SHOULD WE CARE ABOUT ECB INFLATION EXPECTATIONS?

Francesco Roccazzella, Bertrand Candelon

LIDAM Discussion Paper LFIN
2022 / 04

UCLouvain

LFIN | LIDAM
Louvain Institute of Data Analysis and Modeling in economics and statistics

LFIN

Voie du Roman Pays 34, L1.03.01 B-1348 Louvain-la-Neuve

Tel (32 10) 47 43 04

Email: lidam-library@uclouvain.be

https://uclouvain.be/en/research-institutes/lidam/lfin/publications.html

Should we care about ECB inflation expectations? *

Francesco Roccazzella[†] Bertrand Candelon[‡]

Université catholique de Louvain, Lidam - LFIN

June 15, 2022

Abstract

We use optimal combination of forecasts to introduce a novel forecast encompassing test to evaluate time-series and institutional inflation projections in the euro area. Combination weights reveal which forecasts are the most informative. Although, ECB is the most informative forecaster on average, it does not encompass its competitors and its weight varies over time. Macro-financial conditions and monetary policy actions explain this variability. The greater the uncertainty surrounding inflation and the difference between current and the 2% inflation target, the less informative ECB's forecasts are. The more contractionary the monetary policy, the more informative they are. ECB's declining weight and the relation with its determinants raise a warning flag: the potential loss of informativeness damages ECB's leading role at anchoring inflation expectations and questions whether the goal of preserving financial stability is compatible with the inflation targeting objective.

Keywords: Forecast combinations ; Forecast evaluation ; Inflation ; euro area ; ECB

*Bertrand Candelon has conducted as part of the research program entitled "Financial and Extra-financial Risks Modeling" under the aegis of the Europlace Institute of Finance, a joint initiative with Insti7. The usual disclaimer applies. Francesco Roccazzella gratefully acknowledges the hospitality of Universitat Pompeu Fabra where part of this project was completed.

[†]Corresponding author - *Address:* Voie du Roman Pays 34, B-1348 Louvain-la-Neuve, Belgium. *E-mail address:* francesco.roccazzella@uclouvain.be

[‡]*E-mail address:* bertrand.candelon@uclouvain.be

Introduction

Forecasting inflation is essential for policymakers (central banks or treasury offices) to fix their policies but it also presents a broader interest for investors, labor market actors during the wage negotiation processes or households and firms, desiring to evaluate their purchasing power, their saving capacities or planning their inventory. Therefore, it is not surprising that a multitude of inflation forecasts are available.

We can find multiple time series-based forecasts (Stock and Watson, 1999). These models constitute the reduced forms of well-known relationships (Phillip's curve, slope of the yield curve,...) and provide the best prediction at time t given the information available. Institutional forecasts are also widely available. These forecasts rely both on quantitative forecast and on the qualitative feeling of markets and experts. The Survey of Professional Forecasters (SPF) and other international institutions such as the European Central Bank (ECB), the International Monetary fund (IMF), the European Commission (EC) and the Organization for the Economic Cooperation and Development (OECD) are proposing regular projection exercises to forecast inflation future dynamics.

Stock and Watson (1999) highlighted how researchers and policymakers have focused on inflation forecasts for decades, however their conclusion that SPF's inflation forecasts for the U.S. were dominated by simple reduced form of the structural Phillips curve forecast has been widely criticized. Studies have shown that this result was not robust to the sample data and the evaluations techniques were inadequate to potential out-of-sample prediction. For example, Ang et al. (2007) proposed a combination forecast exercise including both SPF and the time series indicators. The authors showed that Survey of Professional Forecasters is more informative on US inflation, i.e., SPF either encompasses or displays significantly greater weight than the competing time-series forecasts when included in a combination. However, other institutional forecasts are omitted from this study and, their relevance for forecasting inflation in the euro area (EA) remains to be explored. Moreover, analysts are not only interested in whether one forecaster is better than another *on average*, but also on when this occurs (Giacomini and White, 2006; Li et al., 2021). This consideration is important when competing forecasts display similar average performance, but behave very differently conditional on certain economic states.

In a combination of forecasts, weights indicate how informative forecasts are (Diebold, 1989) and Diebold and Shin (2019) bring evidence that optimal weights for EA inflation projections produced by individual forecasters within the survey of professional forecasters vary over time. This indicates that also the informational value embedded in such forecasts is equally non-constant over time. It still remains unknown whether this results holds for a broader set of forecasters, including institutions as central banks or the IMF.

We propose to shade new light on how informative inflation forecasts in the euro area are by constructing a novel optimal combination of forecasts between these existing indicators. We constrain the weights to lay in the unit simplex in the spirit of Conflitti et al. (2015) and design a forecast encompassing test to assess whether the weights are statistically significant. This paper extends this literature in three ways. First, together with the time series indicator and the surveys, it includes the inflation forecasts published by international institutions. We provide additional insights on the relative importance of well-known international organizations *vis-à-vis* of the traditional time series and surveys for inflation forecasting. Second, this paper adapt statistical test under constrained parameter space to tests for forecast encompassing. By explicitly considering that weights are constrained in the unit simplex, the proposed test proposes adequate size properties and it is more powerful than a competitive test that ignores such constraints. Furthermore, the testing procedure can be extended to frameworks in which the number of competing forecasts outclasses the number of available realizations of the target variable. Third, while understanding the determinants the weights' dynamics is interesting for macroeconomic analysis *per se*, studying in what economic states ECB's inflation forecasts bring more information also helps policy analysts interested in evaluating ECB's intents at monitoring inflation. In fact, the role of the ECB is very peculiar: it provides regular inflation forecasts but also conducts the monetary policy to maintain price stability. Therefore, ECB's inflation forecasts should naturally be the most informative. Despite the fact that the primary mandate of the ECB is monitoring inflation dynamics of the Euro, it also aims at preserving financial stability. A declining information content of ECB's inflation expectation might signal that the objective of financial stability can challenge the inflation targeting mandate. We therefore investigate whether macro-financial conditions and monetary policy actions affect ECB's forecasts informativeness.

We find that ECB is the most informative forecaster on average. Clearly, as monetary policy led by ECB relies on their internal projection, ECB's forecasts must have an intrinsic value. However, no individual forecasts of inflation (ECB included) encompasses the competing forecasts in the euro area. We also bring evidence that optimal weights greatly vary over time. This indicates that the informational value embedded in the forecast of interest is equally non-constant over time. This creates an ideal laboratory to study whether uncertainty, monetary policy or macro-financial conditions in the euro area explain its variability. Finally, ECB's declining weight and its relation with its determinants are warning flags, damaging ECB's leading role at anchoring inflation expectations and signaling possible incompatibility between preserving financial stability and the inflation targeting goal in a high inflation environment.

The remainder of the paper is organized as follows. We introduce the optimal combination of forecasts and forecast encompassing test in Section 1. We move to the analysis of inflation forecasts in the euro area in Section 2 and study the dynamics of weights in Section 3. We conclude in Section 4.

1 Optimal combination and forecast encompassing

Forecast combination and the forecast encompassing literature are closely interconnected and Diebold (1989) made the first attempt to formally reconcile these two streams of literature. Forecast encompassing tests aim to detect whether a forecast combination incorporates more information than the individual forecasts (Clements and Hendry, 1998).

In this perspective, forecast combination is not an end on itself, but it is rather motivated by overcoming the difficulties related to the combination of information sets when time and resource constraints are large. Assuming unbiased forecasts and under quadratic loss function, an analyst may seek to minimize the variance of forecast errors. [1] Similarly to finance's portfolio theory, the weights reflect forecasts' marginal contribution to the reduction of portfolio's forecast error variance (Roccazzella et al., 2022). Therefore, when looking at forecast encompassing and forecasts combination from a joint perspective, weights can be naturally interpreted as the marginal contribution of each forecaster in term of independent information.

In the following sections, we introduce the ex post optimal combination of forecasts and the testing procedure to evaluate their significance in a constrained parameter space.

1.1 Estimating the optimal weights

Let $\boldsymbol{y}$ be the column vector containing n observations of the target variable. Let $\boldsymbol{f}_i$ be an unbiased forecast of $\boldsymbol{y}$ and define $\boldsymbol{F}$ the $n \times d$ matrix containing the d available forecasts of $\boldsymbol{y}$. Similarly, we can define $\boldsymbol{V}$ as the $n \times d$ matrix of forecast errors and $\boldsymbol{S}$ to be an estimator of the covariance matrix of prediction errors. We denote the population covariance matrix of prediction errors with $\boldsymbol{\Sigma}$. We aim at estimating the linear combination of forecasts that minimizes the variance of the aggregate prediction error constraining the weights to be nonnegative and sum to one. Formally, denoting the unit simplex with d vertices with $\Delta^{d-1} = \{\beta_i \geq 0, \text{ with } i = 1, ..., d \text{ and } \sum_{i=1}^{d} \beta_i = 1\}$, the optimal nonnegative weights solve

$$\boldsymbol{\beta}^* := \underset{\boldsymbol{\beta} \in \Delta^{d-1}}{\operatorname{argmin}} \, \boldsymbol{\beta}^T \boldsymbol{S} \boldsymbol{\beta} \,, \tag{1}$$

The nonnegativity constraint offers three advantages: a) it rules out extreme solutions that result from estimation error when dealing with the minimum-variance objective function in small samples (Jagannathan and Ma, 2003); b) it identifies noise forecasts, i.e., predictions that individually do not help to predict the target, but whose inclusion in the combination of forecasts is only justified by the objective of minimizing the *in-sample* variance of prediction errors[2]; and c) in this context, the estimator of the weights turns

[1] The inclusion of an intercept allows biased forecasts to be evaluated (Timmermann, 2006).

[2] It is well known that the inclusion of such forecasts can negatively impact the stability of the weights (Bunn, 1981).

into a special cases of the Lasso (Tibshirani, 1996), which permits us to extend the test for forecast encompassing to high dimensional problems by encouraging sparse portfolios (i.e., portfolios with only few active positions) (Brodie et al., 2009). Indeed, Fan et al. (2012) employed this correspondence to estimate approximate weights for the minimum variance portfolio under gross-exposure constraints when the number of securities under consideration is (potentially) larger than the sample size.

1.2 Testing for forecast encompassing

We next turn to design a test for multiple forecast encompassing. Noticing that weights are constrained to sum to one, we can rewrite (1) as a linear regression problem where the forecast error corresponding to the i^{th} forecast becomes the dependent variables and the remaining $\boldsymbol{V}_{\cdot,z}$ with $z = 1, 2, ..., d-1$ become the explanatory variables, i.e.,

$$V_d = \sum_{i=1}^{d-1} \beta_i \boldsymbol{V}_{\cdot,i} + \boldsymbol{u}, \quad \text{where} \quad \mathbb{E}[\boldsymbol{u}|\boldsymbol{V}_1, \boldsymbol{V}_2, ..., \boldsymbol{V}_{d-1}] = 0, \quad \mathbb{E}[\boldsymbol{u}\boldsymbol{u}^T] < \infty \;, \tag{2}$$

or alternatively,

$$V_i = \boldsymbol{V}_{\setminus i} \boldsymbol{\beta}_{\setminus i} + \boldsymbol{u} \;. \tag{3}$$

where $\boldsymbol{V}_{\setminus i}$ is the $n \times d-1$ matrix containing the remaining forecast errors and $\boldsymbol{\beta}_{\setminus d}$ is the $d-1$ vector of weights contains all but the weight corresponding to the forecast i^{th}. The weight associated to the i^{th} forecast can be trivially retrieved: $\beta_i = 1 - \boldsymbol{\iota}^T \beta_{\setminus i}$.

In the spirit of Harvey and Newbold (2000), we test whether the marginal contribution of forecast i to the aggregate ex post optimal forecast is statistically significant. However, this test must explicitly consider the constraints on the weights, i.e., we state the following hypothesis

$$H_0 : \theta = 0 \quad vs \quad H_1 : \theta > 0, \quad \theta_i \in \mathbb{R}^{+(r)} \;, \tag{4}$$

where we consider the following partition $\beta = (\theta, \gamma)$, with θ being the $r-$ dimensional parameter of interest and γ denoting the nuisance parameters (with dimension $d-r$). This corresponds to test whether the population ex post optimal weights lay on the boundary set (under H_0) or in the cone generated by the nonnegativity and sum to one constraints (under H_1). In this paper, we assume that $\epsilon \sim \mathscr{N}(0, \sigma_\epsilon)$ and we denote the sample negative log likelihood function as

$$\ell(\boldsymbol{\beta}) = -\frac{1}{n}\sum_{i=1}^{n} \log\mathcal{L}_i(\boldsymbol{\beta}) = \frac{n}{2} log(\pi\sigma_\epsilon) + \frac{1}{2}\frac{[\boldsymbol{y} - \boldsymbol{F\beta}]^T [\boldsymbol{y} - \boldsymbol{F\beta}]}{\sigma_\epsilon} \ , \tag{5}$$

with the corresponding likelihood ratio statistics defined by

$$T_{LR} = 2n\left(\ell(\widehat{\boldsymbol{\beta}}_{H_0}) - \ell(\widehat{\boldsymbol{\beta}}_{H_1})\right) \ , \tag{6}$$

or the asymptotically equivalent score statistics

$$T_S = \left[\nabla\ell(\widehat{\theta}_{H_0}) - \nabla\ell(\widehat{\theta}_{H_1})\right]^T \widetilde{\mathscr{F}}_{\theta|\gamma}^{-1} \left[\nabla\ell(\widehat{\theta}_{H_0}) - \nabla\ell(\widehat{\theta}_{H_1})\right] \ . \tag{7}$$

Under these assumptions, Kudô (1963) proved that, under the null hypothesis, the LR and score statistics associated to (4) are distributed as a mixture of chi-squared distributions. Determining the mixing weights is essential to compute the cumulative tail probability of the test statistics for a given significance level a. When the constrained space is the nonnegative orthant (or more generally when the constraints defining C are linear and independent), Silvapulle and Sen (2004) propose a simple quadratic programming problem to estimate ω (see proposition 3.6.1). When θ is univariate ($r = 1$), the weights are known in analytical form and the chi-bar-squared statistics reduces to

$$\text{Prob}\left[\overline{\chi}^2(\widetilde{\mathscr{F}}_{\theta|\gamma}, \mathbb{R}^+) > c_a\right] = .5 \ \text{Prob}\left[\chi^2(0) > c_a\right] + .5 \ \text{Prob}\left[\chi^2(1) > c_a\right] \ . \tag{8}$$

Finally, given a significance level a and its corresponding critical value c_a, we compute cumulative tail probability as:

$$\text{Prob}\left[\overline{\chi}^2(\widetilde{\mathscr{F}}_{\theta|\gamma}, \mathbb{R}^{+r}) > c_a\right] = \sum_{i=0}^{r} \omega_i \text{Prob}\left[\chi^2(i) > c_a\right] \ ,$$

and we reject H_0 if

$$\text{Prob}\,[T_S > c] > a\ , \quad \text{Prob}\,[T_{LR} > c] > a\ .$$

Large data set, heterogeneous modeling techniques and, different data pre-processing procedures originate a wide array of forecasts, whose size can potentially outclass the number of available realizations of the target variable. Therefore, a "modern" test for forecast encompassing should also deal with potential high dimensional frameworks. This test can accommodate these features using the decorrelated score procedure of Ning and Liu (2017) as described by Yu et al. (2019).

In the remainder of the paper, to ease presentation and simplify notation, we focus on the likelihood framework in a low dimensional setting and on the uni-variate parameter of interest θ ($r = 1$) corresponding to the j^{th} forecast. We refer to Appendix A and B for extension of the test in the high dimensional case and Appendix F for the power analysis and the study of the size properties of the test.

2 Analysis of inflation forecasts in the euro area

We next turn to our empirical study of the dynamics of yearly inflation forecasts in the euro area for the years ranging from 2009 to 2021. We have a total of ten forecasts: five are produced by institutions or professional forecasters, one based on a random walk (RM), one relies on autoregressive integrated moving average ($ARIMA$) models and three bi-variate vector-autoregressive (VAR) models. To ease the presentation, we classify forecasts into institutional and model-based. We follow Patton and Timmermann (2010) and use the December-on-December change of the realized value of HICP to measure yearly inflation. For each quarter, we consider inflation forecasts of the same calendar year. We avoid longer term forecasts (1 and 2 calendar years ahead) because institutional forecasters fail to systematically update their predictions beyond one calendar year ahead (Andrade and Le Bihan, 2013; Easaw and Golinelli, 2021).

Institutional forecasts data are retrieved from *European Central Bank staff macroeco-*

nomic projection for the euro area[3] and consist of forecasts produced by ECB staff, OECD , SPF, IMF and EC. Compared to standard forecasting exercises, in the case of institutional forecasts, the horizon is not fixed, e.g., 1 year ahead. In fact, forecasters predict HICP growth rate of a certain calendar year, e.g., yearly inflation rate in 2018. Institutional forecasters publish their forecasts at different dates within each quarter.[4] This may lead to overstating the relevance of the most recent forecasts. Therefore, we estimate the optimal weights and their corresponding statistics at each month within a quarter and then average them. This insures to observe each at least one publication of the inflation forecast every quarter while handling asynchronicity of the publication process.

Second, we introduce our ARIMA, VARs and RW forecasts. Following the standard notation for the ARIMA model (Lütkepohl, 2007), we denote with p, d and q the order of the autoregressive component, the differencing order and the moving average component. Formally:

$$\left(1-\sum_{i=1}^{p}\alpha_i L^i\right)(1-L)^d y_t = \left(1-\sum_{i=1}^{q} b_i L^i\right)\epsilon_t, \quad \epsilon_t \sim \; i.i.d.\; \mathcal{N}(0,\sigma^2)\,. \tag{9}$$

We consider three bi-variate VARs: a) we model the joint dynamics of monthly year-on-year (y-o-y) HICP growth rates and industrial production; and b), y-o-y HICP growth rates and y-o-y differences of unemployment rates; and c) y-o-y HICP growth rates and realized term spread of the euro area AAA government bonds at 3 months and 10 years maturity.[5] Specifications a) and b) reflect strategies that rely on the formulation of the generalized Phillips curve in term of measures of real activity or unemployment rate to forecast inflation in the spirit of Stock and Watson (1999). Specification c) implicitly assumes that forward looking asset prices and, specifically, the term structure of nominal interest rates contains relevant information to forecast inflation (Stock and Watson, 2003). The underlying economic rationale is that investment decisions rely on the real interest rate and, therefore, for nominal rates, a downward sloping yield curve at longer maturities

[3] Available at https://www.ecb.europa.eu/pub/projections/html/all-releases.en.html

[4] For example, during Q2, the European Commission publishes its forecasts in May while ECB in June.

[5] We approximate y-o-y growth rates using log differences. This simplifies the aggregation process from monthly y-o-y growth rates to the yearly inflation measure by simply computing the arithmetic average.

reflects expectations of a declining rate of inflation, while a steeply upward sloping yield curve indicates expectations of a rising rate of inflation (Mishkin, 1990, 1991). Data is retrieved from Eurostat. The inflation dynamics obtained from the bi-variate VAR(p) is:

$$\boldsymbol{y}_t = \boldsymbol{\alpha} + \sum_{i=1}^{p} \boldsymbol{A}_i \boldsymbol{y}_{t-i} + \boldsymbol{\epsilon}_t \sim \ i.i.d.\ \mathcal{N}(0, \boldsymbol{\Sigma}) \ , \tag{10}$$

where $\boldsymbol{y}$ is the 2×1 vector containing inflation rates and either y-o-y industrial production growth rate or y-o-y differences of unemployment rate, $\boldsymbol{\alpha}$ is the 2×1 vector of the intercept and $\boldsymbol{A}_i$, $i = 1, 2, ..., p$ are 2×2 matrices of coefficients.

We estimate the ARIMA and bi-variate VARs via maximum likelihood and considering a rolling window of 48 months.[6] For each of the considered windows, we control for the stationarity of y-o-y HICP growth rate and select the appropriate differencing order d by performing the augmented Dickey Fuller test. Similarly, we select the appropriate (p, q) order of the ARIMA and p order of the VAR using Akaike information criterion (AIC). We simulate a real-time forecasting exercise with ARIMA and VARs model. For each month t, we assume that we observe the realization of the variables of interest at $t-1$ and forecast yearly inflation by compounding the realized values of y-o-y inflation for the year of reference with forecasts of the $12 - t + 1$ months ahead. For example, in March 2020, we assume to observe the realization of HICP index in February and January 2020 and forecasts y-o-y inflation for March and the remaining 9 month ahead. Then, we obtain the forecast for yearly inflation by averaging the combined realizations and forecasts of the y-o-y inflation for the year of reference. We construct inflation forecasts using a random walk model in the same manner.

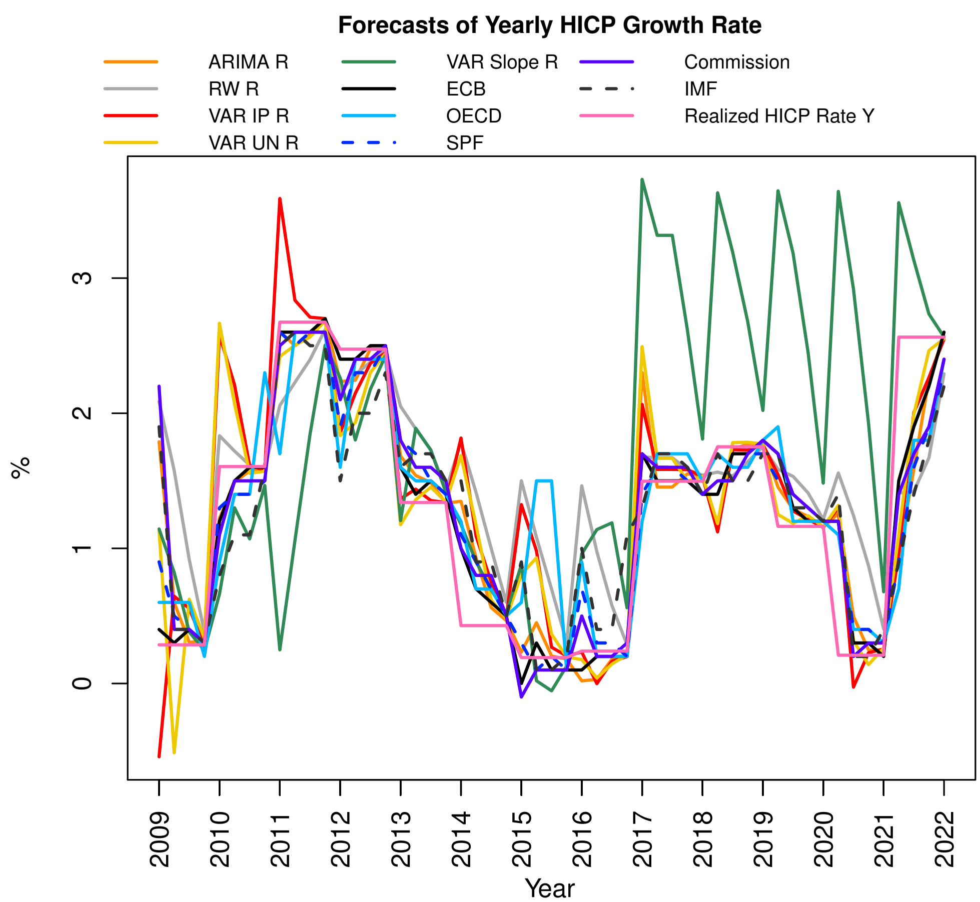
Table 1 reports the summary statistics and Figure 1 depicts the dynamics of the yearly HICP growth rate and its forecasts. On average, both institutional and model-based forecasts are unconditionally unbiased and track the dynamics of yearly inflation rate but,

[6] We have also considered rolling windows of 24, 60 and 72 months as well as expanding windows. We chose the window of 48 because it is the one the offers the best performance in term of root mean square error (RMSE).

Table 1: Summary statistics of the forecasts and realized forecast errors.

	Forecast Errors									
	ARIMA	RW	VAR IP	VAR Un	VAR Slope	ECB	OECD	SPF	Commission	IMF
Min	-1.50	-1.80	-1.39	-1.26	-3.43	-0.99	-1.31	-0.99	-1.91	-1.61
Max	1.62	1.15	1.60	1.46	2.42	1.06	1.86	1.66	1.16	1.66
Median	-0.01	-0.17	-0.03	-0.02	-0.32	0.00	-0.04	-0.01	-0.01	-0.01
Mean	-0.03	-0.25	-0.09	-0.07	-0.51	0.00	-0.04	0.00	-0.03	-0.02
Standard Deviation	0.44	0.61	0.49	0.48	1.06	0.28	0.52	0.39	0.43	0.56
	Forecasts									
	ARIMA	RW	VAR IP	VAR Un	VAR Slope	ECB	OECD	SPF	Commission	IMF
Min	0.02	0.29	-0.54	-0.51	-0.05	0.00	0.10	0.10	-0.10	0.10
Max	2.69	2.62	3.59	2.68	3.73	2.70	2.60	2.60	2.60	2.60
Median	1.45	1.58	1.37	1.34	1.73	1.40	1.40	1.40	1.50	1.40
Mean	1.30	1.52	1.36	1.33	1.77	1.27	1.31	1.27	1.30	1.28
Standard Deviation	0.80	0.57	0.88	0.82	1.10	0.83	0.71	0.74	0.79	0.68

Figure 1: Dynamics of realized and forecasted HICP yearly growth rate.



looking at Figure 1, it is worth to underline three exceptions. First, forecasts based on the bi-variate VAR including industrial production tend to be more volatile, especially at the beginning of each year. This is particularly evident for VAR Slope that relies on the term spread of the nominal yields: its forecast is biased in the 2017-2021 period. Second, most forecasts tend to overestimate yearly inflation from 2015 to 2017. Third, with the exception of VAR Slope, forecasts tend to either jointly overestimate or jointly underestimate inflation in the euro area, implying forecast errors to be cross-sectionally dependent.

2.1 Testing and handling cross-sectional dependence

Strong cross-sectional dependence of forecast error may signal potentially multi-collinear forecast errors and, consequently, poor estimates of the optimal weights. To quantify the degree of cross-sectional dependence of forecast errors, we borrow the $CD-$test (Pesaran, 2021) from the panel econometrics literature, which is defined as:

$$CD = \sqrt{\frac{2T}{N(N-1)}} \sum_{i=1}^{N-1} \sum_{j=i+1}^{N} \hat{\rho}_{ij}, \tag{11}$$

with $\hat{\rho}_{ij}$ being the sample correlation coefficient between the forecast error of model i and j. In Bailey et al. (2016), we can find the order property of the average correlation coefficient:

$$\bar{\rho}_N = \frac{2}{N(N-1)} \sum_{i=1}^{N-1} \sum_{j=i+1}^{N} \rho_{ij} = O(N^{2\alpha-2}), \tag{12}$$

where $\alpha \in [0, 1]$. First, as recommended in Elhorst et al. (2020), we control that the null of weak cross sectional dependence ($\alpha < 0.5$) cannot be rejected through the CD-test ($\hat{\bar{\rho}}$ and CD-statistic's columns in Table 2). We cannot reject weak cross sectional dependence given the high value of $\hat{\bar{\rho}}$. Second, we assess the degree of cross sectional correlation through the two step procedure as in Bailey et al. (2016).

In Table 2 we show that the cross sectional correlation of forecast errors is really strong: the estimated α is larger than .75, clearly pointing at the presence of common components driving the dynamics of forecast errors. We can reach similar conclusions by looking at the loadings of the first Principal Component shown in Table 2: as expected in such a case, they are evenly distributed among the different forecast errors, supporting the evidence of a persistent interdependence trend between inflation forecast errors.

Table 2: Principal component analysis of realized forecast errors and cross sectional dependence test.

Forecaster	PC1	PC2	PC3	PC4	PC5
ARIMA R	0.360	0.002	-0.139	0.409	-0.305
RW R	0.346	0.060	-0.012	-0.433	-0.305
VAR IP R	0.164	-0.687	0.179	0.030	0.502
VAR Un R	0.248	-0.502	0.381	0.074	-0.513
VAR Slope R	0.144	0.457	0.844	0.110	0.144
ECB	0.366	0.015	-0.010	0.084	0.073
OECD	0.338	0.065	-0.031	-0.716	0.053
SPF	0.380	0.027	-0.128	0.103	0.108
Commission	0.353	0.222	-0.241	0.301	-0.069
IMF	0.353	0.096	-0.124	0.082	0.505
Standard deviation	2.58	1.22	0.87	0.58	0.53
Proportion of Variance	66.75%	14.84%	7.56%	3.39%	2.84%
Cumulative Proportion	66.75%	81.59%	89.15%	92.54%	95.38%

$\hat{\bar{\rho}}$	CD-statistic	$\hat{\alpha}$	$\hat{\sigma}_{\hat{\alpha}}$
0.585	28.301	0.8985	0.1944

Notes. We report the estimated average sample correlation forecast errors of model i and j, $\hat{\bar{\rho}}$, the cross-sectional test statistics, $CD-$statistics, as well as $\hat{\alpha}$ and its standard error, $\hat{\sigma}_{\hat{\alpha}}$.

2.2 Estimating the optimal combination weights

We estimate the forecast combination that minimizes the variance of the combined forecast error restricting the weights to be nonnegative, to sum to one and the whole sample. We start by considering the sample maximum likelihood estimator of the covariance matrix in *Specification 1*. Indeed, as underlined in Jagannathan and Ma (2003) and Conflitti et al. (2015), restricting weights to lay in the unit simplex helps reduce the risk in estimated optimal combination of forecasts. However, being a special case of the Lasso (Fan et al., 2012), this estimator often fail to correctly select the best subset of forecasts when forecasts errors are highly correlated (Zhao and Yu, 2006). For this reason, we consider three additional estimators of the combination weights as robustness checks.

Table 2 suggests the presence a strong one-factor structure characterizing the dynamics of inflation forecast errors. Following this observation, in *Specification 2*, we consider the principal orthogonal complement thresholding estimator of the covariance matrix (POET) of Fan et al. (2013) who assumes that, conditional on pre-specified common factors, the residual terms are weakly correlated.

Finally, we can handle the strong cross-sectional dependence by also imposing some structure on the estimator of the optimal weights. Specifically, we employ a consistent shrinkage estimator of the covariance matrix that consist of combining the sample covariance matrix (which can be easily computed and is asymptotically unbiased but prone to estimation error) with a highly structured estimator (that contains relatively little estimation error but potentially misspecified). Being the loadings of the first principal component are roughly the same (Table 2), this factor can further be seen as an equally weighted portfolio of all forecasts (up to a scaling factor). Therefore, it becomes natural to shrink the sample covariance matrix towards a diagonal one (Ledoit and Wolf, 2003) that is representative of this case. We estimate the weights adopting such estimator in *Specification 3*.[7]

Table 3 reports the weights, the T_S score statistics defined in (2) and their corresponding p-values. We find that ECB has the greatest weight across specifications, however its forecasts do not encompass its competitors regardless of the estimation strategy. Indeed, Var IP, Var Un and European Commission's forecasts are always included in the optimal combination of forecasts and their corresponding weights are statistically greater than zero. Instead, ARIMA model, IMF, RW and VAR Slope performs poorly, being either excluded or only marginally (and insignificantly) contributing the minimization of the aggregate prediction error.

However, Table 3 implicitly assumes that forecasts' performance remains constant across the whole sample. We challenge it by re-estimate the forecast combination considering a rolling window of 12 quarters.[8] Figure 2 displays the dynamics of the average weights across

[7] Following Ledoit and Wolf (2003), the weight between the sample and the target matrix is computed by minimizing the Frobenius norm of the difference between the shrinkage estimator and the asymptotic covariance matrix.

[8] As a robustness check, we have considered shorter window of 10 quarters. Results are comparable despite the weights display a more evident instability due to the small sample: in these cases, only ten observations

Table 3: Ex post optimal weights, score statistics and p-values.

		ARIMA	RW	IP	Un	Slope	ECB	OECD	SPF	EC	IMF
Spec. 1	β	0	0	0.140	0.032	0.032	0.651	0	0	0.145	0
	T_S	.	.	4.791	4.992	2.336	5.100	.	.	6.339	.
	$P\bar{\chi}^2$	.	.	0.014	0.013	0.063	0.012	.	.	0.006	.
Spec. 2	β	0.054	0	0.118	0.091	0	0.455	0	0.170	0.088	0.023
	T_S	2.157	.	4.357	5.450	.	3.685	.	2.985	4.646	2.275
	$P\bar{\chi}^2$	0.071	.	0.018	0.010	.	0.027	.	0.042	0.016	0.066
Spec. 3	β	0	0	0.147	0.055	0.033	0.553	0	0.093	0.119	0
	T_S	.	.	4.602	5.197	2.531	3.948	.	2.630	5.561	.
	$P\bar{\chi}^2$	.	.	0.016	0.011	0.056	0.023	.	0.052	0.009	.

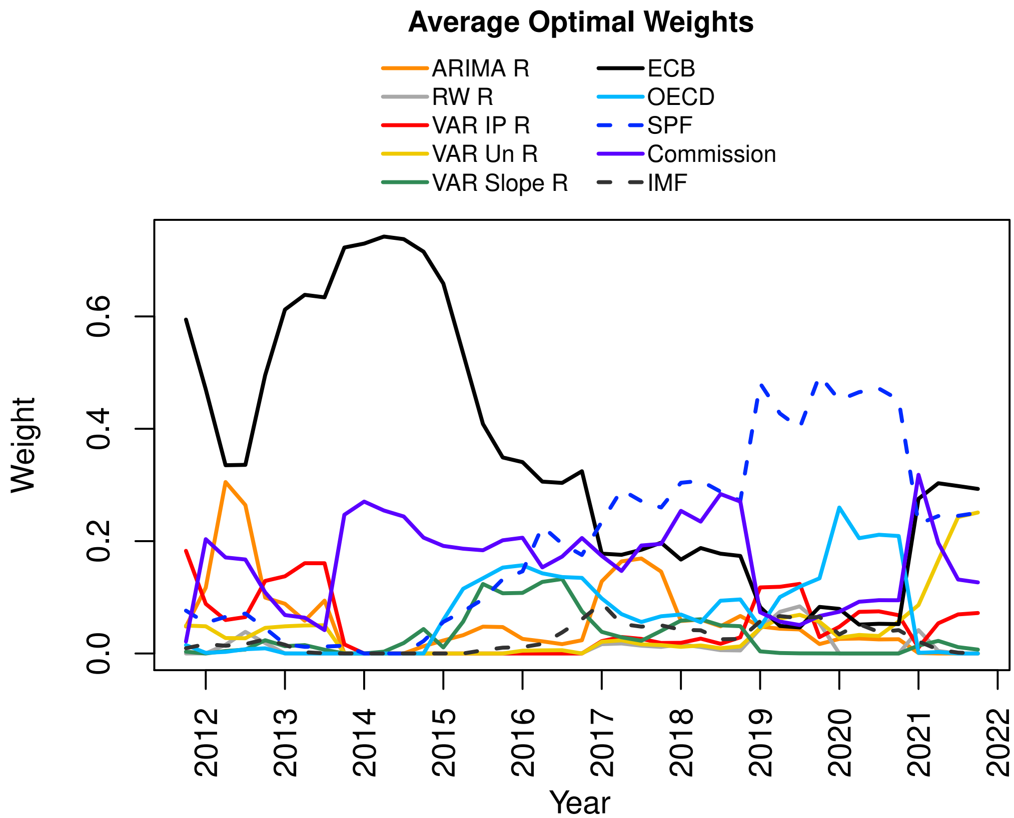
Notes. The sample covers the period 2012 - 2021. Spec 1: Sample Covariance, Spec 2: POET, Spec 3: L&W shrinkage toward diagonal matrix. Score statistics are computed using the realized forecast errors.

specifications in the period Q1 2012 - Q4 2021. The x- axis identifies reference period. For example, the interval 2015-2016 identify the weights corresponding to the yearly inflation in 2015.

In Table 3, we see that on average ECB's weight is greater than its competitors' ones. This is not surprising: ECB's primary objective consists in maintaining price stability by tailoring monetary policy to make sure that inflation remains low, stable and predictable. Therefore, we expect ECB's inflation forecasts to contain credible information about the expected dynamics of inflation, reflecting its commitment to maintain price stability in the medium term. However, Figure 2 points out that weights also greatly vary over time. Despite ECB generally receives the greatest weight, its weight is not consistently greater than all remaining forecasts in each sub period. The weights associated to the other forecasts are also fluctuating over time. These findings underline the importance of the considered reference period and how different conclusions can be drawn when evaluating unconditional forecast encompassing or conditioning on some economic states. The varying underlying process driving HICP in the euro area, the heterogeneity of forecasters' loss functions and heterogeneous information sets can all be plausible explanations on the reason why optimal weights change over time. Therefore, it is compelling to explore whether we can identify these *states* and analyze how the macroeconomic or financial conditions affects

are available to optimally combine ten forecasting methods. Results are available upon request.

Figure 2: Dynamics of average optimal combination weights across specifications.



the informativeness of ECB's forecasts, and therefore ECB's ability at tracking expected inflation.

3 The Dynamics of the ECB's weights

From the forecast encompassing perspective, the weights reflect the contribution of the forecasts to the forecast that minimizes the variance of the aggregate prediction error. Therefore, a time-varying optimal weight indicates that the informational value added by the forecast of interest is equally non-constant over time.

Time varying weights for model-based forecasts could reflect the uncertainty surrounding the forecasting method as, for example, the uncertainty around the choice of (p, d, q) in ARIMA models, the size of the estimation window or the presence of structural breaks in relation between the time-series in the bi-variate VARs. These issues are directly connected to notion of *heterogeneity* of macroeconomic time-series (Clements and Hendry, 1998), i.e., the non-constant underlying process generating of the series to forecast. Heterogeneity also matters for institutional forecasters. Being supported by econometric models and expert judgment, the predictive performance of institutional forecasters can be affected by the forecasting techniques used and the turnover of forecasters within the same institution.

The loss function of monetary authority has also important consequences on the inflation forecasting. For example, Capistrán (2008) argues that the FED's systematic under-prediction and over-predictions of inflation can be explained by the variations of the cost of having inflation above an implicit time-varying target. Furthermore, while all institutional forecasts are certainly useful for consumers and investors, ECB's macroeconomic projections (of which their inflation forecasts are part of) are their primary support for the assessment of economic conjecture and risks to price stability which is the primary objective of ECB's mandate. Therefore, we expect ECB's forecasts to have a direct effect on monetary policy implemented in the euro area: when the expected inflation is low, we expect the monetary authority to try to keep it low to pursue its goal in the medium term. However, policymakers face a dilemma when expected inflation is high: they would like to disinflate, but fear the costs related to a tightening policy (Ball, 1992). These costs do depend on the state of the economy. In sluggish economic growth, high indebtedness and a vulnerable financial system can strongly amplify the negative effects of the recession that would result from such contractionary policies. In that case, the forecasting community does not know how promptly the monetary authority will react and thus the uncertainty surrounding the inflation dynamics will rise.

For these reasons, understanding the determinants of ECB's weight dynamics is compelling for macroeconomic analysis. Studying in what economic states ECB brings more information naturally helps forecast users that aim at monitoring and improving forecast performance but it also helps policy analysts interested in evaluating ECB's intents at monitoring inflation. We consider determinants reflecting both the uncertainty surrounding the HICP dynamics and macro-financial factors that may affect ECB's loss functions.

3.1 The determinants of the combination weights

Identifying the exact determinants of the weights is not an easy task. We do not directly observe whether (and when) the HICP time series has experienced structural breaks and how the current (and expected) economic conjecture and risks to price stability have shaped forecasters' loss functions. Nevertheless, we can identify group of variables that might *indirectly* affect the uncertainty around the expected path of the HICP.

3.1.1 Disagreement and inflation surprise

Following Mankiw et al. (2004) and Manzan (2011), we consider the dispersion of (point) forecasts from the survey of professional forecasters to proxy forecasting uncertainty. Patton and Timmermann (2010) empirically shows how this measure mostly reflects heterogeneity in the priors or underlying forecasting models compared to heterogeneity in information signals. Dovern et al. (2012) confirm that disagreement about inflation forecasts increases with uncertainty about the actual series, but also find that disagreement about prices rises with their level and, in line with the thesis of Rogoff (1985) and Alesina and Gatti (1995), disagreement declines under independent central banks that promote price stability via introduction of clear mandates in terms of price stability or the adoption of more predictable monetary policy with increased and improved communication. To study whether these aspects influence the dynamics of the ex post optimal weights, we include a measure the difference between the the realized y-o-y monthly HICP growth rate and the 2% target (*Difference from 2%*) and the implied dispersion of (point) forecasts from the survey of professional forecasters *Disagreement SPF*. *Disagreement SPF*, is measure as the standard deviation of point forecasts from ECB Survey of Professional Forecasters.[9] We also control for a broader measure of policy-related economic uncertainty at European level, employing the news-based Economic policy uncertainty index of Baker et al. (2016) in our analysis.

Supply shocks, i.e., the sudden change of the supply of a product or commodity, resulting in an unforeseen change in price, e.g., the surge of gas and oil price following the Ukrainian

[9] Results are comparable when we considered alternative measures of dispersion for our measure. Specifically, we also considered interquartile or the 95^{th} - 5^{th} range for the *Disagreement SPF*.

war directly affect inflation forecasting performance and, indirectly, the optimal weights. Some forecasters could better at anticipating or more promptly reacting to recent events because of informative advantages or because facing lower revision costs (Ehrbeck and Waldmann, 1996). We proxy the dynamics of inflation surprise with the first principal component of realized forecast errors (*Inflation surprise*).

3.1.2 Financial conditions and monetary policy

ECB's primary objective is to maintain price stability by aiming for 2% inflation over the medium term. However, other objectives can be pursued "without prejudice to the objective of price stability".[10] These objectives include among the others the stability of the financial system, the risk of financial fragmentation and the risk of euro break up. In this respect, following the outburst of European sovereign debt crisis, the ECB has taken a range of actions beyond monitoring price stability to address bank funding problems, eliminate excessive risk in sovereign markets, and safeguard monetary transmission (Gross and Zahner, 2021). These measures have been successful at monitoring these risks when the euro area was experiencing low (or even negative) inflation (Candelon et al., 2022; De Grauwe and Ji, 2022). However, Russia's invasion of Ukraine, amid an already slowing recovery from the pandemic, raises new financial stability risks, poses questions about the longer-term impact on economies and the surging commodity prices pose challenging trade-offs for central banks (IMF, 2022a). In the current situation, it is still not clear how the ECB would react when facing such dilemma.

We analyze whether the past dynamics of the optimal weights was affected by measures of financial stability in the euro area. Specifically, we consider the *global interdependence* and *intra-European fragility* factor of Candelon et al. (2022). The first factor captures shifts in the level of the entire cross section of long term spreads, and it is related with factors that capture devaluation risk (De Santis, 2019), flight-to-liquidity mechanisms (Monfort and Renne, 2013) and shifts global risk aversion (Favero, 2013), further supporting the use of the first PC as global trend. The second factor measures instead how deep the difference

[10]See the ECB's website for more details `https://www.ecb.europa.eu/mopo/intro/html/index.en.html`

between fragile and financially sounder economies is.[11]

For an exhaustive analysis of financial condition is the euro area, we also include proxies mirroring the monetary policy of ECB. We consider the *Shadow* rate of Wu and Xia (2020) and the *APP Factor* that summarizes Eurosystem holdings corresponding to ECB's Asset Purchase Programme (APP). [12]

3.2 Seasonality of institutional forecasts

When evaluating institutional forecasts, the literature often overlooks that forecasts always refer to the average of the calendar year.[13] In this framework, seasonal patterns in the dynamics of the estimated weights might emerge and ignoring them could result in omitted variable bias when studying the relation with their determinants. Therefore, in the spirit of Lovell (1963) and Saikkonen and Lütkepohl (2000), we complete our set of determinants with three seasonal dummy variables corresponding to the second, third and fourth quarter of the year. This seasonal adjustment implicitly assumes that the intercept of the regression function shifts each quarter and it is convenient to study how the weights are affected when approaching inflation releasing date.

3.3 Controlling for unemployment and industrial production

When studying the dynamics of ECB's weight, we should control for quarterly unemployment and quarterly year-on-year industrial production growth rate. In fact, while both series are undoubtedly important to represent economic conditions in the euro area, they are also commonly employed in the evaluation of inflation dynamics and monetary policy.

On the one hand, despite monetary policy in the euro area is more complicated than what the conventional Taylor rule might suggest, Gorter et al. (2008) find that the ECB

[11]The factors are extracted using principal component analysis using the samples described in Candelon et al. (2022).

[12]We compute the factor by extracting the first principal component of the Eurosystem' holdings purchased under one or more of the asset purchase programmes operated by ECB. Specifically, APP includes corporate sector purchase programme (CSPP), public sector purchase programme (PSPP), asset-backed securities purchase programme (ABSPP) and third covered bond purchase programme (CBPP3). For data and more information, we refer to `https://www.ecb.europa.eu/mopo/implement/app/html/index.en.html`

[13]In the first quarter of the year, the current-year forecast is a four quarters ahead, in the remaining quarters of the year, the horizon becomes three, two and one quarter(s) ahead, respectively.

considers *expected*, i.e., forecasted, inflation and output growth when setting interest rates. [14] On the other hand, it is not uncommon to find reformulation of the generalized Phillips curve in term of industrial production or unemployment rate (Stock and Watson, 1999).

Table 4 summarizes the variables, and the corresponding abbreviations and the sources. Before introducing the modeling framework, it is compelling to inspect the dynamics of the determinants themselves to analyze whether potential issues due to multicollinearity may arise.

Table 4: Names and abbreviations of variable employed for the analysis of the determinants of the ex post optimal weight.

Variable	Measure of	Abbreviation	Source
Difference with 2%	Uncertanty	Dif. 2%	Eurostat
Disagreement SPF	Uncertainty	Dis.	ECB
Economic Policy Uncertainty Index	Uncertainty	EPU	Baker et al. (2016)
Shadow Rate	Monetary Policy	S	Wu and Xia (2020)
Asset Purchasing Program factor	Monetary Policy	APP	ECB
Cross-sectional forecast error factor	Inflation Surprise	Inf. Sur.	Eurostat & ECB
Global Interdependence factor	Financial Conditions	Glob.	Candelon et al. (2022)
Fragility Factor	Financial Conditions	Frag.	Candelon et al. (2022)
Industrial Production	Macroeconomic Conditions	IP	Eurostat
Unemployment rate	Macroeconomic Conditions	Un	Eurostat
Seasonal dummy Q2	Seasonality & Publ. gap	d_{Q2}	·
Seasonal dummy Q3	Seasonality & Publ. gap	d_{Q3}	·
Seasonal dummy Q4	Seasonality & Publ. gap	d_{Q4}	·

Table 5 represents the correlation between the determinants in the sample 2021 Q1 - 2021 Q4. We remark how the dynamics of the shadow rate and unemployment rate are strongly positively correlated (over .9). Similarly, the Economic Policy Uncertainty Index is negatively correlated with the shadow rate and unemployment rate ($-.63$ and $-.54$, respectively). As a result, we expect that including the unemployment rate between the determinants, may be detrimental for the precision of the estimated coefficients.

We next introduce the regression model used to study the dynamics of ECB's weights.

[14] Contrasting with more conventional specifications of the Taylor rule where realized inflation and output are used.

Table 5: Correlation between the determinants in the period Q1 2012 - Q4 2021.

	APP	Un	IP	Dis.	S.	Glob.	Fr.	Dif. 2%	Inf. Sur.
EPU	.25	-.54***	-.17	.15	-.63***	.05	.01	.01	-.14
APP Factor		-.11	.18	.01	-.18	.51***	.08	-.34**	-.19
Unemployment			.09	.18	.97***	-.37**	.22	-.10	.14
IP				-.32***	.09	.45***	.12	.15	-.18
Disagreement					.03	-.10	.23	-.36***	-.01
Shadow						-.43***	.14	.04	.09
Global Inter.							-.17	-.47***	.03
Fragility								.09	-.03
Difference 2%									-.51***

Notes. The marks $***$, $**$, $*$ display 1%, 5% and 10% significant level, respectively.

3.4 The regression framework

We study the relationship between the ECB's weights $\boldsymbol{y}$ and their determinants $\boldsymbol{X}$. Assuming $y_t \sim \mathcal{N}(x_t\beta, \sigma)$, we can estimate the following linear regression model via maximum likelihood.

$$\boldsymbol{y} = \boldsymbol{X}\boldsymbol{\beta} + \boldsymbol{u} \,. \tag{13}$$

Combination weights are, nevertheless, defined in the closed interval $[0, 1]$, therefore predictions arising from a linear regression framework may lead to negative weights or weights greater than unity. Despite it is possible to adapt the standard linear regression framework to accommodate bounded stochastic processes by transforming the original weights to have values in the real line[15], this approach has two main shortcomings. First, the regression parameters are interpretable only in terms of the mean the transformed weights. Second, data generated by bounded stochastic processes are generally skewed and prone to heteroskedasticity.

As robustness check, we also consider a beta regression models, which is instead tailored to model variables in the interval $(0, 1)$.[16] Specifically, we employ the beta regression framework proposed by Ferrari and Cribari-Neto (2004) that assumes a) independent realizations of $y_i \sim \mathscr{B}(\mu_i, \phi_i)$, b) a linear regression model for the transform of the mean

[15] For example via standard *logit* or Box's transformations.

[16] Following Smithson and Verkuilen (2006), we consider the transformation $\frac{y(n-1)+0.5}{n}$ where y is the dependent variable and n is the sample size to adapt the beta regression framework to dependent variables assuming also the extremes 0 and 1.

parameter μ and c) a constant precision parameter ϕ. In this framework, we consider a *logit* link function for the mean of y, leading to

$$\mu_i = \frac{e^{x_i^T \beta}}{1 + e^{x_i^T \beta}} \,. \tag{14}$$

We refer to Appendix C for more details on the beta regression framework.

To tackle potential problems arising from potential heteroschedastic and autocorrelated residuals, we report four-lag Newey-West (NW) standard errors (Newey and West, 1987) and a wild bootstrap procedure as in Mammen (1993) both in the linear and beta regression framework.[17]

3.5 Empirical results

We present the results on ECB's determinants obtained using the linear regression model and beta regression estimated via maximum likelihood in the Table 6. We report the results from the optimal weights obtained in *Specification 2* as the POET estimator of the covariance matrix (Fan et al., 2013) offers better small sample properties compared to the sample covariance matrix.[18]

In panel *a)* we report the main results, while in panel *b)* we also control for unemployment and industrial production. Table 6 displays the estimated coefficients, their standard errors and the p-values implied by testing whether the coefficient of interest is statistically different from zero. The columns NW and B report whether the coefficient of interest is statistically significant at 1% (***), 5% (**), 10% (*) level according to the Newey-West standard errors and the Mammen wild boostrapped quantiles, respectively.

We find that the measures of uncertainty (*Dis.* and *EPU*), how actual inflation differs from the 2% target (*Dif. 2%*) and *Inflation Surprise* and the *Global Inter.* are all important drivers of ECB's weights. They are statistically significant at 5% level also when controlling for unemployment and industrial. Ceteris paribus, the greater these factors, the lower the expected weight.

[17] We refer to Appendix D for a detailed description of the bootstrap algorithm.

[18] We report the results for the sample and shrinkage estimator of the covariance matrix in Appendix E.

The determinants referring to ECB's monetary policy, i.e., the APP factors (APP) and the shadow rate (S) influence ECB's informativeness in an interest way: ceteris paribus, the greater the shadow rate, the greater the weight and the greater the APP factor, the lower the weight. This can be expected: higher interest rates signal the intention of central banks slow down inflation dynamics; while, on the contrary, increasing amount of holdings purchased under the Asset Purchase Programme and the corresponding liquidity injections decrease ECB's weight. The significance of the contribution of ECB's policy to the informativeness of its forecasts is challenged when controlling for unemployment and insutrial production. However, this issue must be expected: despite the estimated coefficients do not vary substantially, the precision of the estimators is strongly affected by the multicollinearity with unemployment and industrial production as noticed in subsection Table 5. Indeed, both IP and $UnEmp$ itself are not statistically different from zero at any significance level. We also notice that including IP and $UnEmp$ does not increase the explanatory power of the model (similar adjusted and pseudo R^2).

3.6 Discussion: forecast informativeness and the goal of monetary policy

In our framework, weights reflect how informative forecasts are: an upward (downward) dynamics of the weight signals that the considered forecaster has become more (less) informative when producing inflation forecasts. Being more informative implies that the public will regard an institution featuring a greater weight as more relevant when forecasting the variable of interest. In subsection 3.5, the greater the difference between current inflation and 2% target, the greater the inflation surprise, the less informative ECB's forecasts are.

Clearly, as a strategy for conducting monetary policy, having informative forecasts is crucial. First, since realized inflation typically responds to changes in monetary policy only with a substantial lag (from one to two years or more), having reliable information about the expected path of price offer a substantial timing advantage when designing monetary policy. Second, Svensson (1997) showed that inflation-targeting also implies expected inflation-targeting: the central bank's inflation forecast becomes an explicit intermediate target. This should simplify both implementation and monitoring of monetary policy, helping the

Table 6: Determinants of the weights constructed using the POET estimator of the covariance matrix of prediction errors. .

	a) Determinants						b) Determinants + Un and IP					
	LM	NW	B	Beta	NW	B	LM	NW	B	Beta	NW	B
b_0	0.879	***	***	2.090	***	***	0.241			-2.296		
	(0.08)			(0.44)			(0.54)			(3.01)		
Dif. 2%	-0.147	***	***	-0.622	***	***	-0.130	***	***	-0.598	***	***
	(0.03)			(0.14)			(0.03)			(0.13)		
Dis.	-0.187	***	***	-1.225	***	***	-0.237	***	***	-1.466	***	***
	(0.04)			(0.25)			(0.07)			(0.26)		
EPU	-0.001	**	***	-0.004	**	**	-0.001	**	***	-0.006	***	*
	(0.0004)			(0.0019)			(0.0005)			(0.0022)		
S	0.063	***	***	0.351	***	***	0.025			0.072		
	(0.01)			(0.04)			(0.03)			(0.18)		
APP	-0.025	*	***	-0.101	*	**	-0.027	*	**	-0.100		*
	(0.01)			(0.06)			(0.01)			(0.06)		
Inf. Sur.	-0.034	***	***	-0.160	***	***	-0.038	***	***	-0.180	***	***
	(0.01)			(0.06)			(0.01)			(0.06)		
Glob.	0.016	**	***	0.106	***	***	0.019			0.072		
	(0.01)			(0.03)			(0.01)			(0.05)		
Fr.	0.019			0.160	**	***	0.015		*	0.113	*	**
	(0.02)			(0.07)			(0.01)			(0.06)		
Un.	.	.	.	.	.	.	0.060			0.392		
							(0.05)			(0.25)		
IP	.	.	.	.	.	.	-0.004			0.018		
							(0.0046)			(0.023)		
d_{Q2}	-0.045			-0.232			-0.047			-0.229		
	(0.03)			(0.16)			(0.03)			(0.16)		
d_{Q3}	-0.060			-0.238			-0.069			-0.226		
	(0.04)			(0.19)			(0.04)			(0.19)		
d_{Q4}	-0.074	**		-0.401	**	*	-0.075			-0.334		
	(0.04)			(0.19)			(0.04)			(0.22)		
Adj. R2	0.84			0.905			0.844			0.903		
N. obs.	38			38			38			38		

Notes. The first columns displays the determinants, b_0 is the intercept. LM and Beta report the estimated coefficients using the linear regression model and the beta regression. While we display Newey-West standard errors in the parenthesis, NW and B report whether the coefficient of interest is statistically significant at 1% (***), 5% (**), 10% (*) level according to the Newey-West standard errors and the Mammen wild boostrap quantiles, respectively.

inflation-targeting central bank and the public to tell whether the monetary policy is indeed fulfilling its mandate to maintain price stability (Bernanke and Mishkin, 1997). However, this could equally have serious adverse consequences for the central bank's accountability and credibility[19] if the central bank's forecasts were believed to be inconsistent with the

[19] The term "credibility" refers to the degree of confidence that the public has in the central bank's determination and ability to meet its announced objectives.

expected dynamics of inflation. In this perspective, low weights relative to their historical trend may be interpreted as warning flags, signaling that ECB's credibility is at stake. As observed in subsection 2.2, we find two main dips of ECB's weight: the first between in 2012, the second between the beginning of 2019 and the end 2020 (see Figure 2).

In 2012, the ECB announced and implemented a set of conventional and unconventional tools to address the difficulties in the transmission of the monetary policy measures, the onset of a credit crunch, and the risk of deflation following the outburst of the European sovereign debt crisis. In fact, ECB's effort on the conventional and unconventional monetary easing carried out, was also motivated by reverting the deflationary trend in the euro area in the period 2014-2018, therefore pushing inflation back to its target.

The period 2019-2021 saw the outbreak of the Covid, the most severe shock to the European economy since WWII with significant disruptions to the labor market, supply chains and posing serious challenges to economic growth and financial stability (IMF, 2022b). Also, the handover from Mario Draghi to Christine Lagarde (November 2019) at the head of the European Central Bank could have brought into question ECB's new policy direction.[20] Nevertheless, the recent economic and geopolitical developments strongly contrasts to the deflationary trend observed in the euro area after the outburst of the European sovereign debt crisis. In fact, despite we can observe a mild rise of ECB's weight in 2021, inflation dynamics had not been yet impacted by the global supply chain disruption following the surge of gas and oil price following the recent war in Ukraine. Indeed, despite this event directly raises expected inflation dynamics, it is not in the data since we do not observe the realized yearly inflation rate for 2022. Eurostat's release of the realized year on year euro area annual inflation rate in April 2022 shows how severe the shock to inflation is: annual inflation was 7.4% in March 2022, up from 5.9% in February. The new release of Survey of Professional Forecasters points in the same direction: HICP inflation expectations are

[20] For example during a press conference on 12 March 2020, Mrs Lagarde suggested it was the responsibility of governments to protect highly indebted eurozone countries rather than the central bank, triggering public's questions on whether the bank would renew the "whatever it takes" policy to protect the eurozone from a recession triggered by the coronavirus outbreak (the verbatim of the speech by Christine Lagarde is available at https://www.ecb.europa.eu/press/govcdec/mopo/html/index.en.html). This resulted in a sudden increase in European sovereign bond spreads, leading ECB's president to walk back her previous comments and restating ECB's fully commitment to avoid any fragmentation in bond markets.

revised upward to 6% for 2022 and 2.4% for 2023 but unchanged at 1.9% for 2024. Nevertheless, ECB's policy rates remain currently unchanged and, following Governing Council's forward guidance, any adjustments to the key ECB interest rates will be gradual and will take place some time after the end of the asset purchasing program expected in Q3 2022.

ECB's decision not to promptly adjust its policy to the surge in forecast inflation in 2022 may seem surprising and strongly contrasts with Bank of England and U.S. Federal Reserve policies that instead vowed tough action on inflation, announcing increasing rates until inflation comes under control. However, as underlined in IMF (2022b) or in the recent talk of the director general of the Bank of International Settlements[21] ECB's policy is currently on the edge and more restrictive measures are required to take actions against the steady increase of price levels. Miranda-Agrippino and Nenova (2022) have shown that ECB's restrictive monetary policy should have similar effect that the one implementing at the moment in the U.S.; however, in the euro area it also may create tensions for the banking sector of peripheral countries (Soenen and Vander Vennet, 2022), generating heterogeneity within euro area financial markets, and in particular in the sovereign bond market. At the end of April 2022, sovereign bond spreads relative to the German Bund are rising again in the euro area: over 175 b.p. for Italy, 220 b.p. for Greece and around 100 b.p. for Spain and Portugal. Fragmentation risk on this market has already reappeared. In this context, the ECB will need to carefully manage the reduction of asset purchases and the implementation of tightening measures to avoid spillover effects which could trigger the financial stability within the euro area. At the same time, however, the normalization of monetary policy cannot be ruled out, should price pressure continue to build up.

Notwithstanding the importance of pursuing financial stability and mitigating fragmentation risk, this also brings into question whether aiming for 2% inflation over the medium term, remains the *primary* objective, potentially undermining the credibility of European Central Bank at maintaining price stability.

[21] On April the 5th, Agustin Carsten has indicated that "Central banks need to adjust to this new environment, not least by raising policy rates to more appropriate levels" https://www.bis.org/speeches/sp220405.htm.

4 Conclusion

This paper proposes a novel method to evaluate optimal combinations of the different forecasts. It designs a forecast encompassing test, which explicitly considers that weights lay in the unit simplex. This test can be readily extended in situation where the number of forecasts to evaluate greatly exceeds the realizations of the variable of interest. Hence, we evaluate institutions and model-based forecasts of inflation in the euro area between 2012 - 2021.

Although, ECB provides the most informative forecast on average, it does not encompass its competitors: SPF and European Commission also significantly contribute to the minimization of the forecast error variance. We also find that weights greatly vary over time when we re-estimate the forecast combination considering a rolling window.

The dynamics of the weights associated to ECB's inflation forecast are particularly interesting. Despite the fact that they appear on average greater than their competitors', they are not consistently remaining at this position in each sub period. In particular, ECB's weight suddenly drops in Q3 2012 and achieved their minimum in Q1 2019. We therefore investigate the fundamental factors behind this dynamics. Macro-financial conditions and monetary policy actions explain this variability. The greater the uncertainty surrounding inflation and the difference between current and the 2% target, the less informative ECB's forecasts are. The more contractionary the monetary policy, the more informative they are.

The decline of ECB's weight and its relation with its determinants are warning flags: the loss of forecast informativeness damages ECB's leading role at anchoring inflation expectations and might signal that preserving financial stability can be a challenge to the inflation targeting goal in high inflation environments.

References

Alesina, A. and R. Gatti (1995). Independent central banks: Low inflation at no cost? *The American Economic Review 85*(2), 196–200.

Andrade, P. and H. Le Bihan (2013). Inattentive professional forecasters. *Journal of Monetary Economics 60*(8), 967–982.

Ang, A., G. Bekaert, and M. Wei (2007). Do macro variables, asset markets, or surveys forecast inflation better? *Journal of Monetary Economics 54*(4), 1163–1212.

Bailey, N., S. Holly, and M. H. Pesaran (2016). A two-stage approach to spatio-temporal analysis with strong and weak cross-sectional dependence. *Journal of Applied Econometrics 31*(1), 249–280.

Baker, S. R., N. Bloom, and S. J. Davis (2016, July). Measuring economic policy uncertainty. *The Quarterly Journal of Economics 131*(4), 1593–1636.

Ball, L. (1992, June). Why does high inflation raise inflation uncertainty? *Journal of Monetary Economics 29*(3), 371–388.

Bernanke, B. S. and F. S. Mishkin (1997, June). Inflation targeting: A new framework for monetary policy? *Journal of Economic Perspectives 11*(2), 97–116.

Brodie, J., I. Daubechies, C. De Mol, D. Giannone, and I. Loris (2009). Sparse and stable markowitz portfolios. *Proceedings of the National Academy of Sciences 106*(30), 12267–12272.

Bunn, D. W. (1981, March). Two methodologies for the linear combination of forecasts. *Journal of the Operational Research Society 32*(3), 213–222.

Busetti, F. and J. Marcucci (2013). Comparing forecast accuracy: A monte carlo investigation. *International Journal of Forecasting 29*(4), 13–27.

Candelon, B., A. Luisi, and F. Roccazzella (2022). Fragmentation in the european monetary union: Is it really over? *Journal of International Money and Finance 122*, 102545.

Capistrán, C. (2008). Bias in federal reserve inflation forecasts: Is the federal reserve irrational or just cautious? *Journal of Monetary Economics 55*(8), 1415–1427.

Clements, M. and D. Hendry (1998). *Forecasting Economic Time Series*. Cambridge University Press.

Conflitti, C., C. D. Mol, and D. Giannone (2015). Optimal combination of survey forecasts. *International Journal of Forecasting 31*(4), 1096–1103.

De Grauwe, P. and Y. Ji (2022). The fragility of the eurozone: Has it disappeared? *Journal of International Money and Finance 120*, 102546.

De Santis, R. (2019). Redenomination risk. *Journal of Money, Credit and Banking 51*(8), 2173–2206.

Diebold, F. X. (1989). Forecast combination and encompassing: Reconciling two divergent literatures. *International Journal of Forecasting 5*(4), 589–592.

Diebold, F. X. and M. Shin (2019). Machine learning for regularized survey forecast combination: Partially-egalitarian lasso and its derivatives. *International Journal of Forecasting 35*(4), 1679–1691.

Dovern, J., U. Fritsche, and J. Slacalek (2012, 11). Disagreement Among Forecasters in G7 Countries. *The Review of Economics and Statistics 94*(4), 1081–1096.

Easaw, J. and R. Golinelli (2021, 09). Professionals Inflation Forecasts: The Two Dimensions Of Forecaster Inattentiveness. *Oxford Economic Papers*. gpab012.

Ehrbeck, T. and R. Waldmann (1996). Why are professional forecasters biased? agency versus behavioral explanations. *The Quarterly Journal of Economics 111*(1), 21–40.

Elhorst, J. P., M. Gross, and E. Tereanu (2020). Cross-sectional dependence and spillovers in space and time: where spatial econometrics and global VAR models meet. *Journal of Economic Surveys*.

Fan, J., Y. Liao, and M. Mincheva (2013). Large covariance estimation by thresholding principal orthogonal complements. *Journal of the Royal Statistical Society. Series B (Statistical Methodology) 75*(4), 603–680.

Fan, J., J. Zhang, and K. Yu (2012). Vast portfolio selection with gross-exposure constraints. *Journal of the American Statistical Association 107*(498), 592–606.

Favero, C. A. (2013). Modelling and forecasting government bond spreads in the euro area: a GVAR model. *Journal of Econometrics 177*(2), 343–356.

Ferrari, S. and F. Cribari-Neto (2004). Beta regression for modelling rates and proportions. *Journal of Applied Statistics 31*(7), 799–815.

Fu, W. and K. Knight (2000). Asymptotics for lasso-type estimators. *The Annals of Statistics 28*(5), 1356 – 1378.

Gambetti, P., G. Gauthier, and F. Vrins (2019). Recovery rates: Uncertainty certainly matters. *Journal of Banking & Finance 106*, 371–383.

Giacomini, R. and H. White (2006). Tests of conditional predictive ability. *Econometrica 74*(6), 1545–1578.

Gorter, J., J. Jacobs, and J. de Haan (2008). Taylor rules for the ecb using expectations data. *The Scandinavian Journal of Economics 110*(3), 473–488.

Gross, J. and J. Zahner (2021). What is on the ecb's mind? monetary policy before and after the global financial crisis. *Journal of Macroeconomics 68*, 103292.

Harvey, D. and P. Newbold (2000). Tests for multiple forecast encompassing. *Journal of Applied Econometrics 15*(5), 471–482.

IMF (2022a, April). Global financial stability report: Financial stability risks have risen as war tests the resilience of the financial system through various channels.

IMF (2022b, April). Regional economic outlook for europe.

Jagannathan, R. and T. Ma (2003, July). Risk reduction in large portfolios: Why imposing the wrong constraints helps. *The Journal of Finance 58*(4), 1651–1683.

Kudô, A. (1963). A multivariate analogue of the one-sided test. *Biometrika 50*(3-4), 403–418.

Ledoit, O. and M. Wolf (2003, dec). Improved estimation of the covariance matrix of stock returns with an application to portfolio selection. *Journal of Empirical Finance 10*(5), 603–621.

Lee, J. D., D. L. Sun, Y. Sun, and J. E. Taylor (2016, June). Exact post-selection inference, with application to the lasso. *The Annals of Statistics 44*(3).

Li, J., Z. Liao, and R. Quaedvlieg (2021, 06). Conditional Superior Predictive Ability. *The Review of Economic Studies 89*(2), 843–875.

Lovell, M. C. (1963). Seasonal adjustment of economic time series and multiple regression analysis. *Journal of the American Statistical Association 58*(304), 993–1010.

Lütkepohl, H. (2007). *New Introduction to Multiple Time Series Analysis*. Springer Berlin Heidelberg.

Mammen, E. (1993). Bootstrap and wild bootstrap for high dimensional linear models. *The Annals of Statistics*, 255–285.

Mankiw, N. G., R. Reis, and J. Wolfers (2004). Disagreement about inflation expectations. In *NBER Macroeconomics Annual 2003, Volume 18*, pp. 209–270. National Bureau of Economic Research, Inc.

Manzan, S. (2011). Differential interpretation in the survey of professional forecasters. *Journal of Money, Credit and Banking 43*(5), 993–1017.

Miranda-Agrippino, S. and T. Nenova (2022). A tale of two global monetary policies. *Journal of International Economics*, 103606.

Mishkin, F. S. (1990). The information in the longer maturity term structure about future inflation. *The Quarterly Journal of Economics 105*(3), 815–828.

Mishkin, F. S. (1991). A multi-country study of the information in the shorter maturity term structure about future inflation. *Journal of International Money and Finance 10*(1), 2–22.

Monfort, A. and J.-P. Renne (2013). Decomposing euro-area sovereign spreads: credit and liquidity risks. *Review of Finance 18*(6), 2103–2151.

Newey, W. K. and K. D. West (1987). A simple, positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix. *Econometrica 55*(3), 703–708.

Ning, Y. and H. Liu (2017). A general theory of hypothesis tests and confidence regions for sparse high dimensional models. *The Annals of Statistics 45*(1), 158–195.

Patton, A. J. and A. Timmermann (2010). Why do forecasters disagree? lessons from the term structure of cross-sectional dispersion. *Journal of Monetary Economics 57*(7), 803–820.

Perlman, M. D. (1969). One-sided testing problems in multivariate analysis. *The Annals of Mathematical Statistics 40*(2), 549–567.

Pesaran, M. H. (2021). General diagnostic tests for cross-sectional dependence in panels. *Empirical Economics 60*, 13–50.

Roccazzella, F., P. Gambetti, and F. Vrins (2022). Optimal and robust combination of forecasts via constrained optimization and shrinkage. *International Journal of Forecasting 38*(1), 97–116.

Rogoff, K. (1985). The optimal degree of commitment to an intermediate monetary target. *Quarterly Journal of Economics 100*, 1169–1189.

Saikkonen, P. and H. Lütkepohl (2000, October). Testing for the cointegrating rank of a VAR process with structural shifts. *Journal of Business & Economic Statistics 18*(4), 451.

Silvapulle, M. and P. Sen (2004, November). *Constrained Statistical Inference: Inequality, Order, and Shape Restrictions*. Wiley.

Smithson, M. and J. Verkuilen (2006). A better lemon squeezer? maximum-likelihood regression with beta-distributed dependent variables. *Psychological methods 11*(1), 54.

Soenen, N. and R. Vander Vennet (2022). Ecb monetary policy and bank default risk. *Journal of International Money and Finance 122*, 102571.

Stock, J. H. and M. W. Watson (1999). Forecasting inflation. *Journal of Monetary Economics 44*(2), 293–335.

Stock, J. H. and M. W. Watson (2003). Forecasting output and inflation: The role of asset prices. *Journal of Economic Literature 41*(3), 788–829.

Svensson, L. E. (1997). Inflation forecast targeting: Implementing and monitoring inflation targets. *European Economic Review 41*(6), 1111–1146.

Tibshirani, R. (1996, January). Regression shrinkage and selection via the lasso. *Journal of the Royal Statistical Society: Series B (Methodological) 58*(1), 267–288.

Timmermann, A. (2006). Chapter 4 forecast combinations. In *Handbook of Economic Forecasting*, pp. 135–196. Elsevier.

Wu, J. C. and F. D. Xia (2020, June). Negative interest rate policy and the yield curve. *Journal of Applied Econometrics 35*(6), 653–672.

Yu, M., V. Gupta, and M. Kolar (2019). Constrained high dimensional statistical inference.

Zhao, P. and B. Yu (2006, December). On model selection consistency of lasso. *Journal of Machine Learning Research 7*, 2541—2563.

A The test for forecast encompassing

In the spirit of Harvey and Newbold (2000), we aim to design a test for multiple forecast encompassing. However, this test must explicitly consider the constraints on the estimated weights. Furthermore, large data set, heterogeneous modeling techniques and, many and different data pre-processing procedures originate a wide array of forecasts, whose size can potentially outclass the number of available realizations of the target variable. Therefore, a "modern" test for forecast encompassing should deal with potential high dimensional frameworks.

High dimensionality and constrained parameter space create, however, two problems. First, in the presence of high dimensional nuisance parameters, the maximum likelihood estimator is no longer consistent and, despite it is possible to identify some maximum penalized likelihood estimators that are consistent under some conditions, they may not have a tractable limiting distribution even in the fixed dimensional case (Fu and Knight, 2000). This invalidates the standard inferential theory for Wald, Lagrange Multiplier (LM) and Likelihood Ratio (LR) (see for an example Ning and Liu, 2017). Second, the nonnegativity and sum to one are inequality constraints that generate a closed and convex *cone* C.[22] In a test for forecast encompassing, we test whether the marginal contribution of forecast i to the aggregate ex post optimal forecast is strictly positive, i.e., we state the following hypothesis

$$H_0 : \theta = 0 \quad vs \quad H_1 : \theta > 0, \quad \theta_i \in \mathbb{R}^{+(r)} , \tag{15}$$

where we consider the following partition $\beta = (\theta, \gamma)$, with θ being the $r-$ dimensional parameter of interest and γ denoting the nuisance parameters (with dimension $d - r$).This corresponds to test whether the population ex post optimal weights lay on the boundary set (under H_0) or in the cone generated by the nonnegativity and sum to one constraints (under H_1). In the low dimensional case, Kudô (1963) and Perlman (1969) proved that, under the null hypothesis, the LM, LR and Wald statistic associated to (15) are distributed as a mixture of chi-squared distributions.

[22]A subset C of $\mathbb{R}^d$ is called a cone (or a positively homogeneous set if $x \in C$ that tx E C for every positive t.

In this paper, we follow the decorrelation procedure of Ning and Liu (2017) to handle the impact of high dimensional nuisance parameters. This approach offers two main advantages: a) not relying post-selection Lee et al. (2016), i.e., considered the conditional inference given the event that some covariates have been selected; and b) the possibility of easily extending the testing procedure under inequality constraints. In fact, following Yu et al. (2019), we adapt the decorrelated score function of Ning and Liu (2017) to extend the test for forecast encompassing for low dimensional parameters to a high dimensional framework.

In this paper, we assume that the residual component $\epsilon \sim \mathcal{N}(0, \sigma_\epsilon)$ and we denote the sample negative log likelihood function as

$$\ell(\boldsymbol{\beta}) = -\frac{1}{n}\sum_{i=1}^{n} \log \mathcal{L}_i(\boldsymbol{\beta}) = \frac{n}{2} log(\pi\sigma_\epsilon) + \frac{1}{2}\frac{\left[\boldsymbol{y} - \boldsymbol{F\beta}\right]^T \left[\boldsymbol{y} - \boldsymbol{F\beta}\right]}{\sigma_\epsilon}\ , \tag{16}$$

and we let $\nabla\ell(\boldsymbol{\beta}) = \nabla\ell(\theta, \boldsymbol{\gamma})$ be the gradient of the negative log likelihood, $\nabla^2\ell(\boldsymbol{\beta})$ be the sample Hessian matrix and $\mathcal{F}(\beta) = \mathbb{E}[\nabla^2\ell(\boldsymbol{\beta})]$ be the population Fisher information matrix, while we denote with $\nabla_\theta\ell(\boldsymbol{\beta})$, $\nabla_\gamma\ell(\boldsymbol{\beta})$, $\mathcal{F}(\beta)_{\theta\theta}$, $\mathcal{F}(\beta)_{\theta\gamma}$, $\mathcal{F}(\beta)_{\gamma\theta}$, $\mathcal{F}(\beta)_{\gamma\gamma}$ the correspondent partitions. We also denote with $\widehat{\boldsymbol{\beta}} = (\widehat{\theta}, \widehat{\gamma})$ the estimated ex post optimal weights.

A.1 The testing procedure

We now introduce the test for forecasts encompassing. To ease the presentation, we split the testing procedure in two steps. In **Algorithm 1**, we estimate the ex post optimal combination of forecasts and estimate the decorrelated score function, parameter(s) of interest and likelihood function. In **Algorithm 2**, we form the corresponding test statistics, compute tail probabilities under the null hypothesis and finally perform inference on the combing weight(s).

The key step in **Algorithm 1** is to estimate ex post optimal weights and the decorrelation operator W. Since the both the weights and the nuisance score functions can be high dimensional (have dimension d and d-1, respectively), we need to impose some assumptions on $\boldsymbol{\beta}$ and W to bound the estimation error when $d > n$. On the one hand, $\boldsymbol{\beta}$ searches for

Algorithm 1: Decorrelation procedure for high dimensional regression frameworks

1 Estimate ex post optimal weights
$\widehat{\boldsymbol{\beta}}_{\setminus d} = \text{argmin}_{\boldsymbol{\beta} \in \mathbb{R}^{+(d-1)}} \left\| V_d - \boldsymbol{V}_{\setminus d} \boldsymbol{\beta}_{\setminus d} \right\|_2^2 \;, \quad \widehat{\beta}_d = 1 - \sum_{i=1}^{d-1} \widehat{\beta}_{\setminus d, i} \;,$

2 Define the partition $\beta = (\theta, \gamma)$

3 Estimate the decorrelation operator W_j
$\widehat{W_j} = \text{argmin}_{W_j \in \mathbb{R}^{d-1}} \|V_j - \boldsymbol{V}_{\setminus j} W_j\|_2^2 + \lambda \|W_j\|_1 \;,$

4 When θ is a r-low-dimensional multi-dimensional parameter of interest corresponding to the forecasts in the set I, W can be solved column by column i.e., $\widehat{\boldsymbol{W}} = \left(\widehat{W}_j, \text{ where } j \in I\right)$

5 Estimate the decorrelated score function:

$$\widetilde{S}(\theta) = \nabla_\theta \ell(\theta, \widehat{\gamma}) - \boldsymbol{W}^T \nabla_\gamma \ell(\theta, \widehat{\gamma}) \;,$$

6 Estimate the decorrelated parameter of interest:

$$\widetilde{\theta} = \widehat{\theta} - \widetilde{\mathscr{F}}_{\theta|\gamma}^{-1} \widetilde{S}, \quad \text{where} \quad \widetilde{\mathscr{F}}_{\theta|\gamma} = \nabla_{\theta\theta}^2 \ell(\widehat{\boldsymbol{\beta}}) - \boldsymbol{W}^T \nabla_{\theta\gamma}^2 \ell(\widehat{\boldsymbol{\beta}}) \;,$$

7 Estimate the decorrelated likelihood function:

$$\widetilde{\ell}(\boldsymbol{\theta}) = \ell(\theta, \widehat{\boldsymbol{\gamma}} - \boldsymbol{W}^T(\theta - \widehat{\theta})) \;.$$

the linear combination of weights that minimizes the variance of the aggregate forecasting error under the nonnegativity and sum to one constraints. When $d > n$, we impose $\boldsymbol{\beta}$ to be sparse. Similarly, when $d > n$, W searches for the best sparse linear combination of the nuisance score functions to approximate the score function.

The key step in **Algorithm 2** is to compute the cumulative tail probability of the considered test statistics for a given significance level a. To do this, we have to compute the chi-bar-squared weights. when the constrained space is the nonnegative orthant (or more generally when the constraints defining C are linear and independent), Silvapulle and Sen (2004) propose a simple quadratic programming problem to estimate ω (see proposition 3.6.1).When θ is univariate ($r = 1$), the weights are known in analytical form and the chi-bar-squared statistics reduces to

$$\text{Prob}\left[\overline{\chi}^2(\widetilde{\mathscr{F}}_{\theta|\gamma}, \mathbb{R}^{+r}) > c_a\right] = .5 \text{ Prob}\left[\chi^2(0) > c_a\right] + .5 \text{ Prob}\left[\chi^2(1) > c_a\right] \;. \tag{17}$$

Algorithm 2: One-sided test for forecasts encompassing

1. Form the decorrelated score test statistics:
   $T_S = \left[\widetilde{S}(\widehat{\theta}_{H_0}) - \widetilde{S}(\widehat{\theta}_{H_1})\right]^T \widetilde{\mathscr{F}}_{\theta|\gamma}^{-1} \left[\widetilde{S}(\widehat{\theta}_{H_0}) - \widetilde{S}(\widehat{\theta}_{H_1})\right]$.
2. Form the decorrelated likelihood ratio (LR) test statistics:
   $T_{LR} = 2n\left(\widetilde{\ell}(\widehat{\boldsymbol{\beta}}_{H_0}) - \widetilde{\ell}(\widehat{\boldsymbol{\beta}}_{H_1})\right)$,
3. Form the decorrelated Wald test statistics:
   $T_W = \left[\widetilde{\theta} - \widehat{\theta}_{H_0}\right]^T \widetilde{\mathscr{F}}_{\theta|\gamma} \left[\widetilde{\theta} - \widehat{\theta}_{H_0}\right] - \left[\widetilde{\theta} - \widehat{\theta}_{H_1}\right]^T \widetilde{\mathscr{F}}_{\theta|\gamma} \left[\widetilde{\theta} - \widehat{\theta}_{H_1}\right]$.
4. Compute the set of mixing weights $\omega_i(r, \widetilde{\mathscr{F}}_{\theta|\gamma}, \mathbb{R}^{+r})$ for the $\overline{\chi}^2$ statistics.
5. Given a significance level a and its corresponding critical value c_a, compute cumulative tail probability :
   $$\text{Prob}\left[\overline{\chi}^2(\widetilde{\mathscr{F}}_{\theta|\gamma}, \mathbb{R}^{+r}) > c_a\right] = \sum_{i=0}^{r} \omega_i \text{Prob}\left[\chi^2(i) > c_a\right] \ .$$
6. Reject H_0 if
   $$\text{Prob}\left[T_S > c\right] > a \ , \quad \text{Prob}\left[T_W > c\right] > a, \quad \text{Prob}\left[T_{LR} > c\right] > a \ .$$

In the remainder of the paper, to ease presentation and simplify notation, we focus on the likelihood framework and on the univariate parameter of interest θ ($r = 1$) corresponding to the j^{th} forecast. Nevertheless, the decorrelated score function can be defined for a multi-dimensional parameter of interest and also when $\boldsymbol{\beta}$ minimizes a loss function other than the negative log-likelihood. (see section 3 and supplementary material of Ning and Liu, 2017, respectively).

B Assumptions of the test when $d > n$

In this section we describe the high-level assumptions required to obtain weak convergence of test under null hypothesis in the high dimensional framework. We closely follow Ning and Liu (2017) and Yu et al. (2019) and refer to their papers for the proof of the results in the case of linear regression models. These conditions can be classified into three main categories: (a) Consistency conditions for initial parameter estimation (Assumption 1, 2 and 3); (b) Concentration of the gradient and Hessian matrix (Assumption 4); (c) Local smoothness on the loss function (Assumption 5). All these assumption are fairly standard in the high dimensional statistics literature and are known to hold for the Lasso and nonnegative least square with sum to one constraint on the weights (Fan et al., 2012).

Assumption 1. *Score condition. The expected value of the score function at the true* $\boldsymbol{\beta}^*$ *is equal to zero, i.e.,* $\nabla \ell(\boldsymbol{\beta}^*) = 0$.

Assumption 2. *Restricted eigenvalue condition for the covariance matrix. Given the positive definite matrix* $\mathscr{F}$ *and* $\nabla \ell^2(\boldsymbol{\beta})$*, let* $\mathcal{S} = supp(\boldsymbol{\beta}^*) \cup supp(\boldsymbol{W}^*)$*. There exists a universal constant* $c, k > 0$ *such that the quadratic form* $\boldsymbol{v}^t \mathscr{F} \boldsymbol{v} \geq k \|\boldsymbol{v}\|_2^2$*, and* $\boldsymbol{v}^T \nabla \ell^2(\boldsymbol{\beta}) \boldsymbol{v}$*, for any* $\boldsymbol{v}$ *in the cone* C*, i.e.,* $\forall \boldsymbol{v} \in C(\mathcal{S}) = \{\boldsymbol{v} : \|\boldsymbol{v}_{\mathcal{S}^c} \leq c \boldsymbol{v}_{\mathcal{S}}\|\}$.

Assumption 3. *Consistency conditions for initial parameter estimation. As* $n \longrightarrow \infty$*, for some sequences* $\nu_1(n)$ *and* $\nu_2(n)$ *converging to* 0*, the following holds*

$$\lim_{n\to\infty} Prob_{\boldsymbol{\beta}^*} \left[\|\widehat{\boldsymbol{\beta}} - \boldsymbol{\beta}^*\|_1 \lesssim \nu_1(n) \right] = 1 \ , \quad and \quad \lim_{n\to\infty} Prob_{\boldsymbol{\beta}^*} \left[\|\widehat{\boldsymbol{W}} - \boldsymbol{W}^*\|_1 \lesssim \nu_2(n) \right] = 1 \ .$$

In lemma 3.1 of the paper, D3 and D4 of the supplementary material, Ning and Liu (2017) show that this assumption holds when considering the Lasso estimator of $\boldsymbol{\beta}$ and $\boldsymbol{W}$ in a linear regression model. In this case with high probability, we have the following

$$\|\widehat{\boldsymbol{\beta}} - \boldsymbol{\beta}^*\|_1 = \mathcal{O}\left(s^* \sqrt{\frac{log(d)}{n}} \right) \ , \quad \|\widehat{\boldsymbol{W}}_j - \boldsymbol{W}_j^*\|_1 = \mathcal{O}\left(s' \sqrt{\frac{log(d)}{n}} \right) \text{ with } j \in I \ ,$$

where $\boldsymbol{\beta}$ and $\boldsymbol{W}$ are defined in **Algorithm 1** and $s^* = \|\boldsymbol{\beta}^*\|_0$ and $s' = \|\boldsymbol{W}^*\|_0$, i.e., the number of non zero elements in $\boldsymbol{\beta}^*$ and $\boldsymbol{W}^*$, respectively.[23]

Assumption 4. *Concentration of the gradient and Hessian. Let* $v^* = (1, -\boldsymbol{W}^{T*})$, *assume that the largest element of the score function is bounded, i.e.,* $\|\nabla \ell(\boldsymbol{\beta}^*)\|_\infty = \mathcal{O}\left(\sqrt{\frac{log(d)}{n}}\right)$ *and*

$$\|v^{T*}\nabla^2 \ell(\boldsymbol{\beta}^* - \mathbb{E}_{\boldsymbol{\beta}^*}\left[v^{T*}\nabla^2 \ell(\boldsymbol{\beta}^*\right]\|_\infty = \mathcal{O}\left(\sqrt{\frac{log(d)}{n}}\right) .$$

Assumption 5. *Local smoothness of the loss function For some constant L, the Hessian matrix* $\nabla^2 \ell(\boldsymbol{\beta})$ *is Lipschitz continuous:*

$$\|\nabla^2 \ell(\boldsymbol{\beta}_1) - \nabla^2 \ell(\boldsymbol{\beta}_2)\|_\infty \leq L\|\boldsymbol{\beta}_1 - \boldsymbol{\beta}_2\|_1 .$$

Lemma 1. *Central limit theorem (CLT) for linear combination of score function. Let* $\boldsymbol{\Sigma}^* = \lim_{n\to\infty} Var_{\boldsymbol{\beta}^*}\left[\sqrt{n}\nabla \ell(\boldsymbol{\beta}^*)\right]$. *Under **Assumption 1-5**, Ning and Liu (2017) and Yu et al. (2019) proved that*[24]

$$\sqrt{n}\boldsymbol{v}^{T*}\nabla \ell(\boldsymbol{\beta}^*)\sqrt{\sigma^*}^{-1} \rightsquigarrow \mathcal{N}(0,1), \quad \textit{where } \sigma^* = \boldsymbol{v}^{T*}\boldsymbol{\Sigma}^*\boldsymbol{v}^* .$$

[23] Remark that $\boldsymbol{W}^* = \mathcal{F}_{\gamma\gamma}^{*-1}\mathcal{F}_{\gamma\theta}^*$.

[24] We refer to Appendix C and D of Ning and Liu (2017) for the proofs that Assumption 1 to 5 hold in a general linear regression model, while we refer to the appendix of Fan et al. (2012) for the proofs related to the nonnegative least squares estimator under sum to one constraint.

C Beta regression framework

The density function of a beta distribution is given by

$$f(y, p, q) = \frac{\Gamma(p+q)}{\Gamma(p)\;\Gamma(q)} y^{p-1}\;(1-y)^{q-1}, \quad \text{with } 0 < y < 1\;, \tag{18}$$

with $p, q > 0$ and $\Gamma(\cdot)$ being the gamma function. However, the beta distribution can be conveniently re-parameterized in term of $0 < \mu < 1$ and $\phi > 0$ that are related to the mean and variance of the dependent variable y,

$$E\left[y_i\right] = \mu_i \text{ and } Var\left[y_i\right] = \frac{\mu_i(1-\mu_i)}{1+\phi}\;. \tag{19}$$

In this framework, the log-likelihood function becomes

$$\ell(\mu, \phi) = \sum_{i=1}^{n} \ell_i(\mu_i, \phi)\;, \tag{20}$$

where

$$\begin{aligned} \ell_i(\mu_i, \phi) = log\left[\Gamma\left(\phi\right)\right] - log\left[\Gamma\left(\mu_i\phi\right)\right] - log\left\{\Gamma\left[(1-\mu_i)\phi\right]\right\} + \\ + \left(\mu_i\phi - 1\right) log\left(y_i\right) + \left[\left(1-\mu_i\right)\phi - 1\right] log\left(1\right)y_i)\;\;. \end{aligned}$$

Using this notation, the standard beta regression model assumes a) independent realizations of $y_i \sim \mathscr{B}\left(\mu_i, \phi_i\right)$, b) a linear regression model for the transform of the mean parameter and c) a constant precision parameter ϕ. Fixing the precision parameter makes the beta regression framework more parsimonious at the cost of a lower flexibility than the variable dispersion beta regression models featured in Gambetti et al. (2019). However, this is a price we have to pay considering that we count only 38 observations in our sample. Nevertheless, we can easily notice that this model can still naturally accommodate heteroskedasticity because of $Var\left[y\right]$ being a also a function of μ_i. Formally,

$$g_1\left(\mu_i\right) = x_i^T\beta \text{ and } g_2\left(\phi_i\right) = \phi\;, \tag{21}$$

with $g_1(\cdot) : (0,1) \to \mathbb{R}$ and $g_2(\cdot) : (0,\infty) \to \mathbb{R}$, with x_i being the vector of regressors, β conformable vectors of regression coefficients, and $g_1(\cdot)$ and $g_2(\cdot)$ being strictly increasing and twice-differentiable link functions. In this paper, we consider a *logit* link function for g_1 and a constant g_2. This results in

$$\mu_i = \frac{e^{x_i^T \beta}}{1 + e^{x_i^T \beta}} \quad \text{and} \quad \phi_i = \phi \, . \tag{22}$$

D Bootstrap procedure

In our framework, we consider bootstrapping procedure to take into account also unknown forms of heteroskedasticity that could emerge in the dynamics of the estimated weights, we proposed the wild bootstrapped procedure as in Mammen (1993). Specifically,

1. Estimate the standard linear regression or beta regression parameters via maximum likelihood as described in the subsection 3.4 (after retrieving the ex post optimal combination weights as outlined in Section 1), and the corresponding residuals u_t^i for each $i = 1, \dots, n$.

2. Draw with replacement the sequence of errors $\{\tilde{u}_i = k_i \hat{u}_i\}_{i=1}^n$ with

$$k_i = \begin{cases} \frac{1+\sqrt{5}}{2}, & \text{with probability } p = \frac{\sqrt{5}-1}{2\sqrt{5}} \\ \frac{1-\sqrt{5}}{2}, & \text{with probability } 1-p. \end{cases}$$

   as in Mammen (1993) with $\{\hat{u}_i\}_{i=1}^n$ being the estimated errors in the beta regression.

3. Generate the bootstrapped dependent variable using $\{\tilde{y}_i = \widehat{y}_i + k_i\}_{i=1}^n$, ¡ith $\{\hat{y}_i\}_{i=1}^n$ being the fitted values of the dependent variable in the beta regression. We map negative values or weights greater than one of the bootstrapped sample $\{\tilde{y}_i\}_{i=1}^n$ to zero and one, respectively. Following Smithson and Verkuilen (2006), we consider the transformation $\frac{y(n-1)+0.5}{n}$ where y is the dependent variable and n is the sample size to adapt the beta regression framework to dependent variables assuming also the extremes 0 and 1.

4. Estimate the beta regression on the bootstrapped data sample.

5. Repeat 2 - 4 a large number of times (in this work, we repeat the procedure $10,000$ times), and then extract the quantiles needed.

E Additional results on the determinants of the weights

Table 7: Determinants of the weights constructed using the sample (Spec. 1) and the shrinkage estimator (Spec. 3) of the covariance matrix of prediction errors. .

	Specification 1			Specification 3		
	LM	NW	B	LM	NW	B
b_0	0.767 (0.0621)	***	***	0.820 (0.0720)	***	***
Dif. 2%	-0.055 (0.0289)	*	*	-0.147 (0.0319)	***	***
Dis.	-0.064 (0.0281)	**	*	-0.079 (0.0306)	**	
EPU	0.001 (0.0004)	**	*	-0.001 (0.0004)	***	***
S	0.117 (0.0055)	***	***	0.072 (0.0039)	***	***
APP	0.000 (0.0103)			-0.019 (0.0115)		
Inf. Sur.	-0.034 (0.0147)	**	***	-0.027 (0.0101)	**	**
Glob.	0.019 (0.0065)	***	***	0.005 (0.0062)		
Fr.	-0.014 (0.0204)			0.000 (0.0195)		
d_{Q2}	-0.077 (0.0368)	**	***	-0.032 (0.0307)		
d_{Q3}	-0.062 (0.0310)	*	***	-0.036 (0.0302)		
d_{Q4}	-0.134 (0.0458)	***	***	-0.050 (0.0253)	*	

Notes. The first columns displays the determinants, b_0 is the intercept. LM report the estimated coefficients using the linear regression model, while we display Newey-West standard errors in the parenthesis, NW and B report whether the coefficient of interest is statistically significant at 1% (***), 5% (**), 10% (*) level according to the Newey-West standard errors and the Mammen wild boostrap quantiles, respectively.

F Size and power analysis

In this section, we study the performance of the the proposed test both in term of empirical size and power. Similarly to Busetti and Marcucci (2013), we design a Monte Carlo experiment where we model the j^{th} forecast using a single factor model

$$f_i = x_i + \epsilon_{i,j}, \quad \text{for} \quad i = 1, 2, ...n \tag{23}$$

and allow for $d > n$. The factor f_i is a i.i.d. $\mathcal{N}(0,1)$ and $\epsilon_{i,j}$ is a idiosyncratic and potentially auto-correlated error term, i.e., $\epsilon_{i,j} = \rho\epsilon_{i-1,j} + z_j,\ z_j \sim \mathcal{N}(0, \sigma_z)$. The i^{th} realization of the variable of interest is obtained from

$$y_i = \boldsymbol{\beta}^T \boldsymbol{f}_i + u_i\ . \tag{24}$$

The systematic error term u_i permits us to study the behavior of the test even in presence of auto-correlated residuals, i.e., $u_i = \rho u_{i-1} + z_i,\ z_i \sim \mathcal{N}(0, \sigma_z)$. $\boldsymbol{\beta}$ is the vector weights that lies in the *d-1* dimensional unit simplex. We partition the vector of weights $\boldsymbol{\beta}$ into two sub-vectors, i.e., $\boldsymbol{\beta} = (\theta, \boldsymbol{\gamma})$, θ is the parameter of interest while $\boldsymbol{\gamma}$ is a sparse vectors of nuisance parameters with d non-zero components. Indeed, we require $\boldsymbol{\gamma}$ to be sparse to insure the consistency conditions on $\widehat{\boldsymbol{\beta}}$ and $\widehat{\boldsymbol{W}}$ when the problem is high-dimensional.

We consider several schemes to assign the weights to the non-zero element i of $\boldsymbol{\gamma}$: a) equal weight, i.e., $\boldsymbol{\gamma}_i = \frac{1-\theta}{d}$; b) random, i.e., $\boldsymbol{\gamma}_i = \frac{r}{k}$, where $r \sim \mathcal{U}(0,1)$ and k is a normalization constant such that $\theta + \boldsymbol{\gamma} = 1$. To avoid repetition , we report the results for the scheme a) with $m = 20$, $d = 5$, $n = \{40, 80, 160\}$, $\rho = \{0, .25, .5, .75\}$ and $\sigma_z = .5$ and $\sigma_u = .25$. [25]

Given n realizations of the target and m forecasts, Table 8 and Table 9 report the results of the simulation study for a nominal size of 5% and 10000 repetitions. Similarly to Ning and Liu (2017), the first column (True weight = 0) shows that the test is slightly undersized

[25] The analysis on empirical size and size-adjusted power in scheme a) and b) are comparable. The parameters $\sigma_z = .5$ and $\sigma_u = .25$ are chosen to obtain an average signal-to-noise ratio ranging from 1.60 (when $\rho = .75$) to 1.8 (when $\rho = 0$) across the simulation study for a sample of 1000 observation. Similarly, we report results for $d = 5$, but other tables with additional robustness checks are available upon request.

in the high dimensional case ($m = 80$ and $n = 40$), but this distortion becomes less evident when $n = 80$ and disappears when $n = 180$.

We confirm usual results about the size-adjusted power curve. Power increases when $n \uparrow$ and when $\theta \to 1$. We also expect that when $\rho \uparrow$, the size-adjusted power $\downarrow$. Indeed, we have specified the dynamics of the ϵ_j as an autoregressive process of order 1 - AR(1) - and, keeping σ_z constant, the long term variance of ϵ_j to tend to infinity as $\rho \to 1$.

Table 8: Empirical and size-adjusted rejection probabilities for the analysis of the empirical size and size-adjusted power.

		Distance from the null						
		0 - Size	.01	.025	.05	.10	.20	.25
$n: 160$, $\rho: 0$								
Nonnegative $\boldsymbol{\beta}$ and $LM \sim \bar{\chi}^2$	Size-adj. power	.050	.064	.113	.259	.720	.998	1
	Rejection probability	.054	.068	.118	.271	.732	.998	1
$\boldsymbol{\beta}$ and $LM \sim \chi^2$	Size-adj. power	.050	.059	.094	.229	.676	.997	1
	Rejection probability	.085	.095	.145	.304	.751	.998	1
$n: 160$, $\rho: .25$								
Nonnegative $\boldsymbol{\beta}$ and $LM \sim \bar{\chi}^2$	Size-adj. power	.050	.059	.102	.249	.697	.997	1
	Rejection probability	.054	.064	.108	.258	.708	.997	1
$\boldsymbol{\beta}$ and $LM \sim \chi^2$	Size-adj. power	.088	.089	.136	.293	.732	.998	1
	Rejection probability	.050	.054	.085	.216	.643	.994	1
$n: 160$, $\rho: .50$								
Nonnegative $\boldsymbol{\beta}$ and $LM \sim \bar{\chi}^2$	Size-adj. power	.050	.060	.089	.209	.606	.989	.999
	Rejection probability	.054	.064	.096	.219	.618	.990	.999
$\boldsymbol{\beta}$ and $LM \sim \chi^2$	Size-adj. power	.086	.095	.120	.252	.643	.991	1
	Rejection probability	.050	.058	.080	.185	.559	.983	.999
$n: 160$, $\rho: .75$								
Nonnegative $\boldsymbol{\beta}$ and $LM \sim \bar{\chi}^2$	Size-adj. power	.050	.054	.080	.149	.411	.914	.983
	Rejection probability	.054	.060	.086	.157	.425	.919	.984
$\boldsymbol{\beta}$ and $LM \sim \chi^2$	Size-adj. power	.082	.088	.112	.186	.456	.926	.985
	Rejection probability	.050	.053	.069	.128	.368	.889	.975
$n: 80$, $\rho: 0$								
Nonnegative $\boldsymbol{\beta}$ and $LM \sim \bar{\chi}^2$	Size-adj. power	.050	.058	.086	.170	.435	.929	.987
	Rejection probability	.057	.068	.098	.185	.459	.936	.989
$\boldsymbol{\beta}$ and $LM \sim \chi^2$	Size-adj. power	.130	.138	.166	.254	.528	.942	.990
	Rejection probability	.050	.051	.068	.127	.351	.875	.971
$n: 80$, $\rho: .25$								
Nonnegative $\boldsymbol{\beta}$ and $LM \sim \bar{\chi}^2$	Size-adj. power	.050	.053	.082	.154	.416	.916	.985
	Rejection probability	.059	.064	.095	.175	.446	.927	.987
$\boldsymbol{\beta}$ and $LM \sim \chi^2$	Size-adj. power	.132	.138	.160	.245	.514	.938	.988
	Rejection probability	.050	.049	.068	.118	.318	.854	.960
$n: 80$, $\rho: .50$								
Nonnegative $\boldsymbol{\beta}$ and $LM \sim \bar{\chi}^2$	Size-adj. power	.050	.055	.073	.137	.354	.848	.958
	Rejection probability	.059	.064	.083	.152	.375	.862	.963
$\boldsymbol{\beta}$ and $LM \sim \chi^2$	Size-adj. power	.129	.130	.149	.216	.442	.881	.967
	Rejection probability	.050	.053	.063	.108	.278	.772	.923
$n: 80$, $\rho: .75$								
Nonnegative $\boldsymbol{\beta}$ and $LM \sim \bar{\chi}^2$	Size-adj. power	.050	.053	.065	.105	.245	.664	.822
	Rejection probability	.060	.063	.078	.121	.272	.688	.841
$\boldsymbol{\beta}$ and $LM \sim \chi^2$	Size-adj. power	.126	.134	.142	.186	.340	.728	.858
	Rejection probability	.050	.053	.059	.088	.191	.573	.746

Notes. We estimate the empirical size by generating the optimal forecast under the null hypothesis, i.e., by enforcing the weight to be tested to be zero in population. We conduct the power analysis by generating the optimal forecast under the alternative for a departure from the null of $\beta = 0$, i.e., by imposing the population weight of the forecast to be tested equal to .01, .025, .05, .10, .20, and .25.

Table 9: Empirical and size-adjusted rejection probabilities for the analysis of the empirical size and size-adjusted power.

		Distance from the null						
		0 - Size	.01	.025	.05	.10	.20	.25
$n: 40,\ \rho: 0$								
Nonnegative $\boldsymbol{\beta}$ and $LM \sim \bar{\chi}^2$	Size-adj. power	.050	.059	.076	.122	.258	.677	.835
	Rejection probability	.067	.080	.097	.149	.297	.714	.860
$\boldsymbol{\beta}$ and $LM \sim \chi^2$	Size-adj. power	.284	.283	.291	.334	.467	.781	.885
	Rejection probability	.050	.053	.054	.075	.131	.430	.608
$n: 40,\ \rho: .25$								
Nonnegative $\boldsymbol{\beta}$ and $LM \sim \bar{\chi}^2$	Size-adj. power	.050	.056	.068	.109	.244	.632	.806
	Rejection probability	.071	.079	.094	.145	.292	.684	.843
$\boldsymbol{\beta}$ and $LM \sim \chi^2$	Size-adj. power	.284	.283	.291	.330	.449	.753	.869
	Rejection probability	.050	.050	.052	.068	.128	.398	.559
$n: 40,\ \rho: .50$								
Nonnegative $\boldsymbol{\beta}$ and $LM \sim \bar{\chi}^2$	Size-adj. power	.050	.060	.074	.108	.226	.585	.744
	Rejection probability	.066	.076	.090	.126	.257	.618	.772
$\boldsymbol{\beta}$ and $LM \sim \chi^2$	Size-adj. power	.274	.279	.282	.310	.420	.708	.821
	Rejection probability	.050	.051	.052	.068	.121	.356	.508
$n: 40,\ \rho: .75$								
Nonnegative $\boldsymbol{\beta}$ and $LM \sim \bar{\chi}^2$	Size-adj. power	.050	.059	.059	.083	.163	.394	.539
	Rejection probability	.069	.079	.080	.109	.199	.445	.586
$\boldsymbol{\beta}$ and $LM \sim \chi^2$	Size-adj. power	.258	.266	.261	.278	.362	.562	.669
	Rejection probability	.050	.051	.052	.061	.098	.238	.344

Notes. We estimate the empirical size by generating the optimal forecast under the null hypothesis, i.e., by enforcing the weight to be tested to be zero in population. We conduct the power analysis by generating the optimal forecast under the alternative for a departure from the null of $\beta = 0$, i.e., by imposing the population weight of the forecast to be tested equal to .01, .025, .05, .10, .20, and .25.

Table 10: Empirical and size-adjusted rejection probabilities for the analysis of the empirical size and size-adjusted power in a large sample and in a high dimensional framework

		Distance from the null						
		0 - Size	.01	.025	.05	.10	.20	.25
n : 1000, d : 20, ρ : 0								
Nonnegative $\boldsymbol{\beta}$ and $LM \sim \bar{\chi}^2$	Size-adj. power	0.050	0.100	0.370	0.898	1.000	1.000	1.000
	Rejection probability	0.051	0.103	0.374	0.900	1.000	1.000	1.000
$\boldsymbol{\beta}$ and $LM \sim \chi^2$	Size-adj. power	0.050	0.097	0.362	0.894	1.000	1.000	1.000
	Rejection probability	0.055	0.104	0.380	0.902	1.000	1.000	1.000
n : 40, d : 80, ρ : 0								
Nonnegative $\boldsymbol{\beta}$ and $LM \sim \bar{\chi}^2$	Size-adj. power	0.050	0.047	0.051	0.071	0.127	0.341	0.461
	Rejection probability	0.054	0.051	0.058	0.076	0.138	0.360	0.479
$\boldsymbol{\beta}$ and $LM \sim \chi^2$	Size-adj. power	0.050	0.046	0.051	0.057	0.096	0.242	0.349
	Rejection probability	0.035	0.033	0.037	0.041	0.071	0.198	0.290

Notes. We study the convergence of the empirical to the theoretical 5% size for a sample size of 1000 observations and the behavior of the proposed test in the high dimensional framework with 40 observations and 80 candidate forecasts. In this latter case, we estimate the unconstrained weights using Lasso. The increase of the sample size was necessary to estimate the penalty intensity via 5-fold cross-validation. We estimate the empirical size by generating the optimal forecast under the null hypothesis, i.e., by enforcing the weight to be tested to be zero in population. We conduct the power analysis by generating the optimal forecast under the alternative for a departure from the null of $\beta = 0$, i.e., by imposing the population weight of the forecast to be tested equal to .01, .025, .05, .10, .20, and .25.