Big Data for Official Statistic Competition
-
WBS University of Warwick Forecast Team

Danilo Cascaldi Garcia*, Kolja Johannsen, Ingomar Krohn,
Ana Beatriz Galvão, Anthony Garratt

December 20, 2016

1 Introduction

The members of Warwick Business School (WBS) University of Warwick Forecast Team are Doctoral Researchers Danilo Cascaldi Garcia, Kolja Johannsen and Ingomar Krohn under guidance of Professors Ana Beatriz Galvão and Anthony Garratt from WBS Economic Modelling and Forecasting group.

WBS Economic Modelling and Forecasting group's main research interests are the theory and practice of economic modelling and forecasting, with an applied macroeconomics emphasis.

As part of the Big Data for Official Statistic Competition, the WBS University of Warwick Forecast Team produces forecasts in two categories: HICP (All items) and HICP (All items excluding Energy), for a set of countries in the Euro Area and UK. Monthly forecasts are produced using (i) a Bayesian vector autoregressive model (ii) various conditional unobserved components stochastic volatility models (iii) a univariate autoregressive model and (iv) equal and log score weight model combinations of (i), (ii) and (iii).

In what follows we describe the modelling approach in detail, outline how forecasts are computed and provide an overview of the data sets used for the Big Data for Official Statistic Competition.

*Corresponding author: Danilo Cascaldi Garcia, Economic Modelling and Forecasting Group - Warwick Business School, University of Warwick, Coventry, CV4 7AL, Danilo.Garcia.14@mail.wbs.ac.uk

2 Model Description

This section describes the modelling and foreasting approach for all models: Bayesian vector autoregressive model (BVAR), autoregressive model (AR(p)) and conditional unobserved component stochastic volatility (UC-SV) models. Lastly, we present how individual forecasts are used in order to construct equal and logarithmic score weight model combinations.

Bayesian vector autoregressive model

The BVAR approach closely follows Carriero *et al.* (2015) and the literature cited within. For each country we define a set of variables $y_t = \{q_{1t}, q_{2t}, ..., q_{Nt}\}'$ such that the VAR(p) can be written as

$$\begin{aligned} y_t &= A_0 + A_1 y_{t-1} + ... + A_p y_{t-p} + u_t \\ u_t &\sim N(0, \Sigma) \end{aligned} \tag{1}$$

In line with Carriero *et al.*(2015) we use a Normal-Inverse/Wishart prior set up of the form

$$\begin{aligned} \alpha|\Sigma &\sim N(\alpha_0, \Sigma \otimes \Omega_0) \\ \Sigma &\sim IW(S_0, v_0) \end{aligned}$$

where $\alpha = vec([A_c, A_1, ..., A_p]')$ and posterior distributions of

$$\begin{aligned} \alpha|\Sigma, data &\sim N(\bar{\alpha}, \Sigma, \otimes\bar{\Omega}) \\ \Sigma|data &\sim IW(\bar{S}, \bar{v}) \end{aligned}$$

for $t = p + 1, ..., T$.

The prior mean and variance assumption follow Minessota-style priors

$$\alpha_0 = E[A_k^{ij}] = \begin{cases} 1 & \text{if } i = j, k = 1 \\ 0 & \text{otherwise} \end{cases}$$

$$\Omega_0 = var[A_k^{ij}] = \begin{cases} \left(\frac{\lambda_1 \lambda_2}{k} \frac{\sigma_i}{\sigma_j}\right)^2, & \text{if } k = 1, ..., p \\ (\lambda_0, \sigma_i)^2, & \text{if } k = 0; \end{cases}$$

Lastly, λ_1 is the shrinkage parameter and values for σ_i and σ_j are obtained using univariate autoregressive models. The prior scale matrix S_0 is assumed to be diagonal with diagonal elements given by

$$S_0^{ii} = (v_0 - N - 1)\sigma_i^2 \text{ and } v_0 = N + 2$$

The data is also augmented by the 'sum of coefficients' and 'initial observation' dummy observations. The sum of coefficients dummy takes into account the belief that the average of the lagged values of a variable ($\bar{Y}_{0n}$) is a good forecast of future observations, with implementation as

$$\mathbf{Y}_{d2} = \begin{bmatrix} diag(\bar{Y}_{0n})/\tau_1 \end{bmatrix} \quad \mathbf{X}_{d2} = \begin{bmatrix} \mathbf{J}_p \otimes diag(\bar{Y}_{0n})/\tau_1 & 0_{n\times 1} \end{bmatrix}.$$

In this case, when $\tau_1 \to 0$ the model tends to a VAR with no cointegration. The initial observation dummy indicates that the values of the variables are set to be the averages of initial conditions, constructed as a single dummy observation as in

$$\mathbf{Y}_{d3} = \begin{bmatrix} \bar{Y}_{0n}/\tau_2 \end{bmatrix} \quad \mathbf{X}_{d2} = \begin{bmatrix} \mathbf{J}_p \otimes \bar{Y}_{0n}/\tau_2 & 1/\tau_2 \end{bmatrix}.$$

Here, when $\tau_2 \to 0$ the model converges to the case where all variables are stationary, or with unit root components without drift terms.

Within this setup the overall prior tightness λ_1 and the optimal lag length are selected to maximise the marginal likelihood

$$\lambda_1 = \underset{\lambda_1}{\operatorname{argmax}} \ln p(Y)$$

over the fixed grid [0.01, 0.025, 0.1, 0.15, 0.2, 0.25, 0.3, 0.35, 0.4, 0.45, 0.5, 0.75, 1, 2, 5]. All other lambda are set equal to 10*λ_1.

Next, forecasts are computed using simulation, following Carriero *et al.*(2015). In fact, posterior draws of the parameters α and Σ are used in order to construct an implied path for $\hat{y}_{T+1}, ..., \hat{y}_{T+h}$. Under the assumption that $\boldsymbol{A} = [A_c, A_1, ..., A_p]$ one obtains a draw for j for all autoregressive coefficients using:

$$(\boldsymbol{A}^{(j)} = \bar{\boldsymbol{A}} * chol(\bar{\Omega}^{(j)}) * V(j) * chol(\boldsymbol{\Sigma}^{(j)}))$$

with $V^{(j)}$ obtained from a standard normal distribution. For a draw of $\boldsymbol{A}^{(j)}$ and $\boldsymbol{\Sigma}^{(j)}$, a sequence of h draws from $N(0, \boldsymbol{\Sigma}^{(j)})$ is drawn in order to obtain by iteration a sequence of forecasts for $\hat{y}_{T+1}, ..., \hat{y}_{T+h}$.

Autoregressive model

The autoregressive model is of the form

$$y_t = \alpha + \sum_{j=1}^{p} \beta_j y_{t-j} + u_t \tag{2}$$

where y_t is the target forecasting variable, α refers to an intercept term, β_j are the autoregressive coefficients and u_t is the residual term, which is assumed to be white noise. The optimal lag length is chosen using Bayesian Information criteria (BIC). In most cases, the information criteria are minimised setting $j = 1$.

Given the data $y_1, y_2, ..., y_T$ the one-period ahead forecast is computed by

$$\begin{aligned} y_{T+1,T} &= \mathbb{E}(y_{T+1}|\Omega_T) \\ &= \hat{\alpha} + \hat{\beta}_j y_t \end{aligned}$$

and in similar fashion the two-period ahead forecast is obtained via backward substitution.

Conditional UC-SV Model

Further we estimate a conditional unobserved component stochastic volatility (UC-SV) model, following a two-step approach. Firstly, we follow the modelling strategy by Stock and Watson (2007) and estimate a UC-SV model of the form.

$$\begin{aligned} y_t &= \tau_t + \eta_t, && \text{where } \eta_t = \sigma_{\eta,t}\zeta_{\eta,t} \\ \tau_t &= \tau_{t-1} + \epsilon_t, && \text{where } \epsilon_t = \sigma_{\epsilon,t}\zeta_{\epsilon,t} \\ \ln\, \sigma^2_{\eta,t} &= \ln\, \sigma^2_{\eta,t-1} + \upsilon_{\eta,t} \\ \ln\, \sigma^2_{\epsilon,t} &= \ln\, \sigma^2_{\epsilon,t-1} + \upsilon_{\epsilon,t} \end{aligned}$$

where $\zeta_t = (\zeta_{\eta,t}, \zeta_{\epsilon,t})$ is i.i.d $N(0, I_2)$, $\upsilon_t = (\upsilon_{\eta,t}, \upsilon_{\epsilon,t})$ is i.i.d $N(0, \gamma I_2)$, ζ_t and υ_t are independently distributed, and γ is a scalar parameter (Stock and Watson, 2007).

Secondly, in line with Stock and Watson (2010) we extract the obtained trend parameter

τ_{t-1} and employ it in an augmented ARDL model such that

$$y_t = \alpha + \beta_1 \tau_{t-1} + \sum_{j=1}^{q} \gamma_j x_{t-j} + u_t \tag{3}$$

where y_t is the target forecasting variable, α the intercept term, β_j is the coefficient of the trend component, which we obtained from the UC-SV model, γ_j are the coefficients of the lagged terms of the exogenous variable and u_t is the residual term, which is assumed to be white noise. It is worth noting that we only use the first lagged trend component as explanatory variable. The number of lags of the additional exogenous variables x_{t-j}, however, may be different, depending on which variables is used. We use the BIC in order to determine the optimal lag length for x_t. Lastly, the set of exogenous variables varies across countries, as outlined in the Data section.

Next, forecasts are constructed in similar fashion to a conventional autoregressive model such that the h-period ahead forecast $\hat{y}_{t+h}$ is based on the information set available at the point in time Ω_t:

$$\begin{aligned} y_{T+1,T} &= \mathbb{E}(y_{T+1}|\Omega_T) \\ &= \hat{\alpha} + \hat{\beta}_1 \tau_t + \sum_{j=0}^{q-h} \hat{\gamma}_j x_{t-j} \end{aligned}$$

Model Combinations

In addition to the individual point forecasts obtained by the models described above, we compute forecasts using equal and logarithmic score based combinations of the outlined models. The approach is as follows:

Firstly, based on the forecasts for the mean and standard deviation of the targeted series, we compute density forecasts for all our estimated models. Secondly, we compute the combined density and derive a combination-based forecast, which allocates different weights to different models. In case of the equal weight, this approach is straight forward since each model is dedicated the same weight.

In case of the logarithmic score based approach the weights across models can be different and are dependent on the models' individual forecasting performance. The weights are determined following the approach by Garratt *et al.* (2014). Firstly, we determine the predictive power of all models by conducting an out-of-sample forecasting exercise. Estimations are based on a rolling window approach and $N = 36$ forecasts. Secondly, we construct a combined density of the form

$$p(y_t) = \sum_{i=1}^{N} w_{i,\tau} g(y_t, I_{i,\tau}) \text{ , with } \tau = \underline{\tau}, ..., \overline{\tau} \tag{4}$$

where $g(y_t|I_{i,t})$ are the one-step ahead forecast densities from model i, $i = 1, ...N$, for the target forecasting series, conditional on the information set $I_{i,t}$. The impact of each model on the overall density may change with each recursion in the evaluation period $\tau = \underline{\tau}, ..., \overline{\tau}$.

The weights $w_{i,\tau}$ are non-negative and sum up to 1. They are calculated for each recursion in the evaluation period based on the fit of the individual component forecasting density. Generally, the logarithmic score measures the density fit of each model through the evaluation period. It means that a high score is given to a density forecast that assigns a high probability to the realised value. In line with Garratt *et al.* (2014) we define the recursive weights for the one-period ahead densities as

$$w_{i,t} = \frac{\exp\left[\sum_{\underline{\tau}-\kappa}^{\tau-1} \ln g(y'_{\tau|I_{i,\tau}})\right]}{\sum_{i=1}^{N} \exp\left[\sum_{\underline{\tau}-\kappa}^{\tau-1} \ln g(y'_{\tau|I_{i,\tau}})\right]} \tag{5}$$

where $\underline{\tau}-\kappa$ to $\underline{\tau}-1$ comprises the training period used to initialise the weights and ln $g(y'_t|I_{i,\tau})$ is the logarithm of the probability density function $g(.|I_{i,\tau})$, evaluated at the forecast, y'_t.

Overall, we obtain up to 17 individual and combined forecasts for each country and target series. The exact number of forecasts varies across countries and depends on the size of the dataset and the number of exogenous variables used in the BVAR and conditional UC-SV. The data for each country is outlined in greater detail in the next section.

3 Data

The WBS University of Warwick Forecast Team computes monthly forecasts for two target series: HICP (All items) and HICP (excluding Energy). Forecasts are submitted for the Euro Area and for the countries France, Germany, Italy and United Kingdom. A detailed outline of the set of country variables used, can be found below.

In order to estimate the models described in the previous section, we first construct the annual change in price by using the 12-month difference:

$$y_t = 100 * \frac{q_t - q_{t-12}}{q_{t-12}}$$

where q_t refers to the respective price level data for each country and y_t is the level of inflation. Next, we estimate all models using y_t and construct a one- and two-period ahead forecast, respectively.[1]. Lastly, using backward construction, we obtain a point forecast for the price level for each model by computing

$$q_{t+1} = (1 + (y_{t+1})/100) * q_t \tag{6}$$

With regard to the forecast of the standard deviation, our approach follows the same rationale. We first construct the variance of the underlying inflation series y_t. Second, we derive the standard deviation from the target series, such as

$$\sigma_{q,t+1} = \sqrt{\frac{\sigma^2_{y,t+1}}{10000} * q^2_{t-12}} \tag{7}$$

where $\sigma_{q,t+1}$ is the forecast of the target series' standard deviation, $\sigma_{y,t+1}$ is the standard deviation of inflation and q_{t-12} is last period's price level.

With regard to the source of Data, we extract all variables from *Datastream*. For each country, the series of interest are HICP (All Items: 2005=100 NADJ) and HICP (excluding Energy, 2005=100 NADJ). Moreover, for the BVAR and conditional UC-SV models, we use additional exogenous macroeconomic variables which serve as explanatory variables. The approach follows closely Carriero *et al.* (2015) and set of variables differ across countries, as listed below.[2]

For the Euro Area we use the following variables: EK Industrial Production: Manufacturing (EA18) VOLA (EKIPMAN.G), EK Unemployment (EA18) VOLA (EKESTUNPO), EK Industrial Prodn. - Consumer non durables (%MOM) (EA18) SADJ (EKESICN%Q), EK New Residential Buildings - Cost Index (EA18) NADJ (EKECEIBCF), EURO Stoxx - Price Index (DJEURST), BD Discount Rate / Short term Euro Repo Rate (BDPRATE.), EK Real Effective Exchange Rates - CPI Based VOLN (EKOCC011), EK Economic Sentiment Indicators (EA18) VOLA (EKEUSESIG), WD Commodity Prices: Crude Oil NADJ

[1] We construct two forecasts in order to take into account that in some months the publication of the most recent data entities is after the monthly submission deadline

[2] Datastream codes are given in parentheses

(WDI76AADF), EK Construction Survey: Employment Expectations (EA) SADJ (EK45. 4BSQ), BD Long Term Government Bond Yield 9-10 Years (BDGBOND.), BD 3-Month Fibor (BDOIR076R).

For France we use FR Industrial Production VOLA (FRIPTOT.G), FR Unemployment Rate, Total SADJ (FRESUNEMO), FR Unemployment (Harmonized): Total TRND (FRESQ T8JT), FR Banque de France SVY: Business Sentiment Indicator (CAL ADJ) (FRSURCBSQ), FR Survey: MFG Output - Order Book & Foreign Demand SADJ (FRSURFMPQ), FR Survey: MFG Output - Finished Good Inventories SADJ (FRSURSMPQ), FR Share Price Index - SBF 250 NADJ (FRSHRPRCF), WD Commodity Prices: Crude Oil NADJ (WDI76AADF), BD HWWA Index of World Market Prices of Raw Mats, EURO Area NADJ (BDHWWAINF), FR Real Effective Exchange Rates - CPI Based Voln (FROCC011), FR Average Cost of Funds For Banks/ Euro Repo Rate (FRPRATE.), FR Government Guaranteed Bond Yield (EP) NADJ (FRGBOND.) , FR Capital Market Yields 13-Week Treasury Bills, Mo.Wght.Avg. (FRGBILL3)

For Germany the variables BD Industrial Production including construction (Cal Adj) VOLA (BDIPTOT.G), BD Cnstr.Ind: Capacity Utilization SADJ (BDIFDCTNQ), BD Unemployment Rate, Total SADJ (BDESUNEMO), BD Employment Duration - Short-term Workers VOLN (BDEMPSTWP), BD New Orders to Manufacturing - Domestic: Consumer Goods VOLA (BDDCNORDG), BD Dax Share Price Index, EP NADJ (BDSHRPRCF), BD Real Effective Exchange Rates - CPI Based VOLN (BDOCC011), BD Construction Orders Received - Residential Construction VOLA (BDHOUSE.G), BD Discount Rate / Short term Euro Repo Rate (BDPRATE.), BD PPI: Indl. Products, Total, sold on the domestic market NADJ (BDPROPRCF), BD Consumer Confidence Indicator - Germany SADJ (BDCNFCONQ), BD HWWI Index of World MKT.PRC.OF Raw Mats, Euro Area: excl. Energy (BDIUW501F), BD Long Term Government Bond Yield 9-10 Years (BDGBOND.), BD 3 Month Fibor (BDOIR076R) are used.

For Italy we use IT Industrial Production VOLA (ITIPTOT.G), IT Unemployment rate, Total SADJ (ITESUNEMO), IT Unemployment VOLA (ITESTUNPO), IT Ind.: Overall-Empl expect SADJ (ITTTA7BSQ), IT New Residential Buildings - Cost Index NADJ (ITECEIBCF), IT Discount Rate / Short Term Euro Repo Rate (ITPRATE.), IT Milan Comit General Share Price Index (EP) NADJ (ITSHRPRCF), IT Real Effective Exchange Rates - CPI Based VOLN (ITOCC011), IT PPI Linked & Rebased NADJ (ITPROPRAF), IT Economic Sentiment Indicator VOLA (ITEUSESIG), WD Commodity Prices: Crude Oil NADJ

(WDI76AADF), IT Government Bond Gross Yield (Rendistato) (EP) (ITGBOND.), Italy T-Bill Auct. Gross 3-Month Middle Rate (ITBT03G).

For the UK we use UK Index of Production - All Production Industries VOLA (UKIP-TOT.G), UK LFS: In Emp.: Aged 16+: Annual = Spring Quarter (Mar-May) Vola (UKM-GRZ..O), UK GFK Consumer Confidence Index NADJ (UKGFKCCNR), UK LFS: UNEMPLOYMENT RATE, ALL, AGED 16 & OVER SADJ (UKUN%O16Q), UK LSL/ ACAD Average House Price CURA (UKFTHPI.B), UK c (UKSHRPRCF), UK Real Effective Exchange Rates - CPI based VOLN (UKOCC011), UK Interbank Rate - 3 Month (Month Average) (UKINTER3), UK PPI - Output of Manufactured Products (Home Sales) NADJ (UKPROPRCF), UK Bank of England Base Rate (EP) (UKPRATE.), UK Yield 10-Year (UKOIR080R), UK Discount Rate 3-Month Treasury Bills (UKOIR077R), UK RPI - All Items excluding Mortgage Interest NADJ (UKRPAXMIF) and WD Commodity Prices: Crude Oil NADJ (WDI76AADF).

4 Estimation/Replication procedure

The forecasting procedure can be differentiated in three different steps. Firstly, raw data has to be updated. The data is collected in 10 different Excel (csv) files. For HICP, these files are 'EA18model,csv', 'FRAmodel.csv', 'GERmodel.csv', 'ITAmodel.csv' and 'UK-model.csv'. For HICP excluding Energy, the files are 'EA18model_e,csv', 'FRAmodel_e.csv', 'GERmodel.csv_e', 'ITAmodel_e.csv' and 'UKmodel_e.csv'.

Secondly, it is necessary to replace missing values. We use the BVAR model in order to conduct an imputation procedure and execute estimations for each country/EA18 using the MATLAB files 'Imputation_EA18.m', 'Imputation_FRA.m', 'Imputation_GER.m', 'Imputation_ITA.m' and 'Imputation_UK.m'. These procedures will also create the auxiliary Excel files, which are used for the forecasting procedure.

Finally, we use the MATLAB procedure 'Forecast_comp.m' to conduct the forecast estimations. We execute each combination of country/EA18 and HICP/HICP excluding Energy by changing its associated model code in line 12 ('Model'). We conduct forecasts one-step ahead, except for February, when January HICP data was not available by the submission date. In that case, we adopted a two-step ahead forecast, which can be displayed by changing line 23 ('horizon2') from 0 to 1. The output follows the structure of final template provided by BDCOMP.

5 Results

We begin the evaluation of model performance by calculating the mean square error (MSE) for each model.[3] We label the model with the smallest MSE as "Best" and the model with the largest MSE as "Worse" performing model for each country, as shown in Table (1). For both indices, using simple AR models results in the worst performance for nearly all countries, except the UK for which the Conditional UC-SC with LSL as explanatory variable shows a higher MSE.

The best performance for the HICP index is obtained with the Bayesian VAR approach for the United Kingdom, while a log score based model combination performs best for France and EA18. For Germany and Italy different UC-SV models appear to be the best model choice. For the HICP excluding Energy, conditional UC-SV with different explanatory variables as additional regressors, such as e.g. unemployment rate (UK), PPI (Italy) and or discount rate (EA18) outperform other models, such as e.g. the BVAR.

Table 1: Performance by Country: Best and Worse Model (by MSE)

	HICP		**HICP excl. Energy**	
Country	**Best**	**Worst**	**Best**	**Worst**
UK	BVAR	Cond. UC-SV LSL	Cond. UC-SV Unemployment rate	AR
DE	Cond. UC-SV Unemployment rate	AR	Cond. UC-SV Manufacturing	AR
FR	log score MC	AR	Cond. UC-SV HWWA index	AR
IT	Cond. UC-SV Comm. prices	AR	Cond. UC-SV PPI	AR
EA18	log score MC	AR	Cond. UC-SV discout rate	AR

As additional measure for model evaluation, we report the relative mean square error

[3]Mean square error is calculated as $MSE = \frac{1}{N}\sum_{i=1}^{N}(R_i - F_i)^2$ where R_i refers to the true value and the F_i is the forecasted value for each model.

(RMSE) for all countries and modelling approaches in Tables (2) and (3).[4] For each country, the smallest RMSE is marked in bold.

For HICP including all items (Table (2)), both RMSE and MSE identify the same models as best approach (e.g. for the UK, Approach 1: BVAR). Concerning the HICP excluding Energy prices, the inferences are less distinct. As revealed in Table (3), for nearly all countries (exception: France) more than one modelling approach have the same smallest RMSE. For these countries, it is difficult to identify one approach which uniquely captures price dynamics.

Table 2: RMSE - HICP

	UK	DE	FR	IT	EA18
Approach 1	**0.26**	1.31	1.39	6.02	2.38
Approach 2	0.30	1.82	2.62	6.55	2.74
Approach 3	0.92	1.10	0.72	0.73	0.61
Approach 4	0.77	1.04	1.15	0.76	0.55
Approach 5	0.79	**0.98**	0.71	0.76	0.66
Approach 6	1.14	1.09	0.92	0.76	0.66
Approach 7	1.25	1.16	0.84	0.74	0.78
Approach 8	0.99	1.16	0.89	0.73	0.64
Approach 9	0.97	1.16	1.05	0.74	0.71
Approach 10	1.14	1.13	1.29	0.78	0.70
Approach 11	0.48	1.09	0.76	0.69	1.01
Approach 12	1.20	1.09	0.96	0.80	0.67
Approach 13	1.04	1.05	0.82	**0.64**	0.49
Approach 14	0.59	1.12	0.75	0.74	0.70
Approach 15	1.14	1.01	0.78	0.91	**0.41**
Approach 16	0.71	1.06	**0.57**	4.05	
Approach 17	0.67	1.08			

Table 3: RMSE - HICP excl. Energy

	UK	DE	FR	IT	EA18
Approach 1	0.30	0.70	0.87	6.08	1.46
Approach 2	0.33	0.69	0.92	6.55	1.59
Approach 3	0.22	0.42	**0.12**	0.40	0.14
Approach 4	0.21	0.42	**0.12**	0.40	0.16
Approach 5	**0.20**	0.40	**0.12**	0.41	0.14
Approach 6	0.25	0.43	**0.12**	0.41	0.14
Approach 7	0.25	**0.38**	**0.12**	0.40	**0.13**
Approach 8	0.23	0.44	**0.12**	0.40	**0.13**
Approach 9	0.30	0.42	0.13	0.43	0.14
Approach 10	0.27	0.42	**0.12**	**0.39**	0.14
Approach 11	0.21	0.43	**0.12**	**0.39**	**0.13**
Approach 12	0.29	0.41	0.14	0.40	0.14
Approach 13	0.25	0.43	**0.12**	**0.39**	0.14
Approach 14	**0.20**	0.41	**0.12**	0.40	0.20
Approach 15	0.24	0.43	0.15	0.68	0.14
Approach 16	0.21	0.44	**0.12**	0.41	
Approach 17	0.23	0.40			

Turning towards the forecasted standard deviation, we begin with showing the ranking (best and worse) of average standard deviations in Table (4) for each country. Best refers to the model with the lowest standard deviation in its forecast, i.e. the model which is most confident in its mean forecast. Similar to the previous Table, AR models appear as worst. This is consistent as a worse performance in terms of MSE is in line with less confidence in a models forecast. In contrast, conditional UC-SV models show the smallest standard deviation for HICP including all items while log-score based model combinations perform well for the HICP excluding Energy index.

In order to evaluate the density forecasts, we report the modified likelihood in Table (5) and (6).[5] The highest value for each country is marked in bold. In contrast to RMSE in the previous table, we can identify one model which outperforms all other models for each

[4]RMSE is calculated following the BDCOMP guidelines: $RMSE = 1/N \times \sum_{i=1}^{N}((F_i - R_i)/R_i)^2)$. RMSE is re-scaled by the factor 1/100000 in order to allow for comparability.

[5]Following the BDCOMP guidelines, we calculate the modified as $L = (\prod_{i=1}^{N} P_i)^{(1/N)}$.

Table 4: Performance by Country: Best and Worse (Average Standard deviaton)

	HICP		HICP excl. Energy	
Country	**Best**	**Worst**	**Best**	**Worst**
UK	Cond. UC-SV RPI	AR	log score MC	AR
DE	Cond. UC-SV PPI	AR	Cond. UC-SV Manufacturing	BVAR
FR	Cond. UC-SV HWWA	equal weights MC	log score MC	AR
IT	Cond. UC-SV PPI	Cond. UC-SV Comm. Prices	log score MC	equal weights MC
EA18	Cond. UC-SV Comm. Prices	equal weights MC	log score MC	AR

country when the modified likelihood is used. Looking at the HCIP including all items, the BVAR shows the highest modified likelihood measure for the UK and Germany (0.097 and 0.064). For France and EA18, the log-score weighted model combinations have the best model fit (0.109 and 0.100), while only for Italy (0.074) the conditional UC-SV performs best. Concerning HICP excluding Energy (Table (6)), conditional UC-SV models with different explanatory variables outperform all other modelling approaches. For example, for the UK a conditional UC-SV with the unemployment rate as additional regressor on the right-hand side appears to perform best.

Table 5: Likelihood HICP

	UK	DE	FR	IT	EA18
Approach 1	**0.097**	**0.064**	0.077	0.011	0.038
Approach 2	0.090	0.057	0.045	0.013	0.047
Approach 3	0.058	0.056	0.075	**0.074**	0.080
Approach 4	0.065	0.059	0.051	0.072	0.087
Approach 5	0.062	0.062	0.077	0.071	0.076
Approach 6	0.049	0.056	0.064	0.072	0.073
Approach 7	0.045	0.053	0.069	0.073	0.066
Approach 8	0.056	0.053	0.066	0.073	0.077
Approach 9	0.056	0.053	0.057	0.073	0.071
Approach 10	0.049	0.054	0.037	0.071	0.072
Approach 11	0.084	0.056	0.068	0.068	0.045
Approach 12	0.048	0.053	0.062	0.068	0.075
Approach 13	0.054	0.058	0.070	0.071	0.091
Approach 14	0.075	0.055	0.075	0.073	0.081
Approach 15	0.050	0.060	0.078	0.072	**0.100**
Approach 16	0.071	0.062	**0.109**	0.058	
Approach 17	0.070	0.056			

Table 6: Likelihood HICP excl. Energy

	UK	DE	FR	IT	EA18
Approach 1	0.126	0.093	0.096	0.009	0.031
Approach 2	0.126	0.092	0.094	0.008	0.030
Approach 3	0.149	0.110	0.203	0.120	0.205
Approach 4	0.154	0.110	0.202	0.119	0.194
Approach 5	**0.155**	0.114	0.202	0.119	0.208
Approach 6	0.144	0.109	0.201	0.119	0.200
Approach 7	0.143	**0.120**	0.202	0.119	0.213
Approach 8	0.149	0.107	0.204	0.119	**0.214**
Approach 9	0.134	0.111	0.197	0.115	0.205
Approach 10	0.141	0.110	0.202	**0.122**	0.204
Approach 11	0.152	0.110	0.202	0.121	0.210
Approach 12	0.138	0.113	0.193	0.120	0.205
Approach 13	0.145	0.109	**0.205**	0.121	0.202
Approach 14	0.155	0.112	0.201	0.120	0.196
Approach 15	0.147	0.110	0.196	0.120	0.203
Approach 16	0.150	0.111	0.202	0.115	
Approach 17	0.151	0.112			

Furthermore, for both indices we compare the performance of the BVAR, of model combinations (equal and log-score weighted) and the best-performing model relative to the AR process.[6] The results are shown in Figure (1) and Figure (2) for the HICP including all items and HICP excluding Energy, respectively. The graphs are constructed in the following way. First, we calculate the MSE for each model and second, we take the ratio between each model's MSE and the MSE of a simple AR process. We refer to this measure as "relative MSE" and display the measure in the Figures below. If the relative MSE is bigger than one, it indicates that the respective model performs worse than a simple AR process. The model with the smallest relative MSE performs the best.

Focusing on Figure (1), it is prevalent that for the UK both model combinations perform comparably bad for the HICP, as indicated by the large relative MSE. In contrast, the BVAR shows a relative MSE smaller than one (outperforming an AR process) and is the best performing model. For the other four countries the combination of models performs nearly as good as the best model and, as in the case of France or Italy, the log score-weighted modelling approach shows the smallest forecasting error.

Concerning HICP excluding Energy index (Figure (2)), the overall good performance of the model combinations stands out. While the best model performances for all countries is often obtained by a conditional UC-SV model, the log-score weighted combined model performs nearly as good as the best model (as indicated by the small different between grey and yellow bars).

In a final evaluation exercise, we control how often our forecasts lie outside of a one standard deviation error band. As we consider a small sample of only 11 forecasting submissions (January to November), we would expect that 3.3 forecast lie within the error bands (approx. 68%). This exercise is posed to give an idea of how reliable the standard deviations measures of the models are. Tables (7) and (8) show that for nearly all models the number of forecasts out of one error bands is higher than 3.3. This is particularly the case for UC-SV models. Further, while it appears like the AR models are performing relatively well, one has to keep in mind that the models' standard deviation are also very large (as indicated in Tables (4)). Therefore, it is little surprising that many forecasts lie within one-standard deviation error bands. This implies that the AR model indicates its limits in a more reliable form than the UC-SV models. Overall, we believe that the model combinatios perform particular well, given their small standard deviations and their appropriate number of forecasts outside of the error bands.

[6] We choose to compare these models against each other since we use all these models for all countries. However, it is worth mention that depending on data availability the BVAR model includes a different number of variables across countries. Performance is measured using MSE.

As a concluding remark, it is important to emphasize that 11 forecast points are not enough to have a final conclusion on which model is the best predictor for each country. One issue with this comparison is the February case. The BVAR and AR models performed particularly badly in this month, due to a much lower two-step ahead forecast power than one-step ahead[7]. This outlier explains the substantially high MSE for approaches 1 and 2 in Tables (2) and (3) for France, Italy and EA18. When excluding February, the outperform observed in Figures (1) and (2) for these countries and models is less intense.

Figure 1: MSE Relative to AR process (HICP all items)

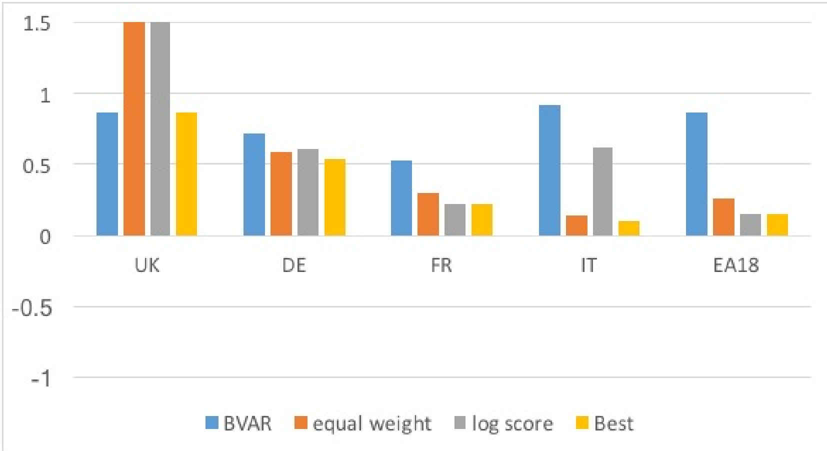

Figure 2: MSE Relative to AR process (HICP excl. Energy)

.

[7]As pointed out before, January data was not available by the time of the submission date for February forecast.

Table 7: HICP out of bounds

	UK	DE	FR	IT	EA18
BVAR	2	3	1	4	6
AR	1	4	3	3	2
UC-SV 1	5	8	6	5	8
UC-SV 2	5	7	6	5	8
UC-SV 3	6	8	6	4	7
UC-SV 4	5	8	6	6	8
UC-SV 5	7	7	6	5	8
UC-SV 6	5	8	6	5	7
UC-SV 7	6	8	6	4	8
UC-SV 8	7	8	8	4	8
UC-SV 9	4	8	7	4	8
UC-SV 10	7	8	6	4	7
UC-SV 11	6	8	6	5	7
UC-SV 12	5	8	5	5	
UC-SV 13	7	8			
equal weights MC	5	7	5	4	6
log score MC	5	8	2	5	4

Table 8: HICP (excl. Energy) out of bounds

	UK	DE	FR	IT	EA18
BVAR	1	3	2	3	3
AR	2	3	2	3	3
UC-SV 1	3	5	1	4	4
UC-SV 2	3	5	2	4	4
UC-SV 3	2	5	3	4	4
UC-SV 4	3	5	3	3	4
UC-SV 5	4	3	2	4	3
UC-SV 6	4	5	2	3	3
UC-SV 7	4	5	2	4	4
UC-SV 8	4	5	2	3	4
UC-SV 9	3	5	1	4	3
UC-SV 10	4	5	2	3	4
UC-SV 11	4	5	3	3	4
UC-SV 12	2	5	3	3	
UC-SV 13	4	5			
equal weights MC	3	4	2	2	3
log score MC	4	5	3	3	4

Table 9: Euro Area

	Model	Exogenous Variable
1. Approach	BVAR	
2. Approach	AR	
3. Approach	Conditional UC-SV	Industrial Production: Manufacturing (EA18) VOLA
4. Approach	Conditional UC-SV	Unemployment (EA18) VOLA
5. Approach	Conditional UC-SV	Industrial Prodn. - Consumer non durables (%MOM) (EA18) SADJ
6. Approach	Conditional UC-SV	New Residential Buildings - Cost Index (EA18) NADJ
7. Approach	Conditional UC-SV	EURO Stoxx - Price Index
8. Approach	Conditional UC-SV	Discount Rate / Short term Euro Repo Rate
9. Approach	Conditional UC-SV	Real Effective Exchange Rates - CPI Based VOLN
10. Approach	Conditional UC-SV	Economic Sentiment Indicators (EA18) VOLA
11. Approach	Conditional UC-SV	Commodity Prices: Crude Oil NADJ
12. Approach	Conditional UC-SV	Construction Survey: Employment Expectations (EA) SADJ
13. Approach	Conditional UC-SV	Long Term Government Bond Yield 9-10 Years - 3-Month Fibor
14. Approach	Model Combination (equal weights)	
15. Approach	Model Combination (logarithmic score)	

Table 10: France

	Model	Exogenous Variable
1. Approach	BVAR	
2. Approach	AR	
3. Approach	Conditional UC-SV	Industrial Production VOLA
4. Approach	Conditional UC-SV	Unemployment Rate, Total SADJ
5. Approach	Conditional UC-SV	Unemployment (Harmonized): Total TRND
6. Approach	Conditional UC-SV	Banque de France SVY: Business Sentiment Indicator (CAL ADJ)
7. Approach	Conditional UC-SV	Survey: MFG Output - Order Book & Foreign Demand SADJ
8. Approach	Conditional UC-SV	Survey: MFG Output - Finished Good Inventories SADJ
9. Approach	Conditional UC-SV	Share Price Index - SBF 250 NADJ
10. Approach	Conditional UC-SV	WD Commodity Prices: Crude Oil NADJ
11. Approach	Conditional UC-SV	HWWA Index of World Market Prices of Raw Mats, EURO Area NADJ
12. Approach	Conditional UC-SV	Real Effective Exchange Rates - CPI Based Voln
13. Approach	Conditional UC-SV	Average Cost of Funds For Banks/ Euro Repo Rate
14. Approach	Conditional UC-SV	Government Bond Yield (EP) NADJ minus 13-Week Treasury Bills
15. Approach	Model Combination (equal weights)	
16. Approach	Model Combination (logarithmic score)	

Table 11: Germany

	Model	Exogenous Variable
1. Approach	BVAR	
2. Approach	AR	
3. Approach	Conditional UC-SV	Industrial Production including construction (Cal Adj) VOLA
4. Approach	Conditional UC-SV	Cnstr.Ind: Capacity Utilization SADJ
5. Approach	Conditional UC-SV	Unemployment Rate, Total SADJ
6. Approach	Conditional UC-SV	Employment Duration - Short-term Workers VOLN
7. Approach	Conditional UC-SV	New Orders to Manufacturing - Domestic: Consumer Goods VOLA
8. Approach	Conditional UC-SV	Dax Share Price Index, EP NADJ
9. Approach	Conditional UC-SV	Real Effective Exchange Rates - CPI Based VOLN
10. Approach	Conditional UC-SV	Construction Orders Received - Residential Construction VOLA
11. Approach	Conditional UC-SV	Discount Rate / Short term Euro Repo Rate
12. Approach	Conditional UC-SV	PPI: Indl. Products, Total, sold on the domestic market NADJ
13. Approach	Conditional UC-SV	Consumer Confidence Indicator - Germany SADJ
14. Approach	Conditional UC-SV	HWWI Index of World MKT.PRC.OF Raw Mats, Euro Area: excl. Energy
15. Approach	Conditional UC-SV	Long Term Government Bond Yield 9-10 Years minus 3 Month Fibor
16. Approach	Model Combination (equal weights)	
17. Approach	Model Combination (logarithmic score)	

Table 12: Italy

	Model	Exogenous Variable
1. Approach	BVAR	
2. Approach	AR	
3. Approach	Conditional UC-SV	Industrial Production VOLA
4. Approach	Conditional UC-SV	Unemployment rate, Total SADJ
5. Approach	Conditional UC-SV	Unemployment VOLA
6. Approach	Conditional UC-SV	Ind.: Overall-Empl expect SADJ
7. Approach	Conditional UC-SV	New Residential Buildings - Cost Index NADJ
8. Approach	Conditional UC-SV	Discount Rate / Short Term Euro Repo Rate
9. Approach	Conditional UC-SV	Milan Comit General Share Price Index (EP) NADJ
10. Approach	Conditional UC-SV	Real Effective Exchange Rates - CPI Based VOLN
11. Approach	Conditional UC-SV	PPI Linked & Rebased NADJ
12. Approach	Conditional UC-SV	Economic Sentiment Indicator VOLA
13. Approach	Conditional UC-SV	Commodity Prices: Crude Oil NADJ
14. Approach	Conditional UC-SV	Government Bond Gross Yield (EP) minus Gross 3-Month Middle Rate
15. Approach	Model Combination (equal weights)	
16. Approach	Model Combination (logarithmic score)	

Table 13: United Kingdom

	Model	Exogenous Variable
1. Approach	BVAR	
2. Approach	AR	
3. Approach	Conditional UC-SV	Index of Production
4. Approach	Conditional UC-SV	LFS: In Emp.: Aged 16+: Annual =Spring Quarter (Mar-May) VOLA
5. Approach	Conditional UC-SV	Unemployment Rate SADJ
6. Approach	Conditional UC-SV	GFK Consumer Confidence Index NADJ
7. Approach	Conditional UC-SV	LSL/ ACAD Average House Price CURA
8. Approach	Conditional UC-SV	FT All Share Index (EP)
9. Approach	Conditional UC-SV	Real Effective Exchange Rates - CPI Based VOLN
10. Approach	Conditional UC-SV	Interbank Rate - 3 Month (Month Avg)
11. Approach	Conditional UC-SV	PPI - Output of Manufactured Products (Home Sales) NADJ
12. Approach	Conditional UC-SV	Bank of England Base Rate
13. Approach	Conditional UC-SV	Yield 10-year Treasury Bill Rate minus Discount Rate 3-Month Trasury Bill
14. Approach	Conditional UC-SV	RPI - All Items Excluding Mortgage Interest NADJ
15. Approach	Conditional UC-SV	WD Commodity Prices: Crude Oil NADJ
16. Approach	Model Combination (equal weights)	
17. Approach	Model Combination (logarithmic score)	

6 References

Carriero, A., Galvão, A.B. & Kapetanios, G. (2015). A comprehensive evaluation of macroeconomic forecasting methods. *Working Paper*.

Garratt, A., Mitchell, J. & Vahey, S.P. (2014). Measuring output gap nowcast uncertainty. *International Journal of Forecasting*, 30, 268-279.

Stock, J.H. & Watson, M.W. (2007). Why Has U.S. Inflation Become Harder to Forecast?. *Journal of Money, Credit and Banking*, 39(1), 3- 33.

Stock, J.H. & Watson, M.W. (2010). Modeling Inflation after the Crisis. Federal Reserve Bank of Kansas City, *Proceedings - Economic Policy Symposium - Jackson Hole*, 173-220.