EuroMInd-$\mathcal{D}$: A Density Estimate of Monthly Gross Domestic Product for the Euro Area

Tommaso Proietti
Università di Roma "Tor Vergata", Italy, and CREATES, Denmark

Martyna Marczak
University of Hohenheim, Germany

Gianluigi Mazzi
Eurostat, Luxembourg

Draft version September 22, 2014

Abstract

EuroMInd-$\mathcal{D}$ is a density estimate of monthly gross domestic product (GDP) constructed according to a bottom–up approach, pooling the density estimates of eleven GDP components, by output and expenditure type. The components density estimates are obtained from a medium-size dynamic factor model of a set of coincident time series handling mixed frequencies of observation and ragged–edged data structures. They reflect both parameter and filtering uncertainty and are obtained by implementing a bootstrap algorithm, for simulating from the distribution of the maximum likelihood estimators of the model parameters, and conditional simulation filters, for simulating from the predictive distribution of GDP. Both algorithms process sequentially the data as they become available in real time. The GDP density estimates for the output and expenditure approach are combined using alternative weighting schemes and evaluated with different tests based on the probability integral transform and by applying scoring rules.

Keywords: Density Forecast Combination and Evaluation; Mixed–Frequency Data; Dynamic Factor Models; State Space Models.

JEL Classification: C32, C52, C53, E37

1 Introduction

The recent developments in the analysis of economic time series and in data production and dissemination have made available a number of high frequency, timely and representative indicators of the state of the economy. As far as the euro area is concerned, we mention New Eurocoin (NE), the Economic Sentiment Index (ESI), Euro–Sting and EuroMInd. New Eurocoin is a monthly coincident indicator of economic growth for the euro area (Altissimo et al., 2010), published monthly by CEPR (www.cepr.org) and the Bank of Italy. The European Commission (Directorate-General for Economic and Financial Affairs) compiles the Economic Sentiment Indicator (ESI), a composite coincident indicator for the timely assessment of socio–economic situation in the euro area. Euro–Sting (Camacho and Perez–Quiros, 2010) is a monthly indicator of the euro area Gross Domestic Product (GDP), based on a parametric dynamic factor model with mixed frequency data, constructed as an extension of the model described in Mariano and Murasawa (2003).

Another recent development is a steadily growing research on the probabilistic forecasts of macroeconomic time series; see Tay and Wallis (2000) for an early survey. While inflation has been the traditional focus, density forecasting of GDP is also prominent, see Aastveit et al. (2014) and Mazzi et al. (2014) for recent references, with a lot of attention being paid to density forecast combination.

This article tries to join these strands of literature; its primary objective is to introduce EuroMInd-$\mathcal{D}$, a density nowcast and forecast for the euro area GDP and its main components. Our methodology is based on EuroMInd, a monthly indicator of the euro area GDP that is constructed according to the bottom-up approach outlined in Frale et al. (2011). The focus is on the breakdown of GDP at market prices by output and expenditure type into 11 components. From the output side GDP is decomposed as follows:

Label	*Value added of branch*	
A–B	Agriculture, hunting, forestry and fishing	+
C–D–E	Industry, incl. Energy	+
F	Construction	+
G–H–I	Trade, transport and communication services	+
J–K	Financial services and business activities	+
L–P	Other services	=
	Total Gross Value Added	+
TlS	Taxes less subsidies on products	=
	GDP at market prices	

The breakdown of total GDP from the expenditure side is the following:

Label	*Component*	
FCE	Final consumption expenditure	+
GCF	Gross capital formation	+
EXP	Exports of goods and services	-
IMP	Imports of goods and services	=
	GDP at market prices	

The time series for the GDP components are quarterly and are available from the National Quarterly Accounts compiled by Eurostat. A set of coincident monthly indicators is also available for each component: for instance, in the case of the industry sector (C–D–E) we can consider the monthly index of industrial production and hours worked. For each individual GDP component, the monthly indicators and the quarterly GDP estimates are jointly modelled according to a single index dynamic factor model formulated at the monthly frequency and customised to handle ragged-edge data structures and temporal aggregation. The model parameters are estimated by maximum likelihood and signal extraction is performed by a suitable implementation of the Kalman filter and smoother handling sequential processing of the data as they become available according to their production schedule.

Monthly density estimates (nowcasts, forecasts and backcasts) are obtained for the GDP components by implementing a bootstrapping procedure that resamples from the distribution of the maximum likelihood estimator of the parameters, and a conditional simulation filter that draws from the distribution of the monthly indicators of each individual component, given the observed time series. The density estimates reflect parameter estimation uncertainty and filtering uncertainty, due to the fact that monthly GDP is unobserved and has to be distilled from a set of monthly indicators and the actual measurements available at the quarterly frequency of observation.

In the next stage, the components' density estimates are pooled into two aggregate GDP density estimates, EuroMInd-$\mathcal{D}_o$ and EuroMInd-$\mathcal{D}_e$, respectively from the output and the expenditure side. For that purpose, we apply a pooling procedure that is consistent with the national accounts standard, implementing the so–called *annual overlap technique*; see Bloem et al. (2001). The procedure takes into account that components are expressed in chained volumes but are additive only when they are expressed at the average prices of the previous years. The two can be combined into a single EuroMInd-$\mathcal{D}$ estimate with pooling weights reflecting estimation accuracy.

Combined densities are also considered in the pseudo–realtime exercise, aiming at obtaining three predictive densities for quarterly GDP with a lead of respectively 3, 2, and 1 months with respect to the official release by Eurostat. For the combination, we employ the linear pool based and use varying weights based on two scoring rules, the logarithmic score and the continuous ranked probability score (CRPS) (Gneiting and Raftery, 2007). Calibration of the density forecasts is validated by means of a comparison with the published quarterly national accounts estimates. More specifically, it is evaluated using the so–called probability integral transform (PIT) (see Diebold et al., 1998), and the transformed PIT. The accuracy of density forecasts is compared by using the aforementioned scoring rules.

This article contributes to the literature on the probabilistic forecast of GDP for taking a bottom–up approach leading to two different density estimates that are subsequently combined, along with offering density predictions and nowcasts for the 11 GDP components. Second, the pooling scheme scheme that we propose is consistent with the method advocated by the IMF for the construction of aggregate GDP measures, see Bloem et al. (2001) (known as the *annual overlap technique*). The resulting density estimates are consistent in aggregation with the totals published by Eurostat, in that the draws from the empirical distribution function for the three months making up the quarter sum up to the published quarterly totals. Third, from the methodological point of view, we implement algorithms for bootstrapping and conditional simulation based on the sequential processing of the information available in real time. Fourth, the use of a mixed frequency model featuring the quarterly national accounts estimates enables us to assess the role of the quarterly releases on the calibration of the densities. One of the findings is that the density predictions of GDP and its components referring to a particular quarter become calibrated only after the quarterly national accounts concerning the previous quarter are released by Eurostat. The predictive ability improves with the availability of the monthly indicators in the course of the quarter, but still the information content of the GDP release is very important.

The plan of the paper is the following. In Section 2 we expose the methodology for the construction of EuroMInD-$\mathcal{D}$. We start from the specification of the single index dynamic factor model for the GDP components, and we deal with its state space representation, and how the latter is modified to handle handle mixed frequency time series arising from temporal aggregation. Section 3 deals with the main state space methods for point estimation and likelihood evaluation. In particular, it provide details on filtering the observations in a sequential way as they become available in realtime, according to their production schedule, and on the essential algorithms that will constitute the key ingredients for drawing samples from the required conditional densities. The methodology for obtaining density estimates (historical, forecasts and nowcasts) of the GDP components is exposed in Section 4, which provides details on the bootstrap and the conditional simulation sampler. The draws from the conditional distribution of the GDP components (given varying information sets) are combined according to the pooling procedure outlined in Section 4.3. Section 5 exposes the methods used for the evaluation of the two density forecasts resulting from the output and the expenditure approaches, and how the two are combined into the EuroMInd-$\mathcal{D}$ estimate. Section 6.3.2 is devoted to the empirical results: we describe the monthly GDP historical density estimates using the complete data and we discuss the calibration and optimal combination of the density predictions and nowcast of total GDP, using a pseudo real–time assessment exercise. We offer some conclusions in Section 7.

2 Model Specification

For each of the 11 GDP components, the model is specified at the monthly observations frequency and assuming a set of complete observations. The resulting state space model is then modified so as to accommodate the temporal aggregation of the GDP components. This section illustrates the specification of the dynamic factor model for a particular component. Beyond the technical aspects, it aims at presenting the basic structure of the dynamic factor model. Then the state space representation of the model is derived (Section 2.2), and modified to take into account the observational constraints arising from the temporal aggregation of the GDP components to the quarterly frequency of observation (Section 2.3).

2.1 The single index dynamic factor model

Let $\mathbf{y}_t = [y_{1t}, \ldots, y_{it}, \ldots, y_{Nt}]'$, $t = 1, \ldots n$, denote an $N \times 1$ vector of time series, that we assume to be integrated of order one, or $I(1)$, so that $\Delta y_{it}, i = 1, \ldots, N, t = 2, \ldots, n$, has a stationary and invertible representation. The elements of the vector $\mathbf{y}_t$ include a set of monthly coincident indicators and the relevant GDP component. For instance, for the component of GDP referring to sector C–D–E (Industry, including Energy), $N = 6$; the first four series, $i = 1, 2, 3, 4$ are the monthly industrial production indices for the four euro area largest economies (Germany, France, Italy and Spain), the fifth series is the monthly series of hours worked is the industrial sector, and the last series is the national accounts' value added of the sector CDE (that will be subject to temporal aggregation). For a complete account of the indicator series used for the other components we refer to Frale et al. (2011).

The first differences of $\mathbf{y}_t$ can be expressed as a lagged linear combination of $K < N$ stationary common factors $\boldsymbol{\chi}_t$, plus a stationary random vector with components that are cross–sectionally independent:

$$\Delta \mathbf{y}_t = \mathbf{m} + \boldsymbol{\theta}(L)\boldsymbol{\chi}_t + \boldsymbol{\chi}_t^*.$$

Here $\mathbf{m}$ is an $N \times 1$ vector of drifts, such that $\mathrm{E}(\Delta \mathbf{y}_t) = \mathbf{m}$. Letting $\boldsymbol{\theta}_j, j = 0, \ldots, J$, denote $N \times K$ matrices of factor loadings, we write $\boldsymbol{\theta}(L) = \boldsymbol{\theta}_0 + \boldsymbol{\theta}_1 L + \cdots + \boldsymbol{\theta}_J L^J$.

In all the components models for EuroMInd there is a unique common factor, with a stationary AR(p) representation and $J \leq 1$. The idiosyncratic components are also modelled as independent stationary AR(p_i). The orders $p, p_i, i = 1, \ldots, N$, are typically small. In particular, we set

$$\phi(L)\chi_t = \eta_t, \quad \eta_t \sim \mathrm{NID}(0, \sigma^2_\eta),$$

where $\phi(L)$ is an autoregressive polynomial of order p with stationary roots:

$$\phi(L) = 1 - \phi_1 L - \cdots - \phi_p L^p,$$

and

$$\mathbf{D}(L)\boldsymbol{\chi}_t^* = \boldsymbol{\eta}_t^*, \quad \boldsymbol{\eta}_t^* \sim \mathrm{NID}(\mathbf{0}, \boldsymbol{\Sigma}_{\eta^*}),$$

where the matrix polynomial $\mathbf{D}(L)$ is diagonal:

$$\mathbf{D}(L) = \mathrm{diag}\left[d_1(L), d_2(L), \ldots, d_N(L)\right],$$

with $d_i(L) = 1 - d_{i1}L - \cdots - d_{ip_i}L^{p_i}$ and $\boldsymbol{\Sigma}_{\eta^*} = \mathrm{diag}(\sigma^2_1, \ldots, \sigma^2_N)$. The disturbances η_t and $\boldsymbol{\eta}_t^*$ are mutually uncorrelated at all leads and lags.

For filtering and signal extraction under temporal aggregation, it is preferable to set up the model in terms of the level of the variables. For this purpose, let us define the single index $\mu_t = \mu_{t-1} + \chi_t, \mu_0 = 0$, and the idiosyncratic component $\boldsymbol{\mu}_t^*$, such that $\boldsymbol{\mu}_t^* = \mathbf{m} + \boldsymbol{\mu}_{t-1}^* + \boldsymbol{\chi}_t^*$, with $\boldsymbol{\mu}_0^* = \mathbf{0}$ representing an $N \times 1$ vector of initial values such that $\boldsymbol{\mu}_1^* - \boldsymbol{\mu}_0^* = \mathbf{m} + \boldsymbol{\chi}_1^*$. Hence, when referred to the levels, the above factor model decomposes $y_{it}, i = 1, \ldots, N$, into a common nonstationary component with ARIMA(p, 1, 0), representation and an idiosyncratic component with ARIMA(p_i, 1,0) representation. This specification has been adopted by Stock and Watson (1991) for extracting an index of coincident indicators for the US economy, as the common factor in a four-variate time series consisting of industrial production, disposable income, retail sales and employment. The common cyclical trend is often termed the *single index*.

The model can be extended to account for the presence of regression effects, common to the N time series equations, leading to the following specification (which assumes $J = 1$):

$$\begin{array}{lcll} \mathbf{y}_t & = & \boldsymbol{\theta}_0\mu_t + \boldsymbol{\theta}_1\mu_{t-1} + \boldsymbol{\mu}_t^* + \mathbf{B}\mathbf{x}_t, t = 1, ..., n, & \\ \phi(L)\Delta\mu_t & = & \eta_t, & \eta_t \sim \mathrm{NID}(0, \sigma_\eta^2), \\ \mathbf{D}(L)\Delta\boldsymbol{\mu}_t^* & = & \boldsymbol{\delta} + \boldsymbol{\eta}_t^*, & \boldsymbol{\eta}_t^* \sim \mathrm{NID}(\mathbf{0}, \boldsymbol{\Sigma}_{\eta^*}), \end{array} \tag{1}$$

where $\mathbf{x}_t$ is a $k \times 1$ vector of explanatory variables that are used to incorporate calendar effects (trading day regressors, Easter, length of the month) and intervention variables (level shifts, additive outliers, etc.), and $\mathbf{B}$ is an $N \times k$ matrix of coefficients. The drift term is related to the mean vector $\mathbf{m}$ by $\boldsymbol{\delta} = \mathbf{D}(1)^{-1}\mathbf{m}$. The model assumes a zero drift for the single index. We further assume that $\sigma_\eta^2 = 1$ as an identification restriction.

2.2 State space representation

Model (1) can be represented in state space form (SSF). We start from the first order Markovian representation of the the single index, $\phi(L)\Delta\mu_t = \eta_t$, considering the SSF of the stationary AR(p) model for $\Delta\mu_t$, which can be written as:

$$\Delta\mu_t = \mathbf{e}_{1,p}'\mathbf{g}_t, \quad \mathbf{g}_t = \mathbf{T}_{\Delta\mu}\mathbf{g}_{t-1} + \mathbf{e}_{1p}\eta_t,$$

where $\mathbf{e}_{1,p} = [1, 0, \ldots, 0]'$ and

$$\mathbf{T}_{\Delta\mu} = \left[\begin{array}{cc} \phi_1 & \\ \vdots & \mathbf{I}_{p-1} \\ \phi_{p-1} & \\ \phi_p & \mathbf{0}_{1,p-1} \end{array}\right].$$

Hence, $\mu_t = \mu_{t-1} + \mathbf{e}_{1p}'\mathbf{g}_t = \mu_{t-1} + \mathbf{e}_{1p}'\mathbf{T}_{\Delta\mu}\mathbf{g}_{t-1} + \eta_t$, and defining

$$\boldsymbol{\alpha}_{\mu,t} = \left[\begin{array}{c} \mu_t \\ \mathbf{g}_t \end{array}\right], \quad \mathbf{T}_\mu = \left[\begin{array}{cc} 1 & \mathbf{e}_{1p}'\mathbf{T}_{\Delta\mu} \\ 0 & \mathbf{T}_{\Delta\mu} \end{array}\right],$$

the Markovian representation of the model for μ_t becomes $\mu_t = \mathbf{e}_{1,p+1}'\boldsymbol{\alpha}_{\mu,t}$, $\boldsymbol{\alpha}_{\mu,t} = \mathbf{T}_\mu\boldsymbol{\alpha}_{\mu,t-1} + \mathbf{H}_\mu\eta_t$, where $\mathbf{H}_\mu = [1, \mathbf{e}_{1,p}']'$.

A similar representation holds for each individual μ_{it}^*, with ϕ_j replaced by d_{ij}, so that, if we let p_i denote the order of the i-th lag polynomial $d_i(L)$, we can write:

$$\mu_{it}^* = \mathbf{e}_{1,p_i+1}'\boldsymbol{\alpha}_{\mu_i,t}, \quad \boldsymbol{\alpha}_{\mu_i,t} = \mathbf{T}_i\boldsymbol{\alpha}_{\mu_i,t-1} + \mathbf{c}_i + \mathbf{H}_i\eta_{it}^*,$$

where $\mathbf{H}_i = [1, \mathbf{e}_{1,p_i}']'$, $\mathbf{c}_i = \delta_i\mathbf{H}_i$ and δ_i is the drift of the $i - th$ idiosyncratic component, and thus of the series, since we have assumed a zero drift for the common factor.

Combining the blocks, we obtain the SSF of the complete model (1)

$$\mathbf{y}_t = \mathbf{Z}\boldsymbol{\alpha}_t + \mathbf{X}_t\boldsymbol{\beta}, \quad \boldsymbol{\alpha}_t = \mathbf{T}\boldsymbol{\alpha}_{t-1} + \mathbf{W}\boldsymbol{\beta} + \mathbf{H}\boldsymbol{\epsilon}_t, \tag{2}$$

where the state vector $\boldsymbol{\alpha}_t = [\boldsymbol{\alpha}'_{\mu,t}, \boldsymbol{\alpha}'_{\mu_1,t}, \ldots, \boldsymbol{\alpha}'_{\mu_N,t}]'$, has dimension $m = \sum_i (p_i + 1) + p + 1$, $\boldsymbol{\epsilon}_t = [\eta_t, \eta^*_{1t}, \ldots, \eta^*_{Nt}]'$ and the system matrices are given below:

$$\begin{aligned} &\mathbf{Z} = \left[\boldsymbol{\theta}_0 \vdots \boldsymbol{\theta}_1 \vdots \mathbf{0}_{N,p-1} \vdots \text{diag}(\mathbf{e}'_{1,p_1}, \ldots, \mathbf{e}'_{1,p_N})\right], \quad \mathbf{T} = \text{diag}(\mathbf{T}_\mu, \mathbf{T}_1, \ldots, \mathbf{T}_N), \\ &\mathbf{H} = \text{diag}(\mathbf{H}_\mu, \mathbf{H}_1, \ldots, \mathbf{H}_N). \end{aligned} \tag{3}$$

The vector of initial values is written as

$$\boldsymbol{\alpha}_1 = \mathbf{W}_1\boldsymbol{\beta} + \mathbf{H}\boldsymbol{\epsilon}_1,$$

so that $\boldsymbol{\alpha}_1 \sim \text{N}(\mathbf{0}, \mathbf{W}_1\mathbf{V}\mathbf{W}'_1 + \mathbf{H}\text{Var}(\boldsymbol{\epsilon}_1)\mathbf{H}')$, $\text{Var}(\boldsymbol{\epsilon}_1) = \text{diag}(1, \sigma^2_1, \ldots, \sigma^2_N)$.

The vector $\boldsymbol{\beta}$ has $2N + K$ elements; the first $2N$ elements are the pairs $\{(\mu_{i0}, \delta_i), i = 1, \ldots, N\}$, i.e. the starting values at time $t = 0$ of the idiosyncratic components and the constant drifts δ_i. Recall that the common component, μ_t, has a fixed and known initial value, $\mu_0 = 0$, and that the variance of its disturbance, σ^2_η, is assumed to be equal to 1, so that $\mu_1 = \eta_t \sim \text{N}(0, 1)$. The last K elements are the nonzero elements of $\text{vec}(\mathbf{B}')$, where $\mathbf{B}$ is matrix of regressor effects in model (1).

The regression matrix is $\mathbf{X}_t = [\mathbf{0}_{N,2N}, \ \ \mathbf{X}^*_t]$. The zero block has dimension $N \times 2N$ and corresponds to the elements of $\boldsymbol{\beta}$ that are used for the initialisation and other fixed effects. $\mathbf{X}^*_t$ is a $N \times K$ matrix containing the values of the exogenous variables. The elements of $\boldsymbol{\beta}$ are taken as diffuse, i.e. it is assumed that $\boldsymbol{\beta} \sim \text{N}(\mathbf{0}, \mathbf{V})$ where $\mathbf{V}^{-1}$ converges to a zero matrix; see de Jong (1991).

For $t = 2, \ldots, n$, the matrix $\mathbf{W}_t$ is time invariant and selects the drift δ_i for the appropriate state element:

$$\mathbf{W} = \begin{bmatrix} \mathbf{0}_{p+1,2N} & \mathbf{0}_{p+1,K} \\ \text{diag}(\mathbf{C}_1, \ldots, \mathbf{C}_N) & \mathbf{0}_{\sum_i(p_i+1),K} \end{bmatrix}, \mathbf{C}_i = [\mathbf{0}_{p_i+1,1} \vdots \mathbf{H}_i],$$

and for $t = 1$,

$$\mathbf{W}_1 = \begin{bmatrix} \mathbf{0}_{p+1,2N} & \mathbf{0}_{p+1,K} \\ \text{diag}(\mathbf{C}^*_1, \ldots, \mathbf{C}^*_N) & \mathbf{0}_{\sum_i(p_i+1),K} \end{bmatrix}, \mathbf{C}^*_i = [\mathbf{e}_{1,p_i+1} \vdots \mathbf{H}_i].$$

2.3 Temporal aggregation

Suppose that $\mathbf{y}_t$ comprises elements of different type: stock variables, time–averaged stock variables and flow variables. Further suppose that not all flow variables are observed monthly – for some of them quarterly aggregates are available only. To take this into account, vector $\mathbf{y}_t$ can be partitioned into blocks such that $\mathbf{y}_t = [\mathbf{y}'_{1t}, \mathbf{y}'_{2t}]'$. The first block gathers all elements observed at monthly frequency. The second block consists of flow variables for which observations at every two months between adjacent quarters are missing, and at all months corresponding to the respective quarter end we observe

$$\mathbf{y}_{2,3\tau} + \mathbf{y}_{2,3\tau-1} + \mathbf{y}_{2,3\tau-2}, \qquad \tau = 1, 2, \ldots, [n/3], \tag{4}$$

with τ denoting quarters and $[\cdot]$ being integer division.

Our treatment of temporal aggregation draws on Harvey (1989), who introduced the so–called cumulator variable, $\mathbf{y}_{2t}^c$, constructed as follows:

$$\mathbf{y}_{2,t}^c = \rho_t \mathbf{y}_{2,t-1}^c + \mathbf{y}_{2t}, \tag{5}$$

where

$$\rho_t = \begin{cases} 0, & \text{if } t = 3(\tau - 1) + 1, \\ 1, & \text{otherwise.} \end{cases}$$

Replacing for $\mathbf{y}_{2t} = \mathbf{Z}_2\boldsymbol{\alpha}_t + \mathbf{X}_{2t}\boldsymbol{\beta}$ into (5), where the matrices $\mathbf{Z}_2$ and $\mathbf{X}_{2t}$ correspond to the second block of variables, $\mathbf{y}_{2t}$, and thus have dimensions $N_2 \times m$ and $N_2 \times K$, respectively, and using (2), gives

$$\mathbf{y}_{2,t}^c = \rho_t \mathbf{y}_{2,t-1}^c + \mathbf{Z}_2\mathbf{T}\boldsymbol{\alpha}_{t-1} + (\mathbf{X}_{2t} + \mathbf{Z}_2\mathbf{W})\boldsymbol{\beta} + \mathbf{Z}_2\mathbf{H}\boldsymbol{\epsilon}_t.$$

Notice that at times $t = 3\tau$, the cumulator variables coincide with the (observed) aggregated series, otherwise they contain the partial cumulative value of the aggregates in the particular quarter.

The vector $\mathbf{y}_{2t}^c$ is used to create new augmented state and observation vectors, $\boldsymbol{\alpha}_t^*$ and $\mathbf{y}_t^\dagger$, respectively:

$$\boldsymbol{\alpha}_t^* = \begin{bmatrix} \boldsymbol{\alpha}_t \\ \mathbf{y}_{2t}^c \end{bmatrix}, \quad \mathbf{y}_t^\dagger = \begin{bmatrix} \mathbf{y}_{1t} \\ \mathbf{y}_{2t}^c \end{bmatrix}$$

where the $\boldsymbol{\alpha}_t^*$ has dimension $m^* = m + N_2$. The measurement and transition equation are given by:

$$\mathbf{y}_t^\dagger = \mathbf{Z}^*\boldsymbol{\alpha}_t^* + \mathbf{X}_t\boldsymbol{\beta}, \quad \boldsymbol{\alpha}_t^* = \mathbf{T}_t^*\boldsymbol{\alpha}_{t-1}^* + \mathbf{W}_t^*\boldsymbol{\beta} + \mathbf{H}^*\boldsymbol{\epsilon}_t, \tag{6}$$

with starting values $\boldsymbol{\alpha}_1^* = \mathbf{W}_1^*\boldsymbol{\beta} + \mathbf{H}^*\boldsymbol{\epsilon}_1$, and system matrices:

$$\begin{aligned} \mathbf{Z}^* &= \begin{bmatrix} \mathbf{Z}_1 & \mathbf{0}_{N_1,N_2} \\ \mathbf{0}_{N_2,m} & \mathbf{I}_{N_2} \end{bmatrix}, \quad \mathbf{T}_t^* = \begin{bmatrix} \mathbf{T} & \mathbf{0}_{m,N_2} \\ \mathbf{Z}_2\mathbf{T} & \rho_t\mathbf{I}_{N_2} \end{bmatrix}, \\ \mathbf{W}_t^* &= \begin{bmatrix} \mathbf{W} \\ \mathbf{Z}_2\mathbf{W} + \mathbf{X}_{2,t} \end{bmatrix}, \quad \mathbf{H}^* = \begin{bmatrix} \mathbf{I}_m \\ \mathbf{Z}_2 \end{bmatrix}\mathbf{H}. \end{aligned} \tag{7}$$

The state space model (6)–(7) is linear and, assuming that the disturbances have a Gaussian distribution, the unknown parameters can be estimated by maximum likelihood, using the prediction error decomposition, performed by the Kalman filter. Given the parameter values, the Kalman filter and smoother (KFS) will provide the minimum mean square estimates of the states $\boldsymbol{\alpha}_t^*$. See Harvey (1989), Durbin and Koopman (2012), and the next section for details.

The KFS provides the best linear estimate of the sequence $\{\mathbf{y}_{2t}^c, t = 1, \ldots, n\}$, given the available observed time series. The latter can be then "decumulated", using $\mathbf{y}_{2t} = \mathbf{y}_{2t}^c - \rho_t\mathbf{y}_{2,t-1}^c$, so as to be converted into estimates of $\mathbf{y}_{2t}$, i.e. the monthly indicator of the GDP component. In order to provide the estimation standard error of $\mathbf{y}_{2t}$, however, an augmented state space form must be considered by including $\mathbf{y}_{2t}$ in the state vector, so that $\boldsymbol{\alpha}_t^* = \left[\boldsymbol{\alpha}_{t'}, \mathbf{y}_{2t}^{c'}, \mathbf{y}_{2t}'\right]'$, and augmenting the transition equation by the following recursive formula:

$$\begin{aligned}\mathbf{y}_{2,t} &= \mathbf{Z}_2\boldsymbol{\alpha}_t + \mathbf{X}_{2,t}\boldsymbol{\beta}, \\ &= \mathbf{Z}_2\mathbf{T}\boldsymbol{\alpha}_{t-1} + (\mathbf{X}_{2,t} + \mathbf{Z}_2\mathbf{W})\boldsymbol{\beta} + \mathbf{Z}_2\mathbf{H}\boldsymbol{\epsilon}_t.\end{aligned}$$

3 Estimation and Signal Extraction Filters

Statistical inference for the state space model (6)–(7) entails estimating the unknown parameters, optimal estimation of the unobserved components and the disaggregate GDP series based on both a real time and the full sample, as well as predicting future monthly GDP. Diagnostic checking requires the computation of the so–called innovations.

In this section we present a set of algorithms that are customized to perform all these tasks taking into account the data generating process, that is temporal aggregation and ragged–edge data structures. When missing data are present in a multivariate time series model, the option is to use sequential processing Anderson and Moore (1979). We also need to be able to entertain nonstationary as well as regression effects, which is done by assuming a diffuse prior on initial conditions and the effects of the explanatory variables (see de Jong, 1991).

Once the monthly components are estimated at chain–linked volumes, they can be aggregated into total GDP measures according to the procedure outlined in Section 4.3.

3.1 Univariate treatment of filtering and smoothing for multivariate models

The univariate statistical treatment of a multivariate state space model was considered by Anderson and Moore (1979), who refer to it as sequential processing. It provides a very flexible and convenient device for filtering and smoothing and for handling missing values. Our treatment is prevalently based on Koopman and Durbin (2000). However, for the treatment of regression effects and initial conditions we adopt the augmentation approach by de Jong (1991).

The multivariate vectors $\mathbf{y}_t^\dagger$, $t = 1, \dots, n$, (which can be partially observed at given times, e.g. in the first and second month of each quarter as far the component $\mathbf{y}_{2t}^c$ is concerned), are stacked one on top of the other to yield a univariate time series $\{y_{t,i}^\dagger, i = 1, \dots, N, t = 1, \dots, n\}$, whose elements are processed sequentially. The ordering according to the index i reflects the timing of the release of the series, with the more timely series being processed before.

The state space model for the univariate time series $\{y_{t,i}^\dagger\}$ is constructed as follows. The measurement equation for the i-th element of the vector $\mathbf{y}_t^\dagger$ is:

$$y_{t,i}^\dagger = \mathbf{z}_i^{*'}\boldsymbol{\alpha}_{t,i}^* + \mathbf{x}_{t,i}'\boldsymbol{\beta}, \quad t = 1, \dots, n, \quad i = 1, \dots, N, \tag{8}$$

where $\mathbf{z}_i^{*'}$ and $\mathbf{x}_{t,i}'$ denote the i-th rows of $\mathbf{Z}^*$ and $\mathbf{X}_t$, respectively. When the time index is kept fixed the transition equation is the identity:

$$\boldsymbol{\alpha}_{t,i}^* = \boldsymbol{\alpha}_{t,i-1}^*, i = 2, \dots, N,$$

whereas, for $i = 1$,

$$\boldsymbol{\alpha}^*_{t,1} = \mathbf{T}^*_t \boldsymbol{\alpha}^*_{t-1,N} + \mathbf{W}^*_t \boldsymbol{\beta} + \mathbf{H}^* \boldsymbol{\epsilon}_{t,1},$$

with $\boldsymbol{\epsilon}_{t,1} \sim \mathrm{N}(\mathbf{0}, \mathrm{Var}(\boldsymbol{\epsilon}_{t,1}))$, $\mathrm{Var}(\boldsymbol{\epsilon}_{t,1}) = \mathrm{diag}(1, \sigma_1^2, \ldots, \sigma_N^2) = \boldsymbol{\Sigma}_\epsilon$.

The state space form is completed by the specification of the moments of the distribution of the initial state vector, that is written as $\boldsymbol{\alpha}^*_{1,1} = \mathbf{a}^*_{1,1} + \mathbf{W}^*_1 \boldsymbol{\beta} + \mathbf{H}^* \boldsymbol{\epsilon}_{1,1}$, where $\mathbf{a}^*_{1,1}$ is a fixed and known vector, $\boldsymbol{\beta}$ is a diffuse random vector, and $\mathrm{Var}(\boldsymbol{\epsilon}_{1,1}) = \mathrm{Var}(\boldsymbol{\epsilon}_{t,1})$.

The augmented Kalman filter, accounting for the presence of missing values, is given by the following definitions and recursive formulae. The initial conditions are set equal to

$$\mathbf{a}^*_{1,1} = \mathbf{0},\ \mathbf{A}^*_{1,1} = \mathbf{W}^*_1,\ \mathbf{P}^*_{1,1} = \mathbf{H}\boldsymbol{\Sigma}_\epsilon \mathbf{H}',\ q_{1,1} = 0,\ \mathbf{s}_{1,1} = \mathbf{0},\ \mathbf{S}_{1,1} = \mathbf{0},\ cn = 0,\ d_{1,1} = 0.$$

Then, for $t = 1, \ldots, n$, $i = 1, \ldots, N-1$, if $y^\dagger_{t,i}$ is observed:

$$\begin{array}{lcllcl}
\nu_{t,i} & = & y^\dagger_{t,i} - \mathbf{z}^{*'}_i \mathbf{a}^*_{t,i}, & \mathbf{V}'_{t,i} & = & -\mathbf{x}'_{t,i} - \mathbf{z}^{*'}_i \mathbf{A}^*_{t,i}, \\
f_{t,i} & = & \mathbf{z}^{*'}_i \mathbf{P}^*_{t,i} \mathbf{z}^*_i, & \mathbf{K}_{t,i} & = & \mathbf{P}^*_{t,i} \mathbf{z}^*_i / f_{t,i}, \\
\mathbf{a}^*_{t,i+1} & = & \mathbf{a}^*_{t,i} + \mathbf{K}_{t,i} \nu_{t,i}, & \mathbf{A}^*_{t,i+1} & = & \mathbf{A}^*_{t,i} + \mathbf{K}_{t,i} \mathbf{V}'_{t,i}, \\
\mathbf{P}^*_{t,i+1} & = & \mathbf{P}^*_{t,i} - \mathbf{K}_{t,i} \mathbf{K}'_{t,i} f_t, & & & \\
q_{t,i+1} & = & q_{t,i} + \nu^2_{t,i} / f_{t,i}, & \mathbf{s}_{t,i+1} & = & \mathbf{s}_{t,i} + \mathbf{V}_{t,i} \nu_{t,i} / f_{t,i} \\
\mathbf{S}_{t,i+1} & = & \mathbf{S}_{t,i} + \mathbf{V}_{t,i} \mathbf{V}'_{t,i} / f_{t,i}, & d_{t,i+1} & = & d_{t,i} + \ln f_{t,i}, \\
cn & = & cn + 1. & & &
\end{array} \tag{9}$$

Else, if $y^\dagger_{t,i}$ is missing, which occurs for the second block of variables $\mathbf{y}^c_{2,t}$ systematically for $t \neq \delta\tau$:

$$\begin{array}{lcl}
\mathbf{a}^*_{t,i+1} & = & \mathbf{a}^*_{t,i}, \quad \mathbf{A}^*_{t,i+1} = \mathbf{A}^*_{t,i}, \\
\mathbf{P}^*_{t,i+1} & = & \mathbf{P}^*_{t,i}, \\
q_{t,i+1} & = & q_{t,i}, \quad \mathbf{s}_{t,i+1} = \mathbf{s}_{t,i}, \quad \mathbf{S}_{t,i+1} = \mathbf{S}_{t,i}, \quad d_{t,i+1} = d_{t,i}.
\end{array} \tag{10}$$

Then for $i = N$

$$\begin{array}{lcllcl}
\nu_{t,N} & = & y^\dagger_{t,N} - \mathbf{z}^{*'}_N \mathbf{a}^*_{t,N}, & \mathbf{V}'_{t,N} & = & -\mathbf{x}'_{t,N} - \mathbf{z}^{*'}_N \mathbf{A}^*_{t,N}, \\
f_{t,N} & = & \mathbf{z}^{*'}_N \mathbf{P}^*_{t,N} \mathbf{z}^*_N, & \mathbf{K}_{t+1,1} & = & \mathbf{T}^*_{t+1} \mathbf{P}^*_{t,N} \mathbf{z}^*_N / f_{t,N} \\
\mathbf{a}^*_{t+1,1} & = & \mathbf{T}^*_{t+1} \mathbf{a}^*_{t,N} + \mathbf{K}_{t+1,1} \nu_{t,N}, & \mathbf{A}^*_{t+1,1} & = & \mathbf{W}^*_{t+1} + \mathbf{T}^*_{t+1} \mathbf{A}^*_{t,N} + \mathbf{K}_{t+1,1} \mathbf{V}'_{t,N}, \\
 & & & \mathbf{P}^*_{t+1,1} & = & \mathbf{T}^*_{t+1} \mathbf{P}^*_{t,N} \mathbf{T}^{*'}_{t+1} + \mathbf{H}^* \boldsymbol{\Sigma}_\epsilon \mathbf{H}^{*'} - \mathbf{K}_{t+1,1} \mathbf{K}'_{t+1,1} f_{t,N}, \\
q_{t+1,1} & = & q_{t,N} + \nu^2_{t,N} / f_{t,N}, & d_{t+1,1} & = & d_{t,N} + \ln f_{t,N}. \\
\mathbf{s}_{t+1,1} & = & \mathbf{s}_{t,N} + \mathbf{V}_{t,N} \nu_{t,N} / f_{t,N}, & \mathbf{S}_{t+1,1} & = & \mathbf{S}_{t,N} + \mathbf{V}_{t,N} \mathbf{V}'_{t,N} / f_{t,N},
\end{array} \tag{11}$$

Here, $\mathbf{V}_{t,i}$ is a vector with $2N + k$ elements, $\mathbf{A}^*_{t,i}$ is $m \times (2N + k)$, cn counts the number of observations.

The diffuse estimate of the vector $\boldsymbol{\beta}$, and its covariance matrix, are, respectively,

$$\hat{\boldsymbol{\beta}} = \mathbf{S}^{-1}_{n+1,1} \mathbf{s}_{n+1,1}, \qquad \mathrm{Var}(\hat{\boldsymbol{\beta}}) = \mathbf{S}^{-1}_{n+1,1}. \tag{12}$$

The diffuse likelihood, based on de Jong (1991) and denoted $\mathcal{L}_\infty$, takes the expression:

$$\mathcal{L}_\infty = -0.5 \left[d_{n+1,1} + (cn - K) \ln(2\pi) + \ln |\mathbf{S}_{n+1,1}| + q_{n+1,1} - \mathbf{s}'_{n+1,1} \mathbf{S}^{-1}_{n+1,1} \mathbf{s}_{n+1,1} \right]. \tag{13}$$

The latter is maximised with respect to the unknown parameters.

3.2 Multivariate innovations, filtered and real time estimates

Diagnostics and goodness of fit are based on the innovations, that are given by $\tilde{\nu}_{t,i} = \nu_{t,i} + \mathbf{V}'_{t,i}\mathbf{S}^{-1}_{t,i}\mathbf{s}_{t,i}$, with variance $\tilde{f}_{t,i} = f_{t,i} + \mathbf{V}'_{t,i}\mathbf{S}^{-1}_{t,i}\mathbf{V}_{t,i}$. The standardised innovations, $\tilde{\nu}_{t,i}/\sqrt{\tilde{f}_{t,i}}$ can be used to check for residual autocorrelation and departure from the normality assumption. The innovations have the following interpretation:

$$\tilde{\nu}_{t,i} = y^{\dagger}_{t,i} - \mathrm{E}(y^{\dagger}_{t,i}|\mathbf{Y}^{\dagger}_{t-1}, y^{\dagger}_{t,j}, j < i),$$

where $\mathbf{Y}^{\dagger}_t$ denotes the information set $\{\mathbf{y}^{\dagger}_1, \dots, \mathbf{y}^{\dagger}_t\}$.

If $y^{\dagger}_{t,i}$ is observed and $i < N$, the filtered, or real–time, estimates of the state vector and the estimation error matrix are computed as follows:

$$\tilde{\boldsymbol{\alpha}}^*_{t,i} = \mathbf{a}^*_{t,i} + \mathbf{A}^*_{t,i}\mathbf{S}^{-1}_{t,i+1}\mathbf{s}_{t,i+1} + \mathbf{P}^*_{t,i}\mathbf{z}^*_i\tilde{\nu}_{t,i}/f_{t,i}, \;\; \tilde{\mathbf{P}}^*_{t,i} = \mathbf{P}^*_{t,i} + \mathbf{A}^*_{t,i}\mathbf{S}^{-1}_{t,i+1}\mathbf{A}^{*'}_{t,i} - \mathbf{P}^*_{t,i}\mathbf{z}^*_i\mathbf{z}'_i\mathbf{P}^*_{t,i}/f_{t,i},$$

where $\tilde{\boldsymbol{\alpha}}^*_{t,i} = \mathrm{E}(\boldsymbol{\alpha}^*_t|\mathbf{Y}^{\dagger}_{t-1}, y^{\dagger}_{t,j}, j \le i)$, $\tilde{\mathbf{P}}^*_{t,i} = \mathrm{Var}(\boldsymbol{\alpha}^*_t|\mathbf{Y}^{\dagger}_{t-1}, y^{\dagger}_{t,j}, j \le i)$. For $i = N$,

$$\tilde{\boldsymbol{\alpha}}^*_{t,N} = \mathbf{a}^*_{t,N} + \mathbf{A}^*_{t,N}\mathbf{S}^{-1}_{t+1,1}\mathbf{s}_{t+1,1} + \mathbf{P}^*_{t,N}\mathbf{z}^*_N\tilde{\nu}_{t,N}/f_{t,N}, \;\; \tilde{\mathbf{P}}^*_{t,N} = \mathbf{P}^*_{t,N} + \mathbf{A}^*_{t,N}\mathbf{S}^{-1}_{t+1,1}\mathbf{A}^{*'}_{t,N} - \mathbf{P}^*_{t,N}\mathbf{z}^*_N\mathbf{z}'_N\mathbf{P}^*_{t,N}/f_{t,N}.$$

For $i < N$, $\tilde{\boldsymbol{\alpha}}^*_{t,i}$ coincides with the one–step–ahead prediction $\tilde{\boldsymbol{\alpha}}^*_{t,i+1} = \mathrm{E}(\boldsymbol{\alpha}^*_t|\mathbf{Y}^{\dagger}_{t-1}, y^{\dagger}_{t,j}, j \le i)$ as the transition equation is

$$\boldsymbol{\alpha}^*_{t,i} = \boldsymbol{\alpha}^*_{t,i-1}, i = 2, \dots, N,$$

whereas, for $i = N$,

$$\begin{aligned} \tilde{\boldsymbol{\alpha}}^*_{t+1,1} &= \mathbf{a}^*_{t+1,1} - \mathbf{A}^*_{t+1,1}\mathbf{S}^{-1}_{t+1,1}\mathbf{s}_{t+1,1}, \\ \tilde{\mathbf{P}}^*_{t+1,1} &= \mathbf{P}^*_{t+1,1} + \mathbf{A}^*_{t+1,1}\mathbf{S}^{-1}_{t+1,1}\mathbf{A}^{*'}_{t+1,1} \end{aligned} \quad (14)$$

$$\begin{aligned} \tilde{\boldsymbol{\alpha}}^*_{t+1,1} &= \mathbf{a}^*_{t+1,1} + \mathbf{A}^*_{t+1,1}\mathbf{S}^{-1}_{t+1,1}\mathbf{s}_{t+1,1} + \mathbf{P}^*_{t+1,1}\mathbf{z}^*_1\tilde{\nu}_{t,N}/f_{t,N}, \\ \tilde{\mathbf{P}}^*_{t+1,1} &= \mathbf{P}^*_{t+1,1} + \mathbf{A}^*_{t+1,1}\mathbf{S}^{-1}_{t+1,1}\mathbf{A}^{*'}_{t+1,1} - \mathbf{P}^*_{t+1,1}\mathbf{z}^*_1\mathbf{z}'_1\mathbf{P}^*_{t+1,1}/f_{t,N}, \end{aligned}$$

are respectively the one–step–ahead prediction $\tilde{\boldsymbol{\alpha}}^*_{t+1,1} = \mathrm{E}(\boldsymbol{\alpha}^*_{t+1}|\mathbf{Y}^{\dagger}_{t-1})$ and the predictive variance $\tilde{\mathbf{P}}^*_{t+1,1} = \mathrm{Var}(\boldsymbol{\alpha}^*_{t+1}|\mathbf{Y}^{\dagger}_{t-1})$. The corresponding expressions for $y^{\dagger}_{t,i}$ missing are straightforward.

3.3 Smoothed estimates

The smoothed estimates are obtained from the augmented smoothing algorithm proposed by de Jong (1988), appropriately adapted here to handle missing values and sequential processing of the observations. Defining $\mathbf{r}_{n,N} = 0, \mathbf{R}_{n,N} = 0, \mathbf{N}_{n,N} = 0$, for $t = n, \dots, 1$, and $i = N, \dots, 1$ if $y^{\dagger}_{t,i}$ is available:

$$\begin{aligned} \mathbf{L}_{t,i} &= \mathbf{I}_m - \mathbf{K}_{t,i}z^{*'}_i \\ \mathbf{r}_{t,i-1} &= \mathbf{z}^*_i\nu_{t,i}/f_{t,i} + \mathbf{L}_{t,i}\mathbf{r}_{t,i}, \;\; \mathbf{R}_{t,i-1} = \mathbf{z}^*_i\mathbf{V}'_{t,i}/f_{t,i} + \mathbf{L}_{t,i}\mathbf{R}_{t,i}, \\ \mathbf{N}_{t,i-1} &= \mathbf{z}^*_i\mathbf{z}^{*'}_i/f_{t,i} + \mathbf{L}_{t,i}\mathbf{N}_{t,i}\mathbf{L}'_{t,i}. \end{aligned}$$

Else, if $y^{\dagger}_{t,i}$ is missing,

$$\mathbf{r}_{t,i-1} = \mathbf{r}_{t,i}, \quad \mathbf{R}_{t,i-1} = \mathbf{R}_{t,i}, \quad \mathbf{N}_{t,i-1} = \mathbf{N}_{t,i}.$$

$$\mathbf{r}_{t-1,N} = \mathbf{T}^{*\prime}_{t+1}\mathbf{r}_{t,i}, \quad \mathbf{R}_{t,i-1} = \mathbf{T}^{*\prime}_{t+1}\mathbf{R}_{t,i}, \quad \mathbf{N}_{t,i-1} = \mathbf{T}^{*\prime}_{t+1}\mathbf{N}_{t,i}\mathbf{T}^{*}_{t+1}.$$

The smoothed estimates are obtained as

$$\begin{array}{rcl} \tilde{\boldsymbol{\alpha}}^{*}_{t|n} &=& \mathbf{a}^{*}_{t,1} + \mathbf{A}^{*}_{t,1}\tilde{\boldsymbol{\beta}} + \mathbf{P}^{*}_{t,1}(\mathbf{r}_{t-1,N} + \mathbf{R}_{t-1,N}\tilde{\boldsymbol{\beta}}) \\ \mathbf{P}^{*}_{t|n} &=& \mathbf{P}^{*}_{t,1} + \mathbf{A}^{*}_{t,1}\mathbf{S}^{-1}_{n+1}\mathbf{A}^{*\prime}_{t,1} - \mathbf{P}^{*}_{t,1}\mathbf{N}_{t-1,N}\mathbf{P}^{*}_{t,1}. \end{array} \tag{15}$$

4 Density Estimation, Boostrapping and Conditional Simulation

Let y_t denote a generic monthly GDP component; the aim is drawing samples $y^{(r)}_{t+l}, t = 1, \ldots, n,\ r = 1, \ldots, R$, and $l \geq 0$, from the conditional distribution $f(y_{t+l}|\mathbf{Y}^{\dagger}_{s})$. When $s = t$ and $l = 0$, the estimation of $f(y_t|\mathbf{Y}^{\dagger}_{s})$ can be referred to as density nowcasting. When $s = n > t$ and $l = 0$, we shall refer to historical estimation or smoothing, and finally when $s = t$ and $l > 0$, we deal with density forecasting.

If $\boldsymbol{\psi}$ denotes the vector containing the hyperparameters of the model, we let $\tilde{\boldsymbol{\psi}}$ denote its maximum likelihood estimator (MLE). Assume that $\tilde{\boldsymbol{\psi}}$ has a distribution represented by a density function $f(\tilde{\boldsymbol{\psi}}_s)$. The required density is obtained as

$$f(y_{t+l}|\mathbf{Y}^{\dagger}_{s}) = \int f(y_{t+l}|\tilde{\boldsymbol{\psi}}, \mathbf{Y}^{\dagger}_{s}) f(\tilde{\boldsymbol{\psi}}) \mathrm{d}\tilde{\boldsymbol{\psi}}. \tag{16}$$

We can draw samples $y^{(r)}_{t+l}, r = 1, \ldots, R$, from (16) by the method of composition (Tanner, 1996), using the following algorithm:

1. Obtain a bootstrap sample $\tilde{\boldsymbol{\psi}}^{(r)} \sim f(\tilde{\boldsymbol{\psi}})$ using the bootstrap algorithm outlined in the next subsection.

2. Obtain an independent sample $y^{(r)}_{t+l} \sim f(y_{t+l}|\tilde{\boldsymbol{\psi}}^{(r)}, \mathbf{Y}^{\dagger}_{s})$, by the simulation smoother, see Durbin and Koopman (2012), described in Section 4.2.

The density (16) is estimated nonparametrically by kernel methods from the samples $y^{(r)}_{t+l}, r = 1, \ldots, R$. A Rao–Blackwellised estimate of the required density is also possible as, given $\tilde{\boldsymbol{\psi}}$, $f(y_{t+l}|\tilde{\boldsymbol{\psi}}, \mathbf{Y}^{\dagger}_{s})$ is Gaussian.

The bootstrap sample obtained in the first step could be replaced by sample from the multivariate normal approximation to the distribution of the MLE. However, we do not pursue this strategy, as the normal approximation may not suitable in small samples. The remainder of this section deals with the methods for boostrapping and simulating from the predictive distribution of the monthly GDP components.

4.1 The bootstrap algorithm

The bootstrap algorithm obtains simulated series from the innovations form of the univariate representation of the multivariate state space form, conditional on the MLE $\tilde{\boldsymbol{\psi}}$.

Let k denote the number of elements of $\boldsymbol{\beta}$ and let t_0 denote an initial time such that the number of available observations (among Nt_0) is larger than k. The bootstrap sample of $y^{\dagger}_{t,i}$ will be denoted by $y^{(b)}_{t,i}$. The first observations up to time t_0 are obtained from the observed sample; the subsequent are obtained form the following algorithm:

1. Resample with replacement the non–missing multivariate standardised innovations $\tilde{\nu}_{t,i}/\sqrt{\tilde{f}_{t,i}}$, $t = t_0 + 1, \dots, n$, $i = 1, \dots, N$, and denote the obtained sample by $\{\xi_{t,i}\}$.

2. Run the augmented Kalman filter (see Section 3.1) for observations $t = 1, \dots, t_0$, and obtain $\tilde{\boldsymbol{\alpha}}^*_{t_0+1,1}$ and $\tilde{\boldsymbol{\beta}}_{t_0+1,1} = \mathbf{S}^{-1}_{t_0+1,1}\mathbf{s}_{t_0+1,1}$.

3. Set $y^{(b)}_{t,i} = y^{\dagger}_{t,i}$ for the initial stretch $t = 1, \dots, t_0, i = 1, \dots, N$.

4. For $t = t_0 + 1, \dots, n$, obtain a bootstrap sample $y^{(b)}_{t,i}$ by the following algorithm: for $i = 1, \dots, N$, if $y^{\dagger}_{t,i}$ is observed, set

$$y^{(b)}_{t,i} = \mathbf{z}^{*\prime}_i \tilde{\boldsymbol{\alpha}}^*_{t,i} + \mathbf{x}'_{t,i}\tilde{\boldsymbol{\beta}}_{t,i} + \sqrt{\tilde{f}_{t,i}}\xi_{t,i}, \tag{17}$$

   and compute for $i = 1, \dots, N$,

$$\begin{array}{rcl} \mathbf{a}^{(b)}_{t,i+1} & = & \mathbf{a}^{(b)}_{t,i} + \mathbf{K}_{t,i}\sqrt{\tilde{f}_{t,i}}\xi_{t,i}, \\ \tilde{\boldsymbol{\beta}}_{t,i+1} & = & \tilde{\boldsymbol{\beta}}_{t,i} + \mathbf{S}^{-1}_{t,i}\mathbf{V}_{t,i}\frac{1}{f_{t,i}+\mathbf{V}'_{t,i}\mathbf{S}^{-1}_{t,i}\mathbf{V}_{t,i}}\sqrt{\tilde{f}_{t,i}}\xi_{t,i}, \\ \tilde{\boldsymbol{\alpha}}^*_{t,i+1} & = & \tilde{\mathbf{a}}^{(b)}_{t,i+1} + \mathbf{A}^*_{t,i+1}\tilde{\boldsymbol{\beta}}_{t,i+1} \end{array} \tag{18}$$

   The quantities $\mathbf{K}_{t,i}$, $\mathbf{V}_{t,i}$, $\mathbf{A}^*_{t,i+1}$ are available from the Kalman filter (9)–(10). The recursion for $\tilde{\boldsymbol{\beta}}_{t,i}$ originates from the recursive estimation of the fixed effects vector, so that only real–time information is used.
   Else, if $y^{\dagger}_{t,i}$ is missing, set $y^{(b)}_{t,i}$ to missing and let

$$\mathbf{a}^{(b)}_{t,i+1} \;=\; \mathbf{a}^{(b)}_{t,i}, \quad \tilde{\boldsymbol{\beta}}_{t,i+1} = \tilde{\boldsymbol{\beta}}_{t,i}, \quad \tilde{\boldsymbol{\alpha}}^*_{t,i+1} = \tilde{\boldsymbol{\alpha}}^*_{t,i+1} + \mathbf{A}^*_{t,i+1}\tilde{\boldsymbol{\beta}}_{t,i+1}.$$

   Then for $i = N$, if $y^{\dagger}_{t,i}$ is observed, compute

$$\begin{array}{rcl} \mathbf{a}^{(b)}_{t+1,1} & = & \mathbf{T}^*_{t+1}\mathbf{a}^{(b)}_{t,N} + \mathbf{K}_{t,N}\sqrt{f_{t,N}}\xi_{t,N}, \\ \tilde{\boldsymbol{\beta}}_{t+1,1} & = & \tilde{\boldsymbol{\beta}}_{t,N} + \mathbf{S}^{-1}_{t,N}\mathbf{V}_{t,N}\frac{1}{f_{t,N}+\mathbf{V}'_{t,N}\mathbf{S}^{-1}_{t,N}\mathbf{V}_{t,N}}\sqrt{f_{t,N}}\xi_{t,N}, \\ \tilde{\boldsymbol{\alpha}}^*_{t+1,1} & = & \mathbf{a}^{(b)}_{t+1,1} + \mathbf{A}^*_{t+1,1}\tilde{\boldsymbol{\beta}}_{t+1,1}. \end{array} \tag{19}$$

   Else, if $y^{\dagger}_{t,N}$ is missing, set $y^{(b)}_{t,N}$ to missing and let

$$\mathbf{a}^{(b)}_{t+1,1} \;=\; \mathbf{T}^*_{t+1}\mathbf{a}^{(b)}_{t,i}, \quad \tilde{\boldsymbol{\beta}}_{t,i+1} = \tilde{\boldsymbol{\beta}}_{t,i}, \quad \tilde{\boldsymbol{\alpha}}^*_{t,i+1} = \mathbf{a}^{(b)}_{t+1,1} + \mathbf{A}^*_{t,i+1}\tilde{\boldsymbol{\beta}}_{t,i+1}$$

5. Estimate the model (6) using the simulated observations $y_{t,i}^{(b)}, t = 1, \ldots, n, i = 1, \ldots, N$, by maximum likelihood. The resulting estimate, $\tilde{\boldsymbol{\psi}}^{(r)}$, is the required bootstrap sample from $f(\tilde{\boldsymbol{\psi}})$.

4.2 The conditional simulation sampler

Conditional on the parameter vector $\tilde{\boldsymbol{\psi}}^{(r)}$, we draw samples from the conditional distribution of the states, and hence of the monthly GDP component, given the past (filtered distribution), the current information (real–time distribution), and all the available information (smoothed distribution). The conditional simulation sampler is obtained from the following algorithm:

- Simulate the complete monthly data $\mathbf{y}_t^+, t = 1, \ldots, n$, from the state space model (2). This is achieved as follows:
  - Draw $\boldsymbol{\alpha}_1^+ \sim \mathrm{N}\left(\mathbf{W}_1\hat{\boldsymbol{\beta}}, \mathbf{H}\mathrm{Var}(\boldsymbol{\epsilon}_1)\mathbf{H}'\right)$, $\mathrm{Var}(\boldsymbol{\epsilon}_1) = \mathrm{diag}\,(1, \sigma_1^2, \ldots, \sigma_N^2)$, where $\hat{\boldsymbol{\beta}}$ is kept fixed at the value given in (12) and $\sigma_i^2, i = 1, \ldots, N$, are the parameter values corresponding to $\tilde{\boldsymbol{\psi}}^{(r)}$.
  - For $t = 1, \ldots, n$, simulate the complete observations and the next states according to:

$$\mathbf{y}_t^+ = \mathbf{Z}\boldsymbol{\alpha}_t^+ + \mathbf{X}_t\hat{\boldsymbol{\beta}}, \quad \boldsymbol{\alpha}_{t+1}^+ = \mathbf{T}\boldsymbol{\alpha}_t^+ + \mathbf{W}\boldsymbol{\beta} + \mathbf{H}\boldsymbol{\epsilon}_{t+1}^+, \boldsymbol{\epsilon}_{t+1}^+ \sim \mathrm{N}\left(\mathbf{0}, \mathrm{diag}(1, \sigma_1^2, \ldots, \sigma_N^2)\right).$$

- Aggregate the last N_2 series using the cumulator $\mathbf{y}_{2,t}^{c+} = \rho_t \mathbf{y}_{2,t-1}^{c+} + \mathbf{y}_{2t}^+$.

- Delete the observations corresponding to missing values in the original data (in the case of $\mathbf{y}_{2,t}^{c+}$ this correspond to systematically sampling the end of quarter value).

- Run the Kalman filter and smoother to the simulated series so as to obtain the conditional means of the states (predicted, real–time and smoothed), $\tilde{\boldsymbol{\alpha}}_{t|s}^+ = \mathrm{E}(\boldsymbol{\alpha}_t|\mathbf{Y}_s^+)$, where $\mathbf{Y}_s^+$ is the set of simulated observations up to and including time s obtained at the previous step (i.e. having the same pattern of missingness as the original data).

- The required draw from the conditional distribution of the states given the original series is thus $\tilde{\boldsymbol{\alpha}}_{t|s} + (\boldsymbol{\alpha}_t^+ - \tilde{\boldsymbol{\alpha}}_{t|s}^+)$.

Application of this algorithm yields the required draw from the conditional distribution of the monthly GDP component given the (real–time, historical) data. The latter can be aggregated according to the procedure outlined in the next subsection.

4.3 Pooling the density estimates of the GDP components

The methods outlined in the previous section yield estimates for the monthly GDP components in chain–linked volumes, with the year 2000 as the reference year. As it is well known, chain–linked volume measures are not consistent in aggregation, e.g. the sum of the value added of the six sectors plus taxes less subsidies would not deliver GDP at market prices, unless they are first expressed at the average prices of the previous years. Hence, to obtain our total monthly GDP estimates at market prices and chained volumes, we resort to a multistep procedure proposed by Frale et al. (2011), that enforces the so–called annual overlap method advocated by Bloem et al. (2001). This consists essentially of three main steps: 1. *Dechaining*, aiming at expressing the draws at the average prices of the previous years. 2. *Aggregation*, which computes total GDP for the output and expenditure approaches, expressed at the average prices of the previous year, according to the two identities:

$$\begin{aligned} \text{GDP at market prices} &= \textstyle\sum_i \text{Value added of branch } i + \text{TlS}, \\ & \quad i = \text{A–B, C–D–E, F, G–H–I, J–K, L–P} \\ \text{GDP at market prices} &= \text{FCE} + \text{GCF} + \text{EXP} - \text{IMP} \end{aligned}$$

3. *Chain linking*, aiming at expressing GDP at chain–linked volumes with reference year 2000. Appendix A provides a detailed description of the above three steps and also deals with the incorporation of the flash estimate of total GDP at market prices compiled by Eurostat.

Application of this procedure results in the density estimates of monthly GDP from the output and expenditure approaches, labelled EuroMInd-$\mathcal{D}_o$ and EuroMInd-$\mathcal{D}_e$. These estimates arise as a linear opinion pool with known fixed aggregation weights of the conditional densities of the GDP components at the prices of the previous years (step 1 converted the density at chained volumes into densities at the prices of the previous years), that is then converted at chain–linked volumes.

5 Evaluation and Combination of the Density Forecasts

5.1 Evaluation

Density forecasts obtained in the context of pseudo real–time forecasting and nowcasting of quarterly GDP will be evaluated using commonly used statistical methods which are based on the so–called "prequential principle" (Dawid, 1984). According to this principle, the assessment of the adequacy depends on the forecasts and realized outcomes only. The observability of quarterly GDP values permits the use of the prequential approach. Evaluation can be performed along two dimensions. The first one concerns the correct specification of each single predictive density while the second one pertains to the comparison across different approaches.

Correct specification of density forecasts is in the literature usually tested using the so–called probability integral transform (PIT). The concept of the PIT goes back to Rosenblatt (1952) but has been popularized by Dawid (1984) and Diebold et al. (1998). Given a density forecast in the form of a density $f(u)$ with the corresponding cumulative distribution function (CDF) $F(u)$, the PIT, p_τ, corresponding to an observed outcome y_τ is defined as follows:

$$p_\tau = \int_{-\infty}^{y_\tau} f(u)\mathrm{d}u$$

We also consider transformed PITs by defining the normal or z–scores as $z_\tau = \Phi^{-1}(p_\tau)$, where $\Phi(\cdot)$ denotes the standard normal CDF. A density forecast is said to be probabilistically calibrated if the PITs are uniformly distributed on the unit interval, i.e. $U(0,1)$ (Gneiting and Ranjan, 2013), or, equivalently, if the z–scores are standard normal (Berkowitz, 2001). Further, we say that a one-step-ahead density forecast is completely calibrated if the PITs and the z–scores are independently and identically distributed (as a uniform and as a standard normal, respectively).

Probabilistic calibration is tested by performing tests of the distributional assumptions. Following Diebold et al. (1998), we employ the Kolmogorov–Smirnov (KS) and the Cramér–von–Mises (CvM) test. We complement them with the Anderson–Darling (AD) test which, as found by Noceti et al. (2003) in their simulation exercise, seems to be more powerful in detecting misspecifications compared to the KS and the CvM tests. For a detailed description of the aforementioned tests as well as their properties, we refer to Noceti et al. (2003). Additionally, we consider the test based on the Pearson's χ^2 goodness–of–fit statistic advocated by Wallis (2003) for evaluation of density forecasts. For testing the normality of the transformed PITs we apply the Bowman–Shenton (BS) test. As regards the independence assumption, Diebold et al. (1998) suggests to use tests on zero–autocorrelation. We perform the Ljung–Box test based on autocorrelation coefficients up to the fourth lag. We also examine complete calibration with the joint test proposed by Berkowitz (2001). This test formulates the null hypothesis: $z_\tau \sim$ IID $\mathrm{N}(0,1)$ which can be tested using an AR(1) model for z_τ:

$$z_\tau - \mu = \rho(z_{\tau-1} - \mu) + \varepsilon_\tau,$$

with μ, ρ and ε denote the mean, the first–order autocorrelation, and an error term, respectively. This representation allows for reformulating the null as: $\mu = 0$, $\rho = 0$ and $\mathrm{Var}(\varepsilon_\tau) = 1$. The test employs the likelihood ratio statistic, which is distributed as a $\chi^2(3)$ random variable under the null.

The assumption of identical distribution is usually not tested in the literature. An attempt to take into account possible instabilities of the data and changes over time by testing on the correct specification in the sub–samples is provided by Rossi and Sekhposyan (2013).

As regards the comparison across different competing density forecasts, scoring rules can serve as a suitable tool for this purpose, as they attach a numerical score representing a quality measure. This quality measure addresses two aspects: calibration and sharpness referring to the dispersion of the predictive distribution (Gneiting et al., 2007).

A scoring rule is proper if the expected value of the score is maximized for an observation drawn from the distribution being the same as the one the forecasts are issued from. See Gneiting and Raftery (2007) for a comprehensive review of different scoring rules. In this article, we adopt two proper scores: the logarithmic score (LogS), proposed by Good

(1952), and the continuous ranked probability score (CRPS), introduced by Matheson and Winkler (1976). The log score at the realized outcome is given as:

$$\text{LogS}(y_\tau) = \log f(y_\tau)$$

LogS is associated with the Kullback–Leibler information criterion (KLIC):

$$\text{KLIC} = \text{E}[\log g(y_\tau) - \log f(y_\tau)],$$

with $g(\cdot)$ denoting the true density, in that maximizing the mean LogS minimizes the KLIC. Even though the logarithmic score exhibits desirable properties and is widely used in the evaluation of density forecasts, its drawback is that it is not robust, for example in presence of outliers. Complete failure of a probabilistic prediction for a single observation would, according to the LogS, disqualify a model even if its overall forecasting performance at all other observations might be good. This property is also called hypersensitivity (Selten, 1998). A more robust and indulgent alternative is the CRPS which assigns high probabilities for values close, but not necessarily equal, to the realized one. CRPS penalizes deviations of the predictive CDF from the true one for a particular time point. More formally,

$$\text{CRPS}(y_\tau) = -\int [F(u) - I(y_\tau)]^2 \text{d}u, \tag{20}$$

where $I(\cdot)$ is an indicator variable taking value 1, if $u > y_\tau$, and 0 otherwise. Eq. (20) can be also written as (Gneiting and Raftery, 2007):

$$\text{CRPS} = \frac{1}{2}\,\text{E}_F|Y - Y'| - \text{E}_F|Y - y_\tau|, \tag{21}$$

where E_F is the expectation with respect to the forecast distribution F, and Y and Y' are independent random draws from F. For the simulated forecast sample, we follow Panagiotelis and Smith (2008), who provide computational algorithms for approximating expressions in (21).

5.2 Combination

Combination of density forecasts is a relatively new area in macroeconomic research. In contrast, combination of point forecasts has a long history (Bates and Granger, 1969). Since the works by Mitchell and Hall (2005) and Hall and Mitchell (2007) justifying forecast density combination and emphasizing the importance of future developments, forecasting research of the recent years pays a lot of attention to this topic.

There are two central issues associated with density combination: the method of aggregation and the choice of weights. The most prominent aggregation method in empirical macroeconomic applications is the linear combination, also known as the "linear opinion pool" (Stone, 1961; Bacharach, 1974). Some studies also apply the "logarithmic opinion pool" proposed by (Winkler, 1968). Kascha and Ravazzolo (2010) analyze properties of combined densities following from these two aggregation methods. They point out that, in contrast to the typically multimodal linear combination, the logarithmic pool is unimodal and retains the symmetry of the individual densities. There is, however, no agreement on which pooling method performs better (Kascha and Ravazzolo, 2010; Bjørnland et al., 2011). Ranjan and Gneiting (2010) and Gneiting and Ranjan (2013) stress that the success of a particular pool depends on the dispersion of the single densities. In general, densities can be underdispersed, oversdispersed or neutrally dispersed, if the variance of the PITs is greater, less or equal to 1/12, respectively. It can be shown that linear combinations are more dispersed than the least dispersed of the components, and overdispersed if the components are neutrally dispersed. The good performance of the linear approach may indeed result from the aggregation of underdispersed components.

As far as the weighing scheme is concerned, the approaches vary from simple ones, like equal weights proposed by Hendry and Clements (2004) and Wallis (2005) to more elaborate ones. Similarly as in point forecasting, weights can be constructed using the mean square error (MSE). Another common weighing method is based on scoring rules which seems intuitively compelling as it takes features of the forecast densities directly into account. A great bulk of literature applies weights constructed with the logarithmic scores; see, among others, Amisano and Giacomini (2007), Hall and Mitchell (2007), Kascha and Ravazzolo (2010), Jore et al. (2010), Bache et al. (2011), Garratt et al. (2014), Aastveit et al. (2014). Weights based on the CRPS are also popular in macroeconomic forecasting (Bjørnland et al., 2011; Groen et al., 2013; Ravazzolo and Vahey, 2014).

The novelty of this article is that, instead of combining forecast densities obtained with different models, we link predictive densities from the output and the expenditure approaches for computing GDP. In particular, for the forecasting period between $\underline{\tau}$ and $\overline{\tau}$, we employ the linear opinion pool:

$$c_\tau(u) = w_{o,\tau}\, f_\tau(u|I_{o,\kappa}) + w_{e,\tau}\, f_\tau(u|I_{e,\kappa}), \qquad \tau = \underline{\tau}, \dots, \overline{\tau}$$

where $c_\tau(u)$ denotes the pooled forecast density at quarter τ, $w_{o,\tau}$ and $w_{e,\tau}$ are weights assigned to the single densities corresponding to the output and the expenditure approach, respectively. Further, $I_{i,\kappa}, i = o, e$, is the information set used in the respective approach at time point κ to produce the density forecast $f_\tau(u)$ for quarter τ. More details on the information set will be given in the description of the forecasting exercise in Section 6.3.1. We use recursive weights based on the scoring rules described in the previous section:

$$w_{i,\tau} = \frac{s_{i,\tau}}{\sum_{i\in\ o,e} s_{i,\tau}}, \qquad \tau = \underline{\tau}, \dots, \overline{\tau}, \quad i = o, e$$

In the case of the log score, $s_{i,\tau} = \exp\left[\sum_{\underline{\tau}}^{\tau} \mathrm{LogS}_\tau(y_\tau|I_{i,\kappa})\right]$. For the CRPS, it holds that $s_{i,\tau} = \left[\sum_{\underline{\tau}}^{\tau} \mathrm{CRPS}_\tau(y_\tau|I_{i,\kappa})\right]$. Additionally, we consider two non–recursive weighing schemes ($w_{i,\tau} = w_i$): equal weights ($w_i = 1/2$) and weights based on the MSE ($s_i = 1/\mathrm{MSE}_i$), as they proved successful in aggregating point forecasts (Clemen, 1989).

It is to be noted that in the empirical application of this article, evaluation by means of methods outlined in the previous subsection as well as combination of predictive densities is conducted using estimated (smoothed) densities (unless otherwise stated). We estimate the predictive densities by smoothing the simulated samples (empirical densities) at every quarterly observation with the Gaussian kernel at 401 equally spaced points using a rule–of–thumb bandwidth.

6 Empirical evidence

6.1 The data

The time series were downloaded from the Europa dataset made available electronically by Eurostat on the web site `http://epp.eurostat.ec.europa.eu/`. The series concerning the value added of the sectors and the GDP components are released with 65 days of delay with respect to the end of each quarter, according to a release calendar available at the beginning of each year[1]. Among the monthly indicators, the financial aggregates are compiled by the ECB with a publication delay of around 30 days from the closing of the reference month. The data for the Industry sector and retail (such as the Industrial Production Index and Car registrations) are released about 45 days after the end of the reference month. Other indicators, such as those for construction (index of production and building permits), the series for the agricultural sector and the labour market (employment and hours worked) have a publication delay of about 70 days. A total of 36 time series is used to estimate EuroMInd-$\mathcal{D}$. Table 1 provides the complete list by GDP component. The sample period starts in January 1995 and ends in June 2014 for most of the monthly indicators. The GDP quarterly components are available up to the first quarter of 2014; for the second quarter, only the flash estimate of total GDP is available.

[1] A preliminary estimate of total GDP at market prices, the so-called flash estimate, is released in advance (45 days after the end of the reference quarter), with the aim of providing a more timely assessment of the level of economic activity

Table 1: List of the time series used for the estimation of EuroMInd, by GDP component.

Time series	Frequency	Delay
A–B: Agriculture, hunting and fishing		
Production of milk	m	60
Bovine meat production in tons	m	60
Value added (chain–linked volumes) AB	q	65
C–D–E: Industry, incl. energy		
Index of Industrial Production Germany	m	45
Index of Industrial Production France	m	45
Index of Industrial Production Italy	m	45
Index of Industrial Production Spain	m	45
Volume of work done (hours worked)	m	60
Value added (chain–linked volumes) CDE	q	65
F: Construction		
Monthly production index	m	70
Building permits	m	70
Volume of work done (hours worked)	m	70
Value added (chain–linked volumes) F	q	65
G–H–I: Trade, transport and communication services		
Monthly production index for consumption goods	m	45
Index of deflated turnover	m	35
Car registrations	m	15
Value added (chain–linked volumes) GHI	q	65
J–K: Financial services and business activities		
Monetary aggregate M3 (deflated)	m	27
Loans of MFI (deflated)	m	27
Value added (chain–linked volumes) JK	q	65
L–P: Other services		
Debt securities issued by central government (deflated)	m	27
Value added (chain–linked volumes) LP	q	65
TlS: Taxes less subsidies on products		
Index of Industrial Production for the euro area	m	45
Index of deflated turnover, retail sector	m	35
Taxes less subsidies (chain–linked volumes)	q	65
FCE: Final consumption expenditure		
Monthly production index for consumption goods	m	45
Index of deflated turnover, retail sector	m	35
Car registrations	m	15
Final consumption expenditure (chain–linked volumes)	q	65
GCF: Gross capital formation		
Monthly production index (CDE) euro area	m	45
Monthly production index for capital goods	m	45
Building permits	m	70
Gross capital formation (chain–linked volumes)	q	65
EXP: Exports of goods and services		
Monthly Export volume index	m	42
Monthly production index for intermediate goods	m	45
Exports of goods and services (chain–linked volumes)	q	65
IMP: Imports of goods and services		
Monthly Import volume index	m	42
Monthly production index for intermediate goods	m	45
Imports of goods and services (chain–linked volumes)	q	65

6.2 The EuroMInd-$\mathcal{D}$ Density Estimates

We illustrate the monthly GDP density estimates, arising from the methodology outlined in the previous sections, conditional on the complete dataset available at the time of writing (mid July 2014). Our density estimates are based on $M = 5,000$ draws from the conditional distribution of total GDP, obtained from pooling the draws of the components, according to the procedure outlined in Section 4.

The EuroMInd indicator by Frale et al. (2011) combines the two point estimates of monthly GDP obtained from the output and the expenditure approaches using weights that are proportional to the average estimation error variance, which is computed from the relevant elements of the matrices $\mathbf{P}_{t|n}$ in (15), modified according to the multistep procedure in the Appendix. In the combined EuroMInd estimate, the GDP estimate arising from the output approach receives a weight of about 75%. The weights used for EuroMind could be adopted to combine the density forecasts as well, so that EuroMInd-$\mathcal{D}$ = 0.75 × EuroMInd-$\mathcal{D}_o$ + 0.25 × EuroMInd-$\mathcal{D}_e$.

Figure 1 presents the density estimates EuroMInd-$\mathcal{D}$ as a fan chart, with shaded regions corresponding to 50%, 70% and 95% highest probability density regions, respectively; the solid line is the median of the distribution. The spread of the densities has a periodic feature, as it is smaller at the times at which the quarterly estimate of GDP becomes available. The end of sample estimates, referring to second quarter of 2014, reflects the additional uncertainty arising from the unavailability of the GDP components for that quarter, although they incorporate the flash estimate of GDP. The distribution of monthly GDP has a tendency to be leptokurtic and positive skewness is sometimes observed. The Jarque–Bera normality test conducted on the draws rejects the null in 18% of the cases. However, no particular clusters of volatility are observed.

Figure 2 displays the density estimates of the annual growth rates from the output approach (fan chart 2a), the expenditure approach (fan chart 2b), and the combined estimates (fan chart 2c). The density estimates of the annual growth rates from the output approach are generally more concentrated.

We can assess the relative contribution of parameter uncertainty by comparing the standard errors of the density estimates with the estimation standard error of EuroMInd. The ratio of the two variability measures is around 1.05 for the sector C–D–E and Import and Export, so that parameter uncertainty is not a major source of variation for this source; for sectors A–B, F, G-H–I, J–K, Taxes less subsidies, consumption (FCE) and investments (GCF), the average ratio ranges from 122 to 136 and finally reaches 169 for sector L–P.

The next session aims at assessing the predictive accuracy of the EuroMInd-$\mathcal{D}$ density estimate.

6.3 Pseudo real–time experiment

6.3.1 Design of the experiment

As it has been mentioned in Section 5, evaluation of density forecasts is solely based on the density forecast – observation pairs. Therefore, calibration and sharpness are judged

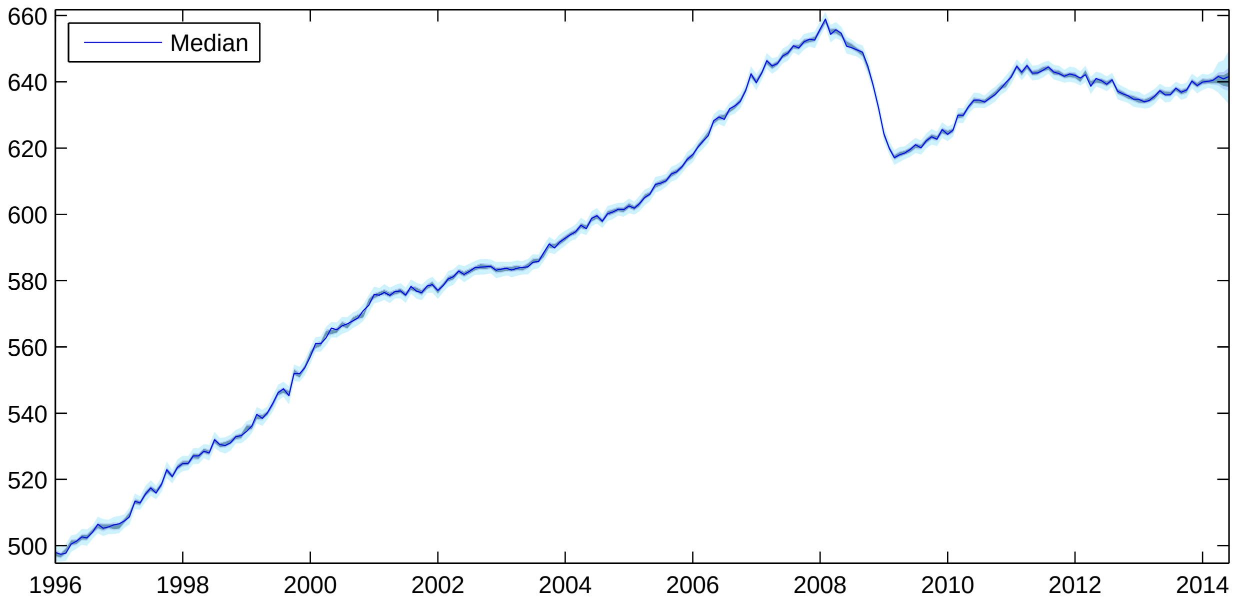
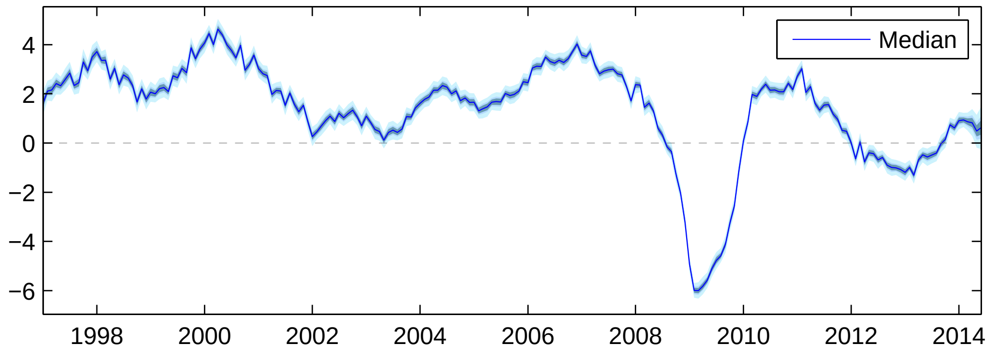


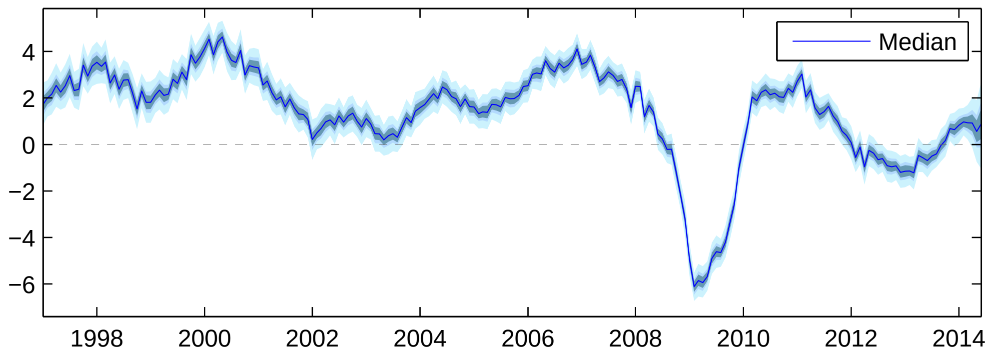
Figure 1: EuroMInd-$\mathcal{D}$. Shaded regions correspond to 50%, 70% and 95% probability bands, respectively.

by confronting the density forecasts of quarterly GDP with the observed quarterly GDP values.

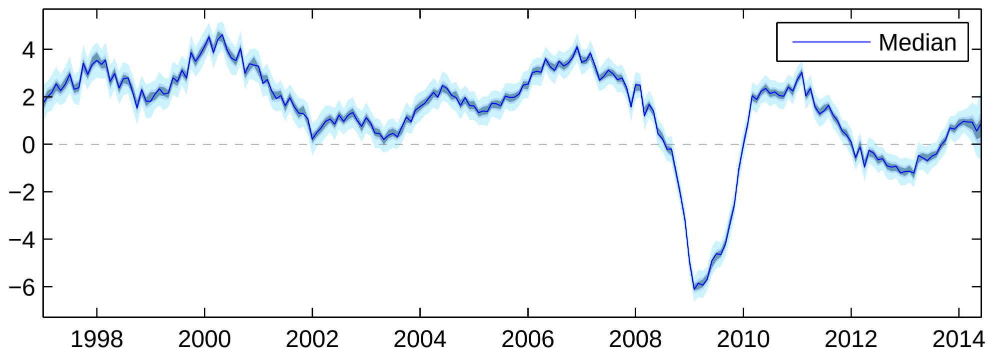
For the forecasting exercise, we divide the whole sample into the training sample from 1995.Q1 to 2000.Q4 and the evaluation sample from 2001.Q1 to 2014.Q1. This exercise is considered as a pseudo real–time experiment as we do not dispose of a real–time database. Hence, we cannot assess the impact of data revision on the reliability of the density estimates. Nevertheless, we can exploit the within–quarter information and take into account publication delays while forecasting GDP. In particular, we focus on three GDP density estimates for quarter τ, that we make available respectively 3, 2, and 1 months in advance with respect to the published GDP figure compiled by Eurostat. The latter becomes available 65 days after the closing of the reference quarter τ, i.e. at time (in months) $3\tau + 2.5$. Our estimates are computed respectively after the closing of quarter $\tau - 1$ at times $3(\tau - 1) + 1.5, 3(\tau - 1) + 2.5$ and $3(\tau - 1) + 3.5$. In providing density forecasts for quarter τ, we thus distinguish between three information sets. The first density estimate is conditional on the information until $3(\tau - 1)$ and the monthly indicators available in $3(\tau - 1) + 1.5$. The second information set additionally includes the monthly indicators in $3(\tau - 1) + 2.5$ and GDP for quarter $\tau - 1$. The last density forecasts is conditioned on the third information up to and including period $3(\tau-1)+3.5$. Depending on a particular information set, forecast samples are first obtained for every month of the considered quarter τ and are subsequently aggregated to quarterly density forecasts. Density forecasts for every GDP component are based on $M = 1,000$ draws from the respective conditional distribution. Aggregation of component densities to total GDP from either the output or the expenditure side is achieved, as in the EuroMind-$\mathcal{D}$



(a) Output approach



(b) Expenditure approach



(c) EuroMind

Figure 2: EuroMInd-$\mathcal{D}$ in growth rates. Shaded regions correspond to 50%, 70% and 95% probability bands, respectively.

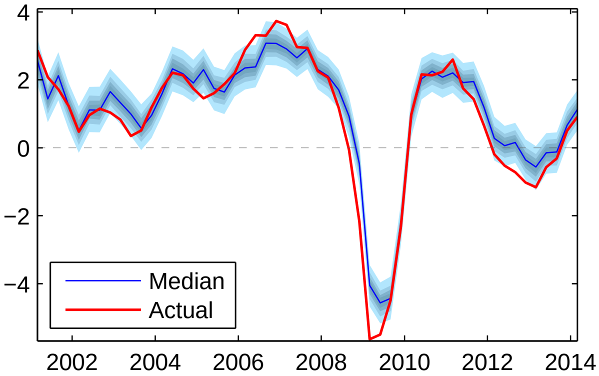
case, with the procedure outlined in Section 4.3.

6.3.2 Findings

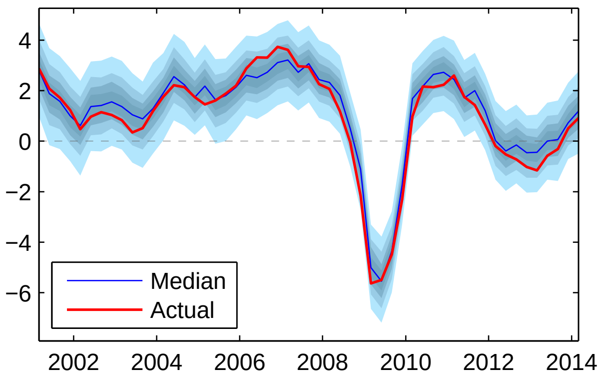
Figure 3 depicts the quarterly density forecasts of annual GDP growth rates, obtained with the output and the expenditure approach. It is apparent that, for each information set, probability bands associated with the expenditure approach are wider than those related to the output approach. This indicates higher uncertainty associated with the expenditure approach.

It can be observed that for both approaches uncertainty is the highest if the forecast densities for the quarterly growth rates are evaluated at the beginning of every quarter. As time progresses during a quarter and more information emerges, uncertainty diminishes considerably. The improvement is especially visible for the second information set, i.e. 2.5 months after the previous quarter closing. At this time point, the previous quarter GDP value is released, thereby providing the most important piece of information to forecast GDP in the actual quarter.

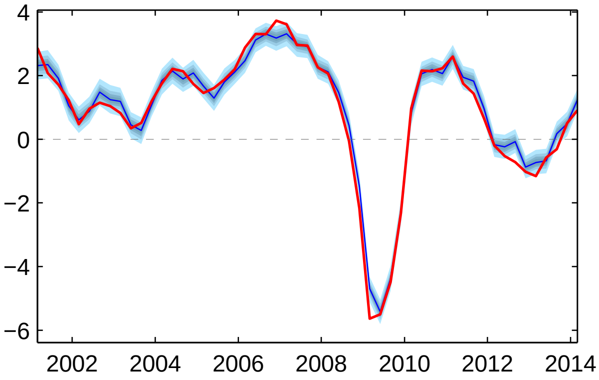
It is worth noting that, despite the high uncertainty in the case of the expenditure approach, the point prediction represented by the median of the distribution is almost in line with the realized GDP growth rates. The output approach, on the other hand, yields predictions more strongly diverging from the outcome, at least as far as the first information set is concerned. Figure 3a makes clear that the median heavily overestimates the downturn in 2009. Even though the median meets the observed outcome to a larger degree as soon as more information becomes available, the trough in 2009 remains outside the considered intervals.



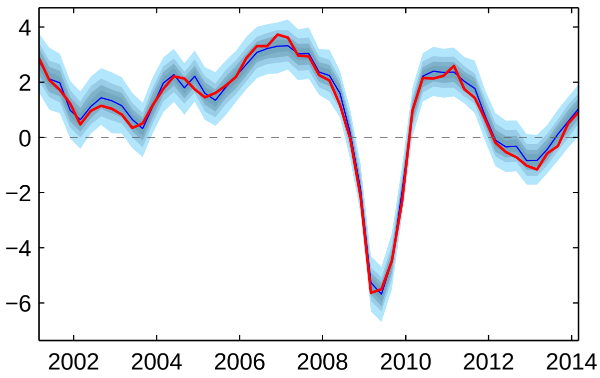
(a) Output approach: information set 1



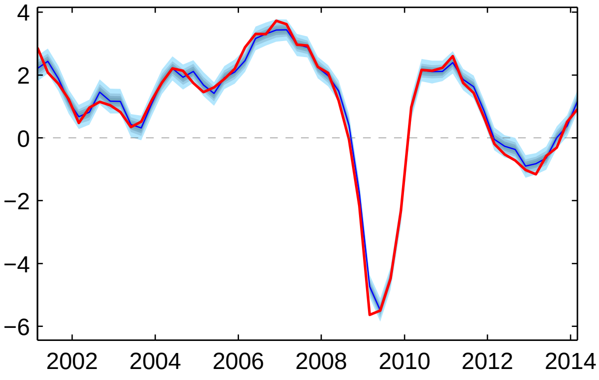
(b) Expenditure approach: information set 1



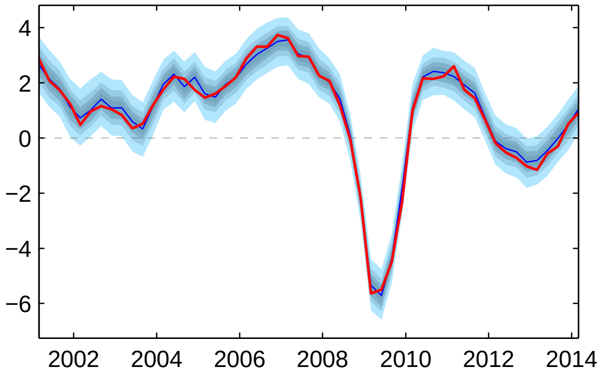
(c) Output approach: information set 2



(d) Expenditure approach: information set 2



(e) Output approach: information set 3

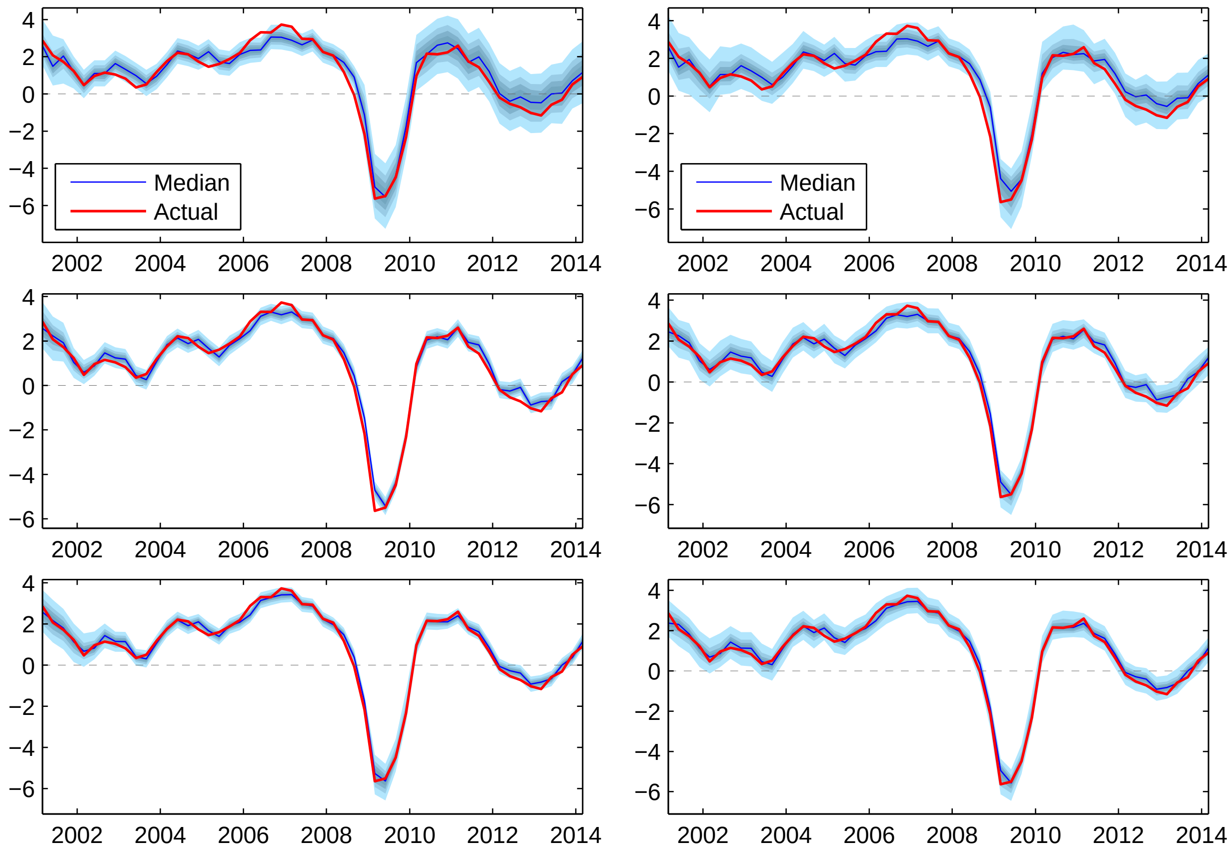


(f) Expenditure approach: information set 3

Figure 3: Real–time quarterly density forecasts of the euro area annual GDP growth for the evaluation period 2001.Q1 – 2014.Q1. Forecast densities related to the output and expenditure (columns) approach are evaluated with three different types of information sets (rows). The shaded areas represent 30%, 50%, 70% and 90% probability bands, respectively.

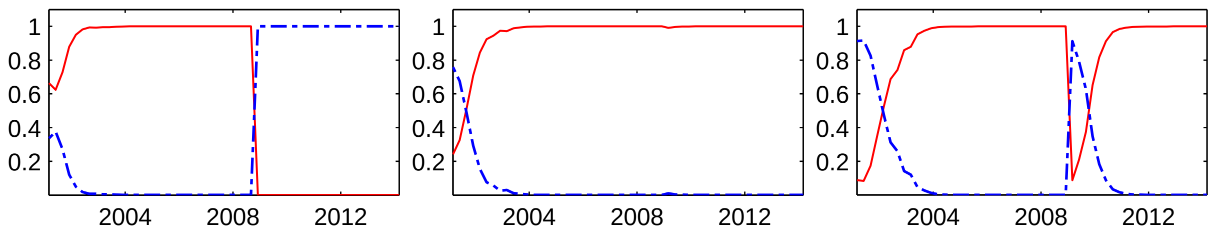
The figures in Appendix C, illustrating density forecasts for the different GDP components, give an intuition in what way individual components can explain the overall predictive accuracy of the output and the expenditure approach.[2] While the growth rates of value added in sectors C–D–E (Figure C.1) are predicted with high accuracy, with the median nearly coincident with the observations, probability bands related to other sectors are wider. Interestingly, the results for sectors G–H–I (Figure C.2) resemble the overall performance of the forecasts of the total value added. Failure in the correct prediction of the trough for sectors G–H–I may be thus responsible for the missed trough in the GDP growth from the output approach. As for the expenditure approach, high uncertainty of the consumption forecasts (Figure C.3) to a large extent contributes to the total GDP forecasts uncertainty. In contrast, export predictions (Figure C.4) are very precise, particularly those based on the second or third information set.

[2]Results for the sectors A–B and L–P, and for Taxes less Subsidies are not presented as these components play a relatively small role for the total GDP.

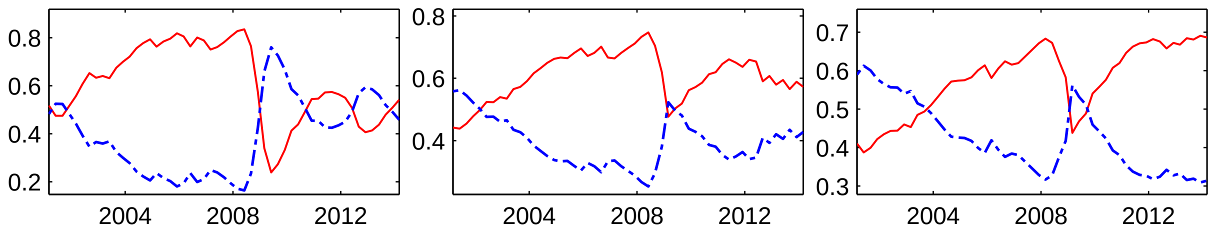


(a) Combined (log score): information set 1, 2, 3, respectively (rows)

(b) Combined (CRPS): information set 1, 2, 3, respectively (rows)



(c) Recursive log score weights: information set 1, 2, 3, respectively (columns); solid line: output approach, dash–dot line: expenditure approach



(d) Recursive CRPS weights: information set 1, 2, 3, respectively (columns); solid line: output approach, dash–dot line: expenditure approach

Figure 4: Combined real–time quarterly density forecasts of the Euro area annual GDP growth rates 2001.Q1 – 2014.Q1. (a),(b): the shaded areas represent 30%, 50%, 70% and 90% probability bands, respectively.

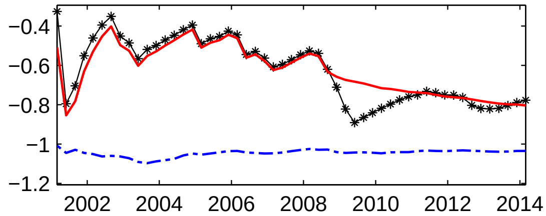
Figures 4a and 4b present the forecast densities combined with weights based on the log score and CRPS weights, respectively. Both combinations seem to provide an improvement compared to the individual density forecasts. First, uncertainty decreases relative to the case of the expenditure approach.[3] Second, the median of both combinations is more in accordance with the observed GDP growth than the median related to the output approach, especially if the forecasts are conditional on the first information set.

Although both weighing procedures yield a similar outcome in these respects, the combination with CRPS weights exhibits a slightly greater dispersion. The similarities and differences of both combinations reflect the evolution and the distribution of weights across the two approaches.

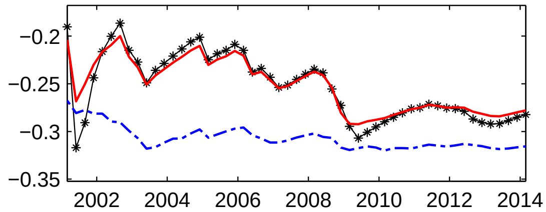
The recursive log score and CRPS weights are displayed in Figures 4c and 4d. It is interesting to note that, for a respective information set, both weights display some similarities in their evolution over time. At the beginning of the evaluation period, relatively more weight is assigned to the expenditure approach, which is particularly conspicuous in the case of the second and the third information set. Then, the output approach starts to play the dominating role which explains the overall higher concentration of the aggregated densities. For the first information set, a switch in the relative importance of two approaches occurs in 2009. This translates to more accuracy in the point prediction of the combinations after the outbreak of the economic and financial crisis. The switch is not present, if the second information set is used for forecasting. If the forecast densities are issued with the complete information available in every quarter, the change towards higher relative importance of the expenditure approach in 2009 is of a temporary nature only. From the comparison of the distribution of the two considered recursive weight types, it is evident that the log score weights are more restrictive. More strongly balanced CRPS weights express greater tolerance for mistakes of the CRPS measure, a property alluded to in Section 5.1.[4]

[3] In general, variance of the PITs = 0.083 (neutrally dispersed distribution), $>$ 0.083 (underdispersed distribution), $<$ 0.083 (overdispersed distribution). Variance of the PITs for the expenditure side (respectively for the three information sets): 0.021, 0.02, 0.016; output side: 0.091, 0.092, 0.083; combination (log scores): 0.059, 0.084, 0.069; combination (CRPS): 0.055, 0.053, 0.043

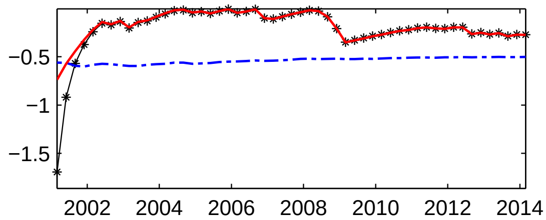
[4] Additionally, we also consider equal weights and weights based on the MSE. However, it turns out that the latter come down to equal weights. Since equal weighing produces dispersed density combinations showing poor performance in evaluation tests, we do not present the results in this article. They can be made available upon request.



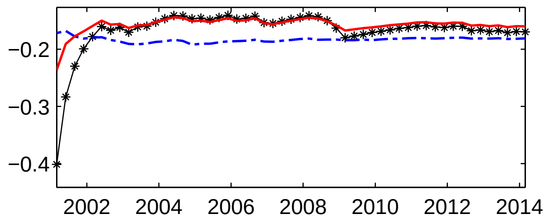
(a) Average log scores: information set 1



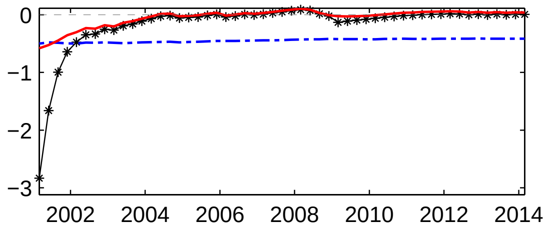
(b) Average CRPS: information set 1



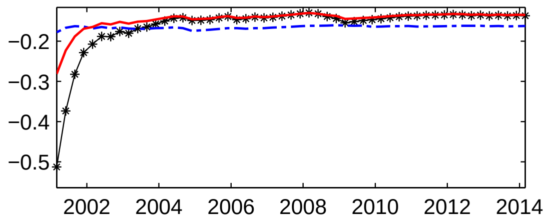
(c) Average log scores: information set 2



(d) Average CRPS: information set 2



(e) Average log scores: information set 3



(f) Average CRPS: information set 3

Figure 5: Average log scores and average CRPS for the real–time quarterly density forecasts of the Euro area annual GDP growth rates. Line with markers: output approach; dash–dot line: expenditure approach; solid line: combination with log score weights (left) and CRPS weights (right)

Figure 5, displaying the average log scores and CRPS, gives additional picture of the forecasting performance of the output and the expenditure approach as well as of the combinations. The averages are computed at every time point, which allows for tracking down the overall forecasting performance within the evaluation interval. Several observations emerge from examining the figures. Precision (as measured in scores) corresponding to the output approach (line with markers) is almost always higher than for the expenditure approach (dash–dot line). At the beginning of the considered time span, however, the output approach proves to be clearly inferior as regards the second and the third information set. This may reflect the fact that the median misses the realized GDP growth rate in the first quarters (see Figures 3c and 3e). Another observation is that the distance between the average score values corresponding to both approaches diminishes as more information becomes available. In such a situation, both approaches can be thus considered as equivalent alternatives for forecasting GDP, especially according to the CRPS. Moreover, precision of the combined densities (solid line) accords with the precision associated with the output approach (line with markers). Only at times when expenditure approach gains in importance, combinations outperform the output approach.

To sum up, combined density forecasts provide a quality improvement measured in scoring rules relative to those related to the output and the expenditure approach. Combinations absorb good quality of both individual densities – good predictive power of the expenditure approach and lower uncertainty of the output approach. Aggregation also implies a trade–off between adopting good and bad features. The outcome depends on the degree in which the bad ones are penalized. The CRPS–based combinations display higher dispersion than those obtained with the log score weights.

Next, we discuss results of evaluation tests for forecast densities of GDP components as well as both combinations. Tests described in Section 5.1 are applied to PITs and transformed PITs (z–scores) computed for each of the three information sets. All findings are reported in Table 2. As regards value added at different sector levels, uniformity tests (KS test, CvM test, AD test, χ^2 test) cannot reject the null of $U(0,1)$ distribution of PITs. For most of the sectors, Ljung–Box Q(4) statistic is insignificant once more information is used for forecasting. Similar observation can be made for the Berkowitz test which involves testing on zero–autocorrelation of the transformed PITs. Nonuniformity and autocorrelation problems arise in the case of the component Taxes less Subsidies (TlS). Misspecification related to the individual GDP components is passed to the total GDP, thereby explaining a relatively poor calibration for the first information set. Disregarding the first information set, however, density forecasts of GDP from the output side can be seen as well calibrated.

Less satisfactory results are obtained for GDP from the expenditure approach. Unlike in the output approach, no improvement in calibration can be ascertained after the first information set. This is due to the overall bad specification of predictive densities for consumption (FCE) in conjunction with the fact that consumption expenditure makes the largest contribution to the total expenditure compared to other expenditure components. Densities of investments (GCF), imports (IMP) and exports (EXP) are, on the other hand, well specified. Only in some cases specification suffers from autocorrelation of PITs which is reflected by a significant Q(4) statistic of the Ljung–Box test and LR statistic of the Berkowitz test.

As regards combined forecast densities, both weighing methods based on scoring rules lead to similar test results. Since in both cases more weight is placed on the densities from the output approach, it is not surprising that the findings for the combinations resemble findings related to GDP from the output approach. Whereas distributional assumption for the PITs seems to be satisfied, at least for the second and third information set, independence of the PITs cannot be established.[5]

[5]Note that normality of transformed PITs tested using the BS statistic can never be rejected. Therefore, this assumption seems can be seen as the weakest requirement for density forecasts to be calibrated.

Table 2: Probability integral transform (PIT) tests for evaluation real–time density forecasts of quarterly GDP and its components

Component	Information set[a)]	Tests on PITs[b)] KS	CvM	AD	χ^2	Q(4)	Tests on transformed PITs[c)] BS	Berkowitz
A–B	1	1.012	0.307	1.365	8.132	15.109*	0.107	15.163*
	2	1.147	0.395	1.902	10.849	3.934	1.197	1.330
	3	1.086	0.371	1.791	9.642	4.122	1.145	1.272
C–D–E	1	0.983	0.317	1.379	8.132	9.283	0.156	9.023*
	2	0.807	0.221	0.999	3.906	3.420	0.449	2.671
	3	1.026	0.420	2.032	19.000*	9.967*	0.819	6.986
F	1	1.257	0.363	1.873	14.170*	19.338*	0.243	15.099*
	2	1.054	0.248	1.495	7.528	16.805*	0.566	17.728*
	3	0.774	0.187	0.948	5.415	4.926	0.071	2.232
G–H–I	1	1.330	0.276	1.727	13.566	19.172*	0.920	29.494*
	2	0.716	0.163	0.810	7.226	6.898	0.176	4.341
	3	0.662	0.153	0.682	3.302	3.726	0.391	4.972
J–K	1	0.997	0.273	1.374	7.226	26.705*	0.018	25.838*
	2	0.657	0.167	0.716	6.321	5.209	0.008	1.620
	3	0.691	0.172	0.712	5.415	4.604	0.007	1.512
L–P	1	0.890	0.259	1.076	4.811	5.512	0.047	8.547*
	2	0.522	0.122	0.358	4.208	6.255	0.151	6.348
	3	0.618	0.131	0.453	2.396	6.491	0.264	6.859
TlS	1	1.326	0.596*	2.886*	15.981*	8.412	0.001	15.456*
	2	1.063	0.342	1.422	10.849	17.535*	0.046	16.886*
	3	0.873	0.265	1.104	7.226	19.538*	0.041	17.431*
GDP: output approach	1	1.411*	0.731*	6.009*	22.924*	43.768*	4.165	92.475*
	2	1.091	0.400	2.524*	9.943	12.763*	0.033	17.429*
	3	0.897	0.299	1.853	11.151	8.102	3.628	15.272*
FCE	1	2.202*	1.919*	9.430*	18.698*	44.753*	0.030	35.184*
	2	2.241*	1.421*	6.067*	18.698*	15.641*	0.035	11.502*
	3	2.181*	1.413*	6.067*	16.887*	13.051*	0.016	11.597*
GCF	1	1.160	0.396	1.986	10.849	8.545	0.032	12.599*
	2	1.226	0.384	1.380	8.736	15.093*	0.014	15.088*
	3	1.225	0.336	1.123	10.245*	14.786*	0.009	15.102*
IMP	1	1.177	0.453	2.063	4.811	17.669*	0.057	13.611*
	2	1.097	0.331	1.345	5.415	3.039	0.044	3.074
	3	1.030	0.253	0.966	7.528	4.131	0.137	3.485
EXP	1	0.553	0.106	0.266	2.698	32.707*	0.206	15.669*
	2	0.880	0.185	0.764	5.415	6.675	0.034	3.282
	3	0.719	0.161	0.598	11.453	10.139*	0.006	8.763
GDP: exp. approach	1	2.390*	1.602*	7.986*	50.094*	42.405*	0.020	82.544*
	2	1.943*	1.485*	7.641*	50.698*	16.428*	0.044	61.766*
	3	2.061*	1.668*	8.486*	61.566*	8.881	0.009	69.108*
GDP: combined (log score)	1	1.645*	0.650*	2.633*	18.094*	31.999*	0.098	27.385*
	2	0.971	0.407	2.458	9.943	16.145*	0.020	15.590*
	3	1.116	0.303	1.437	7.528	10.085*	0.014	10.457*
GDP: combined (CRPS)	1	1.410*	0.733*	5.466*	20.811*	46.246*	1.020	73.607*
	2	0.967	0.400	2.504*	8.736	15.148*	0.063	17.383*
	3	1.117	0.306	1.675	6.925	10.204*	0.594	9.059*

[a)] Real–time forecasts of quarterly values of the respective GDP component in a particular quarter take into account publication lags of the indicators and are based on the information set growing with the progress of the quarter; 1: information encompassing all indicators available until the end of the previous quarter and monthly indicators available after the first month of the considered quarter; 2: information additionally including monthly indicators available after the second month of the considered quarter, as well as the flash estimate of the quarterly component; 3: information including all indicators of the considered quarter

[b)] KS: Kolmogorov–Smirnov statistic; CvM: Cramér–von–Mises statistic; AD: Anderson–Darling statistic; χ^2: statistic of the Pearson chi–squared test proposed by Wallis (2003); Q(p): Ljung–Box statistic based on the first p standardized innovations; For KS, CvM and AD tests we use critical values provided by Stephens (1974); * indicates statistical significance at the 5% level.

[b)] BS: Bowman–Shenton normality statistic; Berkowitz: statistic of a test proposed by Berkowitz (2001); * indicates statistical significance at the 5% level.

7 Conclusions

The paper has presented EuroMInd-$\mathcal{D}$, a density estimate of monthly GDP in chain-linked volumes based on the pooling of the density estimates for 11 GDP components. The density predictions and nowcasts appear to be well calibrated, when they are conditional on an information set that includes at least the release of the quarterly national accounts for the previous quarter. Moreover, the sharpness of the probabilistic estimates renders EuroMInd-$\mathcal{D}$ a useful tool for the assessment of macroeconomic conditions in the euro area. While the current paper concentrated on the evaluation of its predictive accuracy, we think that EuroMInd-$\mathcal{D}$ can serve well the purpose of characterising the business cycle via the probabilistic detection of turning points and the decomposition of output into trends and cycles.

The density estimates reflect parameter and filtering uncertainty. They do not incorporate model uncertainty, as the specification of the model is taken as given and capitalises upon the indicator selection and specification search (number of factors and their lags, autoregressive orders) performed in Frale et al. (2011) Further research should be directed towards the incorporation of business survey variables, often referred to as soft indicators, in the information set. Our preliminary experimentation, based on the two factors specification considered in Frale et al. (2010), led to reject their inclusion on the grounds of the lack of sharpness and overdispersion of the density estimates, due to the contribution of parameter uncertainty.

A The aggregation of the GDP components

This appendix provides a detailed illustration of the three steps of the multistep procedure leading to the monthly GDP estimates. Let us index the month of the year by $m, m = 1, \dots, 12$ and the year by j, $j = 0, \dots, J = [n/12] - 1$, so that the time index is written $t = m + 12j, t = 1, \dots, n$. For a particular estimated monthly GDP component (e.g. value added of the industry sector) let us denote by y_{mj} the value at current prices of month m in year j, and by $\overline{y}_j = \sum_{m=1}^{12} y_{mj}/12$ the annual average (the annual and quarterly figures are available from the national accounts, compiled by Eurostat). The chain–linked volume estimate with reference year s (the year 2000 in our case) will be denoted $\hat{y}_{mj}^{(s)}$. The following multistep procedure enables the computation of volume measures (expressed at the prices of the previous year) that are additive, also cross–sectionally.

1. *Dechaining*

   (a) Transform the monthly estimates into Laspayres type quantity indices with reference year $j-1$ (the previous year), by computing

   $$I_{mj}^{(j-1)} = \frac{\hat{y}_{mj}^{(s)}}{\overline{y}_{j-1}^{(s)}}, \qquad m = 1, \dots, 12, \quad j = 1, \dots, J,$$

   where

   $$\overline{y}_{j-1}^{(s)} = \frac{1}{12} \sum_{m=1}^{12} \hat{y}_{m,j-1}^{(s)}, \qquad j = 1, \dots, J$$

   (b) Compute the series at the average prices of the previous year as:

   $$\hat{y}_{mj}^{(j-1)} = I_{mj}^{(j-1)} \overline{y}_{j-1}, \qquad m = 1, \dots, 12, \quad j = 1, \dots, J,$$

   where it should be recalled that $\overline{y}_{j-1}$ is the annual average at current prices; the annual and quarterly totals are available from the national accounts compiled by Eurostat.

2. *Aggregation step*
   Let $\mathbf{y}_{mj}^{(j-1)}$ denote the vector containing the eleven disaggregate GDP components expressed at the average prices of the previous year. Using the original estimates and the dechaining procedure we can assume that, at least approximately,

   $$\mathbf{y}_{mj}^{(j-1)} \sim \mathrm{N}\left(\hat{\mathbf{y}_{mj}}^{(j-1)}, \hat{\mathbf{V}_{mj}}^{(j-1)}\right),$$

   where the first and second moments are given by the sequential constrained estimates produced by the Kalman filter and smoother outlined in the main text, modified to take into account the dechaining procedure.

   Letting $\mathbf{a}_o$ denote an 11×1 aggregation vector, so that $\mathbf{a}_o'\mathbf{y}_{mj}^{(j-1)} = \tilde{Y}_{mj,o}^{(j-1)}$ is GDP at market prices for the output approach, and $\mathbf{a}_e$ the 11×1 vector, so that $\mathbf{a}_e'\mathbf{y}_{mj}^{(j-1)} = \tilde{Y}_{mj,e}^{(j-1)}$ is GDP at market prices for the expenditure approach, then we have, e.g. $\tilde{y}_{mj,o}^{(j-1)} \sim \mathrm{N}\left(\mathbf{a}_o'\hat{\mathbf{y}}_{mj}^{(j-1)}, \mathbf{a}_o'\hat{\mathbf{V}}_{mj}^{(j-1)}\mathbf{a}_o'\right)$.

3. *Chain–linking* (annual overlap).

   Let $\tilde{Y}_{mj}^{(j-1)}$ denote the GDP estimate for month m, year j, at the average prices of year $j-1$ (for the sake of notation we omit the indices indicating the output or expenditure approach):

   (a) Convert the aggregated volume measures into Laspeyres–type quantity indices with respect to the previous year:

   $$\mathcal{I}_{mj}^{(j-1)} = \frac{\tilde{Y}_{mj}^{(j-1)}}{\overline{Y}_{j-1}}, \qquad m = 1, \ldots, 12, \quad j = 1, \ldots, J.$$

   where $\overline{Y}_{j-1} = \sum_m Y_{m,j-1}/12$ is the average of the previous year at current prices.

   (b) Chain–link the indices using the recursive formula (the first year, $j = 0$, is the reference year):

   $$\mathcal{I}_{mj}^{(0)} = \mathcal{I}_{mj}^{(j-1)} \overline{\mathcal{I}}_{j-1}^{(0)}, \qquad m = 1, \ldots, 12, \quad j = 1, \ldots, J,$$

   where $\overline{\mathcal{I}}_{0}^{(0)} = 1$ and $\overline{\mathcal{I}}_{j-1}^{(0)} = \sum_m \mathcal{I}_{m,j-1}^{(0)}/12$.

   (c) If $s > 0$ then change the reference year to year s:

   $$\mathcal{I}_{mj}^{(s)} = \frac{\mathcal{I}_{mj}^{(0)}}{\overline{\mathcal{I}}_{s}^{(0)}} \qquad m = 1, \ldots, 12, \quad j = 1, \ldots, J$$

   (d) Compute the chain–linked volume series with reference year s:

   $$\tilde{Y}_{mj}^{(s)} = \mathcal{I}_{mj}^{(s)} \overline{Y}_s \qquad m = 1, \ldots, 12, \quad j = 1, \ldots, J,$$

   where $\overline{Y}_s = \sum_m Y_{ms}/12$ is the annual average value of GDP (at basic or market prices) at current prices of the reference year.

The multistep procedure just described enables to obtain monthly estimates in volume such that the values $\tilde{Y}_{mj}^{(j-1)}$ expressed at the average prices of the previous year add up to their quarterly and annual totals published by Eurostat (and are consistent with the flash estimate, see below). On the contrary, as a result of the chaining procedure, the chain–linked volumes $\tilde{Y}_{mj}^{(s)}$ expressed at the prices of the common reference year s (the year 2000) are consistent only with the temporal aggregation constraints; however, their estimates are more reliable since they have been combined with the estimates of other related variables.

B Incorporating the flash estimate

The above framework is suitable also for the incorporation of the flash estimate of total GDP at market prices for the euro area. The flash estimate is released by Eurostat in

February, May, August and November, around 45 days after the end of the reference quarter.

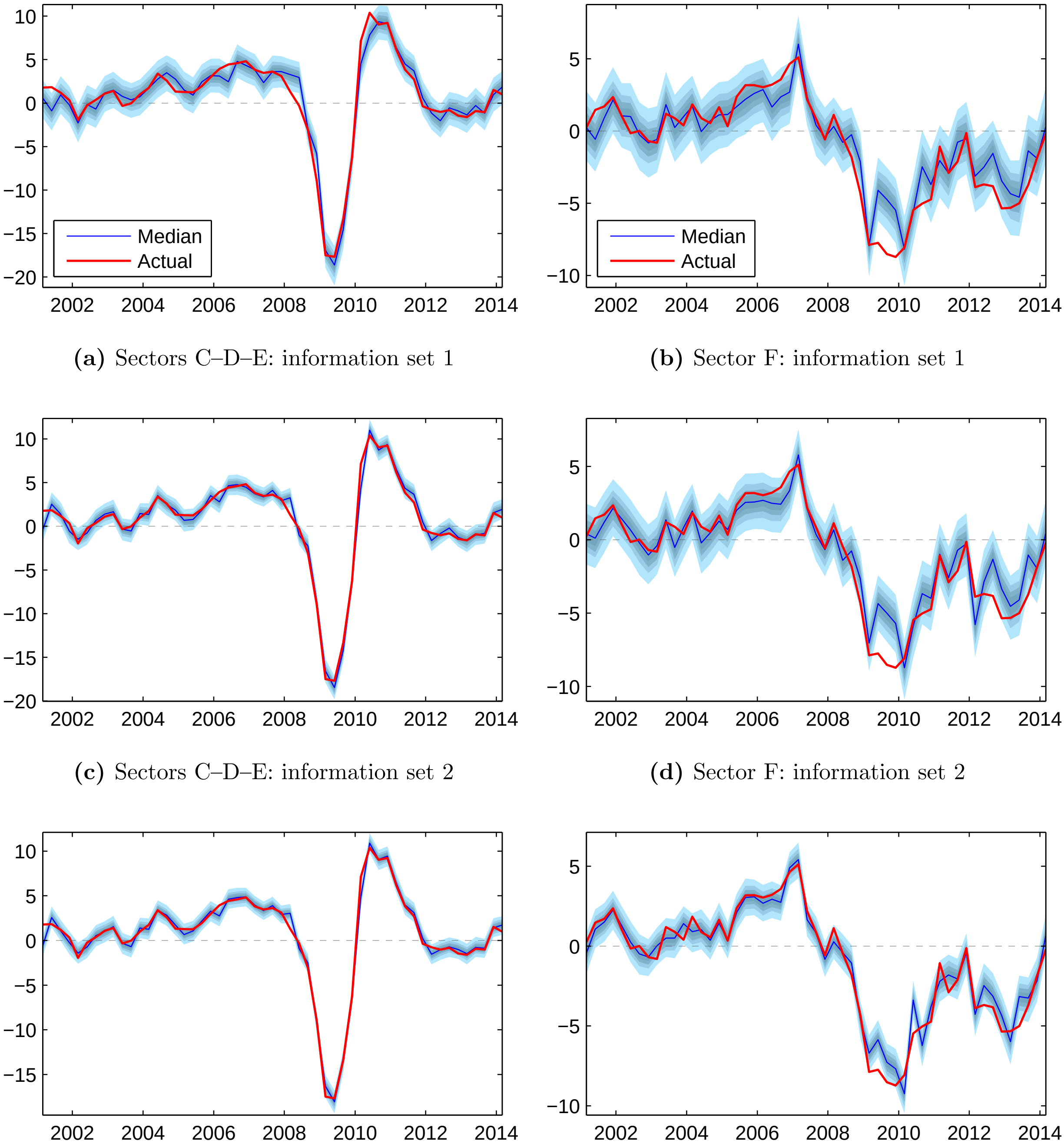
The target variable is the quarterly growth rate on the previous quarter of GDP for the euro area and the European Union, seasonally adjusted and working day corrected, in volume measures as published by Eurostat. Hence, the flash estimate pertains to aggregate GDP at market prices and it is not available for the breakdown of GDP according to the output and the expenditure approach.

The incorporation of the flash estimate of GDP can be carried out by customizing the second step of the above multistep procedure. Let $\mathbf{Q}$ denote the 2×11 matrix $\mathbf{Q} = [\mathbf{a}_o, \mathbf{a}_e]'$ and let $\mathbf{q}$ be the 2×1 vector containing the common value of the flash estimate at the time t for which it is available. The modified estimates of the GDP components that comply with the flash and their MSE matrix are given respectively by

$$
\begin{aligned}
\tilde{\mathbf{y}}_{mj}^{(j-1)} &= \widehat{\mathbf{Y}_{mj}}^{(j-1)} + \hat{\mathbf{V}}_{mj}^{(j-1)}\mathbf{Q}'(\mathbf{Q}\hat{\mathbf{V}}_{mj}^{(j-1)}\mathbf{Q}')^{-1}(\mathbf{q} - \mathbf{Q}\widehat{\mathbf{y}_{mj}}^{(j-1)}) \\
\tilde{\mathbf{V}}_{mj}^{(j-1)} &= \widehat{\mathbf{V}_{mj}}^{(j-1)} - \hat{\mathbf{V}}_{mj}^{(j-1)}\mathbf{Q}'(\mathbf{Q}\hat{\mathbf{V}}_{mj}^{(j-1)}\mathbf{Q}')^{-1}\mathbf{Q}\widehat{\mathbf{V}_{mj}}^{(j-1)}.
\end{aligned}
$$

The discrepancy between the GDP quarterly estimates and the aggregation of the monthly estimates arising from the output approach is distributed across the months using the covariance matrix of the estimates as detailed above. The new balanced estimates are now ready to be aggregated and expressed at the average prices of reference year s, using the chain linking step of the above procedure.

C Figures



(a) Sectors C–D–E: information set 1

(b) Sector F: information set 1

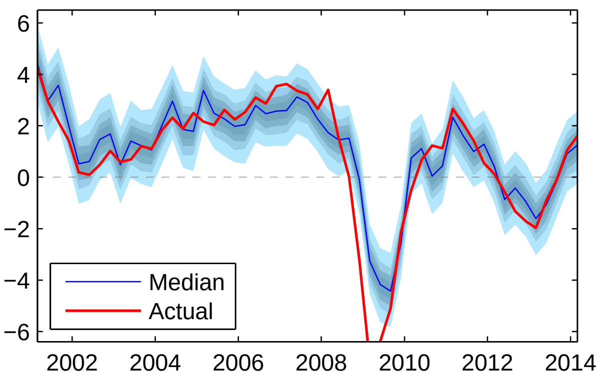
(c) Sectors C–D–E: information set 2

(d) Sector F: information set 2

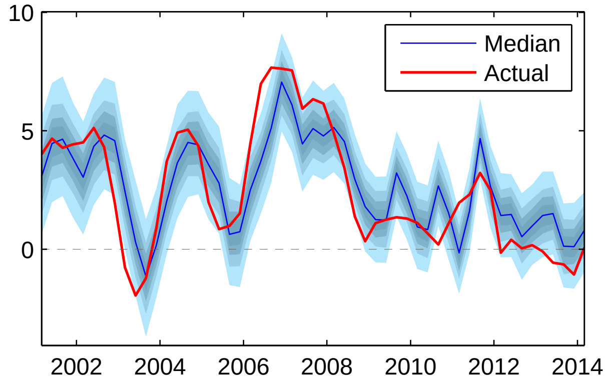
(e) Sectors C–D–E: information set 3

(f) Sector F: information set 3

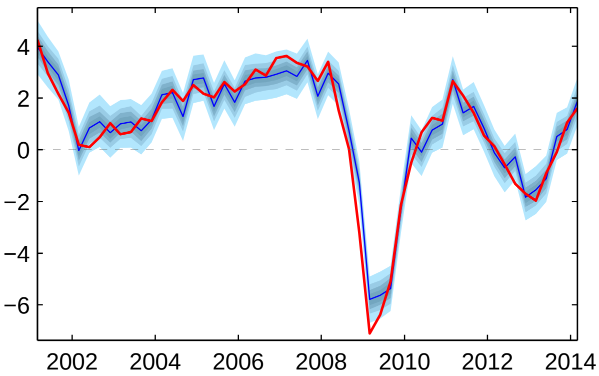
Figure C.1: Real–time quarterly density forecasts of the euro area annual value added growth in the sectors C–D–E and F for the evaluation period 2001.Q1 – 2014.Q1. Forecast densities related to the sectors C–D–E and F (columns) approach are evaluated with three different types of information sets (rows). The shaded areas represent 30%, 50%, 70% and 90% probability bands, respectively.



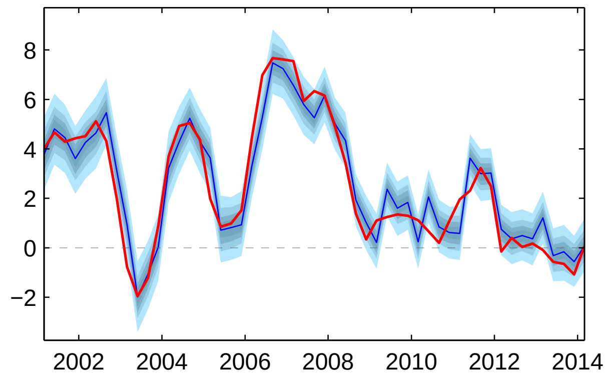
(a) Sector G–H–I: information set 1



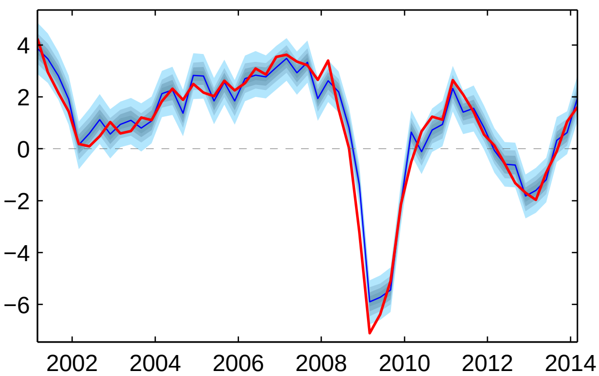
(b) Sector J–K: information set 1



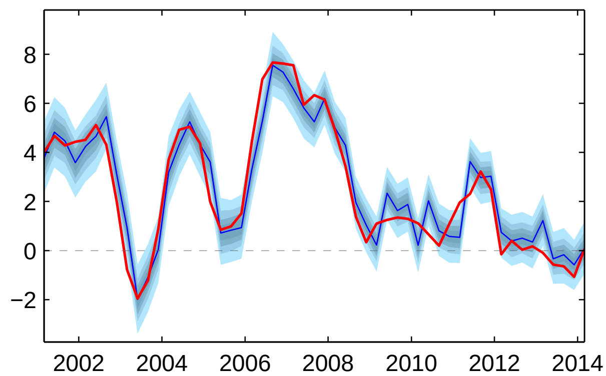
(c) Sector G–H–I: information set 2



(d) Sector J–K: information set 2

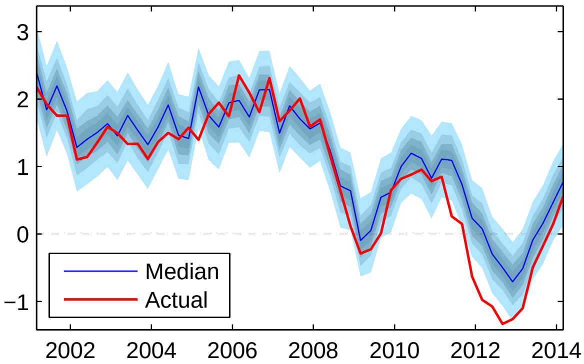


(e) Sector G–H–I: information set 3

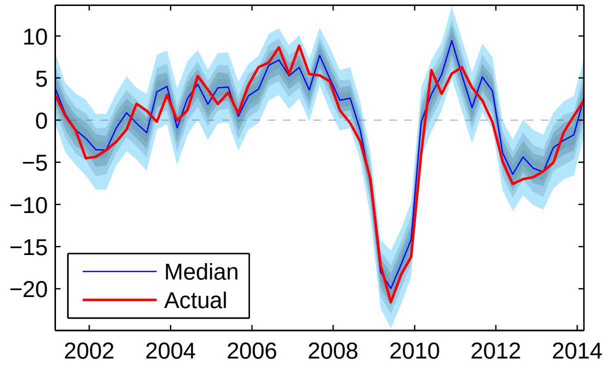


(f) Sector J–K: information set 3

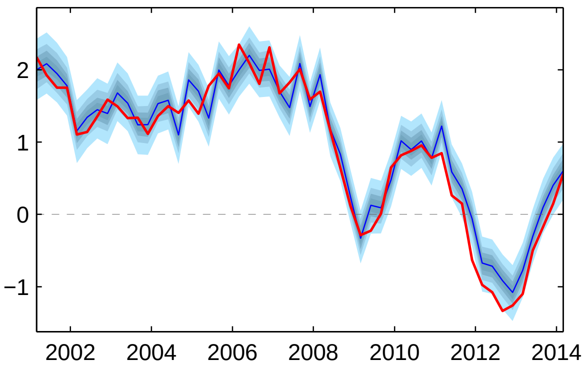
Figure C.2: Real–time quarterly density forecasts of the euro area annual value added growth in the sectors G–H–I and J–K for the evaluation period 2001.Q1 – 2014.Q1. Forecast densities related to the sectors G–H–I and J–K (columns) approach are evaluated with three different types of information sets (rows). The shaded areas represent 30%, 50%, 70% and 90% probability bands, respectively.



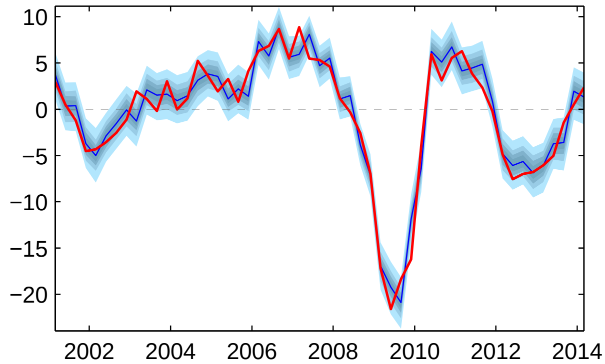
(a) Consumption: information set 1



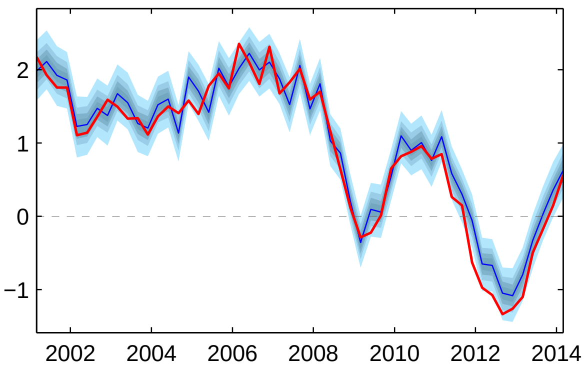
(b) Investment: information set 1



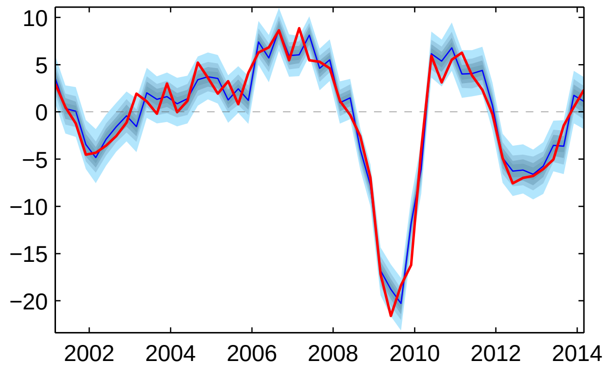
(c) Consumption: information set 2



(d) Investment: information set 2

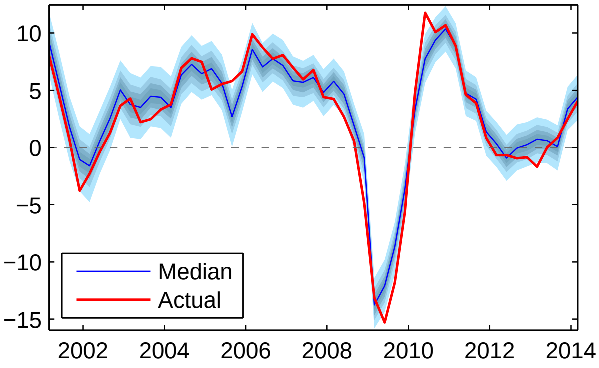


(e) Consumption: information set 3

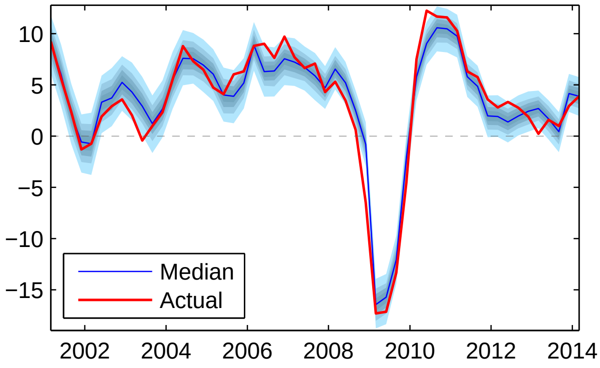


(f) Investment: information set 3

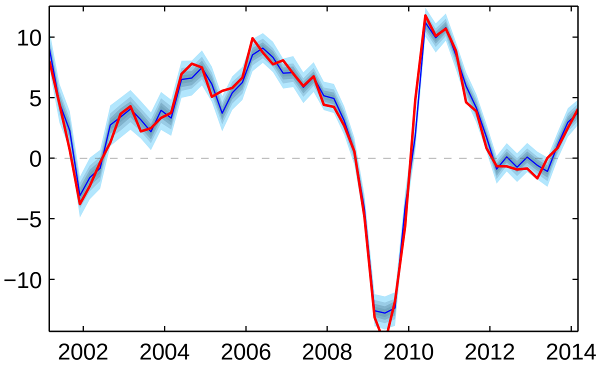
Figure C.3: Real–time quarterly density forecasts of the Euro area annual consumption and investment growth for the evaluation period 2001.Q1 – 2014.Q1. Forecast densities related to consumption and investment (columns) approach are evaluated with three different types of information sets (rows). The shaded areas represent 30%, 50%, 70% and 90% probability bands, respectively.



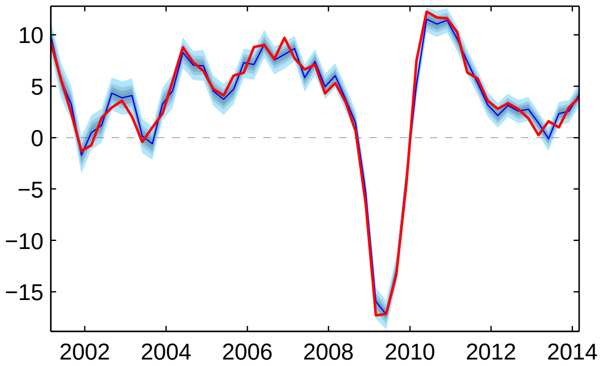
(a) Imports: information set 1



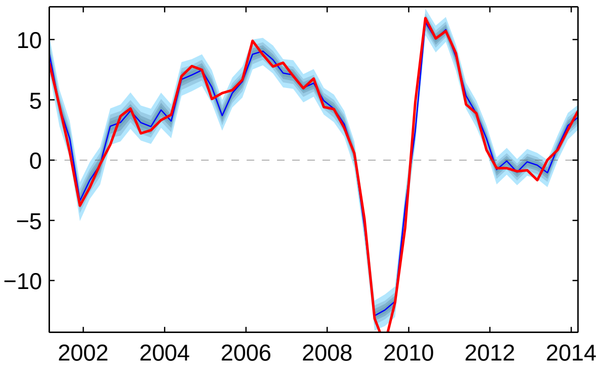
(b) Exports: information set 1



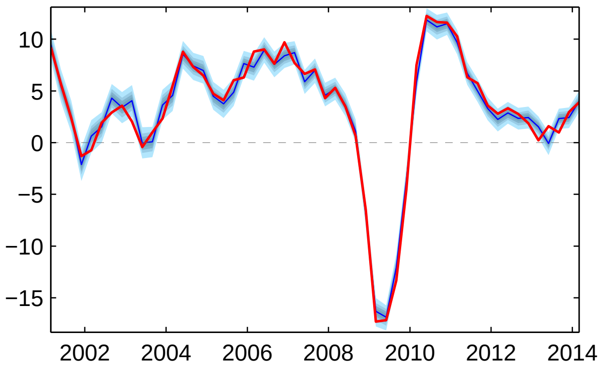
(c) Import: information set 2



(d) Exports: information set 2



(e) Imports: information set 3



(f) Exports: information set 3

Figure C.4: Real–time quarterly density forecasts of the euro area annual consumption and investment growth for the evaluation period 2001.Q1 – 2014.Q1. Forecast densities related to consumption and investment (columns) approach are evaluated with three different types of information sets (rows). The shaded areas represent 30%, 50%, 70% and 90% probability bands, respectively.

References

Aastveit, K. A., Gerdrup, K. R., Jore, A. S., and Thorsrud, L. A. (2014). Nowcasting GDP in Real Time: A Density Combination Approach. *Journal of Business & Economic Statistics*, *32*(1), 48–68.

Altissimo, F., Cristadoro, R., Forni, M., and Veronese, G. (2010). New Eurocoin: Tracking Economic Growth in Real Time. *Review of Economics and Statistics*, *92*, 1024–1034.

Amisano, G., and Giacomini, R. (2007). Comparing Density Forecasts via Weighted Likelihood Ratio Tests. *Journal of Business & Economic Statistics*, *25*(2), 177–190.

Anderson, B. D. O., and Moore, J. B. (1979). *Optimal Filtering.* Englewood Cliffs: Prentice Hall.

Bacharach, J. (1974). *Bayesian Dialogues* (Unpublished Manuscript). Christ Church College, Oxford University.

Bache, I. W., Jore, A. S., Mitchell, J., and Vahey, S. P. (2011). Combining VAR and DSGE Forecast Densities. *Journal of Economic Dynamic & Control*, *35*, 1659–1670.

Bates, J. M., and Granger, C. W. J. (1969). The Combination of Forecasts. *Operational Research Quarterly*, *20*(4), 451–468.

Berkowitz, J. (2001). Testing Density Forecasts, With Applications to Risk Management. *Journal of Business & Economic Statistics*, *19*(4), 465–474.

Bjørnland, H. C., Gerdrup, K. R., Jore, A. S., Smith, C., and Thorsrud, L. A. (2011). Weights and Pools for a Norwegian Density Combination. *North American Journal of Economics and Finance*, *22*, 61–76.

Bloem, A., Dippelsman, R. J., and Maehle, N. O. (2001). *Quarterly National Accounts Manual: Concepts, Data Sources, and Compilation.* Washington DC: International Monetary Fund.

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

Clemen, R. T. (1989). Combining Forecasts: A Review and Annotated Bibliography. *International Journal of Forecasting*, *5*, 559–583.

Dawid, A. P. (1984). Statistical Theory: The Prequential Approach. *Journal of the Royal Statistical Society: Series A*, *147*(2), 278–292.

de Jong, P. (1988). A Cross–Validation Filter for Time Series Models. *Biometrika*, *75*(3), 594–600.

de Jong, P. (1991). The Diffuse Kalman Filter. *The Annals of Statistics*, *19*, 1073–1083.

Diebold, F. X., Gunther, T. A., and Tsay, A. S. (1998). Evaluating Density Forecats wih Applications to Financial Risk Management. *International Economic Review*, *39*(4), 863–883.

Durbin, J., and Koopman, S. J. (2012). *Time Series Analysis by State Space Methods* (2nd ed.). Oxford: Oxford University Press.

Frale, C., Marcellino, M., Mazzi, G., and Proietti, T. (2010). Survey Data as Coicident or Leading Indicators. *Journal of Forecasting*, *29*, 109–131.

Frale, C., Marcellino, M., Mazzi, G., and Proietti, T. (2011). EUROMIND: A Monthly Indicator of the Euro Area Economic Conditions. *Journal of the Royal Statistical Society: Series A*, *174*, 439– 470.

Garratt, A., Mitchell, J., and Vahey, S. P. (2014). Measuring Output Gap Uncertainty. *International Journal of Forecasting*, *30*, 268–279.

Gneiting, T., Balabdaoui, F., and Raftery, A. E. (2007). Probabilistic Forecats, Calibration and Sharpness. *Journal of the Royal Statistical Society: Series B*, *69*, 243–268.

Gneiting, T., and Raftery, A. E. (2007). Strictly Proper Scoring Rules, Prediction, and Estimation. *Journal of the American Statistical Association*, *102*(477), 359–378.

Gneiting, T., and Ranjan, R. (2013). Combining Predictive Distributions. *Electronic Journal of Statistics*, *7*, 1747–1782.

Good, I. J. (1952). Rational Decisions. *Journal of the Royal Statistical Society: Series B*, *14*(1), 1107–114.

Groen, J. J. J., Paap, R., and Ravazzolo, F. (2013). Real–Time Inflation Forecasting in a Changing World. *Journal of Business & Economic Statistics*, *31*(1), 29–44.

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

Harvey, A. C. (1989). *Forecasting, Structural Time Series Models and the Kalman Filter.* Cambridge: Cambridge University Press.

Hendry, D. F., and Clements, M. P. (2004). Pooling of Forecasts. *Econometrics Journal*, *7*, 1–31.

Jore, A. S., Mitchell, J., and Vahey, S. P. (2010). Combining Forecast Densities from VARs with Uncertain Instabilities. *Journal of Applied Econometrics*, *25*, 621–634.

Kascha, C., and Ravazzolo, F. (2010). Combining Inflation Density Forecasts. *Journal of Forecasting*, *29*, 231–250.

Koopman, S. J., and Durbin, J. (2000). Fast Filtering and Smoothing for Multivariate State Space Models. *Journal of Time Series Analysis*, *21*, 281–296.

Mariano, R., and Murasawa, Y. (2003). A New Coincident Index of Business Cycles Based on Monthly and Quarterly Series. *Journal of Applied Econometrics*, *18*, 427–443.

Matheson, J. E., and Winkler, R. L. (1976). Scoring Rules for Continuous Probability Distributions. *Management Science*, *22*(10), 1087–1096.

Mazzi, G., Mitchell, J., and Montana, G. (2014). Density Nowcasts and Model Combination: Nowcasting Euro–Area GDP Growth over the 2008-09 Recession. *Oxford Bulletin of Economics and Statistics*, *76*(2), 233–356.

Mitchell, J., and Hall, S. G. (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*, 995–1033.

Noceti, P., Smith, J., and Hodges, S. (2003). An Evaluation of Tests of Distributional Forecasts. *Journal of Forecasting*, *22*, 447–455.

Panagiotelis, A., and Smith, M. (2008). Bayesian Density Forecasting of Intraday Electricity Prices Using Multivariate Skew t Distributions. *International Journal of Forecasting*, *24*, 710–727.

Ranjan, R., and Gneiting, T. (2010). Combining Probability Forecasts. *Journal of the Royal Statistical Society: Series B*, *72*, 71–91.

Ravazzolo, F., and Vahey, S. P. (2014). Forecast Densities for Economic Aggegates from Disaggregate Ensembles. *Studies in Nonlinear Dynamics & Econometrics*, *18*(4), 367–381.

Rosenblatt, M. (1952). Remarks on a Multivariate Transformation. *Annals of Mathematical Statistics*, *23*, 470–472.

Rossi, B., and Sekhposyan, T. (2013). Conditional Predictive Density Evaluation in the Presence of Instabilities. *Journal of Econometrics*, *177*, 199–212.

Selten, R. (1998). Axiomatic Characterization of the Quadratic Scoring Rule. *Experimental Economics*, *1*, 43–62.

Stephens, M. A. (1974). EDF Statistics for Goodness of Fit and Some Comparisons. *Journal of the American Statistical Association*, *69*(347), 730–737.

Stock, J. H., and Watson, M. W. (1991). A Probability Model of the Coincident Economic Indicators. In K. Lahiri and G. H. Moore (Eds.), *Leading Economic Indicators* (pp. 63–86). Cambridge University Press.

Stone, M. (1961). The Opinion Pool. *Annals of Mathematical Statistics*, *32*, 1339–1342.

Tanner, M. A. (1996). *Tools for Statistical Inference: Methods for the Exploration of Posterior Distributions and Likelihood Functions* (3rd ed.). New York: Springer.

Tay, A. S., and Wallis, K. F. (2000). Density Forecasting: A Survey. *Journal of Forecasting*, *19*, 235–254.

Wallis, K. F. (2003). Chi–squared Tests of Interval and Density Forecasts, and the Bank of England's Fan Charts. *International Journal of Forecasting*, *19*, 165–175.

Wallis, K. F. (2005). Combining Density and Interval Forecasts: A Modest Proposal. *Oxford Bulletin of Economics and Statistics*.

Winkler, R. L. (1968). The Consensus of Subjective Probability Distributions. *Management Science*, *15*(2), B61–B75.