Nowcasting the Business Cycle in an Uncertain Enviroment*

Knut Are Aastveit† Francesco Ravazzolo‡ Herman Van Dijk§

Preliminary version: December 27, 2013

Abstract

We introduce 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. By combining predictive densities from a set of dynamic factor models, using combination weights that are time-varying, depend on past predictive forecasting performance and other learning mechanisms that are incorporated in a Bayesian Sequential Monte Carlo method, we are able to weight 'soft' and 'hard' data uncertainty, parameter uncertainty, model uncertainty and uncertainty in the combination of weights in a coherent way. Using experiments with simulated data our results show that soft data contain useful information for nowcasting even if the series is generated from the hard data. Moreover, a carefully combination of hard and soft data, as in the proposed approach, improves density nowcasting. For empirical analysis we use U.S. real-time data and obtain as results that our CDFM approach yields more accurate nowcasts of GDP growth and more accurate prediction of NBER Business cycle turning points than other combination strategies. Interestingly, the CDFM performs particularly well, relative to other combination strategies, when focusing on the tails and it delivers timely and accurate probabilities of high growth and stagnation..

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

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

---

*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), 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 (2010) 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. (2013) 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).

In this paper, we introduce a Combined Density Factor Model (CDFM) approach that accounts for time varying factor model uncertainty in order to provide more accurate nowcasts of predictive densities. By combining predictive densities from a set of dynamic factor models, using combination weights that are time-varying and depend on past predictive forecasting performance and other learning mechanisms and by making use of a Bayesian Sequential Monte Carlo method, we are able to weight "soft" and "hard" data uncertainty, parameter uncertainty, model uncertainty 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, similar to Aastveit et al. (2013).

We allow for data and model feature uncertainty by varying the data set and the number of factors. The former is motivated by Banbura and Rünstler (2011) dividing the data into "soft data" and "hard data". The soft data include financial data and surveys and reflect market expectations while the hard data include measures of certain components of GDP (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 considered to contain a more precise signal for the measurement of GDP. We apply a factor model to three different data sets: soft data, hard data and all data series (i.e. including both hard and soft data) and let the number of factor vary from 1 to 4 factors. In total we consider 12 different factor models.

Our approach to combine density nowcasts from 12 factor models leads to a Combined Density Factor Model (CDFM) where the combined density is a convolution of the set of individual model densities. The algorithm that we use is an extension of Billio et al. (2013) to the case

[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. (2013) for a nowcasting application.

of dynamic factor models with data uncertainty. A Sequential Monte Carlo method is used to approximate the filtering and predictive densities. 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 12 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 a simulation experiment to compare soft and hard data and analyze the performance of the CDFM. The results illustrate that soft data contain useful information due to being timely available and increase both point and density nowcast performance even when the true data are generated from hard database. Furthermore, CDFM with optimal learning based on density nowcasting provides better density nowcasts than any of the individual models.

Next, we show the usefulness of the CDFM for nowcasting GDP growth and business cycle turning points 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. We repeat that our CDFM includes 12 different dynamic factor models: 4 models are based on hard data; 4 models are based on soft data; and 4 models are based on all data. In each group, we consider 1 to 4 factors, resulting in 4 specifications. 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 CDFM outperforms all individual models in terms of log score (LS) and cumulative rank probability score (CRPS)for all blocks and results. Interestingly, the favorable nowcasting properties from the CDFM also applies when focusing on the tails of the predictive distribution. Moreover, the CDFM also outperform the strategy of selecting the models with the highest realized cumulative log score as well as Bayesian model averaging based on predictive likelihood. Finally, we show that a real-time indicator based on the Bry and Boschan (1971) (BB) rule and nowcasts from our model are more accurate in terms of concordance statistics than those given by the alternative methods. Probabilities of high growth and stagnation given by the CDFM are timely and accurate.

The structure of the paper is as follows. Section 2 introduces our CDFM 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 Model

2.1 Combined Density Factor Model: Overview of state space model and density convolution

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 (2010). 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.

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, $F_m = \begin{pmatrix} f_{1t_m}, & \ldots, & f_{rt_m} \end{pmatrix}'$ is the static common factors and $\epsilon_{t_m} = \begin{pmatrix} \epsilon_{1t_m}, & \ldots, & \epsilon_{nt_m} \end{pmatrix}'$ 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 $u_m \sim WN(0, I_s)$, 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.[3]

Lastly, predictions of quarterly GDP growth, y_{t_q}, are obtained by using a bridge equation. The monthly factors F_{t_m} are first forecasted over the remainder of the quarter using equation (2). To obtain quarterly aggregates of the monthly factors, ($F_{t_q} = F_{t_m}^{(3)}$), we use the same approach as Giannone et al. (2008) and Aastveit et al. (2013). Prior to estimating equation (1) and (2), we transform 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 $x_{t_q} = x_{t_m}^{(3)} = (1 - L_m^3)(1 + L_m + L_m^2)logZ_{t_m}$.

The nowcast of quarterly GDP growth (y_{t_q}), can then be expressed as a linear function of the expected common factors:

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

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 the information set (X). This can potentially result in M 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. Then, we can report tail probabilities of such features as high, low an even negative growth. For the sake of brevity, we define the K specifications as K different initial conditions. The predictive density that can be derived using equations (1), (2) and (4)

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

and using a general density function of the time varying weights. This yields (see for details 2.2) :

$$p(y_{t_q+h}) = \int p(y_{t_q+h}|\widetilde{y}_{t_q+h}, w_{t_q+h}, F_{t_q+h}, K)p(\widetilde{y}_{t_q+h}|F_{t_q+h}, K)p(w_{t_q+h}|w_{t_q})p(F_{t_q+h}|K)dK \quad (5)$$

where $p(F_{t_q+h}|K)$ is the predictive densities for the factors given by equation (2) with K different initial conditions; $p(\widetilde{y}_{t_q+h}|F_{t_q+h}, K)$ is a set of K predictive densities for the variable y_{t_q+h} following equation (4) with K different initial conditions; and $p(y_{t_q+h}|\widetilde{y}_{t_q+h}, w_{t_q+h}, F_{t_q+h}, K)$ is the combination scheme for the K different predictive densities with combination weights distributed as $p(w_{t_q+h}|w_{t_q})$. The function $p(y_{t_q+h}|\widetilde{y}_{t_q+h}, w_{t_q+h}, F_{t_q+h}, K)$ is based on a convolution mechanism that produces a modified predictive density for y_{t_q+h} of the K original $p(\widetilde{y}_{t_q+h}|F_{t_q+h}, K)$ densities.

Furthermore, we follow Billio et al. (2013), and propose a Gaussian combination, with logistic-Gaussian weights with learning based on past predictive performances which allows for model incompleteness, i.e., the "true" model is not a part of the model space. This is done via the following specification:

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

where $\tilde{y}_{t_q+h}$ is a matric containing the K predictive densities $p(\widetilde{y}_{t_q+h}|F_{t_q+h}, k)$, $k = 1, ..., K$; and where the weights $w_{t_q+h} = (w_{1,t_q+h}, ..., w_{K,t_q+h})$ are logistic transforms with K models

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

and

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

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, see section 2.3

for possible scoring functions of predictive densities. We define

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

where LS is defined in equation 13 below, 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 $\tau = 1$. The convolution applies in equation (6), and it precisely creates the following combined density:

$$(\tilde{y} * w)_{t_q+h} = \int_{-\inf}^{+\inf} (\tilde{y}_{t_q+h,1}(\xi) * \ldots * \tilde{y}_{t_q+h,K}(\xi)) * (w_{t_q+h,1}(d-\xi) * \ldots * w_{t_q+h,K}(d-\xi)) d\xi \quad (7)$$

where $\tilde{y}_{t_q+h} = (\tilde{y}_{t_q+h,1} * \ldots * \tilde{y}_{t_q+h,K})$; $w_{t_q+h} = (w_{t_q+h,1} * \ldots * w_{t_q+h,K})$. For each value of ξ, the convolution formula can be described as a weighted average of the function $\tilde{y}_{t_q+h}(\xi)$ with weight $w_{t_q+h}(d-\xi)$. As d changes, the weighting function emphasizes different parts of the input function. We follow Billio et al. (2013) and use different draws from the K individual predictive densities as values for ξ.

Convolution has several important mathematical properties, see for example Damelin and Miller (2011), that we exploit to derive equation (7):

- Property 1: $(\tilde{y} * w)_{t_q+h} = (w * \tilde{y})_{t_q+h}$.
- Property 2: $\tilde{y} * (w * \gamma)_{t_q+h} = ((w * \tilde{y}) * \gamma)_{t_q+h}$.
- Property 3: $\tilde{y} * (w + \gamma)_{t_q+h} = (\tilde{y} * w)_{t_q+h} + (\tilde{y} * \gamma)_{t_q+h}$.
- Property 4: $\alpha(\tilde{y} * w)_{t_q+h} = \alpha(\tilde{y})_{t_q+h} * w_{t_q+h}$, for any real or complex α.

Property 1 implies that the operation is invariant to the order of the operands (commutative property). Property 2 implies that the order in which the operations are performed does not matter as long as the sequence of the operands is not changed (associative property). Property 3 implies that multiplying a density by a group of added densities yields the same outcome as multiplying each density separately and then adding them together (distributive property). Property 4 implies that associativity holds for any scalar multiplication.

The methodology is very general and allows to convolute predictive densities provided by various sources, e.g. from parametric Bayesian or frequentist models, nonparametric models, given the condition that $p(\widetilde{y}_{t_q+h}|F_{t_q+h}, K)$ are densities. Indeed, in the applications we construct predictive densities using frequentist bootstrapping, but combine them using Bayesian inference; see the next section for details.

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 Factor Model (CDFM) approach.

2.2 Algorithm and parallelization

The main steps of our algorithm 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).

We briefly describe them.

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 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 the sequential Monte Carlo algorithm of Billio et al. (2013). 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.

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).[4]

2.3 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 $e_{k,t+h}$

[4]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

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), Kascha and Ravazzolo (2010), Billio et al. (2013), and Aastveit et al. (2013). The KLIC chooses 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

$$\begin{aligned}\mathrm{KLIC}_{k,t+h} &= \int p(y_{t+h}|y_{1:t}) \ln \frac{p(y_{t+h}|y_{1:t})}{p(\tilde{y}_{k,t+h}|y_{1:t})} dy_{t+h}, \\ &= \mathbb{E}_t[\ln p(y_{t+h}|y_{1:t}) - \ln p(\tilde{y}_{k,t+h}|y_{1:t}))]. \end{aligned} \tag{8}$$

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 $p(\tilde{y}_{k,t+h}|y_{1:t})$:

$$\overline{KLIC}_k = \frac{1}{t^*} \sum_{t=\underline{t}}^{\bar{t}} [\ln p(y_{t+h}|y_{1:t}) - \ln p(\tilde{y}_{k,t+h}|y_{1:t})]. \tag{9}$$

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,

$$LS_k = -\frac{1}{t^*} \sum_{t=\underline{t}}^{\bar{t}} \ln p(\tilde{y}_{k,t+h}|y_{1:t}), \tag{10}$$

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

We further evaluate density forecasts based on the continuous rank probability score (CRPS); see, for example, Gneiting and Raftery (2007), Gneiting and Ranjan (2013), Groen et al. (2013) and Ravazzolo and Vahey (2012). The CRPS for model k measures the average absolute distance between the empirical cumulative distribution function (CDF) of y_{t+h}, which is simply a step function in y_{t+h}, and the empirical CDF that is associated with model k's predictive density:

$$\text{CRPS}_{k,t+1} = \int \left(F(z) - \mathbb{I}_{[y_{t+h},+\infty)}(z)\right)^2 \mathrm{d}z \tag{11}$$

$$= \mathbb{E}_t|\tilde{y}_{t+h,k} - y_{t+h}| - \tfrac{1}{2}\mathbb{E}_t|\tilde{y}_{t+h,k} - y'_{t+h,k}|, \tag{12}$$

where F is the CDF from the predictive density $p(\tilde{y}_{k,t+h}|y_{1:t})$ of model k and $\tilde{y}_{t+h,k}$ and $\tilde{y}'_{t+h,k}$ are independent random variables with common sampling density equal to the posterior predictive density $p(\tilde{y}_{k,t+h}|y_{1:t})$. We report the sample average CRPS:

$$CRPS_k = -\frac{1}{t^*}\sum_{t=\underline{t}}^{\bar{t}} CRPS_{k,t+h}, \tag{13}$$

Smaller CRPS values imply higher precisions and, as for the log score, we report the average $CRPS_k$ for each model k in all tables.

Finally, we assess how our range of models fares when different areas of their predictive densities are emphasized in the forecast evaluation, such as the tails. Gneiting and Ranjan (2013) propose to integrate weighted versions of Gneiting and Raftery (2007) quantile scores (QS), with weights chosen in iorder emphasize a certain area of the underlying forecast density. We will use a discrete approximation to this integration, i.e.,

$$\begin{aligned} \text{QS-T}_k &= \frac{1}{t^*}\sum_{s=\underline{t}}^{\bar{t}}\left(\frac{1}{99}\sum_{j=1}^{99}(2\alpha_j - 1)^2\text{QS}(\alpha_j, k, s+h)\right), \\ \text{QS-L}_k &= \frac{1}{t^*}\sum_{s=\underline{t}}^{\bar{t}}\left(\frac{1}{99}\sum_{j=1}^{99}(1 - \alpha_j)^2\text{QS}(\alpha_j, k, s+h)\right), \end{aligned} \tag{14}$$

where $\alpha_j = j/100$ and

$$\text{QS}(\alpha, k, t+h) = \left(\mathtt{I}\{y_{t+h} \leqq F^{-1}(\alpha, l)\} - \alpha\right)\left(F^{-1}(\alpha, l) - y_{t+h}\right)$$

with $F^{-1}(\alpha, k)$ is the quantile forecast h periods ahead using model k for level $\alpha \in (0, 1)$. Integrating QS measures over $\alpha \in (0, 1)$ results in the CRPS measure in equation (11) (see Gneiting and Ranjan (2013)). In quation (14), QS-T emphasizes the tail, and QS-L the left tail (negative GDP growth) of the predictive density relative to the realization. As for the CRPS, a lower QS implies a more accurate forecast.

3 Data

We consider in total 120 monthly leading indicators to nowcast quarterly business cycle turning points and GDP growth in the United States. Our real-time dataset is similar to the one used in Aastveit et al. (2013).[5] 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 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.[6] 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 forecasting business cycle turning points, our benchmark is the NBER reference cycle. However, when nowcasting GDP growth, the choice of a benchmark for the "actual" measure of GDP is less 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.

[5]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).

[6]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 relase)	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.

4 Simulation Exercise

In this section we implement a simulation exercise to understand what are the differences between the hard and soft data.

We simulate a series y_t assuming that the data generating process (DGP) follows a dynamic factor model with 2 factors from a balanced panel of hard data. Hence, we assume that the hard data contain all relevant information for aggregate GDP. However, in real time that data set is not balanced and soft data are often more timely available than hard data. Several studies have therefore found that soft data capture "sentiments" over future developments. In our simulation exercise, we are interested in studying if the timeliness of soft data improve the forecasts when the true DGP is generated from the hard data set. We also consider a dynamic factor model that uses both hard and soft data, in addition to our combination approach CDFM. As in the empirical exercise, we produce nowcasts (and backcasts) for the 11 blocks, but for the sake of brevity we just report MSPE results for point forecasting and LS results for density forecasting for a subset of blocks.

Table 2 reports out-of-sample results for Block 2 and Block 10. As expected, the DFM_hard is more accurate than DFM_soft. However, since DFM_all increase the LS score and reduce MSPE relative to DFM_hard, soft data contains additional information that is useful for improving the nowcasts. This is due to the soft data being more timely than the hard data. Finally, in our proposed CDFM approach the optimal learning is based on the LS (density forecasting).

Table 2. Simulation exercise

	Block 2	Block 10	Block 2	Block 10
	LS		MSPE	
DFM_all	-1.704	-0.628	**0.222**	**0.113**
DFM_hard	-1.417	-0.731	0.234	0.129
DFM_soft	-2.542	-0.904	0.296	0.269
CDFM	**-0.895**	**-0.337**	0.236	0.123

Interestingly, nowcasts from the CDFM give sizeable improvements, in terms of LS, relative to the three individual models.

5 Empirical Application

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 12 different DFMs: 4 models extract factors from the hard data; 4 models use the soft data; and 4 models use all the data. In each group, we consider 1 to 4 factors, resulting in 4 different DFM specifications for each data group. Our exercise refers to a researcher who construct nowcasts in real time accounting for various forms of uncertainty, including uncertainty related to model specification.[7] 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. CDFM: Our proposed Combined Density Factor Model approach.

Table 3 reports results for the three different strategies at the 11 different points in time (blocks) during the quarter. The table reveals three interesting results. First, the table shows that the CDFM approach provides the best statistics for almost all of the blocks. The only exceptions are CRPS in block 8; and MSPE in blocks 8, 9 and 10, where SEL provides marginally lower scores. Three of these four cases refer to backsting. Second, the gains are substantial when we evaluate density nowcasts based on log scores. Interestingly, this illustrates the gains

[7]Results for individual models are available upon request. CDFM outperforms all individual models in terms of log score for all blocks.

Table 3. Point and density forecasting

	LS	CRPS	QS-T	QS-L	MSPE
Nowcasting					
Block 1					
SEL	-1.663	0.357	0.042	0.059	0.436
BMA	-1.706	0.357	0.042	0.059	0.436
CDFM	**-1.008**	**0.340**	**0.039**	**0.055**	**0.370**
Block 2					
SEL	-1.632	0.366	0.043	0.061	0.458
BMA	-1.650	0.366	0.043	0.061	0.460
CDFM	**-0.981**	**0.339**	**0.039**	**0.055**	**0.367**
Block 3					
SEL	-1.456	0.352	0.041	0.058	0.405
BMA	-1.492	0.356	0.041	0.058	0.413
CDFM	**-0.941**	**0.333**	**0.038**	**0.053**	**0.350**
Block 4					
SEL	-1.251	0.356	0.041	0.058	0.417
BMA	-1.324	0.351	0.040	0.057	0.404
CDFM	**-0.913**	**0.327**	**0.037**	**0.052**	**0.337**
Block 5					
SEL	-1.032	0.330	0.038	0.053	0.329
BMA	-0.974	0.319	0.036	0.051	0.313
CDFM	**-0.844**	**0.310**	**0.035**	**0.048**	**0.303**
Block 6					
SEL	-0.974	0.327	0.037	0.052	0.333
BMA	-0.934	0.319	0.036	0.050	0.314
CDFM	**-0.850**	**0.312**	**0.035**	**0.049**	**0.305**
Block 7					
SEL	-1.110	0.341	0.039	0.056	0.352
BMA	-0.888	0.310	0.035	0.049	0.302
CDFM	**-0.811**	**0.303**	**0.034**	**0.047**	**0.291**
Block 8					
SEL	-0.735	**0.276**	0.031	0.044	**0.235**
BMA	-0.771	0.283	0.032	0.044	0.251
CDFM	**-0.740**	0.287	**0.031**	**0.042**	0.262
Block 9					
SEL	-0.769	0.283	0.032	0.045	**0.246**
BMA	-0.795	0.289	0.033	0.045	0.259
CDFM	**-0.741**	**0.287**	**0.032**	**0.045**	0.262
Backcasting					
Block 10					
SEL	-0.711	**0.268**	0.030	0.042	**0.225**
BMA	-0.755	0.276	0.031	0.042	0.237
CDFM	**-0.693**	0.276	**0.030**	**0.042**	0.242
Block 11					
SEL	-0.727	0.277	0.030	0.042	0.227
BMA	-0.698	0.271	0.030	0.042	0.233
CDFM	**-0.660**	**0.267**	**0.029**	**0.041**	**0.223**

The table shows average log score (LS), cumulative rank probability score (CRPS), quantile score based on tails (QS_T), quantile score based on left tail (QS_L), and mean square prediction error (MSPE) for three different prediction methods: selecting the model with highest recursive score at each point in time (SEL), standard Bayesian model averaging based on predictive likelihood (BMA), and our dynamic factor model combination (CDFM) for different blocks. Bold numbers indicates the most accurate model for different statistics. See Table 1 for information on different blocks.

of using the learning function that we impose in the CDFM. As described in section 2, the learning function is based on past log score performance. Alternatively, we could have imposed a similar learning function based on another density loss function, such as the CRPS, or a more restricted version, such as the MSPE. Finally, when focusing on the tails the CDFM performs better than the alternatives for all blocks. This result holds for both tails (QS_T) or when we just focus on the left tail (QS_L). This is an interesting result, since accurate forecasts during recessions and turbulent times are very hard to obtain.

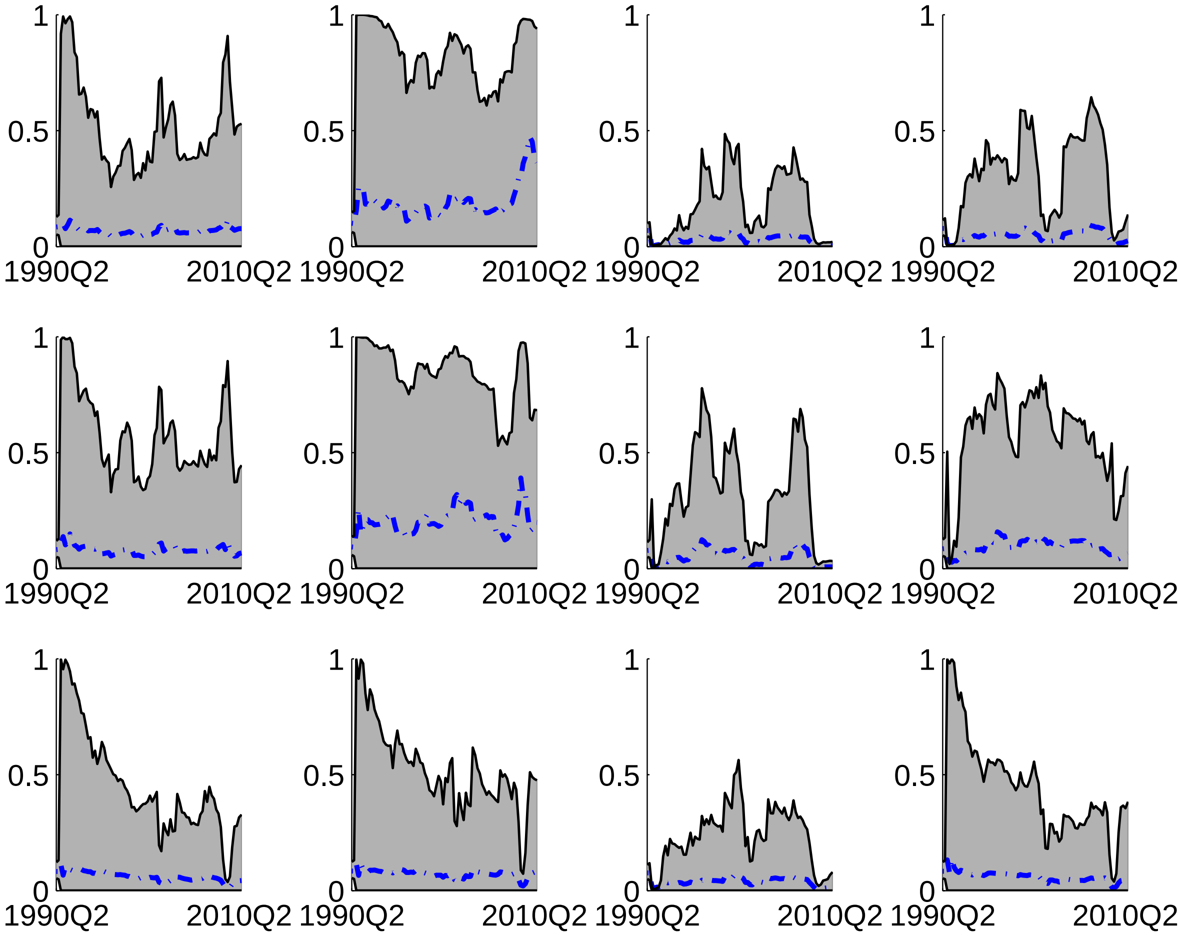
Figure 1 shows the weights associated to the 12 models for block 1. We notice the large uncertainty on the weights, with substantial variation over time. There is not a set of clear dominant models, but the DFMs with two factors have in general a higher probability, The models with three factors receives less weight. Models that use all and hard data have marginally higher weights than the models based on soft data. However, as is evident from the figure, the models that use only the soft data do not have zero weights. Thus, by using a methodology that carefully copes with the uncertainty in the weights attached to each model, we find that the division of hard and soft data provides additional gains, in terms of more accurate forecasts, than by only selecting DFMs that use either all the data or hard data. Results are similar for other blocks and horizons, see additional figures in the appendix. However, successive blocks assign more and more weight to models with only 1 and 2 factors.

5.2 Prediction of the business cycle phases

In the previous section we have shown that the CDFM provides accurate nowcasts when focusing on the entire distribution of GDP growth or on tails of it. In this section we follow Billio et al. (2012), Billio et al. (2013) and Aastveit et al. (2013) and apply an extended Bry and Boschan (1971) (BB) rule in real time to GDP growth in order to predict classical turning points of the US economy.

Following Harding and Pagan (2002), the BB procedure identifies a *potential peak* in a quarter if the value is a local maximum. Correspondingly, we can identify a *potential trough* if the value is a local minimum. Searching for maxima and minima over a window of 5 quarters seems to produce reasonable results. After potential turning points are identified, the choice of final turning points depends on several rules to ensure alternating peaks and troughs and

Figure 1. Posterior weights, block 1



The figures plot the 90% credibility intervals of the model posterior weights and their medians (dotted lines). The first row shows weights for DFM models based on all data and 1, 2, 3 and 4 factors, respectively. The second row shows weights for DFM models based on hard data and 1, 2, 3 and 4 factors, respectively. The third row shows weights for DFM models based on soft data and 1, 2, 3 and 4 factors, respectively.

minimum duration of phases and cycles. The definitions of peaks can formally be written as:

$$\wedge_t = 1\{(y_{t-2}, y_{t-1}) < y_t > (y_{t+1}, y_{t+2})\} \tag{15}$$

and correspondingly for troughs as:

$$\vee_t = 1\{(y_{t-2}, y_{t-1}) > y_t < (y_{t+1}, y_{t+2})\} \tag{16}$$

When forecasting peaks and troughs, the values on the right-hand side of the equations are replaced by forecasts. Formally:

$$\wedge_t = 1\{(y_{t-2}, y_{t-1}) < y_t, Prob(y_{t+1}, y_{t+2}) < y_t) > 0.5\} \tag{17}$$

and

$$\vee_t = 1\{(y_{t-2}, y_{t-1}) > y_t, Prob(y_{t+1}, y_{t+2}) > y_t) > 0.5\} \tag{18}$$

The business cycle can be interpreted as a state S_t, which takes the value 1 in expansions and 0 in recessions. Turning points occur when the state changes. The relationship between the business cycle and the local peaks and troughs can be written as

$$S_t = S_{t-1}(1 - \wedge_{t-1}) + (1 - S_{t-1})\vee_{t-1}) \tag{19}$$

If the economy is in an expansion, $S_{t-1} = 1$. If no peak occurred in (t-1), then $\wedge_{t-1} = 0$ and it follows that the state $S_t = 1$. On the other hand, if there is a peak in (t-1) then $\wedge_{t-1} = 1$ and the state changes to $S_t = 0$. The state will remain at 0 until a trough is detected.

Defining a peak or a trough at time t, where t is the last value in the sample, requires a probabilistic assessment of future values for y_{t+1} and y_{t+2} as the formulas describe. We use density forecasts from the models in the previous section and report results using the median as summary statistics of the predictive densities. Furthermore, we compare our ex-ante definition of the business cycle to ex-post NBER business cycle dating and compute the concordance statistics (CS) to provide insights about which method is more accurate. The concordance statistics counts the proportion of time during which the predicted and the reference turning point series are in the same state and it is defined for model k at time t as:

$$CS_k = \sum_{s=\underline{t}}^{\bar{t}} ((S_{k,s}S_{R,s}) - (1 - S_{k,s})(1 - S_{R,s})) \tag{20}$$

where $S_{k,s}$ is defined in (19) for model k at time s and $S_{R,s}$ refers to the cycle provided by the ex-post NBER reference dating.

Table 4 compares the concordance statistics for the CDFM, SEL and BMA approach 9 different points in time during the quarter.[8] For all the 9 blocks the CDF produce the highest concordance statistics. However, as the table shows, it is a very challenging task to predict recessions in real time, see also the detailed discussion in Hamilton (2011).

The full distribution of the CDFM can also be used to compute probabilities to be in specific phases of the business cycle, and not just an indicator as in previous paragraphs. Considering that we document that CDFM performs well in the tails, see table 3, we focus on two "tail" probabilities: the probability of negative growth (nowcasts below 0) and the probability of quarterly growth higher than 1% (or annual growth higher than 4%). Figures 2 and 3 shows that the probabilities are timely and accurate, in particular for negative growth. New information in the quarter increases predictability accuracy, resulting in more precise probabilities of high or low growth in the later blocks.

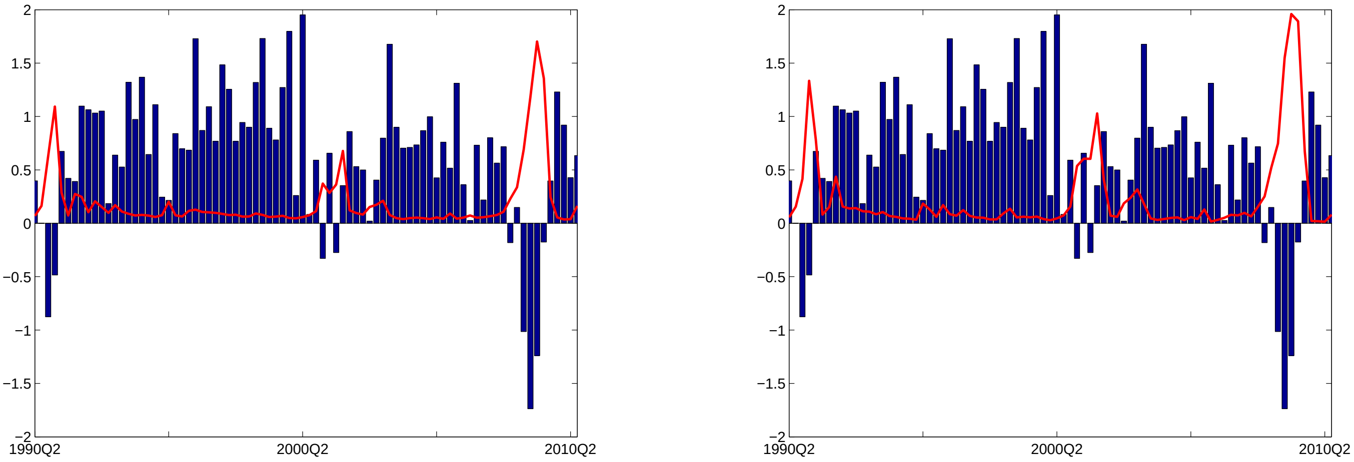
Table 4. Concordance statistics for turning points

Blocks	SEL	BMA	CDFM
1	0.829	0.829	**0.841**
2	0.829	0.829	**0.854**
3	0.829	0.829	**0.841**
4	0.829	0.829	**0.841**
5	**0.841**	**0.841**	**0.841**
6	**0.841**	**0.841**	**0.841**
7	**0.841**	**0.841**	**0.841**
8	0.829	0.829	**0.854**
9	**0.854**	0.841	**0.854**

The table shows the concordance statistics (CS) for three different prediction methods: selecting the model with highest recursive score at each point in time (SEL), standard Bayesian model averaging based on predictive likelihood (BMA), and our dynamic factor model combination (CDFM) for different blocks. Bold numbers indicates the most accurate model. See Table 1 for information on different blocks.

[8]We only provide results for the nowcasts, and not the backcasts.

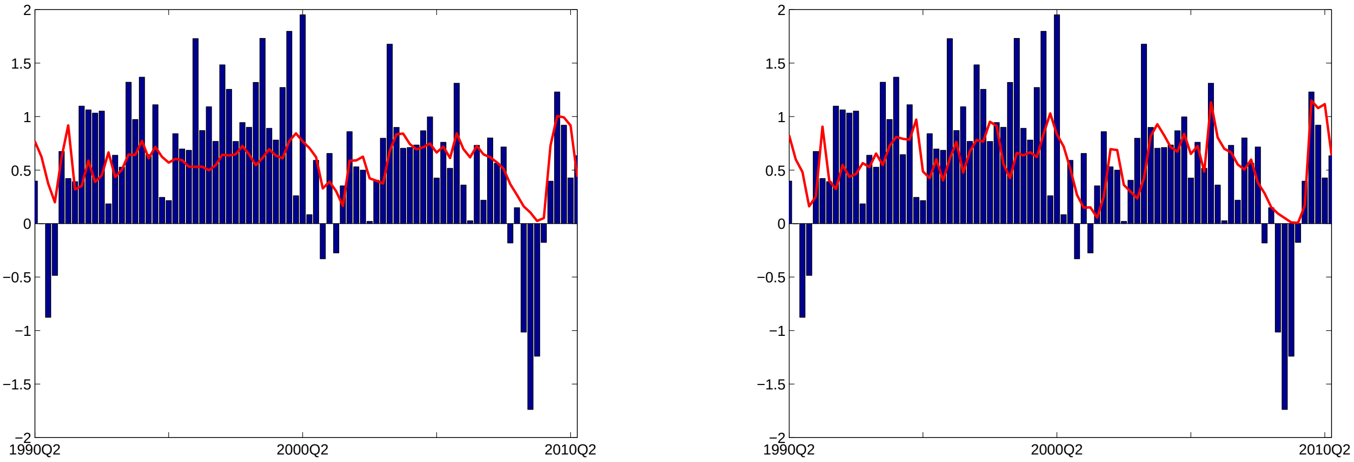
Figure 2. Probabilities of negative growth



Block 1 Block 9

The figures show the probabilities over time of negative quarterly growth given by the CDFM for the first and ninth blocks. The red line plots the probabilities scaled by 2 (therefore covering the interval [0,2]); the bars plot the realization.

Figure 3. Probabilities of negative growth



Block 1 Block 9

The figures show the probabilities over time of quarterly growth higher than 1% (implying annual growth higher than 4%) given by the CDFM for the first and ninth blocks. The red line plots the probabilities scaled by 2 (therefore covering the interval [0,2]); the bars plot the realization.

6 Conclusion

In this paper, we have introduced a Combined Density Factor Model (CDFM) approach that accounts for time varying factor model uncertainty in order to provide more accurate nowcasts of predictive densities. By combining predictive densities from a set of dynamic factor models, using combination weights that are time-varying and depend on past predictive forecasting performance and other learning mechanisms and by making use of a Bayesian Sequential Monte Carlo method, we are able to weight 'soft' and 'hard' data uncertainty, parameter uncertainty, model uncertainty and uncertainty in the combination of weights in a coherent way.

We first implement a simulation experiment to compare soft and hard data and analyze the performance of the CDFM. The results illustrate that soft data contain useful information due to being timely available and increase both point and density nowcast performance even when the true data are generated from the hard database. Furthermore, we show that the CDFM with optimal learning based on density nowcasting provides better density nowcasts than any of the individual models.

We then show the usefulness of the CDFM for nowcasting GDP growth and business cycle turning points 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. The CDFM we use includes 12 different dynamic factor models: 4 models are based on hard data; 4 models are based on soft data; and 4 models are based on all data. In each group, we consider 1 to 4 factors, resulting in 4 specifications. Our approach to combine density nowcasts from 12 factor models leads to the CDFM where the combined density is a convolution of the set of individual model densities. By making use of recent increases in computing power and recent advances in parallel programming technique, it is feasible to use a sequential Monte Carlo method to approximate the filtering and predicative densities and apply non-linear time-varying weights to the 12 factor models at different points in time during the quarter.

We find that the CDFM outperforms all individual models in terms of log score (LS) and cumulative rank probability score (CRPS) for all blocks and results. Interestingly, the favorable nowcasting properties from the CDFM also applies when focusing on the tails of the predictive distribution and it delivers timely and accurate probabilities of high growth and stagnation. Moreover, the CDFM also outperform the strategy of selecting the models with the highest realized cumulative log score as well as Bayesian model averaging based on predictive likelihood. Finally, we show that a real-time indicator based on the Bry and Boschan (1971) (BB) rule and nowcasts from our model are more accurate in terms of concordance statistics than those given by the alternative methods.

References

Aastveit, K., K. Gerdrup, A. Jore, and L. Thorsrud (2013). Nowcasting gdp in real-time: A density combination approach. *Journal of Business Economics and Statistics forthcoming*.

Aastveit, K. A., A. Jore, and F. Ravazzolo (2013). Forecasting recession in real time. Working paper, Norges 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 (2010). Maximum likelihood estimation of factor models on data sets with arbitrary pattern of missing data. Working Paper Series 1189, European Central Bank.

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. van Dijk (2012). Combination schemes for turning point predictions. *Quarterly Review of Economics and Finance 52*(4), 402–412.

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.

Bry, G. and C. Boschan (1971). Cyclical Analysis of Time Series: Selected Procedures and Computer Programs. NBER Technical Paper 20.

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.

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.

Damelin, S. and W. Miller (2011). *The Mathematics of Signal Processing.* Cambridge: Cambridge University Press.

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

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. and A. E. Raftery (2007). Strictly proper scoring rules, prediction, and estimation. *Journal of the American Statistical Association 102*, 359–378.

Gneiting, T. and R. Ranjan (2013). Combining predicitve distributions. *Electronical Journal of Statistics 7*, 1747–1782.

Groen, J. J. J., R. Paap, and F. Ravazzolo (2013). Real-time inflation forecasting in a changing world. *Journal of Business & Economic Statistics 31*(1), 29–44.

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.

Ravazzolo, F. and S. P. Vahey (2012). Forecast densities for economic aggregates from disaggregate ensembles. Working Paper 2010/02, Norges Bank.

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

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.

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.

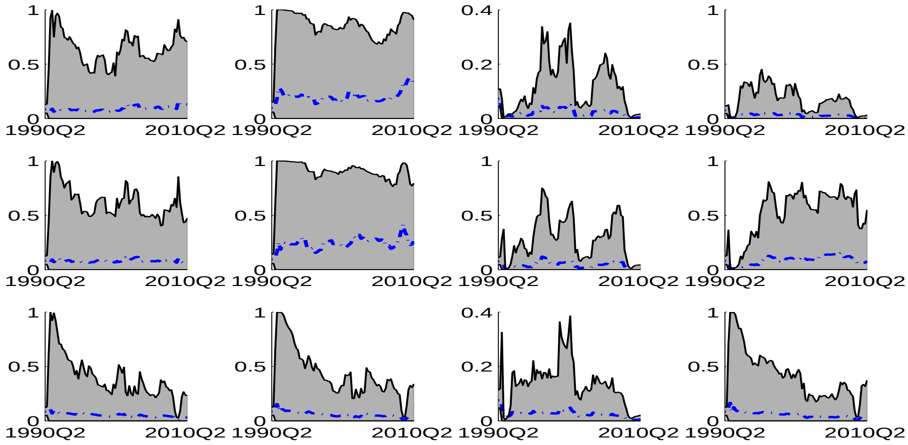
Timmermann, A. (2006). Forecast combinations. In G. Elliott, C. W. J. Granger, and A. Timmermann (Eds.), *Handbook of Economic Forecasting*, Volume 1, pp. 136–96. Amsterdam: Elsevier.

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

Appendix A

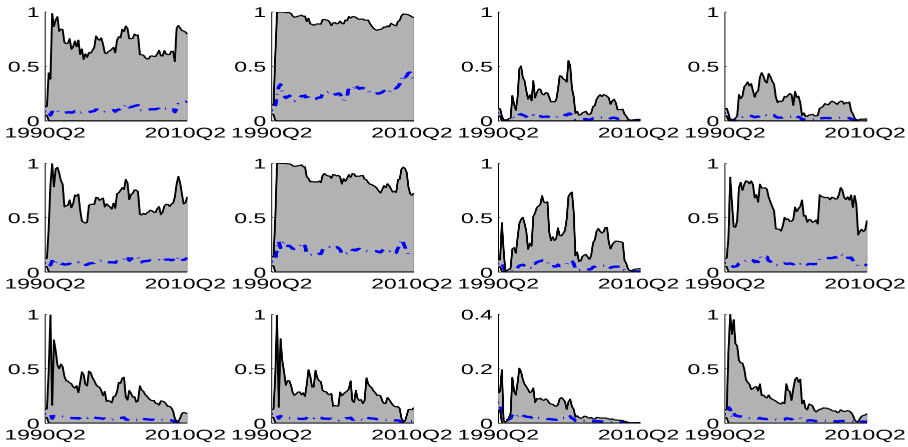
A1. Posterior combination weights for other blocks and horizons

Figure 4. Posterior weights, block 3



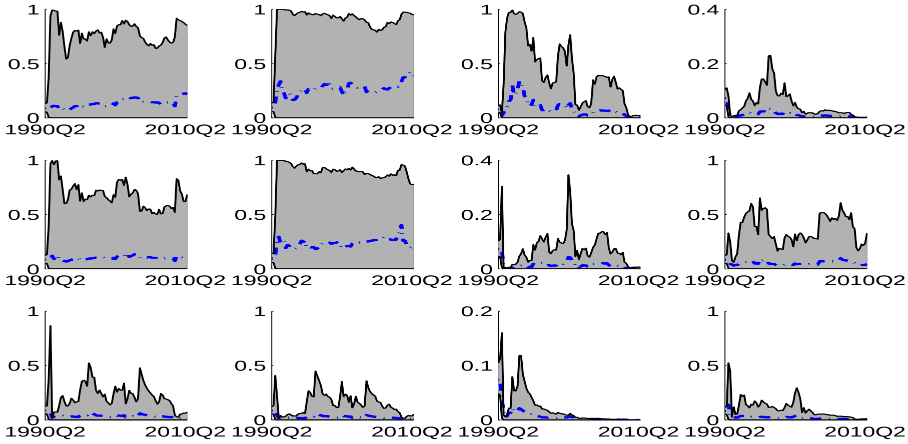
See note in figure 1.

Figure 5. Posterior weights, block 6



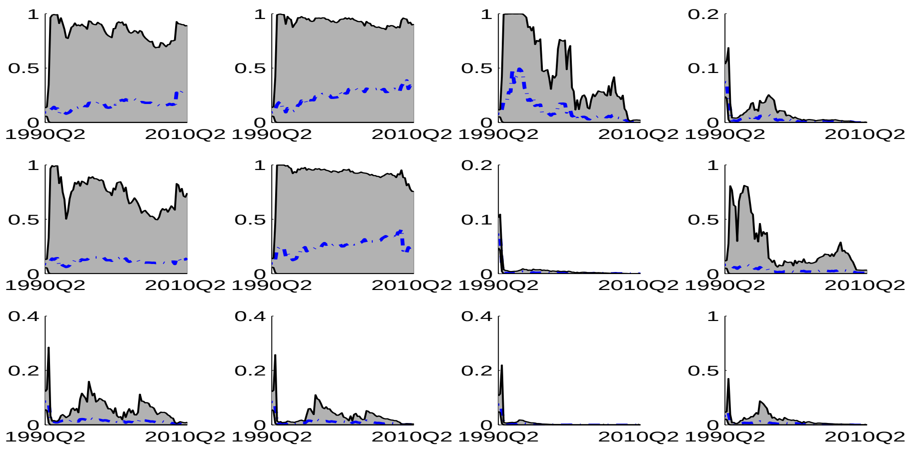
See note in figure 1.

Figure 6. Posterior weights, block 9



See note in figure 1.

Figure 7. Posterior weights, block 11



See note in figure 1.

7 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.