacher (2013). Pooling versus model selection for nowcasting with many predictors: An application to German GDP. *Journal of Applied Econometrics* (forthcoming).

Ley, E. and M. F. J. Steel (2009). On the effect of prior assumptions in bayesian model averaging with applications to growth regression. *Journal of Applied Econometrics 24(4)*, 651–674.

Mitchell, J. and S. G. Hall (2005). Evaluating, comparing and combining density forecasts using the klic with an application to the bank of england and niesr 'fan' charts of inflation. *Oxford Bulletin of Economics and Statistics 67*(s1), 995–1033.

Romer, C. D. and D. H. Romer (2000). Federal reserve information and the behavior of interest rates. *American Economic Review 90*(3), 429–457.

Sala-I-Martin, X., G. Doppelhofer, and R. I. Miller (2004). Determinants of Long-Term Growth: A Bayesian Averaging of Classical Estimates (BACE) Approach. *American Economic Review 94*(4), 813–835.

Stark, T. and D. Croushore (2002). Forecasting with a real-time data set for macroeconomists. *Journal of Macroeconomics 24*(4), 507–531.

Stock, J. and M. Watson (2002a). Forecasting using principal components from a large number of predictors. *Journal of the American Statistical Association 97*, 1167–1179.

Stock, J. and M. Watson (2002b). Macroeconomic forecasting using diffusion indexes. *Journal of Business & Economic Statistics 20*(2), 147–62.

Stock, J. H. and M. W. Watson (2006). *Forecasting with Many Predictors*, Volume 1 of *Handbook of Economic Forecasting*, Chapter 10, pp. 515–554. Elsevier.

Stock, J. H. and M. W. Watson (2014). Estimating turning points using large data sets. *Journal of Econometrics 178*(P2), 368–381.

Strachan, R. W. and H. K. V. Dijk (2013). Evidence on features of a dsge business cycle model from bayesian model averaging. *International Economic Review 54*(1), 385–402.

Timmermann, A. (2006). Forecast combinations. In G. Elliott, C. W. J. Granger, and A. Timmermann (Eds.), *Handbook of Economic Forecasting*, Volume 1, pp. 136–96. Amsterdam: Elsevier.

Wallis, K. F. (1986). Forecasting with an econometric model: The ”ragged edge” problem. *Journal of Forecasting 5*, 1–13.

Data description

Data Group	Description	Transformation	Publication Lag	Start Vintage
Hard	Federal funds rate	1	One month	Last vintage
Hard	3 month Treasury Bills	1	One month	Last vintage
Hard	6 month Treasury Bills	1	One month	Last vintage
Soft	Spot USD/EUR	2	One month	Last vintage
Soft	Spot USD/JPY	2	One month	Last vintage
Soft	Spot USD/GBP	2	One month	Last vintage
Soft	Spot USD/CAD	2	One month	Last vintage
Soft	Price of gold on the London market	2	One month	Last vintage
Soft	NYSE composite index	2	One month	Last vintage
Soft	Standard & Poors 500 composite index	2	One month	Last vintage
Soft	Standard & Poors dividend yield	2	One month	Last vintage
Soft	Standard & Poors P/E Ratio	2	One month	Last vintage
Soft	Moodys AAA corporate bond yield	1	One month	Last vintage
Soft	Moodys BBB corporate bond yield	1	One month	Last vintage
Soft	WTI Crude oil spot price	2	One month	Last vintage
Soft	Purchasing Managers Index (PMI)	1	One month	03.03.1997
Soft	ISM mfg index, Production	1	One month	02.11.2009
Soft	ISM mfg index, Employment	1	One month	02.11.2009
Soft	ISM mfg index, New orders	1	One month	02.11.2009
Soft	ISM mfg index, Inventories	1	One month	02.11.2009
Soft	ISM mfg index, Supplier deliveries	1	One month	02.11.2009
Hard	Civilian Unemployment Rate	1	One month	05.01.1990
Hard	Civilian Participation Rate	1	One month	07.02.1997
Hard	Average (Mean) Duration of Unemployment	2	One month	05.01.1990
Hard	Civilians Unemployed - Less Than 5 Weeks	2	One month	05.01.1990
Hard	Civilians Unemployed for 5-14 Weeks	2	One month	05.01.1990
Hard	Civilians Unemployed for 15-26 Weeks	2	One month	05.01.1990
Hard	Civilians Unemployed for 27 Weeks and Over	2	One month	05.01.1990
Hard	Employment on nonag payrolls: Total nonfarm	2	One month	05.01.1990
Hard	Employment on nonag payrolls: Total Private Industries	2	One month	05.01.1990
Hard	Employment on nonag payrolls: Goods-Producing Industries	2	One month	05.01.1990
Hard	Employment on nonag payrolls: Construction	2	One month	05.01.1990
Hard	Employment on nonag payrolls: Durable goods	2	One month	05.01.1990
Hard	Employment on nonag payrolls: Nondurable goods	2	One month	05.01.1990
Hard	Employment on nonag payrolls: Manufacturing	2	One month	05.01.1990
Hard	Employment on nonag payrolls: Mining and logging	2	One month	05.01.1990

Data Group	Description	Transformation	Publication Lag	Start Vintage
Hard	Employment on nonag payrolls: Service-Providing Industries	2	One month	05.01.1990
Hard	Employment on nonag payrolls: Financial Activities	2	One month	05.01.1990
Hard	Employment on nonag payrolls: Education & Health Services	2	One month	06.06.2003
Hard	Employment on nonag payrolls: Retail Trade	2	One month	05.01.1990
Hard	Employment on nonag payrolls: Wholesale Trade	2	One month	05.01.1990
Hard	Employment on nonag payrolls: Government	2	One month	05.01.1990
Hard	Employment on nonag payrolls: Trade, Transportation & Utilities	2	One month	05.01.1990
Hard	Employment on nonag payrolls: Leisure & Hospitality	2	One month	06.06.2003
Hard	Employment on nonag payrolls: Other Services	2	One month	05.01.1990
Hard	Employment on nonag payrolls: Professional & Business Services	2	One month	06.06.2003
Hard	Average weekly hours of PNW: Total private	2	One month	Last vintage
Hard	Average weekly overtime hours of PNW: Mfg	2	One month	Last vintage
Hard	Average weekly hours of PNW: Mfg	2	One month	Last vintage
Hard	Average hourly earnings:Construction	2	One month	Last vintage
Hard	Average hourly earnings: Mfg	2	One month	Last vintage
Hard	M1 Money Stock	2	One month	30.01.1990
Hard	M2 Money Stock	2	One month	30.01.1990
Soft	Consumer credit: New car loans at auto finance companies, loan-to-value	2	Two months	Last vintage
Soft	Consumer credit: New car loans at auto finance companies, amount financed	2	Two months	Last vintage
Hard	Federal government total surplus or deficit	2	One month	Last vintage
Hard	Exports of goods, total census basis	2	Two months	Last vintage
Hard	Imports of goods, total census basis	2	Two months	Last vintage
Hard	Industrial Production Index	2	One month	17.01.1990
Hard	Industrial Production: Final Products (Market Group)	2	One month	14.12.2007
Hard	Industrial Production: Consumer Goods	2	One month	14.12.2007
Hard	Industrial Production: Durable Consumer Goods	2	One month	14.12.2007
Hard	Industrial Production: Nondurable Consumer Goods	2	One month	14.12.2007
Hard	Industrial Production: Business Equipment	2	One month	14.12.2007
Hard	Industrial Production: Materials	2	One month	14.12.2007
Hard	Industrial Production: Durable Materials	2	One month	14.12.2007
Hard	Industrial Production: nondurable Materials	2	One month	14.12.2007
Hard	Industrial Production: Manufacturing (NAICS)	2	One month	14.12.2007
Hard	Industrial Production: Durable Manufacturing (NAICS)	2	One month	14.12.2007
Hard	Industrial Production: Nondurable Manufacturing (NAICS)	2	One month	14.12.2007
Hard	Industrial Production: Mining	2	One month	14.12.2007
Hard	Industrial Production: Electric and Gas Utilities	2	One month	14.12.2007
Hard	Capacity Utilization: Manufacturing (NAICS)	1	One month	05.12.2002
Hard	Capacity Utilization: Total Industry	1	One month	15.11.1996
Soft	Housing starts: Total new privately owned housing units started	2	One month	18.01.1990

Data Group	Description	Transformation	Publication Lag	Start Vintage
Soft	New private housing units authorized by building permits	2	One month	17.08.1999
Soft	Phily Fed Buisness outlook survey, New orders	1	Current month	Last vintage
Soft	Phily Fed Buisness outlook survey, General business activity	1	Current month	Last vintage
Soft	Phily Fed Buisness outlook survey, Shipments	1	Current month	Last vintage
Soft	Phily Fed Buisness outlook survey, Inventories	1	Current month	Last vintage
Soft	Phily Fed Buisness outlook survey, Unfilled orders	1	Current month	Last vintage
Soft	Phily Fed Buisness outlook survey, Prices paid	1	Current month	Last vintage
Soft	Phily Fed Buisness outlook survey, Prices received	1	Current month	Last vintage
Soft	Phily Fed Buisness outlook survey, Number of employees	1	Current month	Last vintage
Soft	Phily Fed Buisness outlook survey, Average workweek	1	Current month	Last vintage
Hard	Producer Price Index: Finished Goods	2	One month	12.01.1990
Hard	Producer Price Index: Finished Goods Less Food & Energy	2	One month	11.12.1996
Hard	Producer Price Index: Finished Consumer Goods	2	One month	11.12.1996
Hard	Producer Price Index: Intermediate Materials: Supplies & Components	2	One month	12.01.1990
Hard	Producer Price Index: Crude Materials for Further Processing	2	One month	12.01.1990
Hard	Producer Price Index: Finished Goods Excluding Foods	2	One month	11.12.1996
Hard	Producer Price Index: Finished Goods Less Energy	2	One month	11.12.1996
Hard	Consumer Prices Index: All Items (urban)	2	One month	18.01.1990
Hard	Consumer Prices Index: Food	2	One month	12.12.1996
Hard	Consumer Prices Index: Housing	2	One month	Last vintage
Hard	Consumer Prices Index: Apparel	2	One month	Last vintage
Hard	Consumer Prices Index: Transportation	2	One month	Last vintage
Hard	Consumer Prices Index: Medical care	2	One month	Last vintage
Hard	Consumer Prices Index: Commodities	2	One month	Last vintage
Hard	Consumer Prices Index: Durables	2	One month	Last vintage
Hard	Consumer Prices Index: Services	2	One month	Last vintage
Hard	Consumer Prices Index: All Items Less Food	2	One month	12.12.1996
Hard	Consumer Prices Index: All Items Less Food & Energy	2	One month	12.12.1996
Hard	Consumer Prices Index: All items less shelter	2	One month	Last vintage
Hard	Consumer Prices Index: All items less medical care	2	One month	Last vintage
Hard	Real Gross Domestic Product	2	One quarter	28.01.1990
Hard	Real Disposable Personal Income	2	One month	29.01.1990
Hard	Real Personal Consumption Expenditures	2	One month	29.01.1990
Hard	Real Personal Consumption Expenditures: Durable Goods	2	One month	29.01.1990
Hard	Real Personal Consumption Expenditures: Nondurable Goods	2	One month	29.01.1990
Hard	Real Personal Consumption Expenditures: Services	2	One month	29.01.1990
Hard	Personal Consumption Expenditures: Chain-type Price Index	2	One month	01.08.2000
Hard	Personal Consumption Expenditures: Chain-Type Price Index Less Food & Energy	2	One month	01.08.2000
Soft	New one family houses sold	2	One month	30.07.1999

Data Group	Description	Transformation	Publication Lag	Start Vintage
Soft	New home sales: Ratio of houses for sale to houses sold	2	One month	Last vintage
Soft	Existing home sales: Single-family and condos	2	One month	Last vintage
Soft	Chicago Fed MMI Survey	2	One month	Last vintage
Soft	Composite index of 10 leading indicators	1	One month	Last vintage
Soft	Consumer confidence surveys: Index of consumer confidence	1	Current month	Last vintage
Soft	Michigan Survey: Index of consumer sentiment	1	Current month	31.07.1998
Hard	Average weekly initial claims	2	Current month	Last vintage

Note: In column 4, 1 denotes differencing to the initial series and 2 denotes log differencing to the initial series.