ERASMUS UNIVERSITY ROTTERDAM

Erasmus School of Economics
Master Thesis Quantitative Finance

Date: 28-06-2024

Combining Forecasts using Forecaster Characteristics: Leveraging The Wisdom of a Dynamic Crowd

Author: Emma Ligthart
Student ID: 500511

Supervisor: Prof. Dr. D.J.C. van Dijk
Second assessor: Dr. R. Lonn

Abstract

This paper proposes a new forecast combination method called Forecaster Characteristics Combination (FCC), which leverages forecaster characteristics to combine forecasts. A key advantage of this method is its ability to combine a large number of forecasts without increasing the number of unknown parameters. Other advantages include its ability to handle unbalanced panel data and to incorporate numerous forecaster characteristics simultaneously, while remaining computationally inexpensive. In an empirical application, the FCC method is used to construct forecast combinations for three key U.S. macroeconomic variables: real GDP growth, unemployment, and inflation, using expert forecasts from the Survey of Professional Forecasters. The results show that the FCC method generally outperforms several benchmarks, including the equally weighted average. The proposed method achieves lower mean squared forecasting error for both inflation and real GDP growth.

Keywords: forecast combination, expert forecasts, point forecasts, unbalanced panel

Acknowledgements: I thank Prof. Dr. D.J.C. van Dijk for his supervision, feedback and useful suggestions.

Contents

1 Introduction

Combined forecasts are often more accurate than forecasts from a single person or model (Armstrong, 2001). This phenomenon, known as the wisdom of the crowd, suggests that a group's collective knowledge exceeds that of any single person or model. Ideally, one would like to exploit this wisdom. The main challenge of using the wisdom of the crowd lies in effectively combining individual forecasts. This challenge has been studied extensively in the econometric literature, following the seminal paper by Bates and Granger (1969). Initial efforts focused on simple combination methods such as the equally weighted average (Clemen & Winkler, 1986; Palm & Zellner, 1992). Later, more sophisticated methods arose, ranging from principal components combination to multiple temporal aggregation (Diebold & Shin, 2019; Petropoulos & Spiliotis, 2021).

This paper contributes to the field of forecast combination by introducing a new method. In particular, we propose a method called Forecaster Characteristics Combination (FCC), which is directly inspired by the parametric portfolio policy of Brandt et al. (2009). The FCC method combines forecasts by leveraging forecaster characteristics such as accuracy, bias and disconsensus. Specifically, the method calculates forecast combination weights by combining a characteristics-based term with a baseline weight. The characteristics-based term is determined by regressing the realisations of the predicted variable on the forecaster characteristics; a regression that can be estimated conveniently using ordinary least squares. The baseline weight is set to the equally weighted average, due to its simplicity and strong performance in the field of forecast combination (Genre et al., 2013). Apart from the regular FCC method, we propose several extended versions, which employ regularisation techniques such as Ridge, LASSO and Elastic Net.

The proposed FCC method has three main advantages. First and foremost, the FCC method is able to handle a large number of forecasters, without increasing the number of unknown parameters. This results in a significant dimensionality reduction compared to other forecast combination approaches, for which computational complexity typically increases with the number of forecasters. By avoiding the curse of dimensionality, the FCC method greatly reduces estimation uncertainty relative to other methods. Consequently, the FCC method has the potential to address the "forecast combination puzzle," which is discussed in more detail in Section 2.1.

A second advantage of the FCC method arises in the context of combining expert survey forecasts. When dealing with expert survey forecasts, an unbalanced panel emerges: unlike model forecasts, where the number of forecasts at each point in time is fixed, the amount of expert forecasts varies over time, because experts join and leave the panel. Such an unbalanced data structure limits the amount of feasible forecast combination methods, making many regression based and machine learning methods inapplicable (Montero-Manso et al., 2020; Stock & Watson, 2004). A notable advantage of the FCC method is that it is able to handle the highly unbalanced panel structure of expert surveys without needing to discard a substantial part of the data, which is generally required by other forecast combination methods. Specifically, the FCC method can handle varying amounts of forecasts over time due to the specification of weights as a function of forecaster characteristics: the number of characteristics is constant over time.

A third advantage of the FCC method is its ability to leverage numerous forecaster characteristics simultaneously, while remaining computationally inexpensive. Using forecaster characteristics to construct forecast combinations is not new: Bates and Granger (1969) already weighed forecasts inversely to the mean squared forecasting error (MSFE) of the corresponding forecaster. There are numerous other methods that exploit forecaster characteristics in a computationally inexpensive manner (see, *inter alia*, Aiolfi & Timmermann, 2006; Capistrán & Timmermann, 2009; Nowotarski et al., 2014; Pawlikowski & Chorowska, 2020). Yet, these methods exploit only one forecaster characteristic at the time. Recently, machine learning methods that leverage numerous forecaster characteristics simultaneously were employed to construct forecast combinations (see, *inter alia*, Kang et al., 2022; Li et al., 2020; Ma & Fildes, 2021; Montero-Manso et al., 2020).[1] However, compared to these methods, the FCC method is much less computationally expensive, as it employs standard OLS regression instead of intricate machine learning methods.

In an empirical application, the FCC method and its extensions are employed to construct point forecast combinations for three key U.S. macroeconomic variables: (i) real GDP growth, (ii) unemployment, and (iii) inflation. For that purpose, we employ expert panel data from the Survey of Professional Forecasters issued by the Federal Reserve Bank of Philadelphia. We construct separate forecast combinations for two horizons: nowcasts and one-step-ahead forecasts. To evaluate the forecasts combinations, we use first vintage macroeconomic realisation estimates from both the Federal Reserve Bank of Philadelphia and the Federal Reserve Economic Data dataset.

We compare the performance of the FCC method to several benchmarks, namely the (i) equally weighted average, (ii) median, (iii) trimmed mean, (iv) bias-adjusted mean, (v) discounted MSFE, and (vi) partially-egalitarian LASSO. Provided that we compare multiple models that are estimated on a single dataset, significant results may occur by chance−a problem known as the multiple comparison problem. To avoid this problem, we employ the model confidence set procedure by Hansen et al. (2011), which determines a set of best performing models.

The results show that the FCC method generally outperforms the benchmarks: it achieves superior forecasting performance for two of the three macroeconomic variables. Only for the unemployment rate most benchmarks outperform the FCC method. Nevertheless, the FCC method is selected into the model confidence set for the headline CPI inflation rate and the real GDP growth rate for both forecast horizons. The favourable performance of the FCC method can be attributed to its flexibility in varying the coefficients assigned to characteristics over time. As a consequence, the extensions of the FCC method−which employ Ridge, LASSO and ENet regularisation−do not improve the forecasting performance of the regular FCC method: they impose constraints on the model's flexibility. A sensitivity analysis shows that the results are generally robust, for example to the removal of outliers.

The novelty of this research lies in translating a method from the field of portfolio management to the field of forecast combination. In particular, we propose a new forecast combination method that is directly inspired by the parametric portfolio policy of Brandt et al. (2009). To our knowledge, this has not been explored before. In a broader societal context, forecasting mac-

[1] Even though these methods show promising results, they are computationally expensive and not well suited to handle the unbalanced nature of expert surveys. Therefore, they are not further considered in the empirical application of this paper.

roeconomic variables holds relevance, notably for policymakers who rely on these predictions to formulate their policies. Furthermore, improved macroeconomic forecasts benefit businesses and investors by providing more reliable data for planning and risk management.

Much of the econometric literature on forecast combination focuses on econometric model forecasts (Wang et al., 2023). Conversely, this paper focuses on expert forecasts. The reason is that several studies show that expert-based forecast combinations outperform model-based forecast combinations (see, *inter alia*, Ang et al., 2007; Lin et al., 2014; Loungani, 2001; Song et al., 2013; Van Dijk & Franses, 2019). Consequently, in this paper, the term "forecaster" refers to an expert that participates in a forecasting survey. However, the proposed FCC method has broader applicability and can also be used to combine model forecasts, in which case "forecaster" would refer to a forecasting model. In other words, although the empirical application in this paper focuses on the combination of expert forecasts, the FCC method can also be employed to combine model forecasts.

The rest of this paper is structured as follows. Section 2 explains the regular and extended FCC methods and connects them to the literature. Section 3 describes the empirical setting used to evaluate the FCC method. Section 4 elaborates on the forecaster characteristics included in the FCC method. Section 5 discusses the benchmark models and evaluation metrics. Section 6 presents the results, and finally, Section 7 provides the main conclusions and limitations of the paper.

2 Forecaster Characteristics Combination (FCC)

This section proposes a new forecast combination method called Forecaster Characteristics Combination (FCC). We first connect this method to the literature, in Section 2.1, before explaining it in more detail, in Sections 2.2 and 2.3. Implementation details are left to be discussed in Section 2.4.

2.1 Connection to the Literature

The FCC method builds upon the parametric portfolio policy by Brandt et al. (2009), a method from the portfolio management field. To understand why this method applies to the field of forecast combination, it is necessary to recognise the similarities between the two fields. Specifically, both the tasks and challenges encountered in portfolio management and forecast combination are similar. Despite the difference in application, the main task of both fields lies in constructing combination weights that sum to one at each point in time. An important challenge in both fields relates to achieving an optimal bias-variance trade-off.

In the field of forecast combination, one of the main challenges stems from a surprising result: the simple equally weighted average often outperforms more complex methods (Genre et al., 2013). This phenomenon is commonly referred to as the forecast combination puzzle, a term introduced by Stock and Watson (2004). Aiming to address this puzzle, Timmermann (2013) notes that the objective function underlying the combination problem is often mean squared forecasting error (MSFE) loss. In the context of forecast combination, MSFE loss can be decomposed into two parts: (i) squared bias, from the deviation between the estimated weights and the true weights and (ii) variance, from the estimation uncertainty of the weights themselves (Smith & Wallis, 2009). Consequently, to minimise the loss function, forecast combination methods should optimally balance the bias and variance introduced. In general, bias can be reduced by increasing model complexity, but this is often accompanied by an increase in the number of parameters to be estimated and, hence, an increase in estimation uncertainty (variance). Thus, the forecast combination puzzle goes to show that complex methods increase the variance more than they reduce bias, resulting in larger MSFE compared to the simple equally weighted average.

Achieving an optimal bias-variance trade-off is also a key challenge in the field of portfolio management. In portfolio management, bias refers to the systematic error in estimating combination weights, whereas variance is the extent to which these estimates fluctuate or deviate from their expected values. Both factors influence the performance of portfolio construction methods. Same as in forecast combination, the simple equally weighted average often outperforms more complex methods, a problem highlighted in the hallmark paper by DeMiguel et al. (2009). Thus, in the field of portfolio management too, sophisticated combination methods perform worse than the equally weighted average. The reason is the same as in the field of forecast combination: while complex models reduce bias, they increase estimation variance even more.

Given the similarities between the challenges encountered in forecast combination and portfolio management, it is sensible to explore the solutions proposed by the latter field. A prominent solution to estimate portfolio weights is provided by Brandt et al. (2009), who parameterise portfolio weights based on the corresponding asset's characteristics. Hence, in this approach, the number of parameters to be estimated is independent of the number of assets, but rather depends on the number of characteristics included. Provided that the number of characteristics is much smaller than the number of assets, this approach achieves a major dimensionality reduction. Such a dimensionality reduction is accompanied by a decrease in estimation uncertainty, because there are less parameters to be estimated. Consequently, the parametric portfolio policy shows potential to achieve favourable performance. Several empirical applications of the parametric portfolio policy confirm its strong performance compared to numerous benchmarks including the equally weighted average (Behr et al., 2012; Brandt et al., 2009; Hand & Green, 2011).

Both due to the promising results of the parametric portfolio policy by Brandt et al. (2009) and the parallels between portfolio management and forecast combination, it is interesting to translate this method to the field of forecast combination. The parametric portfolio policy parameterises asset combination weights using asset characteristics. Hence, applying the method to forecast combination would involve parameterising forecast combination weights using forecaster characteristics. We call this approach Forecaster Characteristics Combination (FCC) because it leverages forecaster characteristics to combine forecasts.

Similarly to the parametric portfolio policy, the FCC approach achieves a major dimensionality reduction: the number of parameters to be estimated reduces from the number of forecasters to the number of characteristics included, which is much smaller. As such, the FCC method shows potential to address the forecast combination puzzle: it likely exhibits lower estimation uncertainty compared to other methods and might therefore achieve lower MSFE. Additionally, the FCC approach effectively accommodates the dynamic nature of an expert panel. In particular, varying amounts of forecasts over time can be handled due to the parameterisation of weights as a function of forecaster characteristics: the number of characteristics is constant over time. The remainder of this paper focuses on developing the FCC method and analysing whether it improves forecasting performance in an empirical setting, compared to several of benchmarks.

2.2 Introduction to FCC

This section derives the FCC method and its mathematical notation. We start the derivation from the basics of forecast combination. An optimal forecast combination minimises the MSFE of the combined forecast over the weights assigned to the individual forecasts. Hence, a general notation of the forecast combination problem is:

$$\min_{\{w_{i,t,h}\}_{i=1,t=1}^{N_t,T}} \frac{1}{T}\sum_{t=1}^{T}(r_{t+h} - f^c_{t+h|t})^2, \quad r_{t+h}, f^c_{t+h|t} \in \mathbb{R}, \tag{1}$$

$$\text{s.t.} \quad \sum_{i=1}^{N_t} w_{i,t,h} = 1 \quad \forall t = 1,\ldots,T, \quad h \in \mathbb{Z}_{0,+}. \tag{2}$$

In these expressions, r_{t+h} denotes the realisation of the variable of interest for time period $t+h$. Index t indicates time ($t = 1,\ldots T$) and index h denotes the forecast horizon. Index h takes on positive integer values including zero, since forecasts for the current ($h = 0$) and any future horizons may be considered.[2] The problem in equation (1) is optimised for each horizon separately. Furthermore, $f^c_{t+h|t}$ denotes the forecast combination at time t for horizon h, which is equal to $\sum_{i=1}^{N_t} w_{i,t,h} \times f_{i,t+h|t}$. Weight $w_{i,t,h}$ denotes the combination weight assigned to panel member i at time t for horizon h. These combination weights are restricted to sum to one over all forecasters, at each point in time for a given horizon, by the restriction in equation (2). Moreover, $f_{i,t+h|t}$ denotes the individual forecast of panel member i for time period $t+h$ given information up to time t. The number of forecasters runs from $i = 1,\ldots,N_t$, such that the total number of forecasters N_t varies over time. This variation is due to the dynamic nature of the panel: members join and leave.

In sum, the crux of forecast combination is finding combination weights $w_{i,t,h}$ that minimise expression (1). The essence of the parametric portfolio policy by Brandt et al. (2009) is to parameterise portfolio weights as a function of an asset's characteristics. To translate this approach to forecast combination, we propose to parameterise the forecast combination weights $w_{i,t,h}$ as a function of forecaster characteristics:

$$w_{i,t,h} = q(\mathbf{z}_{i,t,h}; \boldsymbol{\theta}_h), \quad \mathbf{z}_{i,t,h} \in \mathbb{R}^Q, \tag{3}$$

in which $\mathbf{z}_{i,t,h}$ denotes a vector of Q characteristics of panel member i at time t for horizon h, $\boldsymbol{\theta}_h$ is a vector of coefficients to be determined, and the function $q(.)$ denotes a function that transforms forecaster characteristics into a combination weight using $\boldsymbol{\theta}_h$. We specify the function $q(.)$ similar to Brandt et al. (2009), who propose combining a characteristics-based term with a

[2]For example, when forecasting macroeconomic variables, it is possible to make forecasts for the current quarter because the realisation estimates are typically published at least one quarter later.

baseline weight. Hence, we specify $q(.)$ as:

$$w_{i,t,h} = \bar{w}_{i,t,h} + \frac{1}{N_t}\boldsymbol{\theta}'_h\hat{\mathbf{z}}_{i,t,h}, \quad \bar{w}_{i,t,h} \in \mathbb{R}, \quad \boldsymbol{\theta}_h, \hat{\mathbf{z}}_{i,t,h} \in \mathbb{R}^Q. \tag{4}$$

In this expression, $\bar{w}_{i,t,h}$ denotes the baseline weight, which is restricted to sum to one over all forecasters, at each point in time, for a given horizon. Furthermore, $\boldsymbol{\theta}'_h$ is the transpose of the vector $\boldsymbol{\theta}_h$, which is set to dimensionality Q. Vector $\hat{\mathbf{z}}_{i,t,h}$ denotes a vector of *standardised* characteristics of panel member i. This vector is created by cross-sectionally standardising the characteristics in vector $\mathbf{z}_{i,t,h}$ to mean zero, for each time period.

This standardisation occurs for two reasons. Firstly, it ensures stationarity of the characteristics over time. Secondly and importantly, the standardisation ensures that the combination weights $w_{i,t,h}$ sum to one over all forecasters i, at each point in time, for a given horizon. It thus ensures that the restriction in equation (2) is met. In more detail, the standardisation ensures that the cross-sectional mean of $\boldsymbol{\theta}'_h\hat{\mathbf{z}}_{i,t,h}$ equals zero at each point in time. Hence, together with the restriction that the baseline weights sum to one, this standardisation ensures that the portfolio weights sum to one at all points in time.

The standardisation has some consequences for the interpretation of the forecaster characteristics: they measure relative rather than absolute values. We further elaborate on this issue in Section 4.9. Furthermore, in our standardisation of forecaster characteristics, we deviate from Brandt et al. (2009), who also standardise the characteristics to standard deviation one. This additional standardisation step removes information about the different variability of the characteristics. To retain more information about the forecaster characteristics, we omit the standardisation to standard deviation one. Nevertheless, we include a sensitivity analysis in Section 6.4 to demonstrate the effects of standardising to standard deviation one.

Besides the standardisation, the normalisation term $\frac{1}{N_t}$ in equation (4) is important to address. This normalisation guarantees that the combination weight function can be applied to an arbitrary and time-varying number of panel members over time. Without this normalisation, a larger number of panel members at one point in time results in larger deviations from the baseline weight. Thus, the normalisation accommodates the dynamic nature of the expert panel.

With the specification of the combination weights in equation (4), we follow Brandt et al. (2009) in defining weights that vary around a baseline weight. To complete the definition of the combination weights, we have to determine the baseline weight $\bar{w}_{i,t,h}$. Brandt et al. (2009) set the baseline weight to the equally weighted average. We follow them in using the equally weighted average as baseline, due to its simplicity and strong performance in the field of forecast combination (Genre et al., 2013). With equal weights as baseline weight $\bar{w}_{i,t,h}$, the specification of combination weights becomes:

$$w_{i,t,h} = \frac{1}{N_t} + \frac{1}{N_t}\boldsymbol{\theta}'_h\hat{\mathbf{z}}_{i,t,h} \tag{5}$$

$$= \frac{1}{N_t}(1 + \boldsymbol{\theta}'_h\hat{\mathbf{z}}_{i,t,h}). \tag{6}$$

Using this expression for the combination weights, the optimisation problem from equation (1) can be written as:

$$\min_{\boldsymbol{\theta}_h} \frac{1}{T} \sum_{t=1}^{T} \left(r_{t+h} - \left(\sum_{i=1}^{N_t} \frac{1}{N_t} (1 + \boldsymbol{\theta}_h' \hat{\mathbf{z}}_{i,t,h}) \times f_{i,t+h|t} \right) \right)^2. \tag{7}$$

This adaptation of the optimisation problem drastically reduces dimensionality. In equation (1) the minimisation is executed over weights $w_{i,t,h}$ for all panel members $i = 1, \ldots, N_t$ and time periods $t = 1, \ldots, T$, resulting in a dimensionality of $\sum_{t=1}^{T} N_t$. Conversely, in the revised problem in equation (7), the optimisation is executed over a vector $\boldsymbol{\theta}_h$ with dimensionality Q, which is the total number of characteristics. Therefore, the computational complexity of the minimisation problem in equation (7) scales with the number of characteristics Q, rather than with the total number of unique forecasters N_t across all time periods $t = 1, \ldots, T$. Provided that Q is much smaller than $\sum_{t=1}^{T} N_t$, the dimensionality of the optimisation problem is substantially reduced. Even when we are dealing with a balanced panel ($N_t = N$ for all t) and constant weights over time ($w_{i,t,h} = w_{i,h}$ for all t), the number of parameters changes from N to Q. This still offers a substantial dimensionality reduction in many cases.

The optimisation problem in equation (7) can be conveniently estimated using ordinary least squares (OLS). A full derivation of the least squares estimator is provided in Appendix A. In short, we rewrite equation (7) in least squares format. To that end, we define:

$$y_t := r_{t+h} - \left(\frac{1}{N_t} \sum_{i=1}^{N_t} f_{i,t+h|t} \right), \quad y_t \in \mathbb{R}, \tag{8}$$

$$\mathbf{x}_t := \frac{1}{N_t} \sum_{i=1}^{N_t} \hat{\mathbf{z}}_{i,t,h} \times f_{i,t+h|t}, \quad \mathbf{x}_t \in \mathbb{R}^Q. \tag{9}$$

In these expressions, y_t represents the dependent variable and $\mathbf{x}_t$ the vector of independent variables in the regression. Using these definitions, the optimisation problem in equation (7) can be rewritten as:

$$\min_{\boldsymbol{\theta}_h} \frac{1}{T} \sum_{t=1}^{T} (y_t - \boldsymbol{\theta}_h' \mathbf{x}_t)^2 \tag{10}$$

No intercept is included as y_t is expected to be centred very close to zero. To see this, note that y_t equals the difference between the realisation r_{t+h} and the equally weighted average forecast. Provided that the equally weighted average forecast approximately centres around the value of the realisation, their difference$-y_t-$centres approximately around zero.

Next, we stack the elements $y_1, \ldots, y_T$ into vector $\mathbf{y}$ of dimension $T \times 1$. Additionally, we stack the transposed vectors $\mathbf{x}_1', \ldots, \mathbf{x}_T'$ into matrix $\mathbf{X}$ of dimension $T \times Q$. We leave the factor $\frac{1}{T}$ out of the minimisation problem in equation (7) since it does not affect the solution. As a result, the minimisation problem from equation (7) can be rewritten as:

$$\min_{\boldsymbol{\theta}_h} \|\mathbf{y} - \mathbf{X}\boldsymbol{\theta}_h\|_2, \quad \mathbf{y} \in \mathbb{R}^T, \quad \mathbf{X} \in \mathbb{R}^{T \times Q}, \quad \boldsymbol{\theta}_h \in \mathbb{R}^Q, \tag{11}$$

in which $\|.\|_2$ denotes the L2 norm. We recognise a standard OLS problem in equation (11).

Consequently, the least squares estimator for $\boldsymbol{\theta}_h$ is:

$$\hat{\boldsymbol{\theta}}_h = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{y}. \tag{12}$$

With this estimator for $\boldsymbol{\theta}_h$, the FCC forecast combinations are calculated as:

$$f^{c,\text{FCC}}_{t+h|t} = \sum_{i=1}^{N_t} \frac{1}{N_t}(1 + \hat{\boldsymbol{\theta}}'_h \hat{\mathbf{z}}_{i,t,h}) \times f_{i,t+h|t}. \tag{13}$$

2.3 Extensions of FCC

The model in equation (7) can be viewed as a "kitchen-sink" model in the sense that it estimates Q variable coefficients simultaneously (Rapach et al., 2010). As a consequence, regularisation on these variables may further improve the bias-variance trade-off of the solution. Therefore, we extend the FCC method using several regularisation techniques: Ridge, LASSO, post-LASSO and Elastic Net. These extensions are described in Section 2.3.1. Additionally, to avoid the "kitchen-sink" effect, we estimate each of the Q characteristic coefficients separately and combine the resulting single characteristic forecasts in three manners. These extensions are described in Section 2.3.2.

2.3.1 Regularisation of the FCC Method

Econometric literature includes a wide range of regularisation techniques. The central idea is to penalise the size of the regression coefficients−in our case the coefficients in $\boldsymbol{\theta}_h$. The purpose of this penalisation is to reduce the estimation variance and thereby improve the bias-variance trade-off of the solution. A well-known regularisation technique is Ridge regression, which was introduced by Hoerl and Kennard (1970), who derived the method from Tikhonov (1943). This method introduces a smooth convex penalty−the L2 norm−to shrink the coefficients towards zero. Under Ridge regression our estimation problem becomes:

$$\min_{\boldsymbol{\theta}_h} \frac{1}{T}\sum_{t=1}^{T}\left(r_{t+h} - \left(\sum_{i=1}^{N_t}\frac{1}{N_t}(1 + \boldsymbol{\theta}'_h \hat{\mathbf{z}}_{i,t,h}) \times f_{i,t+h|t}\right)\right)^2 + \lambda \sum_{q=1}^{Q}(\theta_{h,q})^2, \quad \lambda \in \mathbb{R}, \tag{14}$$

in which $\theta_{h,q}$ denotes the q-th element of $\boldsymbol{\theta}$ and λ is a penalty parameter that controls the degree of shrinkage. We estimate $\hat{\boldsymbol{\theta}}^{\text{Ridge}}_h$ using equation (14) and construct FCC-Ridge combination forecasts as follows:

$$f^{c,\text{Ridge}}_{t+h|t} = \sum_{i=1}^{N_t} \frac{1}{N_t}(1 + \hat{\boldsymbol{\theta}}^{\text{Ridge}\prime}_h \hat{\mathbf{z}}_{i,t,h}) \times f_{i,t+h|t}. \tag{15}$$

A downside of Ridge regression is that it is unable to select characteristics. In the context of this paper, variable selection is desirable because it is interesting to know which forecaster characteristic is relevant and which is not.

A regularisation technique that both shrinks and selects coefficients is LASSO, introduced by Tibshirani (1996). This method is able to select coefficients, by shrinking other coefficients all the way to zero. Therefore, we also perform LASSO regularisation as an extension to the regular FCC method. To shrink the coefficients towards zero, LASSO regularisation introduces an L1 penalty term on the coefficients. In our FCC framework this can be expressed as follows:

$$\min_{\boldsymbol{\theta}_h} \frac{1}{T}\sum_{t=1}^{T}\left(r_{t+h} - \left(\sum_{i=1}^{N_t}\frac{1}{N_t}(1+\boldsymbol{\theta}_h'\hat{\mathbf{z}}_{i,t,h}) \times f_{i,t+h|t}\right)\right)^2 + \lambda\sum_{q=1}^{Q}\left|\theta_{h,q}\right|, \quad \lambda \in \mathbb{R}, \tag{16}$$

in which $|.|$ denotes the absolute value, and all other variables are as defined before. We estimate $\hat{\boldsymbol{\theta}}_h^{\text{LASSO}}$ using equation (16) and construct FCC-LASSO combination forecasts similarly as in equation (15).

In addition to direct LASSO regularisation, we implement a post-LASSO method, following DeMiguel et al. (2020). This post-LASSO approach first estimates $\hat{\boldsymbol{\theta}}_h^{\text{LASSO}}$ and then removes the characteristics with coefficients equal to zero. Subsequently, a reduced optimisation problem is estimated using the least squares estimator from equation (12), but only including the LASSO-selected characteristics. DeMiguel et al. (2020) argue that this re-estimation circumvents consistency concerns raised by Chatterjee and Lahiri (2011). The resulting forecast combination is referred to as $f_t^{c,\text{postLASSO}}$.

A disadvantage of LASSO regularisation is that it struggles with highly correlated variables, often arbitrarily selecting one variable from a group of highly correlated variables (Fan & Li, 2001). In this paper, we construct forecaster characteristics that exhibit high correlations (further discussed in Section 4.9). Therefore, it is important to address this issue and implement a regularisation technique that handles correlated variables well. A well-known technique that fulfils this requirement is Elastic Net regularisation (ENet), introduced by Zou and Hastie (2005). Next to handling correlated variables well, ENet regularisation is still able to select coefficients, similar to LASSO. For that reason, we estimate the FCC method with ENet regularisation. ENet regularisation combines two penalty terms: the LASSO penalty and the Ridge penalty. Consequently, the optimisation problem under Elastic Net regularisation is formulated as follows:

$$\min_{\boldsymbol{\theta}_h} \frac{1}{T}\sum_{t=1}^{T}\left(r_{t+h} - \left(\sum_{i=1}^{N_t}\frac{1}{N_t}(1+\boldsymbol{\theta}_h'\hat{\mathbf{z}}_{i,t,h}) \times f_{i,t+h|t}\right)\right)^2 + \lambda\sum_{q=1}^{Q}\left(\alpha\theta_{h,q}^2 + (1-\alpha)\left|\theta_{h,q}\right|\right), \ \lambda, \alpha \in \mathbb{R}, \tag{17}$$

in which hyperparameter α determines the mixture between the Ridge and LASSO penalty. We estimate coefficient $\hat{\boldsymbol{\theta}}_h^{\text{ENet}}$ using equation (17) and construct FCC-ENet combination forecasts similarly as in equation (15).

2.3.2 Single Characteristic Estimation

Apart from regularising the coefficients in the FCC model, we extend the FCC method with an approach called single characteristic estimation. As mentioned previously, the regular FCC model can be regarded as a "kitchen-sink" model, in which Q coefficients are estimated simultaneously. To mitigate this "kitchen-sink" effect, we estimate each of the Q characteristic coefficients separately, which results in Q single characteristic forecasts. We combine these single characteristic forecasts into a final forecast combination in three manners: by taking the average over (i) all single characteristic forecasts, (ii) the LASSO selected single characteristic forecasts and (iii) the ENet selected single characteristic forecasts.

This single characteristic estimation approach is inspired by Rapach et al. (2010), who find that forecast combinations from single-variable regressions outperform forecast combinations from "kitchen-sink" models, in which all variables are estimated simultaneously. Rapach et al. (2010) propose a forecast combination method that involves estimating single-predictor forecasts from single-predictor regressions. They combine these single-predictor forecasts by taking the mean and find favourable results. In a subsequent paper, Rapach and Zhou (2020) extend this combination method to Combination Elastic Net (C-ENet). For this method, they refine the single-predictor combination step using regularisation: they select a subset of single-predictor forecasts to be combined using Elastic Net regression.

Drawing from this literature, we implement three different single characteristic estimation methods. Firstly, we apply the simple single characteristic combination method from Rapach et al. (2010). Secondly, we implement the C-ENet method from Rapach and Zhou (2020). Provided that the choice of Elastic Net regularisation in the selection step of C-ENet seems arbitrary, we are interested in replacing it with LASSO regularisation. Hence, thirdly, we implement a C-LASSO method, which extends the C-ENet method by using LASSO instead of Elastic Net.

All three single characteristic estimation methods are executed in three steps: a construction, selection and combination step. The first and third step are identical for each method, but the second step differs. The steps are as follows:

Step 1: Construction of Single Characteristic Forecasts

In the first step, we construct forecast combinations for each characteristic separately, which results in Q single characteristic forecasts. Particularly, we estimate coefficient $\theta_{h,q}$ for each characteristic separately by optimising:

$$\min_{\theta_{h,q}} \frac{1}{T} \sum_{t=1}^{T} \left(r_{t+h} - \left(\sum_{i=1}^{N_t} \frac{1}{N_t} (1 + \theta_{h,q} \hat{z}_{i,t,h,q}) \times f_{i,t+h|t} \right) \right)^2, \quad \text{for} \quad q = 1, \ldots, Q, \tag{18}$$

in which $\hat{z}_{i,t,h,q}$ denotes the q-th element of the vector of standardised characteristics $\hat{\mathbf{z}}_{i,t,h}$ and all other variables are as defined before. Note that this optimisation problem is similar to the one in equation (7), but only considers one characteristic at the time. We estimate $\hat{\theta}_{h,q}$ using standard OLS. Consequently, we construct single characteristic forecasts for characteristic q as:

$$f^{c,q}_{t+h|t} = \sum_{i=1}^{N_t} \frac{1}{N_t} (1 + \hat{\theta}_{h,q} \hat{z}_{i,t,h,q}) \times f_{i,t+h|t} \tag{19}$$

Step 2: Selection of Single Characteristic Forecasts

Step 2a: Simple Combination. In the simple single characteristic estimation method, all Q single characteristic forecasts are included in the forecast combination (Rapach et al., 2010).

There is no intermediate selection step.

Step 2b: C-ENet. In the C-ENet method, a subset of single characteristic forecasts is selected using Elastic Net regularisation (Rapach & Zhou, 2020). Specifically, a number of single characteristic forecasts $f_{t+h|t}^{c,q}$ is selected with the following regression:

$$\min_{\boldsymbol{\kappa}_h,\lambda,\alpha} \frac{1}{T}\left(r_{t+h} - \left(\sum_{q=1}^{Q} \kappa_{h,q} \times f_{t+h|t}^{c,q}\right)\right)^2 + \lambda \sum_{q=1}^{Q}(\alpha\kappa_{h,q}^2 + (1-\alpha)\left|\kappa_{h,q}\right|), \quad \boldsymbol{\kappa}_h \in \mathbb{R}^Q, \quad \lambda, \alpha \in \mathbb{R}, \tag{20}$$

in which $\boldsymbol{\kappa}_h$ is a vector of coefficients to be estimated, while λ and α are hyperparameters that determine the degree of shrinkage of $\boldsymbol{\kappa}_h$. The Elastic Net procedure "kills" some of the single characteristic forecasts $f_{t+h|t}^{c,q}$, by setting $\kappa_{h,q}$ to zero. Only single characteristic forecasts whose corresponding coefficient $\kappa_{h,q}$ is not set to zero, are selected to be combined in Step 3.

Step 2c: C-LASSO. The C-LASSO method is based on the C-ENet method by Rapach and Zhou (2020) and selects a subset of single characteristic forecasts using LASSO instead of Elastic Net. Hence, we estimate:

$$\min_{\boldsymbol{\kappa}_h,\lambda} \frac{1}{T}\left(r_{t+h} - \left(\sum_{q=1}^{Q} \kappa_{h,q} \times f_{t+h|t}^{c,q}\right)\right)^2 + \lambda \sum_{q=1}^{Q}\left(\kappa_{h,q}\right)^2, \quad \boldsymbol{\kappa}_h \in \mathbb{R}^Q, \quad \lambda \in \mathbb{R}, \tag{21}$$

in which $\boldsymbol{\kappa}_h$ is a vector of coefficients to be estimated and λ is a hyperparameter. Similar to the ENet procedure, the LASSO procedure "kills" some of the single characteristic forecasts by setting $\kappa_{h,q}$ to zero. Again, only single characteristic forecasts, whose corresponding coefficient $\kappa_{h,q}$ is not set to zero, are selected to be combined in Step 3.

Step 3: Combination of Single Characteristic Forecasts

The single characteristic forecasts selected in step 2 are combined with their equally weighted average. In particular, let $\mathcal{J}_t$ denote the index set of the characteristics selected in Step 2. Then, the final single characteristic combination forecast is computed as:

$$f_{t+h|t}^{c} = \frac{1}{|\mathcal{J}_t|}\sum_{q \in \mathcal{J}_t} f_{t+h|t}^{c,q}, \tag{22}$$

in which $|\mathcal{J}_t|$ denotes the cardinality of $\mathcal{J}_t$.

This approach results in three different forecast combinations: (i) one from simple single characteristics estimation $f_{t+h|t}^{c,\mathrm{C}}$, (ii) one from C-ENet $f_{t+h|t}^{c,\text{C-ENet}}$ and (iii) one from C-LASSO $f_{t+h|t}^{c,\text{C-LASSO}}$. The application of these methods in this paper deviates from Rapach et al. (2010) and Rapach and Zhou (2020), by combining multiple forecast *combinations* into a single forecast combination, rather than combining a number of forecasts into a single forecast combination. Hence, the performance of the methods applied in this manner remains an empirical question.

2.4 Implementation Details

There are several implementation details of the FCC methods that are worth mentioning: (i) the estimation of $\boldsymbol{\theta}_h$, (ii) the estimation of hyperparameters λ and α, and (iii) the construction of the characteristics in $\hat{\mathbf{z}}_{i,t,h}$.

Estimation of $\boldsymbol{\theta}_h$

For the out-of-sample estimation of $\boldsymbol{\theta}_h$, we employ both a moving and an expanding window approach. In both approaches, the coefficients in $\boldsymbol{\theta}_h$ are estimated at each point in time, using only information available at that point in time. Consequently, the combination weights are re-estimated at each point in time. According to Diebold and Pauly (1987), such time-varying weights improve forecasting performance due to their ability to incorporate changes. After iterative experimentation, we choose a window size of 40 time periods, which corresponds to a time span of ten years. It is important to ensure large enough window size to avoid noisy weight estimates (Baumeister & Kilian, 2015).

In the moving window approach, we use a fixed window size of 40 observations. During each time period t, the model parameters are re-estimated using the 40 most recent observations. This approach ensures that the estimation relies only on the most recent data, allowing the model to accommodate potential changes over time. Conversely, in the expanding window approach, only the initial estimation is conducted using the first 40 observations. Then, as new data becomes available, it is added to the estimation dataset. This method leverages the entire history available at time t for parameter estimation. Expanding window estimation thus incorporates more information over time, compared to the moving window approach. However, the expanding window approach is less effective at accommodating changes over time. Which of the estimation procedures performs better in our context remains an empirical question to be answered in Section 6.2.

Estimation of hyperparameters λ and α

All hyperparameters (λ and α) used in the extensions of the regular FCC model are estimated using 5-fold cross validation, which is standard practice in machine learning. The hyperparameters are estimated separately for each FCC model extension. The estimation procedure for the hyperparameters follows the estimation procedure of the coefficients in $\boldsymbol{\theta}_h$. Hence, if the coefficients in $\boldsymbol{\theta}_h$ are estimated with the moving window approach, the hyperparameters are estimated with the moving window approach as well, using the same window of observations. Consequently, the hyperparameters are re-optimised at each point in time.

Estimation of characteristics in $\hat{\mathbf{z}}_{i,t,h}$

The estimation of the characteristics in $\hat{\mathbf{z}}_{i,t,h}$ is elaborately discussed in Section 4. It is important to mention here, that these characteristics remain the same regardless of the estimation approach of $\boldsymbol{\theta}_h$. Hence, also in the moving window approach, the characteristics in $\hat{\mathbf{z}}_{i,t,h}$ may exploit all information available up to time t and not just the 40 most recent observations.

3 Empirical Setting

To evaluate the FCC method, we establish an empirical setting. We evaluate the FCC method by combining forecasts on three key macroeconomic variables. The data employed for this setting consists of two parts. Firstly, we use expert forecasts from the Survey of Professional Forecasters (SPF), described in Section 3.1. Secondly, we employ datasets with macroeconomic realisations from the Federal Reserve Bank of Philadelphia and the Federal Reserve Economic Data database, described in Section 3.2.

3.1 Expert Forecasts

We choose to use expert forecasts from the Survey of Professional Forecasters, because it is the longest-running survey of macroeconomic forecasts in the U.S. The survey originates from a collaboration between the American Statistical Association and the National Bureau of Economic Research, who started the survey in the fourth quarter of 1968. Currently, the Federal Reserve Bank of Philadelphia issues the survey on a quarterly basis. The panel members are professionals whose job responsibilities include forecasting economic variables (Croushore, 1993). They employ statistical models, important indicators and other surveys to construct their forecasts. Their submissions are provided under anonymity and are identified only by an ID-number.

Although the SPF requests forecasts for 32 macroeconomic variables (Sill, 2012), this paper focuses on three key variables: real GDP, unemployment, and inflation, following Genre et al. (2013). Specifically, we use forecasts of the real GDP level, the annualised headline CPI inflation rate, and the average unemployment rate.[3] In the GDP level dataset, forecasts up to 1992 concern real GNP, while forecasts after 1992 concern real GDP. To reduce the effect of this change and to ensure stationarity, we transform the GDP level forecasts to GDP *growth rate* forecasts. Hence, we consider forecasts of the GDP quarterly growth rate instead of the level.

We construct separate forecast combinations for two horizons: (i) forecasts for the current quarter, referred to as "nowcasts", and (ii) forecasts for the next quarter, referred to as "one-step-ahead forecasts." The nowcasts concern the quarter during which the survey is completed. They can still be regarded forecasts because realisations of macroeconomic variables are typically published at least one quarter later. However, nowcasts may contain relevant information that is already available when forecasting, which is why we also consider one-step-ahead forecasts.

Table 3.1 presents the summary statistics of the SPF forecasts included in this paper. Overall, we include forecasts from Q4 1968 to Q4 2023, covering 221 time periods. Only the CPI dataset starts later, and runs from Q3 1981 to Q4 2023, covering 170 time periods. The total number of forecasts included varies for each macroeconomic variable, due to the dynamic nature of the panel. The total number of unique forecasters is highest for real GDP and unemployment (both equal to 457 forecasters) and lowest for CPI inflation (265 forecasters). This difference may be attributed to the longer observation periods available for real GDP and unemployment. For all variables, there are at least 9 forecasts available at each point in time. The maximum number of forecasts at one point in time varies across variables and is highest for real GDP, equal to 83

[3] The SPF datasets can be downloaded from https://www.philadelphiafed.org/surveys-and-data/real-time-data-research/individual-forecast.

forecasts. On average, forecasters provide forecasts for around 20 time periods. However, the total number of forecasts per expert varies considerably between a minimum of 1 forecast and a maximum of 125 forecasts.

Table 3.1: Summary Statistics of SPF Expert Panel Forecast

	GDP	**UMP**	**CPI**
Time Periods Included	Q4 1968 to Q4 2023	Q4 1968 to Q4 2023	Q3 1981 to Q4 2023
Total Number of Nowcasts Included	8,494	8,539	5,932
Total Number of One-Step-Ahead Included	8,493	8,538	5,931
Total Number of Unique Forecasters	457	457	265
Average Number of Forecasters	34.28	38.49	34.75
Minimum Number Forecasters	9	9	9
Maximum Number Forecasters	83	78	53
Average Time Periods per Forecaster	18.51	18.61	22.29
Minimum Time Periods per Forecaster	1	1	1
Maximum Time Periods per Forecaster	125	124	119

Note: This table presents summary statistics of the SPF forecasts included in this study. "GDP" stands for GDP level forecasts, "UMP" for the unemployment rate forecasts and "CPI" for the CPI inflation rate forecasts. "Total Number of Unique Forecasters" is the total number of unique forecasters over the entire forecasting period. "Average Number of Forecasters" is the average number of forecasts available at each point in time. "Minimum and Maximum Number of Forecasters" regard the number of forecasters included at one point in time. "Average Time Periods per Forecaster" is the average number of time periods forecasters provide a forecast. "Minimum and Maximum Time Periods" regard the number of time periods a forecaster provides a forecast.

Figure 3.1 further illustrates the dynamic nature of the panel, presenting the panel size of the real GDP forecasts over time. The panel sizes of unemployment and CPI are nearly identical and therefore not shown. Figure 3.1 clearly highlights the large amount of variation in the number of forecasters over time. The panel size reaches a minimum of 9 forecasts in the second quarter of 1990, but quickly recovers. This drop and steep increase in the number of forecasters can be attributed to the 1990 transfer of the survey from the American Statistical Association to the Federal Reserve Bank of Philadelphia (Croushore, 1993).

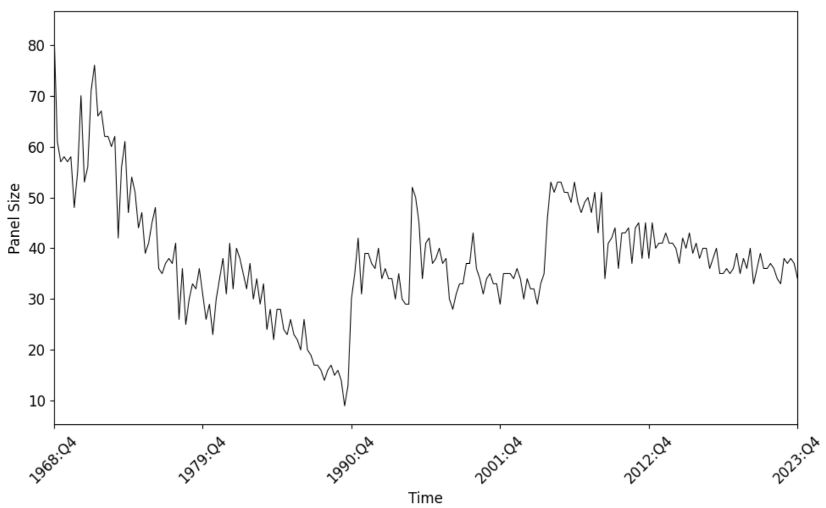


Figure 3.1: SPF Panel Size over Time for Real GDP

3.2 Macroeconomic Realisations

We employ data on macroeconomic realisations to evaluate the forecast combinations. Unfortunately, only estimates of these realisations are available and they are subject to revisions over time. The Federal Reserve Bank of Philadelphia keeps track of these revisions and offers a real-time dataset, which provides vintages from the first up to the most recent one. To illustrate the significance of these revisions, we consider an example: the first vintage estimate of GDP for Q1 1990 was published in Q2 1990 at 4,196, whereas the most recent estimate for Q1 1990, as of Q1 2024, equals 10,091. Hence, the estimate of Q1 1990 GDP has more than doubled due to revisions over time. These revisions occur because of changes in definitions, classifications, statistical methodology, and increased availability of information (Fixler et al., 2021). We minimise the influence of GDP level revisions by evaluating GDP growth rates rather than levels.

The revision of macroeconomic realisations poses a challenge: the choice of vintage affects the evaluation of the forecasts combinations. One may argue that the most recent vintage estimates are more reliable than the older first vintage estimates. However, the survey forecasts are conducted in "real time," such that panel members rely only on information available at the time of the survey. Because of this "real time" nature of the survey forecasts, we evaluate the forecasts based on first vintage estimates, following Genre et al. (2013).

I use datasets with first vintage estimates of the macroeconomic variables under consideration from the real-time dataset for macroeconomists by the Federal Reserve Bank of Philadelphia and the Federal Reserve Economic Data database.[4] Appendix B offers a more detailed description of the datasets and the data transformations.

Table 3.2: Summary Statistics of Macroeconomic Realisations

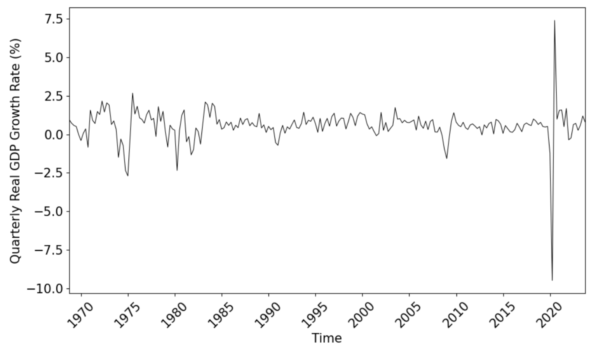
	GDP	**UMP**	**CPI**
Count	221	221	170
Mean	0.57	6.06	2.97
Standard Deviation	1.10	1.72	2.37
First-order Autocorrelation	0.03	0.90	0.42
Minimum	−9.49	3.33	−8.55
25% Quantile	0.31	4.83	1.88
50% Quantile	0.63	5.73	2.94
75% Quantile	0.97	7.17	3.88
Maximum	7.41	13.03	11.92

Note: This table presents the summary statistics of the macroeconomic realisations for all three economic variables. "GDP" represents the real GDP quarterly growth rate, "UMP" the unemployment rate and "CPI" the annualised CPI inflation rate. All three variables are in percentages.

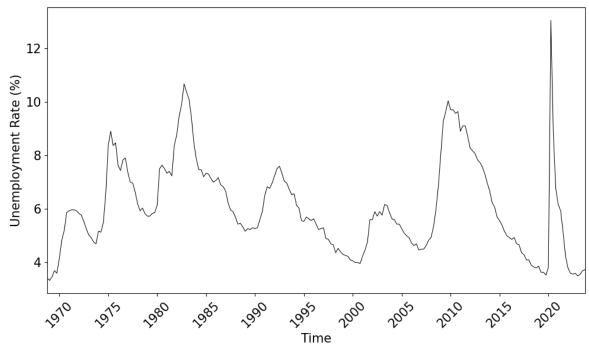
Table 3.2 provides the summary statistics of the realisations for all three macroeconomic variables after data transformations. The standard deviations of the unemployment rate and the real GDP growth rate are similar, whereas the CPI inflation rate exhibits larger variability.

[4]These datasets are available at https://www.philadelphiafed.org/surveys-and-data/real-time-data-research/real-time-data-set-for-macroeconomists and https://alfred.stlouisfed.org/series?seid=CPIAUCSL.

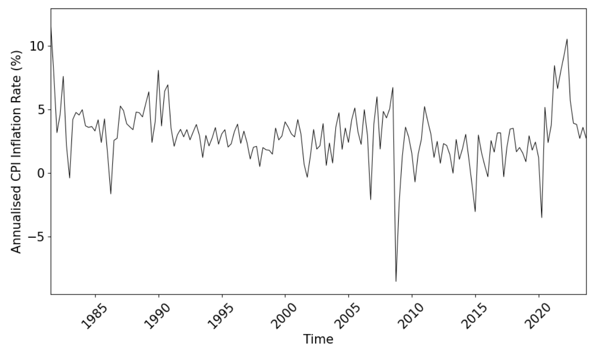
Furthermore, there is considerable variation in the first-order autocorrelation of the economic variables. In particular, the real GDP quarterly growth rate shows almost no autocorrelation, whereas the unemployment rate shows high autocorrelation. This suggests that predicting the GDP quarterly growth rate may be more challenging compared to predicting the unemployment rate. As expected, both the CPI inflation rate and the real GDP growth rate exhibit positive and negative values, whereas the unemployment rate takes on only positive values. All three macroeconomic variables contain some outliers, with a maximum as large as thirteen times the mean for the real GDP growth rate, twice the mean for the unemployment rate and four times the mean for the CPI inflation rate.



(a) Real GDP Growth Rate



(b) Unemployment Rate



(c) CPI Inflation Rate

Figure 3.2: Time Series of Macroeconomic Realisations

To identify these outliers, Figure 3.2 presents time series plots of the three macroeconomic variables. There are several outliers worth mentioning. Firstly, an important outlier occurs during the onset of the COVID-19 pandemic in Q2 and Q3 2020. During this period, all macroeconomic variables reach unusual levels: the unemployment and inflation rates peak, and the GDP growth rate first plummets and then peaks. Figure 3.2b shows that the unemployment rate also peaks during both the 1981-1982 recession and the Great Recession from 2008-2009. During the Great Recession as well, the CPI inflation rate reaches noteworthy negative values, especially during Q4 2008, see Figure 3.2c. The negative inflation rate can be explained by the collapse of several financial institutions in September 2008, which triggered panic in the financial markets. This panic led to a sharp decline in consumer and business spending, resulting in reduced demand for goods and services. In response, businesses lowered prices to attract customers, contributing to a steep decrease in the CPI inflation rate. Although all outliers are included in our main results, we provide a sensitivity analysis omitting several outliers in Section 6.4.

4 Forecaster Characteristics

This section elaborates on the forecaster characteristics included in the FCC method. Table 4.1 provides an overview of all characteristics included in this study and their expected coefficient sign. The SPF identifies panel members only by an ID-number, such that most forecaster characteristics are constructed from the data. As a consequence, almost all characteristics depend on forecast horizon h. The effect of determining characteristics for each horizon separately is examined in the sensitivity analysis in Section 6.4. The following sections describe each of the characteristics: Sections 4.1 to 4.6 describe the regular characteristics, Section 4.7 describes the momentum characteristics, and Section 4.8 describes the spillover characteristics. As explained in Section 2.2, all characteristics are cross-sectionally standardised to mean zero before they are used in the FCC method. Consequently, an analysis of the standardised characteristics is included in Section 4.9.

Table 4.1: Forecaster Characteristics

Characteristic	**Abbreviation**	**Notation**	**Available from**	**Sign**
Panel A: Regular Characteristics				
Industry	Ind	$\pi_{i,t}$	Q2 1990	$+$
Experience	Exp	$\beta_{i,t,v}$	first forecast of variable v	$+$
Disconsensus	Disc	$\psi_{i,t,v,h}$	first+h forecast of variable v	$+$
Accuracy	Acc	$\iota_{i,t,v,h}$	second+h forecast of variable v	$-$
Bias	Bias	$\gamma_{i,t,v,h}$	second+h forecast of variable v	$-$
Consistency	Cons	$\rho_{i,t,v,h}$	third+h forecast of variable v	$+$
Panel B: Momentum Characteristics				
Accuracy Momentum	Acc Mom	$\eta_{i,t,v,h}$	second+h forecast of variable v	$-$
Bias Momentum	Bias Mom	$\mu_{i,t,v,h}$	second+h forecast of variable v	$-$
Consistency Momentum	Cons Mom	$\nu_{i,t,v,h}$	third+h forecast of variable v	$+$
Panel C: Spillover Characteristics				
Spillover Disconsensus	Spill Disc	$\Psi_{i,t,v,h}$	first+h forecast of other variables	$+$
Spillover Accuracy	Spill Acc	$\chi_{i,t,v,h}$	second+h forecast of other variables	$-$
Spillover Bias	Spill Bias	$\Gamma_{i,t,v,h}$	second+h forecast of other variables	$-$
Spillover Consistency	Spil Cons	$\xi_{i,t,v,h}$	third+h forecast of other variables	$+$
Spillover Accuracy Momentum	Spill AM	$\Lambda_{i,t,v,h}$	second+h forecast of other variables	$-$
Spillover Bias Momentum	Spill BM	$\omega_{i,t,v,h}$	second+h forecast of other variables	$-$
Spillover Consistency Momentum	Spill CM	$\Xi_{i,t,v,h}$	third+h forecast of other variables	$+$

Note: This table presents a list of all forecaster characteristics included in this paper. Additionally, it provides their notation, availability, and the expected sign of their coefficient (denoted by "Sign"). The availability of forecaster industry is specific to a certain quarter, whereas all other characteristics are available from either the first, second or third $+h$ time a forecaster provides a forecast. E.g. for forecast horizon $h = 1$, error is available from the third forecast onward.

4.1 Industry

From the second quarter of 1990 onwards, the SPF provides an industry token for each panel member, indicating whether they work for a financial service provider, a nonfinancial service provider, or if their industry is unknown. This characteristic is included in the study through dummy variable $\pi_{i,t}$:

$$\pi_{i,t} = \begin{cases} 1 & \text{if panel member } i \text{ works for a financial service provider at time } t, \\ 0 & \text{if panel member } i \text{ works for a nonfinancial service provider} \\ & \text{or if their industry is unknown at time } t. \end{cases} \tag{23}$$

Thus, industry dummy $\pi_{i,t}$ captures whether panel member i works for a financial service provider at time t. The industry dummy is allowed to vary over time, reflecting possible job changes. Since industry tokens are available only from Q2 1990 onwards, the industry dummy is set to zero for all panel members before this time period. We expect the sign of the coefficient belonging to the industry dummy to be positive because we expect that participants working for financial service providers produce more accurate forecasts and should, therefore, be assigned larger weight.

4.2 Experience

The more experienced a forecaster is, the better their forecasts may be. Therefore, forecaster experience is included as a characteristic. We expect the sign of the experience coefficient to be positive: more experienced forecasters might provide better forecasts and should, therefore, be assigned a larger combination weight. We measure experience by the logarithm of the total number of times a panel member has provided a forecast:

$$\beta_{i,t,v} = \log\left(1 + \sum_{\tau=1}^{t} \mathbb{1}_{i,\tau,v}\right) \tag{24}$$

In this expression, $\beta_{i,t,v}$ denotes forecaster experience of panel member i at time t for macroeconomic variable $v \in \{\text{GDP}, \text{UMP}, \text{CPI}\}$. The function $\mathbb{1}_{i,\tau,v}$ is an indicator function that equals one if panel member i provides a forecast (either a nowcast or a one-step-ahead forecast) for macroeconomic variable v at time τ.

We use the logarithm of the total number of forecasting times because we expect that the optimal forecast combination is not linearly related to experience. Instead, we expect that the initial increases in a forecaster's experience contribute more to their forecasting qualities than later increases. Thus, we apply an increasing and concave function to forecaster experience: the logarithmic function. The argument within the logarithmic function equals one plus the total number of forecasting instances, to ensure that $\beta_{i,t,v} = 0$ if forecaster i has never provided a forecast and $\beta_{i,t,v} > 0$ if forecaster i has provided a forecast at least once.

4.3 Disconsensus

According to Van Dijk and Franses (2019), disconsensus amongst expert forecasters may be one of the key reasons for the favourable performance of combined expert forecasts, compared to combined model forecasts. Based on their conclusion, we include a measure of disconsensus as characteristic. We expect the sign of the corresponding coefficient to be positive, such that forecasters with larger disconsensus are assigned larger weight. Disconsensus is defined as:

$$\psi_{i,t,v,h} = \begin{cases} \left|\zeta^v_{t,h} - f^v_{i,t+h|t}\right| & \text{if} \quad \mathbb{1}_{i,t,v,h} = 1, \\ 0 & \text{otherwise,} \end{cases} \tag{25}$$

in which $\psi_{i,t,v,h}$ denotes the disconsensus characteristic, $\zeta^v_{t,h}$ denotes the median forecast of variable v at time t for forecast horizon h and $|.|$ denotes the absolute value. Note that the disconsensus characteristic only takes on non-zero values in case forecaster i has provided a forecast for variable v at time t for horizon h ($\mathbb{1}_{i,t,v,h} = 1$). This disconsensus metric measures the absolute distance between the forecast of panel member i at time t and the median forecast at time t. The median, rather than the mean, is used as consensus forecast because of its robustness to outliers.

4.4 Accuracy

The more accurate a panel member's forecasts, the larger their combination weight should be. Therefore, forecaster accuracy is included as characteristic. We measure accuracy by the historical MSFE. We expect the sign of the corresponding coefficient to be negative: the larger the historical error of a forecaster, the smaller their combination weight. In detail, we define the accuracy characteristic as the MSFE of all previous forecasts by panel member i for a given horizon. We include the entire error history of a forecaster to incorporate as much information as possible. Thus, forecasting accuracy $\iota_{i,t,v,h}$ is determined as follows:

$$\iota_{i,t,v,h} = \begin{cases} \frac{\sum_{\tau=1}^{t-(1+h)} \mathbb{1}_{i,\tau,v,h} (\epsilon^v_{i,\tau+h})^2}{\sum_{\tau=1}^{t-1} \mathbb{1}_{i,\tau,v}} & \text{if} \quad \sum_{\tau=1}^{t-(1+h)} \mathbb{1}_{i,\tau,v,h} \geq 1, \\ 0 & \text{otherwise.} \end{cases} \tag{26}$$

$$\epsilon^v_{i,\tau+h} = r^v_{\tau+h} - f^v_{i,\tau+h|\tau}. \tag{27}$$

Here, $\iota_{i,t,v,h}$ denotes the accuracy of panel member i at quarter t for macroeconomic variable v and forecast horizon h. The forecasting error from the forecast about time $\tau + h$ by forecaster i for variable v is denoted by $\epsilon^v_{i,\tau+h}$. The variable $r^v_{\tau+h}$ denotes the realisation of macroeconomic variable v at time $\tau+h$, and $f^v_{i,\tau+h|\tau}$ represents the forecast of panel member i for macroeconomic variable v at time τ for horizon h. The indicator function $\mathbb{1}_{i,\tau,v,h}$ equals 1 if forecaster i has provided a forecast for variable v at time τ for forecast horizon h. The use of this indicator function accommodates panel members that provide forecasts for non-consecutive time periods.

To ensure that no information from the future is used to construct the accuracy measure, the sum in expression (26) depends on the horizon. Specifically, the accuracy measure at time t for horizon h only includes errors up to time $t - (1 + h)$. This is because the forecast from one time period later, $f^v_{i,t+h|t}$, can only be evaluated at time $t + h + 1$, when the realisation r^v_{t+h} becomes available. Consequently, $\iota_{i,t,v,h}$ is set to zero if a panel member has never provided a forecast for macroeconomic variable v at time $t - (1 + h)$.

4.5 Bias

The accuracy characteristic is measured by the historical MSFE of a forecaster. As explained in Section 2.1, the MSFE can be decomposed into squared bias and variance. To isolate the effects of these factors, we include squared bias as characteristic as well. A variance characteristic is not included, because it would be too strongly correlated with the accuracy characteristic. Intuitively, a measure of squared bias captures that some panel members consistently over- or underestimate a macroeconomic variable. We refer to the squared bias characteristic as bias for short. We expect the sign of the bias coefficient to be negative: the larger the (squared) bias of a forecaster, the smaller their combination weight should be. Bias is measured as follows:

$$\gamma_{i,t,v,h} = \begin{cases} \left(\frac{\sum_{\tau=1}^{t-(1+h)} \mathbb{1}_{i,\tau,v,h} \epsilon_{i,\tau+h}^{v}}{\sum_{\tau=1}^{t-(1+h)} \mathbb{1}_{i,\tau+h,v}} \right)^2 & \text{if} \quad \sum_{\tau=1}^{t-(1+h)} \mathbb{1}_{i,\tau,v,h} \geq 1, \\ 0 & \text{otherwise,} \end{cases} \tag{28}$$

in which $\gamma_{i,t,v,h}$ denotes the bias characteristic and all other variables are as defined previously.

4.6 Consistency

Apart from forecasting accuracy itself, the consistency of the forecasting accuracy is interesting as well. Some forecasters may provide very consistent forecasts over time, whereas others may be less consistent: they provide poor and accurate forecasts alternately. For that reason, we include a consistency characteristic. We expect the sign of the consistency coefficient to be positive: the more consistent a forecaster is, the larger their weight. We measure consistency by the first-order autocorrelation of the forecasting error. This metric is appropriate because it measures the persistence of errors over time. Specifically, a high first-order autocorrelation indicates that the size of the forecasting errors (whether small or large) tends to be stable, demonstrating a consistent pattern. Conversely, a low first-order autocorrelation suggests that the errors fluctuate significantly, with large and small errors occurring alternately for a specific forecaster. Thus, we define consistency as:

$$\rho_{i,t,v,h} = \begin{cases} \frac{\sum_{\tau=2}^{t-(1+h)} \mathbb{1}_{i,\tau,v,h} \mathbb{1}_{i,\tau-1,v,h} \left(\epsilon_{i,\tau+h}^{v} - \bar{\epsilon}_{i,t,h}^{v} \right) \left(\epsilon_{i,\tau+h-1}^{v} - \bar{\epsilon}_{i,t,h}^{v} \right)}{\sum_{\tau=1}^{t-(1+h)} \mathbb{1}_{i,\tau,v,h} \left(\epsilon_{i,\tau+h}^{v} - \bar{\epsilon}_{i,t,h}^{v} \right)^2}, & \text{if } \sum_{\tau=1}^{t-(1+h)} \mathbb{1}_{i,\tau,v,h} \geq 2, \\ 0 & \text{otherwise,} \end{cases} \tag{29}$$

$$\bar{\epsilon}_{i,t,h}^{v} = \frac{\sum_{\tau=1}^{t-(1+h)} \mathbb{1}_{i,\tau,v,h} \epsilon_{i,\tau+h}^{v}}{\sum_{\tau=1}^{t-(1+h)} \mathbb{1}_{i,\tau,v,h}}. \tag{30}$$

In these expressions $\rho_{i,t,v,h}$ denotes the consistency of panel member i at quarter t for macroeconomic variable v and horizon h. The variable $\bar{\epsilon}_{i,t,h}^{v}$ denotes the average forecasting error of panel member i available at time t. Hence, it measures the average errors up to time $t-(1+h)$, because errors after time $t-(1+h)$ are not yet available at time t (as explained in Section 4.4). To ensure that the denominator in equation (29) does not equal zero, $\rho_{i,t,v,h}$ only takes nonzero values if a panel member has provided at least two forecasts for horizon h by time $t-(1+h)$.

4.7 Momentum Characteristics

We have defined the accuracy, bias and consistency characteristics in terms of a sum over all historical forecasting errors. Apart from characteristics that incorporate the entire error history, it is interesting to include characteristics that incorporate only the most recent errors. For that reason, we include a number of momentum characteristics: accuracy momentum, bias momentum and consistency momentum. These characteristics are defined similarly to their regular counterparts, but include only the most recent forecasting errors, instead of the entire error history. Specifically, we include information from the previous four forecasts only, because this corresponds to the intuitive time span of a year.

For example, accuracy momentum is defined as the MSFE of the previous four forecasts. Hence, we express accuracy momentum as follows:

$$\eta_{i,t,v,h} = \begin{cases} \frac{\sum_{\tau=t-(4+h)}^{t-(1+h)} \mathbb{1}_{i,\tau,v,h}(\epsilon^{v}_{i,\tau+h})^2}{\sum_{\tau=t-(4+h)}^{t-(1+h)} \mathbb{1}_{i,\tau,v,h}}, & \text{if } \sum_{\tau=t-(4+h)}^{t-(1+h)} \mathbb{1}_{i,\tau,v,h} \geq 1, \\ 0 & \text{otherwise.} \end{cases} \tag{31}$$

In this expression, $\eta_{i,t,v,h}$ denotes the accuracy momentum of panel member i for variable v at time t for horizon h, and all other variables are as defined before. If a panel member has not provided any forecasts during the previous four time periods, accuracy momentum is set to zero. Similarly to the expression for accuracy in equation (26), the sum in the accuracy momentum expression depends on the horizon h, such that no information from the future is used to determine accuracy momentum at time t.

We include momentum characteristics for all characteristics that depend on the entire error history of the forecaster. Thus, apart from accuracy momentum, we include momentum characteristics for bias and consistency. The expressions for bias momentum ($\mu_{i,t,v,h}$) and consistency momentum ($\nu_{i,t,v,h}$) are similar to their regular counterparts and are provided in appendix C. We expect the signs of the momentum characteristics to be identical to their regular versions. For instance, the effect of bias and recent bias on the combination weights are expected to be the same.

4.8 Spillover Characteristics

Most of the characteristics described above are specific to a certain macroeconomic variable v. However, it is likely that experts make their forecasts for GDP, unemployment and inflation jointly. Therefore, a characteristic of one variable may be relevant for the forecast combination of another variable. For example, if forecaster i exhibits small error for the unemployment rate, this may signal that the forecaster is also accurate for the CPI inflation rate. Hence, we expect to find correlations in forecaster characteristics across different macroeconomic variables.

To substantiate this intuition, Table 4.2 provides the correlation of the same characteristics across different macroeconomic variables. Panel A provides the correlations for the nowcasts, Panel B for the one-step-ahead forecasts. Overall, the table shows that there is indeed considerable correlation between the same characteristics across different variables. Correlations reach values as high as 0.86. As expected, all characteristics exhibit positive correlation across macroeconomic variables. Hence, for instance, high accuracy in forecasting one variable is associated with high accuracy in forecasting another variable.

Table 4.2: Correlation Between Characteristics Across Macroeconomic Variables

Panel A: Nowcasts														
	Disc		**Acc**		**Bias**		**Cons**		**Acc Mom**		**Bias Mom**		**Cons Mom**	
	UMP	CPI	UMP	CPI	UMP	CPI	UMP	CPI	UMP	CPI	UMP	CPI	UMP	CPI
GDP	0.05	0.08	0.35	0.20	0.22	0.10	0.28	0.28	0.26	0.12	0.43	0.10	0.21	0.19
UMP		0.05		0.08		0.13		0.13		0.06		0.06		0.23
Panel B: One-Step-Ahead Forecasts														
	Disc		**Acc**		**Bias**		**Cons**		**Acc Mom**		**Bias Mom**		**Cons Mom**	
	UMP	CPI	UMP	CPI	UMP	CPI	UMP	CPI	UMP	CPI	UMP	CPI	UMP	CPI
GDP	0.05	0.14	0.86	0.12	0.71	0.16	0.40	0.43	0.59	0.11	0.54	0.10	0.26	0.24
UMP		0.05		0.12		0.16		0.43		0.11		0.10		0.24

Note: This table presents the correlation for the same characteristics across different macroeconomic variables. Panel A presents the correlations for the nowcasts, panel B for the one-step-ahead forecasts. The abbreviations of the characteristics can be found in Table 4.1. Furthermore,"GDP" stands for the real GDP growth rate, "UMP" for the unemployment rate and "CPI" for the annualised CPI inflation rate.

To exploit the correlations between the characteristics across different macroeconomic variables, we include a number of spillover characteristics. We choose a straightforward way to incorporate information from other variables: we define spillover characteristics as the average of the same characteristic for the other variables. For example, spillover accuracy of forecaster i, at time t and horizon h for the real GDP growth rate is defined as the average forecasting accuracy of forecaster i, at time t and horizon h for the unemployment rate and the CPI inflation rate. Thus, for variable v, we include spillover accuracy $\chi_{i,t,v,h}$ as follows:

$$\chi_{i,t,v,h} = \frac{\iota_{i,t,m,h} + \iota_{i,t,n,h}}{2}, \quad m, n \in \{\text{UMP}, \text{CPI}, \text{GDP}\} \setminus v, \quad m \neq n, \tag{32}$$

in which $\iota_{i,t,m,h}$ denotes the accuracy characteristic of variable m.

We include spillover characteristics for all characteristics that differ across macroeconomic variables, except for experience. Most forecasters hand in forecasts for all three variables simultaneously, such that experience differs minimally across variables. Consequently, a spillover characteristic is unnecessary, as it would be almost perfectly correlated with the regular experience characteristic. Overall, we include spillover disconsensus ($\Psi_{i,t,v,h}$), spillover accuracy ($\chi_{i,t,v,h}$), spillover bias ($\Gamma_{i,t,v,h}$), spillover consistency ($\xi_{i,t,v,h}$), spillover accuracy momentum ($\Lambda_{i,t,v,h}$), spillover bias momentum ($\omega_{i,t,v,h}$), and spillover consistency momentum ($\Xi_{i,t,v,h}$). The mathematical expressions for these spillover characteristics are similar to equation 32 and are provided in Appendix C. We expect the signs of the coefficients of the spillover characteristics to be identical to their regular counterparts, since all characteristics exhibit positive correlation across variables.

4.9 Analysis of Standardised Characteristics

All characteristics described in the previous sections are cross-sectionally standardised to mean zero before they are used in the FCC method. As explained in Section 2.2, this standardisation ensures that the combination weights sum to one. The standardisation has an important consequence for the interpretation of the characteristics: the relative size of the characteristics matters more than their absolute value. Thus, the characteristics included in the FCC model can be viewed as relative characteristics, measuring a forecaster's characteristic relative to those of other forecasters during the same time period. For example, at time t, the standardised accuracy characteristic of forecaster i measures the accuracy of forecaster i compared to the accuracy of all other forecasters included at time t.

To provide insight into the interactions between the standardised characteristics, Table 4.3 presents the correlations of the standardised characteristics for horizon $h = 0$ for the GDP growth rate. The correlations for the unemployment and CPI inflation rate are similar and can be found in Appendix D. Taking a closer look at Table 4.3, we find that large correlation occurs between (spillover) accuracy and (spillover) bias, as expected. Similarly, large correlation is found between (spillover) accuracy momentum and (spillover) bias momentum. Unsurprisingly, we also find large correlation between accuracy and accuracy momentum, as well as between bias and bias momentum. Thus, large recent error or bias is associated with large historical error or bias.

Taking a look at the correlations between experience and the other variables, we find a surprising result: experience exhibits positive correlation with accuracy. Hence, more experience is correlated with larger error. Conversely, experience is negatively correlated with bias, such that more experience is associated with smaller squared bias. Provided that accuracy is measured by the MSFE, which is decomposed into squared bias and variance, we infer that more experienced forecasters are associated with larger estimation variance. A possible explanation is that experienced forecasters may rely more on personal judgement, which could result in increased forecasting variance. Regardless, experience shows an anticipated relation with consistency: more experienced forecasters are associated with more consistent forecasts.

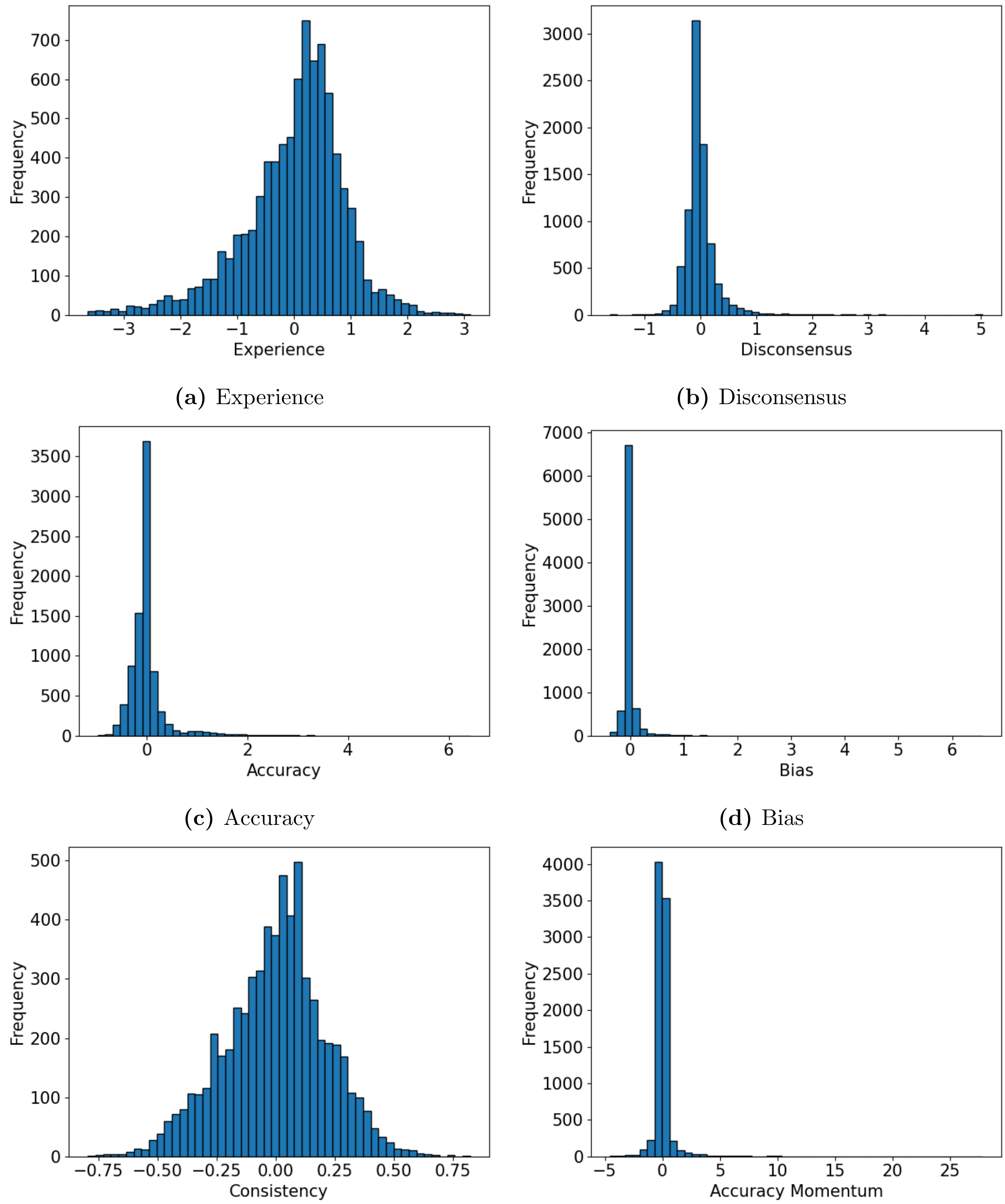
Table 4.3: Correlation between Standardised Characteristics of the Real GDP Growth Rate for Horizon $h = 0$

	Exp	**Disc**	**Acc**	**Bias**	**Cons**	**Acc Mom**	**Bias Mom**	**Cons Mom**	**Spill. Disc**	**Spill Acc**	**Spill Bias**	**Spill Cons**	**Spill AM**	**Spill BM**	**Spill CM**
Ind	−0.12	−0.02	−0.09	−0.02	−0.14	−0.05	−0.05	0.02	−0.04	−0.09	−0.01	−0.07	−0.12	−0.13	0.01
Exp		−0.02	0.04	−0.14	0.25	0.05	0.09	−0.13	−0.01	−0.05	−0.11	0.21	0.02	0.06	−0.14
Disc			0.19	0.11	−0.01	0.11	0.11	0.01	0.08	0.05	0.02	−0.02	0.06	0.08	0.00
Acc				0.58	0.03	0.44	0.46	−0.03	0.06	0.13	0.02	−0.06	0.07	0.08	−0.03
Bias					0.04	0.19	0.14	0.04	0.02	0.03	0.04	−0.05	0.02	0.00	0.05
Cons						0.04	0.07	0.09	−0.01	−0.08	−0.05	0.26	−0.04	−0.07	0.08
Acc Mom							0.74	−0.03	0.05	0.05	0.01	−0.02	0.13	0.10	−0.01
Bias Mom								−0.11	0.07	0.04	0.00	−0.03	0.09	0.15	−0.10
Cons Mom									−0.01	−0.01	0.02	0.09	−0.02	−0.07	0.29
Spill Disc										0.05	0.02	−0.01	0.07	0.12	0.01
Spill Acc											0.79	−0.04	0.64	0.30	−0.01
Spill Bias												−0.04	0.53	0.02	0.02
Spill Cons													−0.01	−0.02	0.13
Spill AM														0.66	−0.01
Spill BM															−0.09

Note: This table presents the correlations between the characteristics for the nowcasts of the GDP growth rate. The abbreviations used for the characteristics can be found in Table 4.1.

To provide further insight into the distributions of the standardised characteristics, Figure 4.1 presents the histograms of several standardised forecaster characteristics with forecasting horizon

$h = 0$ for the real GDP growth rate. As a result of the standardisation, all characteristics are centred around zero. Nevertheless, the histograms show that the dispersion of each characteristic varies. Momentum shows the widest spread, with values between -4.57 to 27.84. Conversely, consistency shows least dispersion, with values between -0.79 to 0.82. Disconsensus, accuracy, bias, and accuracy momentum all exhibit a right skewed distribution, such that there are a small number of observations with large positive values for these characteristics.



(a) Experience

(b) Disconsensus

(c) Accuracy

(d) Bias

(e) Consistency

(f) Accuracy Momentum

Figure 4.1: Histograms of Forecaster Characteristics with Horizon $h = 0$ for the Real GDP Growth Rate

5 Benchmarks

To evaluate the FCC method, we compare its performance to several benchmark methods. For overview, Table 5.1 lists all models used in this paper: the proposed FCC models as well as the benchmarks. Sections 5.1 to 5.5 elaborate on each of the benchmarks and their implementation separately. The evaluation metrics used to compare the performance of the FCC methods with the benchmarks are provided in Section 5.6.

Table 5.1: Model Overview

Model Name	**Abbreviation**	**Notation**
Panel A: FCC models		
Regular FCC	FCC	$f^{c,\text{FCC}}_{t+h\|h}$
FCC with Ridge Regularisation	FCC-Ridge	$f^{c,\text{Ridge}}_{t+h\|h}$
FCC with LASSO Regularisation	FCC-LASSO	$f^{c,\text{LASSO}}_{t+h\|h}$
FCC with post-LASSO Regularisation	FCC-postLASSO	$f^{c,\text{postLASSO}}_{t+h\|h}$
FCC with Elastic Net Regularisation	FCC-ENet	$f^{c,\text{ENet}}_{t+h\|h}$
FCC with Single Characteristic Estimation, no Regularisation	FCC-C	$f^{c,\text{C}}_{t+h\|h}$
FCC with Single Characteristic Estimation, LASSO Regularisation	FCC-C-LASSO	$f^{c,\text{C-LASSO}}_{t+h\|h}$
FCC with Single Characteristic Estimation, Elastic Net Regularisation	FCC-C-ENet	$f^{c,\text{C-ENet}}_{t+h\|h}$
Panel B: Benchmark models		
Equally Weighted Average	EWA	$f^{c,\text{EWA}}_{t+h\|h}$
Trimmed Mean with 5% Trim	Trimmed Mean	$f^{c,\text{Trimmed Mean}}_{t+h\|h}$
Median	Median	$f^{c,\text{Median}}_{t+h\|h}$
Bias-Adjusted Mean Combination	BAM	$f^{c,\text{BAM}}_{t+h\|h}$
Discounted MSFE Combination	dMSFE	$f^{c,\text{dMSFE}}_{t+h\|h}$
Partially-Egalitarian LASSO	peLASSO	$f^{c,\text{peLASSO}}_{t+h\|h}$

Note: This table presents the models included in this paper. Detailed explanation of the FCC models can be found in Section 2. Detailed explanation of the benchmarks models can be found in Sections 5.1 to 5.5.

5.1 Equally Weighted Average

The first and simplest benchmark we include is the equally weighted average (EWA). As mentioned before, this benchmark performs well in the literature (Genre et al., 2013). It is a relevant benchmark because the FCC method employs the EWA weights as baseline. The EWA forecast combination is calculated as:

$$f^{c,\text{EWA}}_{t+h|t} = \frac{1}{N_t} \sum_{i=1}^{N_t} f_{i,t+h|t}. \tag{33}$$

5.2 Trimmed Mean and Median

Individual forecasts may include outliers, which affect the EWA combination forecast. A more robust method is trimmed mean combination, which excludes a portion of the smallest and largest forecasts at each point in time. Following the advice of Armstrong (2001), who suggests the use of trimmed means combination when there are at least 5 forecasts available, we implement trimmed mean combination. In particular, we implement both a symmetric 5% trim, as well as a 50% trim, which is equivalent to using the median.

5.3 Bias-Adjusted Mean

The bias-adjusted mean (BAM) combination method is proposed by Capistrán and Timmermann (2009). This method is a relevant benchmark because of its similarity to FCC in using the equally weighted average as baseline. Specifically, the BAM method regresses the realised value of the variable under consideration on a constant and the equally weighted average forecast combination, as follows:

$$r_{t+h} = c + b\left(\frac{1}{N_t}\sum_{i=1}^{N_t} f_{i,t+h|t}\right), \quad c, b \in \mathbb{R}, \tag{34}$$

in which c and b are parameters to be estimated. The constant c shifts the EWA forecast to eliminate bias. The parameters are estimated using OLS, which results in estimates $\hat{c}$ and $\hat{b}$. Following Genre et al. (2013), we estimate the BAM forecast combinations using moving window estimation, with window size 40 (matching the window size used for FCC estimation). Consequently, the BAM forecast combination can be expressed as:

$$f^{c,\mathrm{BAM}}_{t+h|t} = \hat{c} + \hat{b}\left(\frac{1}{N_t}\sum_{i=1}^{N_t} f_{i,t+h|t}\right), \tag{35}$$

in which all parameters are as defined previously.

5.4 Discounted MSFE

The discounted mean squared forecast error (dMSFE) combination method weighs forecasts inversely to the historical error of a forecaster (Stock & Watson, 2004). This benchmark can be viewed as a simple version of the FCC: it makes use of only a single characteristic, namely (discounted) accuracy. Because of this similarity with the FCC method, dMSFE combination forms a relevant benchmark. The combination forecasts from dMSFE combination are calculated as follows (following notation by Stock and Watson (2004)):

$$f^{c,\mathrm{dMSFE}}_{t+h|t} = \sum_{i=1}^{N_t} \frac{d^{-1}_{i,t+h|t}}{\sum_{j=1}^{N_t} d^{-1}_{j,t+h|t}} f_{i,t+h|t}, \quad \forall\ d_{i,t+h|t},\ d_{j,t+h|t} \neq 0 \tag{36}$$

$$\text{where} \quad d_{i,t+h|t} = \sum_{\tau=1}^{t-(1+h)} \delta^{t-\tau}(r_{\tau+h} - f_{i,\tau+h|t})^2, \quad \delta \leq 1. \tag{37}$$

In these equations, $d_{i,t+h|t}$ is the discounted sum of historical MSFEs up to time $t-(1+h)$ for panel member i. This sum is calculated using discount factor δ such that for $\delta < 1$, recent forecasting errors are assigned larger weight. We implement this method for three different values of $\delta = 0.9, 0.95, 1.0$, following Stock and Watson (2004). A value of $\delta = 1.0$ corresponds to the weighing scheme introduced by Bates and Granger (1969), in which all historical forecasting errors are equally weighed. Hence, a value of $\delta = 1.0$ corresponds to weighing the forecasts inversely to their relative accuracy characteristic (defined in Section 4.4).

5.5 Partially-Egalitarian LASSO

The partially-egalitarian LASSO method (peLASSO) is a relatively recent method, introduced by Diebold and Shin (2019). It first selects forecasts to be combined and subsequently shrinks the combination weights to equal weights using LASSO regularisation. We implement this method as benchmark, for two reasons. Firstly, peLASSO has proven superior to equal weights when applied to expert panel forecasts, making it a challenging benchmark to beat (Diebold & Shin, 2019). Secondly, this benchmark is relevant to the methods proposed in this paper as they include regularisation techniques such as LASSO as well.

The peLASSO method solves the following minimisation problem to determine combination weight vector $\boldsymbol{w}_h$ for forecasting horizon h:

$$\boldsymbol{w}_h^{\text{peLASSO}} = \text{argmin}_{\boldsymbol{w}_h} \left(\sum_{t=1}^{T} \left(r_{t+h} - \sum_{i=1}^{N} w_{h,i} f_{i,t+h|t} \right)^2 + \lambda_1 \sum_{i=1}^{N} \left| w_{h,i} \right| + \lambda_2 \sum_{i=1}^{N} \left| w_{h,i} - \frac{1}{p(\boldsymbol{w}_h)} \right| \right), \tag{38}$$

In this expression, N denotes the total number of forecasters, λ_1 and λ_2 are penalty terms that determine the amount of shrinkage, and $p(\boldsymbol{w}_h)$ is the number of non-zero elements in $\boldsymbol{w}_h$. Thus, the optimisation includes two penalty terms: the first is the standard LASSO penalty, which selects and shrinks weights to zero, effectively removing some of the forecasts. The second penalty term shrinks the "surviving" weights towards equal weights.

The estimation of $\boldsymbol{w}_h$ presents two main challenges. Firstly, the regression model in (38), requires a balanced panel dataset, as the total number of forecasters, N, remains constant over time. To create a balanced dataset from our unbalanced dataset, we implement peLASSO using a moving window approach with window size 40 (matching the window size of the FCC estimation procedure). To balance the dataset, at each point in time, we select the 9 most frequent forecasters within that window. This approach is inspired by Diebold and Shin (2019), who select the 23 most frequent forecasters over their entire dataset. We select 9 forecasters because there are at least 9 forecasts available at each point in time in the dataset. Following Diebold and Shin (2019), we impute missing observations using a linear AR(1) filter.

A second challenge is that analytical optimisation of the problem in equation (38) is difficult due to the discontinuity of the objective function at $w_{h,i} = 0$. Therefore, we follow the two-step implementation proposed by Diebold and Shin (2019):

Step 1. Select a subset of n forecasters from the N forecasters available using LASSO shrinkage on the combination weight.

Step 2. Shrink the n "surviving" combination weights from Step 1 towards $\frac{1}{n}$, using LASSO shrinkage.

Note that Diebold and Shin (2019) recommend LASSO shrinkage for both steps due to favourable results in expert panel forecast combinations. Since the empirical setting in this paper is similar and further research on the peLASSO method is not our focus, we follow their recommendation and use LASSO for both steps. We estimate the hyperparameters in both LASSO steps separately, using 5-fold cross validation. The resulting combination forecast is straightforwardly constructed using weight vector $\hat{\boldsymbol{w}}_h^{\text{peLASSO}}$:

$$f_{t+h|t}^{c,\text{peLASSO}} = \sum_{i=1}^{N} \hat{w}_{h,i}^{\text{peLASSO}} f_{i,t+h|t}, \tag{39}$$

in which $w_{h,i}^{\text{peLASSO}}$ denotes the i-th element of the vector $\hat{\boldsymbol{w}}_h^{\text{peLASSO}}$.

5.6 Evaluation Metrics

To compare the performance of the FCC models with the benchmarks, we need an evaluation metric: a loss function. We use the squared forecasting error and calculate the MSFE as follows:

$$\text{MSFE} = \frac{1}{T-(1+h)} \sum_{t=1}^{T-(1+h)} (r_{t+h} - f_{t+h|t}^{c})^2, \tag{40}$$

in which all variables are as defined previously. Note that we evaluate forecasts up to time $T-(1+h)$ because the realisation for time $T+h$ is available at $T+h+1$. The results section only presents the MSFE of the regular FCC model and presents the *relative* MSFE of the other models. The relative MSFE is defined as the MSFE of the model under consideration over the MSFE of the regular FCC method.

A concern about the evaluation method in this paper relates to the multiple comparison problem. When comparing several forecast models on a single dataset, significant results may be found by chance (Genre et al., 2013). For that reason, it is not appropriate to draw conclusions from pair-wise comparisons of predictive ability. To mitigate this problem, we employ the model confidence set (MCS) approach introduced by Hansen et al. (2011). An MCS is a subset of models that contains the best model with a given level of confidence. Put differently, it can be viewed as a confidence interval for the best performing model. We execute the MCS procedure with the bootstrap implementation and use the squared forecasting error as loss function. Following Hansen et al. (2003), we set the number of bootstrap resamples equal to 1000 and use a confidence level of 95%. Further details about the MCS approach and our implementation are provided in Appendix E.

6 Results

The results in this paper focus on evaluating the forecasting performance of the FCC method. Section 6.1 compares the FCC method's performance to the benchmarks. Subsequently, Section 6.2 compares the performance of the FCC method to its extensions. Section 6.3 analyses the FCC coefficients and their development over time. Finally, Section 6.4 examines the robustness of the results with a sensitivity analysis.

6.1 Benchmark Comparison

This section evaluates the potential of the FCC method by comparing its performance to the benchmarks. Table 6.1 reports the MSFE and MCS results for the regular FCC method, estimated with the moving window approach, as well as the benchmarks. For the FCC method, the MSFE is presented and for the benchmarks, the relative MSFE is presented. Importantly, Table 6.1 demonstrates the strong performance of the FCC method, achieving the lowest MSFE in 5 out of 6 cases. Accordingly, the FCC method is selected to be in the 95% confidence MCS most often, also in 5 out of 6 cases. Only in one instance, the BAM method is selected into the MCS. We conclude that the FCC method may be considered the best performing method.

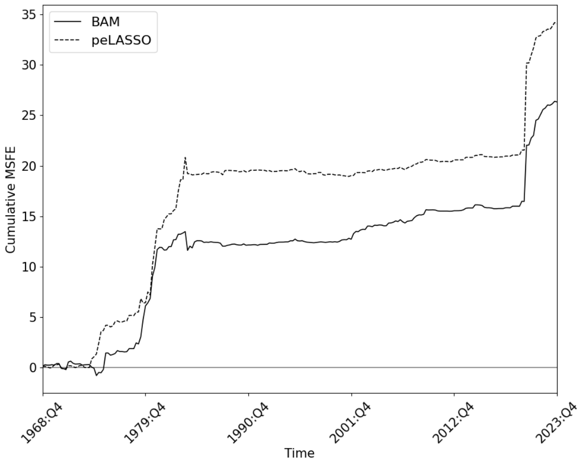
Table 6.1: MSFE Results of the FCC Method and Benchmarks

Combination Method	Nowcast			One-Step-Ahead			MCS Count
	GDP	UMP	CPI	GDP	UMP	CPI	
Panel A: FCC Method							
FCC	***0.110***	*0.130*	***0.543***	***0.442***	***0.397***	***1.604***	5
Panel B: Benchmark Methods							
EWA	2.33	0.53	3.16	2.16	1.68	2.79	0
Median	2.28	0.56	3.18	2.20	1.63	2.83	0
Trimmed Mean	2.35	0.55	3.15	2.16	1.65	2.79	0
BAM	2.08	**0.30**	2.23	1.94	1.36	2.42	1
dMSFE ($\delta = 0.9$)	2.36	2.34	4.71	1.88	1.80	3.33	0
dMSFE ($\delta = 0.95$)	2.28	2.30	4.42	1.93	1.81	3.21	0
dMSFE ($\delta = 1.0$)	2.19	2.27	4.09	1.99	1.81	3.02	0
peLASSO	2.40	0.33	1.71	2.15	1.35	2.54	0

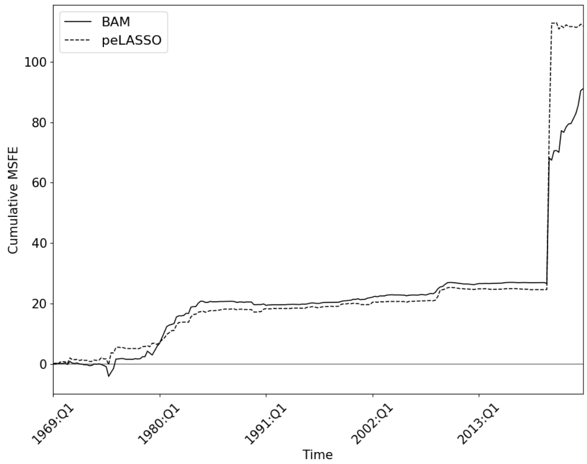
Note: This table presents the MSFE results of the FCC method and the relative MSFE of the benchmarks. Note that the MSFE of the FCC method is presented in italics because it shows the MSFE rather than the relative MSFE. The results for the FCC method are estimated with moving window estimation. The MSFE of the model selected into the MCS with 95% confidence is presented in bold. The column "MCS Count" provides the total number of times a model is selected into the MCS. Furthermore, "GDP" stands for the real GDP growth rate, "UMP" for the unemployment rate and "CPI" for the annualised CPI inflation rate.

To provide insight into the performances of the methods over time, Figure 6.1 plots the difference in cumulative squared forecasting error (SFE) between the BAM and peLASSO benchmarks and the regular FCC method. The larger the value plotted, the better the FCC method performs compared to the benchmarks. Only when a plot reaches values below zero, the benchmarks achieve smaller cumulative SFE compared to the regular FCC method. We examine the forecasting performance for each macroeconomic variable separately.

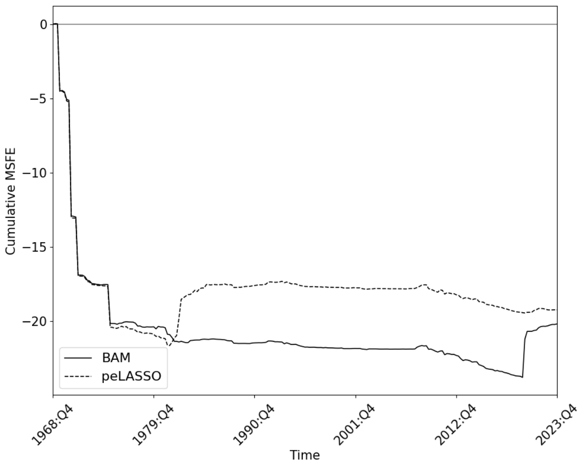
Firstly, for the real GDP growth rate in Figure 6.1a and 6.1b, the regular FCC method consistently achieves lower forecasting error compared to BAM and peLASSO for both horizons. Both benchmarks show a steep increase in cumulative SFE difference around the COVID outlier in Q2 and Q3 2020, which indicates that the FCC method handles the COVID outlier better than the peLASSO and BAM methods. Overall, we conclude that the FCC method outperforms the benchmarks for the GDP growth rate.



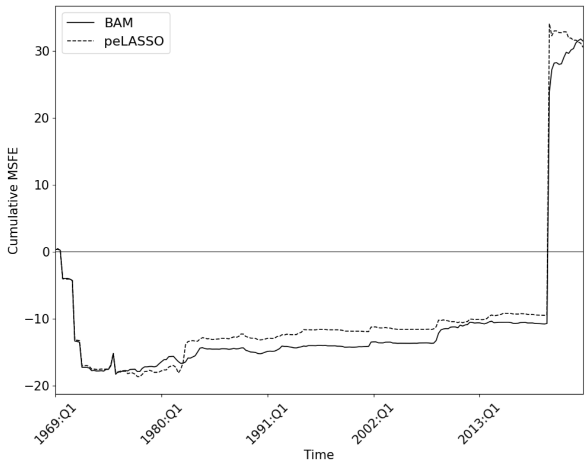
(a) Real GDP Growth Rate Nowcasts



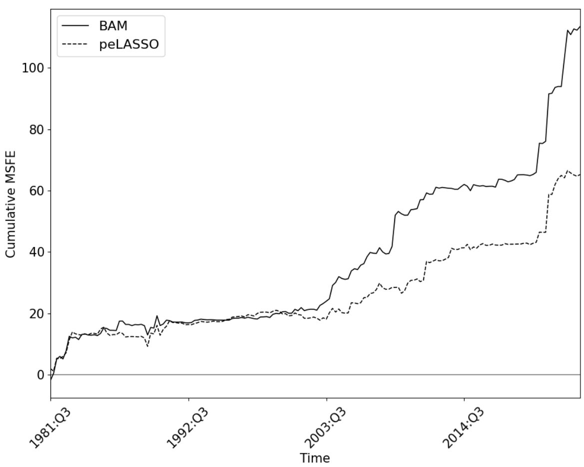
(b) Real GDP Growth Rate One-Step-Ahead



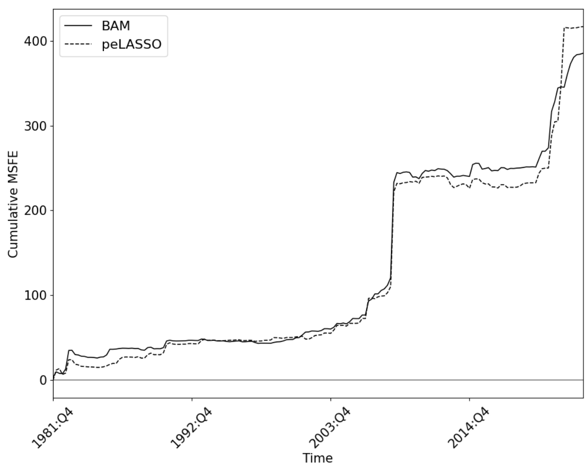
(c) Unemployment Rate Nowcasts



(d) Unemployment Rate One-Step-Ahead



(e) CPI Inflation Rate Nowcasts



(f) CPI Inflation Rate One-Step-Ahead

Figure 6.1: Cumulative Squared Forecasting Error Difference Between the Benchmarks and the FCC method

Figures 6.1c and 6.1d present the cumulative SFE difference for the unemployment rate. From both Table 6.1 and Figure 6.1c it is clear that for the nowcasts, the peLASSO and BAM methods outperform the FCC method. However, for the one-step-ahead forecasts, the forecasting performance of the FCC method is on *average* better than the BAM and peLASSO benchmarks (see Table 6.1). Figure 6.1d shows that this result remains true only on average: for most time periods, the cumulative forecasting error of the BAM and peLASSO methods is lower. It appears that the FCC method achieves lower MSFE only due to the steep increase in cumulative SFE difference around the COVID outlier. Simply put, for the unemployment rate one-step-ahead forecasts, the FCC method achieves lower MSFE compared to the benchmarks *only* due to the COVID outlier. Therefore, we conclude that for unemployment overall, the FCC method does not outperform the benchmarks. We provide an analysis of results without outliers in Section 6.4.

Figure 6.1e and 6.1f show the cumulative squared forecasting error difference for the CPI inflation rate. For both the nowcasts and one-step-ahead forecasts the FCC method consistently outperforms BAM and peLASSO, as the cumulative SFE difference lies above zero during all time periods. For the one-step-ahead forecasts in Figure 6.1f, both benchmarks exhibit a steep increase in cumulative SFE difference during the Great Recession in Q4 2008. This implies that the FCC method achieves lower error during this outlier compared to the benchmarks. Thus, again, the FCC method proves to be better at handling outliers than the benchmarks.

6.2 Estimation Methods and FCC Extensions

We now turn to the comparison of the estimation methods and the performance of the extensions of the FCC method. Table 6.2 presents the MSFE results of the FCC method and its extensions, for both the moving window approach (Panel A) and the expanding window approach (Panel B). Note that for the regular FCC method, the MSFE is presented, whereas for the extensions, the relative MSFE is presented. From Table 6.2, it is clear that for all variables, horizons and all FCC methods, moving window estimation achieves smaller MSFE than expanding window estimation. For that reason, the results focus on the FCC results from moving window estimation.

The superior performance of moving window estimation can be explained by its larger flexibility in allowing for variation in the coefficients, compared to expanding window estimation (as explained in Section 2.4). Provided that the moving window approach results in lower forecasting errors, accommodating variations in the coefficients enhances forecasting performance in our context. This implies that the importance of each forecaster characteristic changes over time. Section 6.3 further analyses the characteristic coefficients and their variation over time.

Turning to the differences between the FCC methods, Table 6.2 demonstrates that the extensions of the regular FCC method do not improve performance. Note that this table also presents results from the 95% confidence MCS procedure on the FCC methods, performed separately for each estimation method. The table shows that the extended methods are never selected into the MCS. All extended methods show larger MSFEs compared to regular FCC. We examine these inferior MSFE results for each extension separately.

Firstly, we analyse the MSFE of the regularised methods: FCC-Ridge, FCC-LASSO, FCC-postLASSO and FCC-ENet. In all cases, the regular FCC method outperforms these regularised methods. We may explain this result as follows. Regularisation methods are developed to prevent overfitting by penalising more complex models. These constraints limit the model's flexibility, which apparently leads to reduced performance in the application of this paper. In our context, the constraints applied by the regularisation techniques are strict: many coefficients are shrunk to zero by both LASSO and ENet during most time periods. To demonstrate this, Appendix F presents the percentage of time periods during which coefficients are set to zero by LASSO and ENet. For many characteristics, these percentages lie above 80%. Hence, in our setting, the benefits expected from regularisation are not fully realised because the regularisation methods are very restrictive. This results in MSFE metrics that are comparable or even worse than those of the unregularised FCC method.

Table 6.2: MSFE Results of the FCC Method and its Extensions

Combination Method	Nowcast			One-Step-Ahead			MCS Count
	GDP	UMP	CPI	GDP	UMP	CPI	
Panel A: Moving Window Estimation							
FCC	***0.110***	***0.130***	***0.543***	***0.442***	***0.397***	***1.604***	6
FCC-Ridge	1.73	1.11	1.87	2.03	1.57	2.02	0
FCC-LASSO	1.62	1.01	1.97	2.07	1.60	2.28	0
FCC-postLASSO	1.55	1.02	1.88	2.00	1.43	2.22	0
FCC-ENet	1.64	1.02	1.93	2.07	1.62	2.33	0
FCC-C-Simple	2.11	1.14	2.81	1.89	1.69	2.42	0
FCC-C-LASSO	2.02	1.13	2.53	1.86	1.57	2.43	0
FCC-C-ENet	2.12	1.14	2.78	1.89	1.69	2.42	0
Panel B: Expanding Window Estimation							
FCC	***0.150***	***0.134***	***1.025***	***0.733***	***0.577***	***2.862***	6
FCC-Ridge	1.30	1.21	1.21	1.26	1.10	1.33	0
FCC-LASSO	1.34	1.00	1.70	1.27	1.11	1.48	0
FCC-postLASSO	1.35	1.01	1.64	1.26	1.08	1.48	0
FCC-ENet	1.65	1.27	1.81	1.27	1.11	1.48	0
FCC-C-Simple	1.72	1.19	1.70	1.26	1.23	1.46	0
FCC-C-LASSO	1.70	1.16	1.62	1.21	1.21	1.49	0
FCC-C-ENet	1.72	1.32	1.69	1.26	1.23	1.46	0

Note: This table presents the MSFE of the FCC method and the relative MSFE of the extensions, for each estimation method and both horizons. The MSFE of the regular FCC method is presented in italics because it shows the MSFE rather than the relative MSFE. The table also presents the results from the MCS procedure on the FCC methods, performed separately for each estimation approach. The MSFE of the model selected into the MCS with 95% confidence is presented in bold. The column "MCS Count" provides the total number of times a model is selected into the MCS. In this table, "GDP" stands for the real GDP growth rate, "UMP" for the unemployment rate and "CPI" for the annualised CPI inflation rate.

Secondly, we examine the relative MSFE of the single characteristic estimation methods: FCC-C-Simple, FCC-C-LASSO, FCC-C-ENet. Although none of the methods outperform regular FCC, among these methods, the FCC-C-LASSO method achieves the lowest MSFE in 5 out of 6 cases. Thus, the C-LASSO method outperforms the original C-Simple and C-ENet methods by Rapach et al. (2010) and Rapach and Zhou (2020). Unfortunately, overall, these methods consistently lead to higher MSFE compared to the regular and regularised FCC methods. Thus, in this context, the use of single coefficient estimation does not improve forecasting performance.

This can be explained as follows. In single characteristic estimation, each characteristic is considered separately. When characteristics are correlated, such a separate estimation may not sufficiently capture interactions between characteristics. Provided that the characteristics in this paper show significant correlation, as discussed in Section 4.9, single characteristic estimation fails to capture the interactions between the characteristics sufficiently. This results in weaker forecasting performance compared to regular FCC. The regular FCC method simultaneously considers all characteristics and therefore provides a more accurate forecast.

Another explanation for the sub-optimal performance of the single characteristic methods is that they were originally designed to combine individual forecasts. Conversely, in our application, we use the methods to combine already combined forecasts. Consequently, the single characteristic estimation methods may not be well-suited to our application. Due to the favourable performance of the regular FCC method compared to its extensions, the remainder of the results focuses on the regular FCC method only.

6.3 Analysis of the FCC Coefficients

To further understand the favourable performance of the FCC method, we examine its coefficients. Panel A of Table 6.3 presents the results of the FCC regression: the average coefficients assigned to each characteristic for both the nowcasts and the one-step-ahead forecasts. The coefficients in this table are averages over all coefficients estimated with the moving window approach. Below each coefficient, a percentage is presented, indicating how often each coefficient is statistically significantly different from zero throughout the entire observation period. This statistical significance was determined by a t-test at a 95% confidence level.

Overall, Table 6.3 shows that the magnitude and sign of the average coefficients vary across variables and forecast horizons. These differences demonstrate that each characteristic requires varying emphasis for different macroeconomic variables. From the significance percentages, it is evident that many coefficients are only significant during a relatively small portion of the observation period. Thus, the importance of characteristics not only varies across macroeconomic variables, but over time as well. Hence, these results highlight the importance of the flexibility of the FCC model.

We now inspect the average coefficients in more detail. The industry dummy shows positive average coefficients for most variables and horizons. Thus, as expected, forecasts from individuals that work in the financial sector are generally assigned larger weight. This is especially true for the GDP growth rate nowcasts and one-step-ahead forecasts. For the unemployment rate, the average coefficient lies closer to zero. The CPI inflation rate exhibits negative average coefficients for the one-step-ahead forecasts, such that for this horizon, inflation forecasts from individuals working in non-financial services are generally assigned larger weight instead.

Surprisingly, experience exhibits mostly negative average coefficients, indicating that forecasts from more experienced individuals are assigned smaller weight. Although this is not in line with our expectation, this result may be explained by the positive correlation between experience and MSFE (as discussed in Section 4.9). Put differently, in our empirical application, more experienced forecasters are associated with larger forecasting errors, which may justify assigning them smaller combination weights.

Table 6.3: Forecaster Characteristic Coefficients from the Regular FCC Method

Variable	Nowcasts			One-Step-Ahead		
	GDP	**UMP**	**CPI**	**GDP**	**UMP**	**CPI**
Panel A: Coefficients						
Panel A.1: Regular Characteristics						
Industry	1.72	0.02	2.26	6.45	0.94	-1.66
	(3.87%)	(0.55%)	(10.77%)	(2.76%)	(4.97%)	(5.38%)
Experience	1.17	-0.25	-0.85	-1.43	-2.24	0.62
	(9.94%)	(2.21%)	(5.38%)	(4.42%)	(19.89%)	(0.77%)
Disconsensus	-1.29	-4.91	1.12	-0.01	-3.07	0.73
	(1.10%)	(29.28%)	(19.23%)	(8.84%)	(8.84%)	(0.00%)
Accuracy	5.60	-7.27	-1.19	-7.03	19.09	1.07
	(2.21%)	(1.66%)	(23.08%)	(11.60%)	(19.34%)	(17.69%)
Bias	-4.95	2.53	0.28	5.50	-19.54	-0.33
	(6.63%)	(7.18%)	(12.31%)	(1.10%)	(2.21%)	(5.38%)
Consistency	4.61	3.40	5.02	5.21	-1.56	-15.46
	(8.84%)	(20.99%)	(16.92%)	(0.55%)	(12.71%)	(27.69%)
Panel A.2: Momentum Characteristics						
Accuracy Momentum	1.07	-21.14	-0.81	7.30	-8.74	0.23
	(3.31%)	(16.57%)	(3.08%)	(2.76%)	(16.02%)	(0.00%)
Bias Momentum	-2.32	-4.45	0.72	-10.60	0.14	-1.30
	(16.57%)	(14.36%)	(6.92%)	(2.76%)	(3.87%)	(11.54%)
Consistency Momentum	3.93	-1.80	-5.08	-3.57	5.92	-10.58
	(4.97%)	(1.10%)	(14.62%)	(9.39%)	(16.57%)	(22.31%)
Panel A.3: Spillover Characteristics						
Spillover Disconsensus	-0.17	1.38	-7.29	5.90	1.35	-11.06
	(0.00%)	(1.10%)	(3.08%)	(7.18%)	(8.84%)	(18.46%)
Spillover Accuracy	-2.53	-0.48	4.98	-1.35	-2.61	-23.91
	(1.66%)	(1.66%)	(4.62%)	(1.66%)	(12.15%)	(4.62%)
Spillover Bias	1.45	-1.82	-18.66	0.67	2.26	-17.30
	(0.00%)	(3.31%)	(3.08%)	(6.08%)	(2.76%)	(0.77%)
Spillover Consistency	3.39	-1.19	-1.90	2.16	-0.05	6.03
	(6.08%)	(0.55%)	(0.77%)	(7.18%)	(19.89%)	(2.31%)
Spillover Accuracy Momentum	3.45	2.69	4.69	-0.83	-1.67	-5.30
	(6.63%)	(4.97%)	(6.15%)	(0.00%)	(12.15%)	(2.31%)
Spillover Bias Momentum	-5.16	-4.18	-4.64	2.70	4.18	-0.19
	(0.00%)	(3.31%)	(3.85%)	(2.76%)	(3.87%)	(0.77%)
Spillover Consistency Momentum	-2.01	-3.10	-4.90	12.97	6.25	4.39
	(5.52%)	(5.52%)	(3.85%)	(12.15%)	(12.15%)	(3.85%)
Panel B: F-test Results						
Null Hypothesis Rejected	11.05%	58.01%	73.85%	12.71%	39.78%	23.08%

Note: This table presents the average coefficients assigned to the forecaster characteristics by the FCC method and the results of the F-test. Panel A presents the average coefficient values from moving window estimation. Below each coefficient, between parentheses, the percentage of time periods during which a coefficient is statistically significant is shown. Statistical significance was determined by a t-test at a 95% confidence level. Panel B provides the results of an F-test at a 95% confidence level for the null hypothesis that all characteristics have coefficients equal to zero. Specifically, it provides the percentage of time periods during which the null hypothesis was rejected.

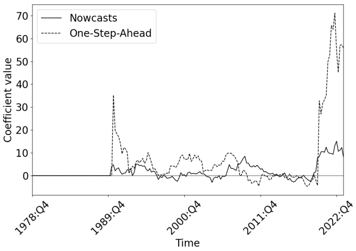
Disconsensus exhibits mostly negative average coefficients. Hence, contrary to our expectation, forecasts from individuals that deviate from the median forecast are generally assigned smaller combination weight. For the CPI inflation rate, the disconsensus effect is opposite: forecasts from individuals that deviate from the consensus forecast are assigned larger weight in general. Thus, only for the CPI inflation rate, our expectation about the disconsensus coefficient is met. Spillover disconsensus shows both positive and negative coefficients, such that no clear conclusions can be drawn about this characteristic.

Surprisingly, accuracy, measured by the historical MSFE, shows both positive and negative average coefficients, such that we cannot conclude that forecasts from individuals with lower historical errors are consistently assigned larger weight. This effect is particularly surprising for the unemployment rate one-step ahead forecasts, for which the average accuracy coefficient reaches a positive value as large as 19.09. Accuracy momentum and spillover accuracy momentum also exhibit both positive and negative average coefficients, such that no definite conclusion can be drawn. Fortunately, spillover accuracy shows more anticipated coefficients: they are generally negative. Hence, as expected, forecasts from individuals that show larger historical error for other variables are assigned smaller weight for the variable under consideration.

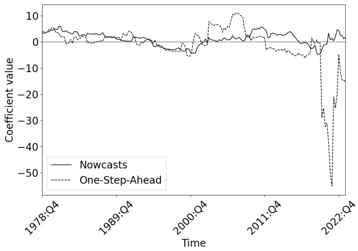
Bias demonstrates both positive and negative coefficients, such that we cannot draw any definite conclusions. The same holds for bias momentum, spillover bias, and spillover bias momentum. As expected, consistency shows mostly positive coefficients. Thus, forecasts from more consistent forecasters are assigned larger weight. This does not hold for consistency momentum, spillover consistency and spillover consistency momentum, for which no clear conclusions can be drawn.

Provided that most coefficients are statistically significantly different from zero during only a relatively small number of periods, it is interesting to analyse how often the vector of coefficients is jointly statistically significantly different from zero. Consequently, an F-test at 95% confidence level was performed on the null hypothesis that all characteristics have coefficients equal to zero. Panel B in Table 6.3 presents the results of the F-test as the percentage of time periods during which the null hypothesis was rejected. Thus, the larger the percentage presented, the more often the vector of coefficients was jointly significantly different from zero.

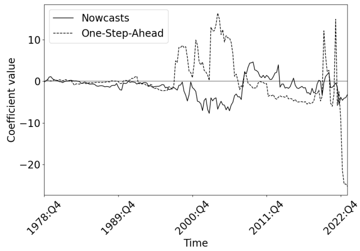
To interpret these percentages, remember that the characteristic coefficients determine the amount of deviation from the equally weighted average combination weights. Thus, when the vector of coefficients is not statistically significantly different from zero, the forecast combination weights are approximately equal to the equally weighted average weights. For the real GDP growth rate nowcasts, the FCC coefficients are only significantly different from zero in 11.05% of cases. For these forecasts, the combination forecasts often lie close to the equally weighted average combinations. Conversely, for the CPI nowcasts, the FCC coefficients are significantly different from zero during most time periods: in 73.85% of cases. Thus, for the CPI nowcasts, the FCC method adjusts the equally weighted average for most time periods. Consequently the FCC CPI forecasts achieve a larger MSFE reduction of the EWA forecasts (68%) than the FCC GDP forecasts, who only reduce the EWA MSFE by 57% (see table 6.1).



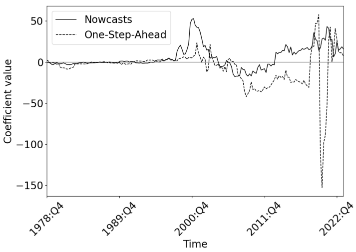
(a) Industry



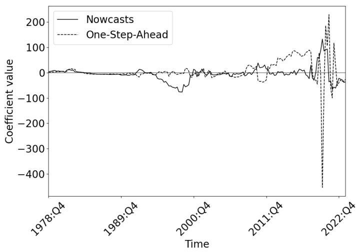
(b) Experience



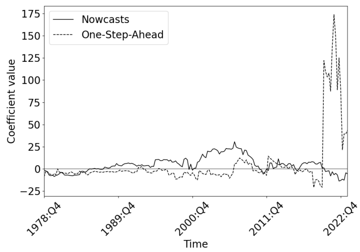
(c) Disconsensus



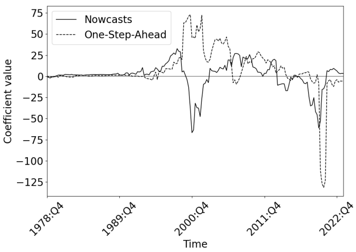
(d) Accuracy



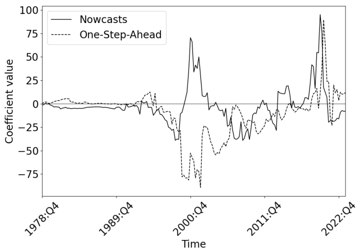
(e) Bias



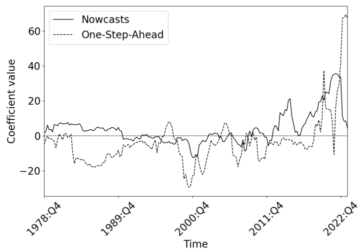
(f) Consistency



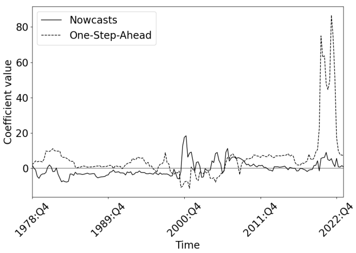
(g) Acc Momentum



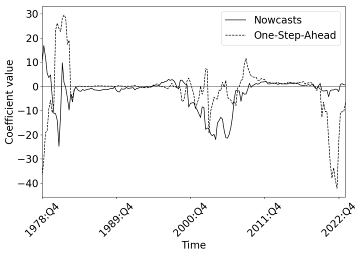
(h) Bias Momentum



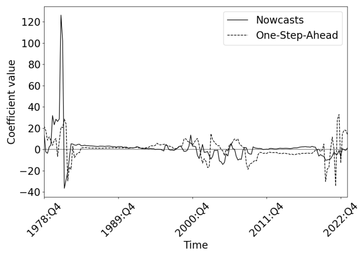
(i) Cons Momentum



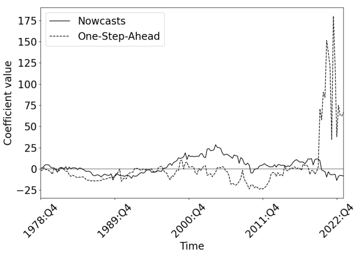
(j) Spillover Disconsensus



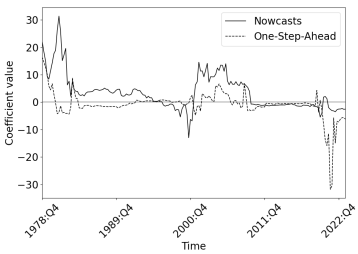
(k) Spillover Accuracy



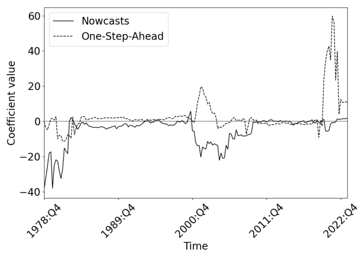
(l) Spillover Bias



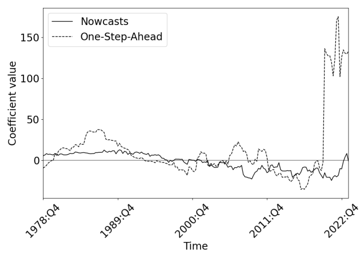
(m) Spillover Consistency



(n) Spillover Acc Momentum



(o) Spillover Bias Momentum



(p) Spillover Cons Momentum

Figure 6.2: Characteristic Coefficients of the GDP Growth Rate.

To provide further insight into the variation of the coefficients over time, Figure 6.2 plots the coefficients of the GDP nowcasts and one-step-ahead forecasts. The coefficients for the unemployment rate and the CPI inflation rate are presented in Appendix G. Overall, this figure shows substantial variation in the coefficients over time. For instance, all coefficients cross the zero axis multiple times during the estimation period. Hence, again, these results demonstrate the flexible nature of the FCC model, which allows for variation of the coefficients over time. The industry dummy exhibits mostly positive coefficients over time. This implies that−in line with our expectation−forecasts from financial service providers are assigned larger weight by the FCC method. For all other variables the results are ambiguous, because coefficients are both positive and negative for a large number of periods. It is worth mentioning that most GDP coefficients reach exceptional values just after the COVID outlier, highlighting the relevance of a sensitivity analysis of results estimated without outliers.

6.4 Sensitivity Analysis

This section provides a sensitivity analysis of the results to three factors: (i) the removal of outliers, (ii) the averaging of characteristics across forecast horizons, and (iii) the standardisation of forecaster characteristics to standard deviation one.

Sensitivity to Outliers

To analyse the effect of removing outliers, results were estimated with the most important outliers removed. Specifically, for both the GDP growth rate and the unemployment rate, the COVID outliers are removed: Q2 and Q3 2020. For the CPI inflation rate, the outlier during the Great Recession is removed: Q4 2008.

Table 6.4: MSFE Results without Outliers for the FCC Methods and Benchmarks

Combination Method	Nowcast			One-Step-Ahead			MCS Count
	GDP	UMP	CPI	GDP	UMP	CPI	
Panel A: FCC Methods							
FCC	***0.120***	*0.133*	***0.627***	***0.238***	*0.178*	***1.823***	4
	(9.09%)	(2.31%)	(15.47%)	(-46.15%)	(-55.16%)	(13.65%)	
Panel B: Benchmark Methods							
BAM	1.66	**0.17**	1.85	1.48	0.80	1.74	1
	(-13.10%)	(-41.03%)	(-4.21%)	(-58.76%)	(-73.52%)	(-18.24%)	
peLASSO	1.83	0.32	1.48	1.43	**0.74**	1.84	1
	(-16.67%)	(-2.33%)	(0.22%)	(-64.21%)	(-75.56%)	(-17.56%)	

Note: This table presents the MSFE of the FCC method and the relative MSFE of the BAM and peLASSO benchmarks, estimated without outliers. For the estimation of these results, time periods 2020 Q2 and 2020 Q3 were removed for the GDP growth rate and the unemployment rate, as well as the time period 2008 Q4 for the CPI inflation rate. The percentages in parentheses represent the percentage change from the original MSFE values estimated with outliers included. The results for the FCC method are estimated with moving window estimation. The table also presents the results from the MCS procedure: the method selected into the MCS with 95% certainty is presented in bold. The column "MCS Count" provides the total number of times a model is selected into the MCS.

Table 6.4 presents results with the aforementioned outliers removed: the MSFE of the regular FCC and the relative MSFEs of the BAM and peLASSO benchmarks. Below each result, the percentual change in MSFE compared to estimation with outliers included is shown in parentheses. Appendix H provides the results estimated without outliers for all of the FCC methods and benchmarks. Overall, most forecasting errors decrease when the outliers are removed. Importantly, the superior performance of the regular FCC method for the GDP growth rate and CPI inflation rate remains. However, as expected from Figure 6.1d, the superior forecasting performance of the FCC method for the unemployment rate one-step-ahead forecasts disappears when the outliers are removed. Instead, the peLASSO method achieves the lowest MSFE and is selected into the MCS. In conclusion, the results appear relatively robust to the deletion of outliers, as the superior performance of the FCC model remains in 4 out of 5 cases.

Sensitivity to Horizon Dependence of Forecaster Characteristics

As explained in Section 4, most of the characteristics included depend on the forecast horizon. That is to say, these characteristics are defined separately for the nowcasts and the one-step-ahead forecasts. However, information from one horizon may be relevant for combining forecasts on another horizon. For example, the forecaster accuracy of the nowcasts may also be useful for the combination of the one-step-ahead forecast. In other words, using characteristics that include information from both horizons may lead to improved forecasting performance.

For that reason, we present results using characteristics that include information from both horizons in Appendix H. These results are constructed with characteristics that average the two horizon dependent characteristics. The results in Appendix H show that the performance of the FCC method is not improved by averaging the characteristics over horizons. Nevertheless, after performing the MCS procedure on these results and the benchmarks, the original superiority of the regular FCC method over the benchmarks remains in all instances. Thus, the results are robust to the horizon dependence of forecaster characteristics.

Sensitivity to Standardisation of Forecaster Characteristics

As explained in Section 2.2, each forecaster characteristic is standardised to mean zero, such that the combination weights sum to one. In the original parametric portfolio policy by Brandt et al. (2009), all characteristics are standardised to a standard deviation of one as well. For that reason, we assess the robustness of the results when the characteristics are also standardised to standard deviation one, apart from mean zero.

Appendix H provides the (relative) MSFE results of the FCC methods with forecaster characteristics additionally standardised to standard deviation one. These results exhibit larger MSFEs for most FCC methods, such that the optimality of the original results remains. This is in line with our expectation: the additional standardisation of the standard deviation removes information about the different variability of the characteristics. Nevertheless, after performing the MCS procedure on the results in Appendix H and the benchmarks, the superiority of the regular FCC method over the benchmark remains in all instances. Hence, the results are robust to standardisation of the characteristics to standard deviation one.

7 Conclusion

This paper develops the Forecaster Characteristics Combination (FCC) method and evaluates its performance against several benchmarks, using expert panel macroeconomic forecasts. We conclude that the FCC method proves to be a promising method. It generally outperforms benchmarks such as the equally weighted average. These results are found with data from the Survey of Professional Forecasters (SPF), focusing on three key macroeconomic indicators: the real GDP growth rate, unemployment rate and CPI inflation rate. Notably, the FCC method demonstrates lower MSFE across two out of three macroeconomic variables: the GDP growth rate and the CPI inflation rate. Only for the unemployment rate, the benchmarks outperform the FCC method.

More detailed analysis reveals that the flexibility of the FCC method is key for its favourable performance. The FCC method varies the level of importance assigned to each forecaster characteristic over time. Accordingly, the coefficients assigned to each characteristic show large variation over time, confirming the flexible nature of the FCC forecast combination method. Consequently, extensions employing regularisation to the regular FCC method do not improve forecasting performance. The constraints inherent to these shrinkage techniques limit model flexibility and, thereby, worsen forecasting performance.

The FCC method's superior performance is even more prominent during atypical periods, such as the Great Recession or COVID-19. During these periods, the forecasting error of the FCC method is significantly lower than the errors of the benchmark methods. Nevertheless, a sensitivity analysis shows that even when outliers are removed, the FCC method mostly maintains its superior performance. Broader economic implications suggest that using the FCC forecast combination method may lead to more reliable macroeconomic forecasts. Although the FCC method proves even more beneficial during atypical periods, the method may improve forecasting performance at all times.

A limitation of the application of the FCC method in this paper is that it focuses solely on expert-based forecasts. However, the FCC method is also applicable to model forecasts. Further research into the performance of the FCC method when combining model forecasts would be an interesting area of further research. Another limitation of this paper is that it evaluates the FCC method using only three macroeconomic variables and two forecast horizons. Although the FCC method performs better for two of the three economic variables for both horizons, it remains uncertain how the method performs over longer horizons and for other macroeconomic variables. Hence, it would be interesting to evaluate the FCC method over more extended horizons and with a broader range of macroeconomic variables. A final limitation of this paper is the potential for altering or expanding the set of characteristics included. Research into determining the optimal set of characteristics for the FCC method would be relevant.

A limitation of the FCC method itself is its assumption that the identified forecaster characteristics remain relevant over the entire forecasting period. It would be worthwhile to investigate whether incorporating new characteristics or periodically re-evaluating the existing ones could further improve forecast accuracy. Also, the FCC method as presented in this paper is suitable for combining point forecasts only. The combination of density forecasts with a method similar to FCC would be an interesting area of further research.

References

Aiolfi, M., & Timmermann, A. (2006). Persistence in forecasting performance and conditional combination strategies. *Journal of Econometrics*, *135*(1), 31–53.

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

Armstrong, J. (2001). Combining forecasts. In J. Armstrong (Ed.), *Principles of forecasting: A handbook for researchers and practitioners* (pp. 417–439). Springer.

Bates, J. M., & Granger, C. W. (1969). The combination of forecasts. *Journal of the Operational Research Society*, *20*(4), 451–468.

Baumeister, C., & Kilian, L. (2015). Forecasting the real price of oil in a changing world: A forecast combination approach. *Journal of Business & Economic Statistics*, *33*(3), 338–351.

Behr, P., Guettler, A., & Truebenbach, F. (2012). Using industry momentum to improve portfolio performance. *Journal of Banking & Finance*, *36*(5), 1414–1423.

Brandt, M. W., Santa-Clara, P., & Valkanov, R. (2009). Parametric portfolio policies: Exploiting characteristics in the cross-section of equity returns. *Review of Financial Studies*, *22*(9), 3411–3447.

Capistrán, C., & Timmermann, A. (2009). Forecast combination with entry and exit of experts. *Journal of Business & Economic Statistics*, *27*(4), 428–440.

Chassot, J. (2024). *Model-confidence-set* (Version 0.21.3). https://github.com/JLDC/model-confidence-set

Chatterjee, A., & Lahiri, S. N. (2011). Bootstrapping LASSO estimators. *Journal of the American Statistical Association*, *106*(494), 608–625.

Clemen, R. T., & Winkler, R. L. (1986). Combining economic forecasts. *Journal of Business & Economic Statistics*, *4*(1), 39–46.

Croushore, D. D. (1993). Introducing: The survey of professional forecasters. *Business Review-Federal Reserve Bank of Philadelphia*, *6*(1), 3–15.

DeMiguel, V., Garlappi, L., & Uppal, R. (2009). Optimal versus naive diversification: How inefficient is the 1/N portfolio strategy? *Review of Financial studies*, *22*(5), 1915–1953.

DeMiguel, V., Martin-Utrera, A., Nogales, F. J., & Uppal, R. (2020). A transaction-cost perspective on the multitude of firm characteristics. *Review of Financial Studies*, *33*(5), 2180–2222.

Diebold, F. X., & Pauly, P. (1987). Structural change and the combination of forecasts. *Journal of Forecasting*, *6*(1), 21–40.

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

Fan, J., & Li, R. (2001). Variable selection via nonconcave penalized likelihood and its oracle properties. *Journal of the American statistical Association*, *96*(456), 1348–1360.

Fixler, D. J., de Francisco, E., & Kanal, D. (2021). The revisions to GDP, GDI, and their major components. *Survey of Current Business*, *101*(1), 1–3.

Genre, V., Kenny, G., Meyler, A., & Timmermann, A. (2013). Combining expert forecasts: Can anything beat the simple average? *International Journal of Forecasting*, *29*(1), 108–121.

Hand, J. R., & Green, J. (2011). The importance of accounting information in portfolio optimization. *Journal of Accounting, Auditing & Finance*, *26*(1), 1–34.

Hansen, P. R., Lunde, A., & Nason, J. M. (2011). The model confidence set. *Econometrica*, *79*(2), 453–497.

Hansen, P. R., Lunde, A., & Nason, J. M. (2003). Choosing the best volatility models: The model confidence set approach. *Oxford Bulletin of Economics and Statistics*, *65*(4), 839–861.

Hoerl, A. E., & Kennard, R. W. (1970). Ridge regression: Biased estimation for nonorthogonal problems. *Technometrics*, *12*(1), 55–67.

Kang, Y., Cao, W., Petropoulos, F., & Li, F. (2022). Forecast with forecasts: Diversity matters. *European Journal of Operational Research*, *301*(1), 180–190.

Li, X., Kang, Y., & Li, F. (2020). Forecasting with time series imaging. *Expert Systems with Applications*, *160*(2), 113–126.

Lin, V. S., Goodwin, P., & Song, H. (2014). Accuracy and bias of experts' adjusted forecasts. *Annals of Tourism Research*, *48*(2), 156–174.

Loungani, P. (2001). How accurate are private sector forecasts? Cross-country evidence from consensus forecasts of output growth. *International Journal of forecasting*, *17*(3), 419–432.

Ma, S., & Fildes, R. (2021). Retail sales forecasting with meta-learning. *European Journal of Operational Research*, *288*(1), 111–128.

Montero-Manso, P., Athanasopoulos, G., Hyndman, R. J., & Talagala, T. S. (2020). FFORMA: Feature-based forecast model averaging. *International Journal of Forecasting*, *36*(1), 86–92.

Nowotarski, J., Raviv, E., Trück, S., & Weron, R. (2014). An empirical comparison of alternative schemes for combining electricity spot price forecasts. *Energy Economics*, *46*(3), 395–412.

Palm, F. C., & Zellner, A. (1992). To combine or not to combine? Issues of combining forecasts. *Journal of Forecasting*, *11*(8), 687–701.

Pawlikowski, M., & Chorowska, A. (2020). Weighted ensemble of statistical models. *International Journal of Forecasting*, *36*(1), 93–97.

Petropoulos, F., & Spiliotis, E. (2021). The wisdom of the data: Getting the most out of univariate time series forecasting. *Forecasting*, *3*(3), 478–497.

Rapach, D. E., Strauss, J. K., & Zhou, G. (2010). Out-of-sample equity premium prediction: Combination forecasts and links to the real economy. *Review of Financial Studies*, *23*(2), 821–862.

Rapach, D. E., & Zhou, G. (2020). Time-series and cross-sectional stock return forecasting: New machine learning methods. In E. Jurczenko (Ed.), *Machine learning for asset management: New developments and financial applications* (pp. 1–33). Wiley.

Sill, K. (2012). Measuring economic uncertainty using the survey of professional forecasters. *Federal Reserve Bank of Philadelphia Business Review*, *92*(1), 16–27.

Smith, J., & Wallis, K. F. (2009). A simple explanation of the forecast combination puzzle. *Oxford Bulletin of Economics and Statistics*, *71*(3), 331–355.

Song, H., Gao, B. Z., & Lin, V. S. (2013). Combining statistical and judgmental forecasts via a web-based tourism demand forecasting system. *International Journal of Forecasting*, *29*(2), 295–310.

Stock, J. H., & Watson, M. W. (2004). Combination forecasts of output growth in a seven-country data set. *Journal of forecasting*, *23*(6), 405–430.

Tibshirani, R. (1996). Regression shrinkage and selection via the LASSO. *Journal of the Royal Statistical Society*, *58*(1), 267–288.

Tikhonov, A. N. (1943). On the stability of inverse problems. *Doklady Akademii Nauk*, *39*(5), 195–198.

Timmermann, A. (2013). *Handbook of economic forecasting* (G. Elliot, C. Granger & A. Timmermann, Eds.). Elsevier.

Van Dijk, D., & Franses, P. H. (2019). Combining expert-adjusted forecasts. *Journal of Forecasting*, *38*(5), 415–421.

Wang, X., Hyndman, R. J., Li, F., & Kang, Y. (2023). Forecast combinations: An over 50-year review. *International Journal of Forecasting*, *39*(4), 1518–1547.

Zou, H., & Hastie, T. (2005). Regularization and variable selection via the elastic net. *Journal of the Royal Statistical Society*, *67*(2), 301–320.

Appendix

A OLS Derivation

This appendix derives the least squares estimator for the vector of coefficients in the FCC method: $\boldsymbol{\theta}_h$. We start from the objective function provided in equation (7), which is as follows:

$$\frac{1}{T}\sum_{t=1}^{T}\left(r_{t+h}-\left(\sum_{i=1}^{N_t}\frac{1}{N_t}(1+\boldsymbol{\theta}_h'\hat{\mathbf{z}}_{i,t,h})\times f_{i,t+h|t}\right)\right)^2, \quad r_{t+h}, f_{i,t+h|t}\in\mathbb{R}, \quad \boldsymbol{\theta}_h,\hat{\mathbf{z}}_{i,t,h}\in\mathbb{R}^Q. \tag{41}$$

This can be rewritten as:

$$\frac{1}{T}\sum_{t=1}^{T}\left(r_{t+h}-\left(\sum_{i=1}^{N_t}\frac{1}{N_t}(f_{i,t+h|t}+\boldsymbol{\theta}_h'\hat{\mathbf{z}}_{i,t,h}\times f_{i,t+h|t})\right)\right)^2, \tag{42}$$

$$\frac{1}{T}\sum_{t=1}^{T}\left(r_{t+h}-\left(\frac{1}{N_t}\sum_{i=1}^{N_t}f_{i,t+h|t}\right)-\left(\frac{1}{N_t}\sum_{i=1}^{N_t}\boldsymbol{\theta}_h'\hat{\mathbf{z}}_{i,t,h}\times f_{i,t+h|t})\right)\right)^2, \tag{43}$$

$$\frac{1}{T}\sum_{t=1}^{T}\left(r_{t+h}-\left(\frac{1}{N_t}\sum_{i=1}^{N_t}f_{i,t+h|t}\right)-\boldsymbol{\theta}_h'\left(\frac{1}{N_t}\sum_{i=1}^{N_t}\hat{\mathbf{z}}_{i,t,h}\times f_{i,t+h|t})\right)\right)^2, \tag{44}$$

$$\frac{1}{T}\sum_{t=1}^{T}\left(y_t-\boldsymbol{\theta}_h'\mathbf{x}_t\right)^2, \quad \mathbf{x}_t\in\mathbb{R}^Q. \tag{45}$$

Note that when moving from equation (44) to (45), we define:

$$y_t := r_{t+h}-\left(\frac{1}{N_t}\sum_{i=1}^{N_t}f_{i,t+h|t}\right), \quad y_t\in\mathbb{R}, \tag{46}$$

$$\mathbf{x}_t := \frac{1}{N_t}\sum_{i=1}^{N_t}\hat{\mathbf{z}}_{i,t,h}\times f_{i,t+h|t}, \quad \mathbf{x}_t\in\mathbb{R}^Q. \tag{47}$$

Next, we stack the elements $y_1,\dots,y_T$ into vector $\mathbf{y}$ of dimension $T\times 1$. Additionally, we stack the transposed vectors $\mathbf{x}_1',\dots,\mathbf{x}_T'$ into matrix $\mathbf{X}$ of dimension $T\times Q$. We leave out the factor $\frac{1}{T}$ from the minimisation problem in equation (45) as it does not affect the solution. Then, the minimisation problem from equation (45) can be rewritten as:

$$\|\mathbf{y}-\mathbf{X}\boldsymbol{\theta}_h\|_2, \quad \mathbf{y}\in\mathbb{R}^T, \quad \mathbf{X}\in\mathbb{R}^{T\times Q}, \quad \boldsymbol{\theta}_h\in\mathbb{R}^Q, \tag{48}$$

in which $\|.\|_2$ denotes the L2 norm, $\mathbf{y}$ is $T\times 1$, $\mathbf{X}$ is $T\times Q$ and $\boldsymbol{\theta}_h$ remains $Q\times 1$, in which Q is the total number of characteristics included. In equation (48), we recognise an ordinary least squares minimisation problem, such that we can use the least squares expression to estimate $\boldsymbol{\theta}_h$:

$$\hat{\boldsymbol{\theta}}_h = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{y}. \tag{49}$$

B Realisation Data and Transformations

This appendix describes the realisation data used for each of the three macroeconomic variables under consideration. Additionally, it provides the transformations used to achieve the correct form of the macroeconomic variables. These transformations are necessary to match the form of the forecasts from the SPF.

GDP Growth Rate

I evaluate the real GDP growth rate forecasts using the first vintage estimates for real GDP level. A seasonally adjusted dataset with GDP level estimates is available in billions of real dollars from the Federal Reserve Bank of Philadelphia. In this dataset, estimates before 1992 concern real GNP and estimates after 1992 concern real GDP. This discrepancy is consistent with the forecasts, such that level estimates from this dataset can be used directly to construct GDP growth rates. First vintage real GDP level estimates are available from the fourth quarter of 1965 up to the fourth quarter of 2023. There is one missing first vintage estimate: for 1995Q4. For this period, we use the second vintage estimate instead. In total, the number of quarterly first vintage estimates of real GDP equals 233. We transform the GDP level estimates to growth rate estimates as follows:

$$r_t^{\text{GDP}} = \frac{\text{GDP}_\text{t} - \text{GDP}_\text{t-1}}{\text{GDP}_\text{t-1} \times 100\%} \tag{50}$$

in which r_t^{GDP} denotes the GDP growth rate realisations for $t = 1, \ldots, T$ and GDP_t denotes the first vintage GDP level estimate during quarter t.

Unemployment Rate

To evaluate the unemployment rate forecasts, we use the first vintage monthly unemployment rate estimates from the real-time dataset for macroeconomists by the Federal Reserve Bank of Philadelphia. Similar to the forecasts, these estimates are seasonally adjusted. The first vintage estimates are available from Q4 1965 to Q4 2023 and thus cover the entire available period of unemployment forecasts from the SPF. The unemployment rate estimates from the real-time dataset for macroeconomists are recorded on a monthly basis. Conversely, the unemployment rate forecasts are quarterly rather than monthly: they capture the quarterly average of the underlying monthly levels. Hence, we transform the monthly unemployment rates to quarterly unemployment rates by averaging the monthly unemployment rates over the three months belonging to a certain quarter:

$$r_t^{\text{UMP}} = \frac{\sum_{m=1}^{3} \text{UMP}_{t_m}^{M}}{3}, \quad t = 1, \ldots, T, \tag{51}$$

in which r_t^{UMP} is the quarterly unemployment rate at quarter t and $\text{UMP}_{t_i}^{M}$ is the monthly unemployment rate retrieved from the Federal Reserve dataset at month $m = 1, \ldots 3$ of quarter t. We use r_t^{UMP} to evaluate the quarterly unemployment forecasts.

CPI Inflation Rate

To evaluate the headline CPI inflation rate forecasts we use first vintage monthly CPI level estimates from Federal Reserve Economic Data (FRED) data base by the Federal Reserve Bank of St. Louis. We use this dataset rather than the Federal Reserve Bank of Philadelphia dataset because the FRED dataset contains earlier vintages for CPI level estimates. The values from these two datasets match during overlapping periods, such that both datasets regard the same CPI. Similar to the CPI forecasts, the FRED estimates are seasonally adjusted. Monthly first vintage estimates are available from July 1972 to March 2024. Thus, the FRED dataset covers the entire forecasting period between Q3 1981 and Q4 2023. Both the estimates and the vintages are updated on a monthly basis. To construct quarterly vintages we average the three monthly estimates belonging to a certain quarter.

Whereas the CPI estimates concern CPI level, the forecasts concern the CPI inflation *rate*, defined as the annualised quarter-over-quarter percentage change of the quarterly CPI level. For that reason, we transform the monthly CPI level estimates to quarterly CPI inflation rate estimates as follows:

$$\mathrm{CPI}_t^Q = \frac{\mathrm{CPI}_{t_1}^M + \mathrm{CPI}_{t_2}^M + \mathrm{CPI}_{t_3}^M}{3}, \quad t = 1, \dots T \tag{52}$$

$$\Delta_{\mathrm{CPI}_t^Q} = \frac{\mathrm{CPI}_t^Q - \mathrm{CPI}_{t-1}^Q}{\mathrm{CPI}_{t-1}^Q}, \tag{53}$$

$$r_t^{\mathrm{CPI}} = ((1 + \Delta_{CPI_t})^4 - 1) \times 100, \tag{54}$$

in which $\mathrm{CPI}_{t_i}^M$ is the monthly CPI level obtained from the Federal Reserve data for months $m = 1, 2, 3$ in quarter $t = 1, \dots T$, CPI_t^Q is the quarterly average CPI level at quarter $t = 1, \dots, T$, $\Delta_{\mathrm{CPI}_t^Q}$ is the CPI quarter-over-quarter percentage change at time t and r_t^{CPI} is the annualised quarter-over-quarter percentage change of the quarterly CPI level at quarter t. Thus, r_t^{CPI} is the realisation estimate used to evaluate the forecasts of the CPI inflation rate.

C Momentum and Spillover Characteristics

This appendix provides the mathematical expressions for the momentum and spillover characteristics. Note that accuracy momentum and spillover accuracy were already defined in Sections 4.7 and 4.8.

Bias Momentum

Bias momentum is calculated as the square of the average forecasting error over the previous four time periods:

$$\mu_{i,t,v,h} = \begin{cases} \left(\frac{\sum_{\tau=t-(4+h)}^{t-(1+h)} \mathbb{1}_{i,\tau,v,h} \epsilon^v_{i,\tau+h}}{\sum_{\tau=1}^{t-(1+h)} \mathbb{1}_{i,\tau+h,v}} \right)^2 & \text{if} \quad \sum_{\tau=t-(4+h)}^{t-(1+h)} \mathbb{1}_{i,\tau,v,h} \geq 1, \\ 0 & \text{otherwise,} \end{cases} \tag{55}$$

in which all variables are as defined before. Hence, bias momentum equals the bias over the previous four time periods, if forecaster i has provided a forecast during any of these time periods. Otherwise, bias momentum is set to zero.

Consistency Momentum

Consistency momentum is calculated as the first-order autocorrelation of the forecasting error over the previous four time periods:

$$\nu_{i,t,v,h} = \begin{cases} \frac{\sum_{\tau=t-(3+h)}^{t-(1+h)} \mathbb{1}_{i,\tau,v,h} \mathbb{1}_{i,\tau-1,v,h} \left(\epsilon^v_{i,\tau+h} - \bar{\epsilon}^v_{i,t,h} \right) \left(\epsilon^v_{i,\tau+h-1} - \bar{\epsilon}^v_{i,t,h} \right)}{\sum_{\tau=1}^{t-(1+h)} \mathbb{1}_{i,\tau,v,h} \left(\epsilon^v_{i,\tau+h} - \bar{\epsilon}^v_{i,t,h} \right)^2}, & \text{if } \sum_{\tau=t-(4+h)}^{t-(1+h)} \mathbb{1}_{i,\tau,v,h} \geq 2, \\ 0 & \text{otherwise,} \end{cases} \tag{56}$$

in which all variables are as defined before. Hence, consistency momentum equals the consistency over the previous four time periods, if forecaster i has provided a forecast during any of these time periods. Otherwise, consistency momentum is set to zero.

Spillover Disconsensus

Spillover disconsensus $\Psi i, t, v, h$ is calculated by taking the average disconsensus of the other macroeconomic variables $m, n \neq v$ for forecaster i at time t for horizon h:

$$\Psi_{i,t,v,h} = \frac{\psi_{i,t,m,h} + \psi_{i,t,n,h}}{2}, \quad m, n \in \{\text{UMP}, \text{CPI}, \text{GDP}\} \setminus v, \quad m \neq n, \tag{57}$$

in which $\psi_{i,t,m,h}$ denotes the disconsensus of forecaster i at time t for macroeconomic variable m and horizon h.

Spillover Bias

Spillover bias $\Gamma_{i,t,v,h}$ is calculated by taking the average bias of the other macroeconomic variables

$m, n \neq v$ for forecaster i at time t for horizon h:

$$\Gamma_{i,t,v,h} = \frac{\gamma_{i,t,m,h} + \gamma_{i,t,n,h}}{2}, \quad m, n \in \{\text{UMP}, \text{CPI}, \text{GDP}\} \setminus v, \quad m \neq n, \tag{58}$$

in which $\gamma_{i,t,m,h}$ is the bias of forecaster i at time t for macroeconomic variable m and horizon h.

Spillover Consistency

Spillover consistency $\xi_{i,t,v,h}$ is calculated by taking the average consistency of the other macroeconomic variables $m, n \neq v$ for forecaster i at time t for horizon h:

$$\xi_{i,t,v,h} = \frac{\rho_{i,t,m,h} + \rho_{i,t,n,h}}{2}, \quad m, n \in \{\text{UMP}, \text{CPI}, \text{GDP}\} \setminus v, \quad m \neq n, \tag{59}$$

in which $\rho_{i,t,m,h}$ is the consistency of forecaster i at time t for macroeconomic variable m and horizon h.

Spillover Accuracy Momentum

Spillover accuracy momentum $\Lambda_{i,t,v,h}$ is the average accuracy momentum of the other macroeconomic variables $m, n \neq v$ for forecaster i at time t for horizon h:

$$\Lambda_{i,t,v,h} = \frac{\eta_{i,t,m,h} + \eta_{i,t,n,h}}{2}, \quad m, n \in \{\text{UMP}, \text{CPI}, \text{GDP}\} \setminus v, \quad m \neq n, \tag{60}$$

in which $\eta_{i,t,m,h}$ denotes the accuracy momentum of forecaster i at time t for macroeconomic variable m and horizon h.

Spillover Bias Momentum

Spillover bias momentum $\omega_{i,t,v,h}$ is the average bias momentum of the other macroeconomic variables $m, n \neq v$ for forecaster i at time t for horizon h:

$$\omega_{i,t,v,h} = \frac{\mu_{i,t,m,h} + \mu_{i,t,n,h}}{2}, \quad m, n \in \{\text{UMP}, \text{CPI}, \text{GDP}\} \setminus v, \quad m \neq n, \tag{61}$$

in which $\mu_{i,t,m,h}$ denotes the bias momentum of forecaster i at time t for macroeconomic variable m and horizon h.

Spillover Consistency Momentum

Spillover consistency momentum $\Xi_{i,t,v,h}$ is the average consistency momentum of the other macroeconomic variables $m, n \neq v$ for forecaster i at time t for horizon h:

$$\Xi_{i,t,v,h} = \frac{\nu_{i,t,m,h} + \nu_{i,t,n,h}}{2}, \quad m, n \in \{\text{UMP}, \text{CPI}, \text{GDP}\} \setminus v, \quad m \neq n, \tag{62}$$

in which $\nu_{i,t,m,h}$ denotes the consistency momentum of forecaster i at time t for macroeconomic variable m and horizon h.

D Correlation between Characteristics

This appendix provides the table of correlations between characteristics for all three macroeconomic variables.

Table D.1: Correlation between Standardised Characteristics for Horizon $h = 0$

	Exp	**Disc**	**Acc**	**Bias**	**Cons**	**Acc Mom**	**Bias Mom**	**Cons Mom**	**Spill. Disc**	**Spill Acc**	**Spill Bias**	**Spill Cons**	**Spill AM**	**Spill BM**	**Spill CM**
Panel A: Real GDP Growth Rate															
Ind	−0.12	−0.02	−0.09	−0.02	−0.14	−0.05	−0.05	0.02	−0.04	−0.09	−0.01	−0.07	−0.12	−0.13	0.01
Exp		−0.02	0.04	−0.14	0.25	0.05	0.09	−0.13	−0.01	−0.05	−0.11	0.21	0.02	0.06	−0.14
Disc			0.19	0.11	−0.01	0.11	0.11	0.01	0.08	0.05	0.02	−0.02	0.06	0.08	0.00
Acc				0.58	0.03	0.44	0.46	−0.03	0.06	0.13	0.02	−0.06	0.07	0.08	−0.03
Bias					0.04	0.19	0.14	0.04	0.02	0.03	0.04	−0.05	0.02	0.00	0.05
Cons						0.04	0.07	0.09	−0.01	−0.08	−0.05	0.26	−0.04	−0.07	0.08
Acc Mom							0.74	−0.03	0.05	0.05	0.01	−0.02	0.13	0.10	−0.01
Bias Mom								−0.11	0.07	0.04	0.00	−0.03	0.09	0.15	−0.10
Cons Mom									−0.01	−0.01	0.02	0.09	−0.02	−0.07	0.29
Spill Disc										0.05	0.02	−0.01	0.07	0.12	0.01
Spill Acc											0.79	−0.04	0.64	0.30	−0.01
Spill Bias												−0.04	0.53	0.02	0.02
Spill Cons													−0.01	−0.02	0.13
Spill AM														0.66	−0.01
Spill BM															−0.09
Panel B: Unemployment Rate															
Ind	−0.12	−0.05	0.00	−0.02	−0.06	−0.02	−0.02	0.02	−0.03	−0.10	−0.01	−0.09	−0.12	−0.13	0.01
Exp		−0.02	−0.07	−0.19	0.26	0.00	0.01	−0.11	−0.02	−0.04	−0.12	0.22	0.03	0.08	−0.16
Disc			0.03	0.04	0.00	−0.03	0.00	0.01	0.08	0.05	0.05	0.00	0.04	0.03	−0.01
Acc				0.83	−0.01	0.34	0.40	0.01	0.01	0.11	0.02	−0.03	0.05	0.06	0.00
Bias					−0.04	0.20	0.27	0.03	0.01	0.18	0.13	−0.06	0.09	0.05	0.03
Cons						−0.01	0.03	0.06	0.02	−0.01	−0.05	0.19	0.01	0.01	0.04
Acc Mom							0.60	0.02	0.02	0.04	0.01	0.00	0.10	0.05	0.02
Bias Mom								−0.02	0.02	0.04	0.00	0.01	0.06	0.13	−0.03
Cons Mom									0.01	0.00	0.03	0.11	−0.02	−0.10	0.30
Spill Disc										0.09	0.03	−0.03	0.10	0.15	0.01
Spill Acc											0.78	−0.06	0.63	0.32	−0.02
Spill Bias												−0.02	0.52	0.03	0.03
Spill Cons													−0.02	−0.03	0.15
Spill AM														0.68	−0.03
Spill BM															−0.12
Panel C: CPI Inflation Rate															
Ind	−0.11	−0.02	−0.10	−0.01	−0.07	−0.12	−0.13	0.00	−0.05	−0.04	−0.03	−0.11	−0.05	−0.05	0.03
Exp		−0.01	−0.05	−0.12	0.25	0.03	0.08	−0.13	−0.01	0.00	−0.25	0.34	0.04	0.08	−0.15
Disc			0.08	0.06	0.00	0.09	0.11	0.00	0.09	0.05	0.00	0.00	0.04	0.07	0.01
Acc				0.82	−0.04	0.68	0.31	−0.01	0.09	0.17	0.14	−0.04	0.07	0.06	0.00
Bias					−0.01	0.61	0.04	0.03	0.05	0.03	0.14	−0.07	0.01	0.01	0.03
Cons						0.00	−0.01	0.12	0.01	0.03	−0.04	0.25	0.02	0.01	0.14
Acc Mom							0.65	−0.01	0.08	0.09	0.08	−0.02	0.12	0.13	−0.03
Bias Mom								−0.10	0.10	0.12	0.04	−0.04	0.11	0.20	−0.11
Cons Mom									−0.01	−0.01	0.03	0.06	0.03	−0.03	0.26
Spill Disc										0.06	0.05	0.02	0.05	0.04	0.00
Spill Acc											0.59	−0.03	0.27	0.30	−0.01
Spill Bias												−0.11	0.08	0.06	0.04
Spill Cons													0.02	0.08	0.10
Spill AM														0.54	0.02
Spill BM															−0.08

Note: This table presents the correlations between characteristics for the nowcasts of all three macroeconomic variables. The abbreviations used for the characteristics can be found in Table 4.1.

E Model Confidence Set Procedure

This appendix provides further details about the MCS approach by Hansen et al. (2011) and our implementation. There are several benefits to the MCS approach. For example, a nice feature of the MCS approach is that it considers the informativeness of the data used. Informative data results in a compact MCS, while uninformative data leads to a larger MCS, potentially including all employed models. Additionally, the MCS approach allows for the possibility that more than one model is superior.

Our implementation of the MCS is as follows. We run the MCS procedure using the squared forecasting error as loss function to determine the best performing model. To implement the MCS approach, we use the bootstrap implementation because of its simplicity: it avoids the estimation of high dimensional covariance matrices (Hansen et al., 2011). The bootstrap MCS algorithm is presented in Algorithm 1. Furthermore, we employ the `model-confidence-set` Python package (Chassot, 2024). Following Hansen et al. (2003) we set the number of bootstrap resamples equal to 1000 and use a confidence level of 95%.

Algorithm 1 MCS Algorithm

- **Step 0: Initialisation**.
  - Set the initial set of models under consideration M to the total set of models M^0. Thus, set $M = M^0$.
  - Define a loss function, we use the squared forecasting error.
  - For each pair of models A and B calculate the loss differential $l_{A,B,t} = \text{SFE}_{A,t} - \text{SFE}_{B,t}$, in which $\text{SFE}_{A,t}$ denotes the squared forecasting error of model A at time t.
- **Step 1: Test Statistic**. Calculate the test statistics for the null hypothesis that all models in the set M have equal predictive ability. As suggested by Hansen et al. (2011), we use the maximum pairwise loss differential statistic, which is calculated as the maximum absolute value of the pairwise loss differentials.
- **Step 2: Bootstrap Procedure**.
  - Generate bootstrap samples from the loss differentials by resampling with replacements.
  - Compute the test statistic for each bootstrap sample.
- **Step 3: Elimination**
  - Calculate the p-values for the test statistics based on the bootstrap distribution.
  - Remove the model with the largest p-value from the M set.
- **Step 4: Iterative Testing** Repeat the bootstrap and elimination steps until no model can be removed from the set M at the required level of confidence.

F Regularisation Results

This appendix provides the percentage of time periods during which a coefficient is set to zero by the regularisation techniques. Note that Table F.1 includes percentages for LASSO (panel A) and Elastic Net (panel B), but not for post-LASSO because the shrinkage percentages of post-LASSO are identical to LASSO.

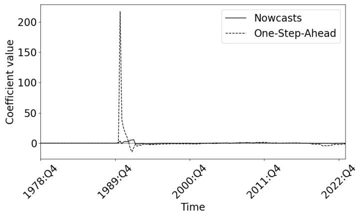
Table F.1: Percentage of Time Periods during Which Coefficients Are Set to Zero by Regularisation

Variable	Nowcasts			One-Step-Ahead		
	GDP	**UMP**	**CPI**	**GDP**	**UMP**	**CPI**
Panel A: LASSO Shrinkage to Zero (in %)						
Financial Dummy	85.08	52.49	68.46	95.03	72.93	92.31
Experience	63.54	57.46	76.92	66.85	53.04	86.15
Disconsensus	78.45	51.38	33.08	75.14	81.22	84.62
Accuracy	95.03	96.69	11.54	93.92	91.71	53.85
Bias	98.90	100.00	73.08	98.90	100.00	76.15
Consistency	86.74	71.27	86.92	97.79	96.13	88.46
Accuracy Momentum	81.77	91.71	38.46	94.48	94.48	53.85
Bias Momentum	83.98	93.92	75.38	96.69	98.90	53.85
Consistency Momentum	100.00	86.74	78.46	99.45	91.16	92.31
Spillover Disconsensus	97.79	83.98	92.31	90.61	87.85	93.08
Spillover Accuracy	74.59	51.38	97.69	88.40	71.27	93.85
Spillover Bias	87.85	60.77	97.69	80.11	74.03	100.00
Spillover Consistency	83.43	95.03	86.15	99.45	92.27	97.69
Spillover Accuracy Momentum	40.88	60.77	97.69	58.56	39.78	87.69
Spillover Bias Momentum	79.01	43.09	96.92	88.40	75.69	97.69
Spillover Consistency Momentum	93.92	95.03	90.77	98.34	95.58	92.31
Panel B: Elastic Net Shrinkage to Zero (in %)						
Financial Dummy	71.27	52.49	52.31	85.08	69.06	90.00
Experience	57.46	53.04	63.08	61.33	53.59	79.23
Disconsensus	61.33	45.30	27.69	69.06	69.06	83.08
Accuracy	72.38	87.29	6.92	86.74	87.29	47.69
Bias	90.06	95.58	60.00	97.24	96.69	76.92
Consistency	83.43	68.51	78.46	86.74	88.40	90.77
Accuracy Momentum	64.09	74.03	30.77	86.74	85.64	43.08
Bias Momentum	69.06	69.06	73.85	90.61	91.16	46.15
Consistency Momentum	85.64	68.51	66.92	90.61	80.66	90.00
Spillover Disconsensus	83.43	68.51	87.69	82.87	81.22	91.54
Spillover Accuracy	65.75	49.17	90.00	81.22	63.54	87.69
Spillover Bias	79.56	51.38	94.62	69.06	66.85	94.62
Spillover Consistency	79.56	85.08	83.08	93.37	77.90	93.08
Spillover Accuracy Momentum	32.60	49.72	90.77	56.91	35.91	85.38
Spillover Bias Momentum	67.40	36.46	96.15	76.24	61.88	90.00
Spillover Consistency Momentum	79.56	75.69	76.92	88.40	83.98	93.08

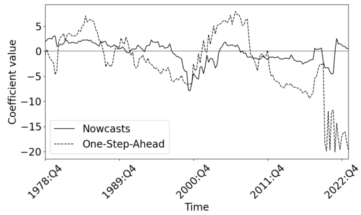
Note: This table presents the percentage of time periods during which a coefficient was set to zero by regularisation. Panel A presents the shrinkage percentages for LASSO regularisation, Panel B for Elastic Net regularisation. “GDP” denotes the GDP growth rate, “UMP” the unemployment rate and “CPI” the CPI inflation rate.

G Coefficients of Unemployment and CPI inflation over Time

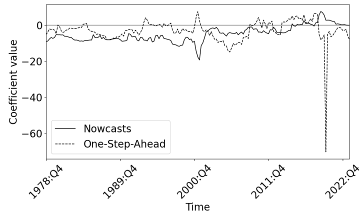
This appendix presents the coefficients over time for the unemployment and CPI inflation rate.



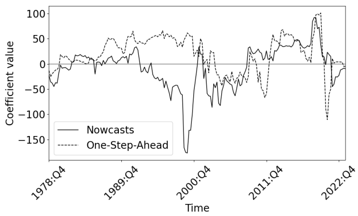
(a) Industry



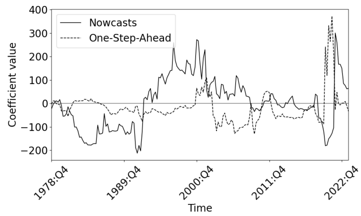
(b) Experience



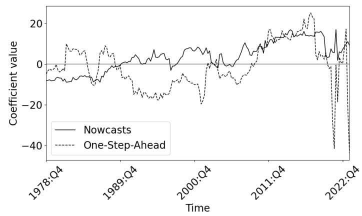
(c) Disconsensus



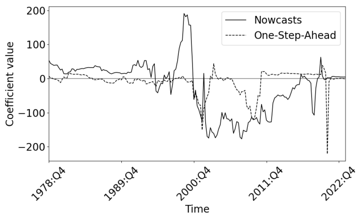
(d) Accuracy



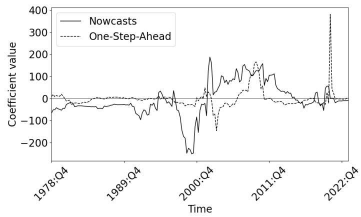
(e) Bias



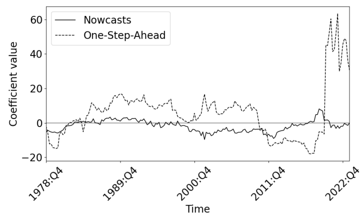
(f) Consistency



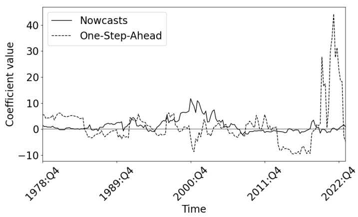
(g) Acc Momentum



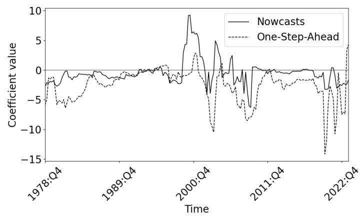
(h) Bias Momentum



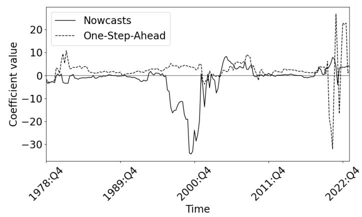
(i) Cons Momentum



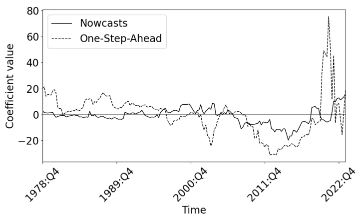
(j) Spillover Disconsensus



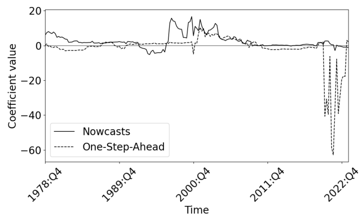
(k) Spillover Accuracy



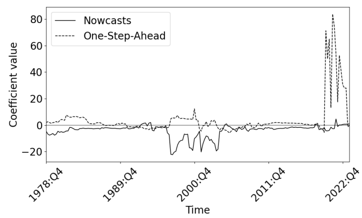
(l) Spillover Bias



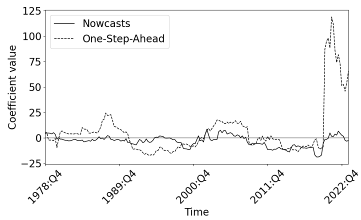
(m) Spillover Consistency



(n) Spillover Acc Momentum

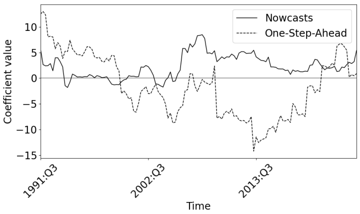


(o) Spillover Bias Momentum

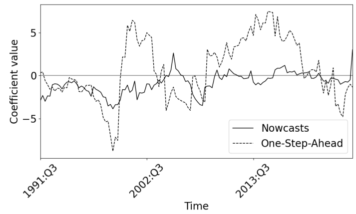


(p) Spillover Cons Momentum

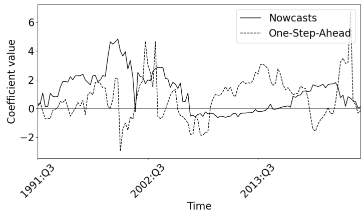
Figure G.1: Characteristic Coefficients of the Unemployment Rate.



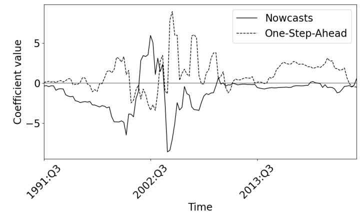
(a) Industry



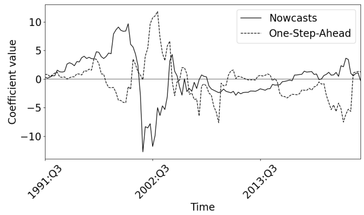
(b) Experience



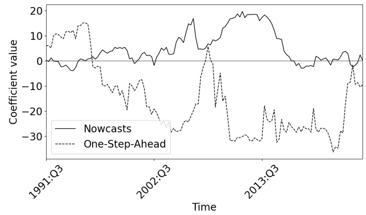
(c) Disconsensus



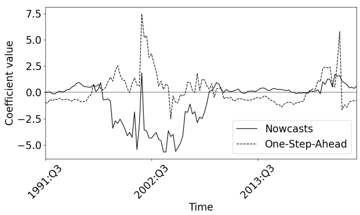
(d) Accuracy



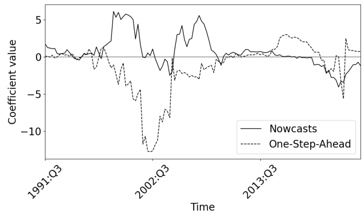
(e) Bias



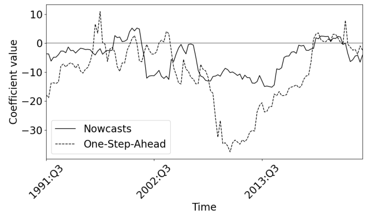
(f) Consistency



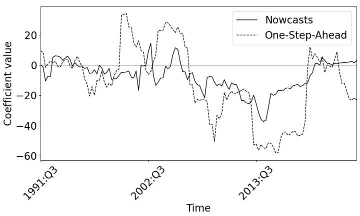
(g) Acc Momentum



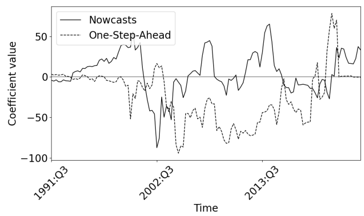
(h) Bias Momentum



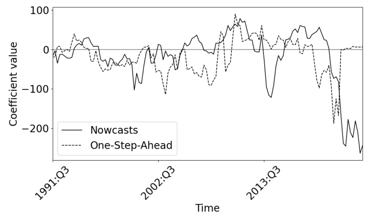
(i) Cons Momentum



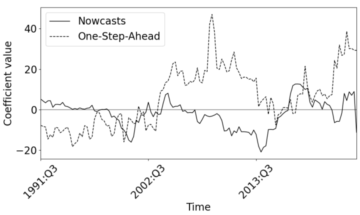
(j) Spillover Disconsensus



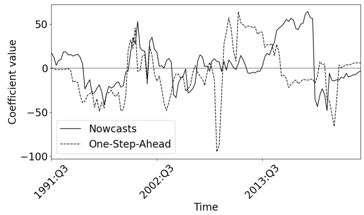
(k) Spillover Accuracy



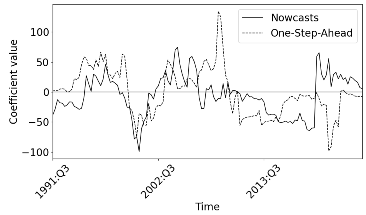
(l) Spillover Bias



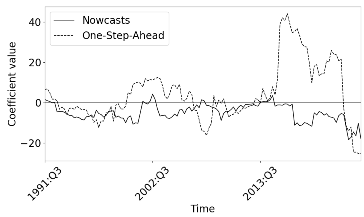
(m) Spillover Consistency



(n) Spillover Acc Momentum



(o) Spillover Bias Momentum



(p) Spillover Cons Momentum

Figure G.2: Characteristic Coefficients of the CPI Inflation Rate.

H Sensitivity Analysis Results

This appendix provides additional results of the sensitivity analysis.

Table H.1: MSFE Results without Outliers for the FCC Methods and Benchmarks

Combination Method	Nowcast			One-Step-Ahead			MCS Count
	GDP	UMP	CPI	GDP	UMP	CPI	
Panel A: FCC Methods							
FCC	***0.120***	*0.133*	***0.627***	***0.238***	*0.178*	***1.823***	4
	(9.09%)	(2.31%)	(15.47%)	(-46.15%)	(-55.16%)	(13.65%)	
FCC-Ridge	1.58	1.03	1.71	1.59	1.40	1.70	0
	(-0.53%)	(-4.86%)	(5.82%)	(-57.86%)	(-60.10%)	(-3.96%)	
FCC-LASSO	1.50	1.02	1.67	1.46	1.30	1.68	0
	(1.12%)	(3.05%)	(-2.15%)	(-62.04%)	(-63.58%)	(-16.25%)	
FCC-postLASSO	1.43	1.03	1.65	1.42	1.17	1.61	0
	(0.59%)	(3.01%)	(1.27%)	(-61.83%)	(-63.27%)	(-17.36%)	
FCC-ENet	1.52	1.02	1.70	1.48	1.33	1.76	0
	(1.11%)	(1.50%)	(1.33%)	(-61.53%)	(-63.41%)	(-14.39%)	
FCC-C-Simple	1.75	1.02	2.22	1.54	1.46	1.81	0
	(-9.48%)	(-7.48%)	(-8.53%)	(-56.12%)	(-61.34%)	(-15.30%)	
FCC-C-LASSO	1.71	1.01	2.10	1.52	1.42	1.87	0
	(-7.66%)	(-9.46%)	(-3.94%)	(-56.07%)	(-59.55%)	(-12.26%)	
FCC-C-ENet	1.75	1.02	2.23	1.54	1.46	1.81	0
	(-9.87%)	(-8.11%)	(-7.04%)	(-56.12%)	(-61.34%)	(-15.30%)	
Panel B: Benchmark Methods							
EWA	1.83	0.20	2.38	1.63	0.94	2.06	0
	(-14.45%)	(-62.32%)	(-12.84%)	(-59.39%)	(-74.96%)	(-16.21%)	
Median	1.78	0.17	2.44	1.63	0.90	2.07	0
	(-14.74%)	(-68.06%)	(-11.52%)	(-60.16%)	(-75.15%)	(-16.60%)	
Trimmed Mean	1.83	0.20	2.38	1.63	0.94	2.06	0
	(-14.73%)	(-63.38%)	(-12.53%)	(-59.43%)	(-74.46%)	(-16.10%)	
BAM	1.66	**0.17**	1.85	1.48	0.80	1.74	1
	(-13.10%)	(-41.03%)	(-4.21%)	(-58.76%)	(-73.52%)	(-18.24%)	
dMSFE ($\delta = 0.9$)	1.88	4.11	4.36	1.63	3.37	2.54	0
	(-13.08%)	(79.93%)	(6.80%)	(-53.31%)	(-15.97%)	(-13.06%)	
dMSFE ($\delta = 0.95$)	1.85	4.10	4.20	1.61	3.34	2.49	0
	(-11.55%)	(82.27%)	(9.83%)	(-55.16%)	(-17.25%)	(-11.55%)	
dMSFE ($\delta = 1.0$)	1.82	4.09	3.94	1.59	3.31	2.36	0
	(-9.54%)	(84.41%)	(11.21%)	(-57.00%)	(-17.83%)	(-11.26%)	
peLASSO	1.83	0.32	1.48	1.43	**0.74**	1.84	1
	(-16.67%)	(-2.33%)	(0.22%)	(-64.21%)	(-75.56%)	(-17.56%)	

Note: This table presents the MSFE of the FCC method and the relative MSFE of the benchmarks, estimated without outliers. For the estimation of these results, time periods 2020 Q2 and 2020 Q3 were removed for the GDP growth rate and the unemployment rate, as well as the time period 2008 Q4 for the CPI inflation rate. The percentages in parentheses represent the percentage change from the original MSFE values estimated with outliers included. The results for the FCC method are estimated with moving window estimation. The table also presents the results from the MCS procedure: the method selected into the MCS with 95% certainty is presented in bold.

Table H.2: MSFE Results with Characteristics Averaged over Forecasting Horizons

Combination Method	**Nowcast**			**One-Step-Ahead**			**MCS Count**
	GDP	**UMP**	**CPI**	**GDP**	**UMP**	**CPI**	
FCC	***0.123***	***0.130***	***0.572***	***0.556***	***0.472***	***1.433***	6
	(11.82%)	(0.00%)	(5.34%)	(25.79%)	(18.87%)	(-10.67%)	
FCC-Ridge	1.41	1.30	1.72	1.47	1.36	2.28	0
	(-8.95%)	(17.36%)	(-3.06%)	(-8.99%)	(2.56%)	(1.21%)	
FCC-LASSO	1.37	1.05	1.65	1.45	1.33	2.99	0
	(-5.06%)	(4.58%)	(-11.59%)	(-11.63%)	(-1.10%)	(17.12%)	
FCC-postLASSO	1.33	1.02	1.54	1.38	1.21	2.98	0
	(-3.53%)	(0.00%)	(-13.63%)	(-13.26%)	(0.70%)	(17.83%)	
FCC-ENet	1.40	1.08	1.73	1.41	1.32	2.97	0
	(-4.44%)	(5.26%)	(-5.52%)	(-14.10%)	(-3.57%)	(13.76%)	
FCC-C-Simple	1.85	1.15	2.71	1.51	1.41	2.75	0
	(-2.16%)	(2.04%)	(1.71%)	(0.48%)	(0.00%)	(1.34%)	
FCC-C-LASSO	1.75	1.11	2.57	1.46	1.45	2.73	0
	(-3.15%)	(-2.70%)	(5.98%)	(-1.21%)	(10.44%)	(0.46%)	
FCC-C-ENet	1.84	1.18	2.64	1.51	1.41	2.75	0
	(-2.92%)	(4.05%)	(1.50%)	(0.48%)	(0.00%)	(1.34%)	

Note: This table presents the MSFE of the FCC method and the relative MSFE of the extensions, for which the characteristics have been averaged over the horizons. The MSFE of the FCC method is presented in italics because it shows the MSFE rather than the relative MSFE. The results are estimated with moving window estimation. Below each result, the percentual difference with the original result from Table 6.1 is presented between parentheses. The table also presents the results from the MCS procedure: the method selected into the MCS with 95% certainty is presented in bold. The column "MCS Count" provides the total number of times a model is selected into the MCS. In this table, "GDP" stands for the real GDP growth rate, "UMP" for the unemployment rate and "CPI" for the CPI inflation rate.

Table H.3: MSFE Results with Characteristics with Standardised Standard Deviation

Combination Method	**Nowcast**			**One-Step-Ahead**			**MCS Count**
	GDP	**UMP**	**CPI**	**GDP**	**UMP**	**CPI**	
FCC	***0.130***	***0.133***	***0.633***	***0.425***	***0.358***	***1.880***	6
	(18.18%)	(2.31%)	(16.57%)	(-3.85%)	(-9.82%)	(17.20%)	
FCC-Ridge	1.60	1.08	1.58	2.05	1.80	2.02	0
	(9.09%)	(2.50%)	(14.70%)	(-2.78%)	(3.67%)	(2.41%)	
FCC-LASSO	1.51	1.00	1.63	2.16	1.75	2.27	0
	(11.18%)	(2.31%)	(3.80%)	(0.66%)	(-1.88%)	(16.30%)	
FCC-postLASSO	1.52	1.01	1.42	1.94	1.75	2.20	0
	(12.35%)	(3.08%)	(8.12%)	(-7.21%)	(-1.05%)	(15.96%)	
FCC-ENet	1.53	1.00	1.60	1.83	1.73	2.22	0
	(10.67%)	(2.31%)	(6.43%)	(-15.08%)	(-4.19%)	(11.48%)	
FCC-C-Simple	1.89	1.12	2.48	1.92	1.86	2.16	0
	(5.98%)	(1.36%)	(11.53%)	(-2.40%)	(-0.45%)	(10.48%)	
FCC-C-LASSO	1.85	1.10	2.18	1.91	1.97	2.11	0
	(4.93%)	(1.46%)	(12.01%)	(-0.40%)	(13.18%)	(7.63%)	
FCC-C-ENet	1.86	1.12	2.40	1.92	1.86	2.16	0
	(6.08%)	(1.54%)	(11.24%)	(-2.40%)	(-0.45%)	(10.48%)	

Note: This table presents the MSFE of the FCC method and the relative MSFE of the extensions, for which the characteristics have been standardised to standard deviation one. Note that the MSFE of the FCC method is presented in italics because it shows the MSFE rather than the relative MSFE. The results are estimated using moving window estimation. Below each result, the percentual difference with the original result from Table 6.1 is provided between parentheses. The table also presents the results from the MCS procedure: the method selected into the MCS with 95% certainty is presented in bold. The column "MCS Count" provides the total number of times a model is selected into the MCS. In this table, "GDP" stands for the real GDP growth rate, "UMP" for the unemployment rate and "CPI" for the CPI inflation rate.