Received: 19 October 2020 | Revised: 22 December 2020 | Accepted: 30 January 2021

DOI: 10.1002/for.2768

RESEARCH ARTICLE

WILEY

Agricultural commodity price dynamics and their determinants: A comprehensive econometric approach

Jesus Crespo Cuaresma[1,2,3,4,5] | Jaroslava Hlouskova[1,6,7] | Michael Obersteiner[1,8]

[1]International Institute for Applied Systems Analysis (IIASA), Laxenburg, Austria

[2]Department of Economics, Vienna University of Economics and Business (WU), Vienna, Austria

[3]Wittgenstein Centre for Demography and Global Human Capital, IIASA, VID/ÖAW, University of Vienna, Vienna, Austria

[4]Austrian Institute of Economic Research (WIFO), Wien, Austria

[5]CESifo, Munich, Germany

[6]Institute for Advanced Studies (IHS), Vienna, Austria

[7]Department of Economics, Faculty of National Economy, University of Economics in Bratislava, Bratislava, Slovakia

[8]Environmental Change Institute, School of Geography and Environment, University of Oxford, Oxford, UK

Correspondence
Jesus Crespo Cuaresma, International Institute for Applied Systems Analysis (IIASA), Laxenburg, Austria.
Email: jcrespo@wu.ac.at

Funding information
H2020 Food, Grant/Award Number: 633692

Abstract

We present a comprehensive modelling framework aimed at quantifying the response of agricultural commodity prices to changes in their potential determinants. The problem of model uncertainty is assessed explicitly by concentrating on specification selection based on the quality of short-term out-of-sample forecasts (1 to 12 months ahead) for the price of wheat, soybeans and corn. Univariate and multivariate autoregressive models (autoregressive [AR], vector autoregressive [VAR] and vector error correction [VEC] specifications, estimated using frequentist and Bayesian methods), specifications with heteroskedastic errors (AR conditional heteroskedastic [ARCH] and generalized AR conditional heteroskedastic [GARCH] models) and combinations of these are entertained, including information about market fundamentals, macroeconomic and financial developments, and climatic variables. In addition, we assess potential non-linearities in the commodity price dynamics along the business cycle. Our results indicate that variables measuring market fundamentals and macroeconomic developments (and, to a lesser extent, financial developments) contain systematic predictive information for out-of-sample forecasting of commodity prices and that agricultural commodity prices react robustly to shocks in international competitiveness, as measured by changes in the real exchange rate.

KEYWORDS

commodity prices, forecast averaging, forecasting, model uncertainty, vector autoregressive models

1 | INTRODUCTION

Due to their underlying consistency with theory, univariate and multivariate time series specifications are often employed to evaluate the determinants of agricultural commodity prices and provide out-of-sample forecasts. Theoretical models of agricultural commodity prices based on fundamentals such as those in Deaton and Laroque (1992, 2003) predict autoregressive (AR) dynamics in price behaviour and lagged adjustment to deviations from the supply–demand equilibrium, thus justifying representations in the form of time series models (see also Ahumada & Cornejo, 2015). Although early modelling efforts aimed at forecasting agricultural commodity prices relied on large-scale econometric models (see, e.g., Just & Rausser, 1981), the use of time series models dominates the current research frontier. Recent contributions to the literature on modelling agricultural commodity prices propose the inclusion of macroeconomic variables in econometric models aimed at forecasting the price of agricultural commodities. The information contained in macroeconomic variables has been shown to improve the predictive ability of econometric specifications for commodity prices (see Gargano & Timmermann, 2014, for instance). Gargano and Timmermann (2014) present evidence that the predictability of commodity prices depends on the state of the economy and that the information contained in macroeconomic variables improves forecast accuracy in models of commodity prices. Chen et al. (2010), on the other hand, show that the inclusion of information on 'commodity currency' exchange rates in econometric models for global commodity prices robustly improves their predictive power. The results in Husain and Bowman (2004), based on the analysis of 15 different commodities, show that statistical models based on futures tend to achieve better predictive ability than those based exclusively on spot price dynamics or on judgement.

The comovement observed across the prices of different agricultural commodities implies that the cross-correlation of price changes can be exploited for prediction. Ahumada and Cornejo (2016) show that the strong correlation observed in the prices of corn, soybeans and wheat can be utilized to significantly improve forecasting accuracy in time series specifications. In addition to univariate and multivariate time series models, the literature on commodity price forecasting has also employed artificial neural networks (Kohzadi et al., 1996), models aimed at modelling the dynamics of the second moment of the commodity price time series (Bernard et al., 2008) for prediction and model averaging schemes (Drachal, 2019). In particular, models that account for the particular dynamics of the volatility of agricultural commodity prices appear to robustly improve probabilistic forecasts of price changes, as shown in Ramirez and Fadiga (2003). The importance of modelling changes in the variance of agricultural commodity prices to improve predictive power has led to research efforts aimed at optimizing the specification of their volatility dynamics (see, e.g., the recent contribution by Degiannakis et al., 2020).

The contributions reviewed above give evidence of the complexity implied in building econometric models for predicting agricultural commodity prices. Furthermore, the specificities of different commodity markets speak against a *one-size-fits-all* modelling framework that can be applied to different agricultural commodities (Brooks & Prokopczuk, 2013). In this contribution, we present a comprehensive econometric framework based on a large battery of univariate and multivariate time series models, as well as their combination, aimed at identifying the determinants of short-term and medium-term dynamics of agricultural commodity prices. Our study contributes to the literature on modelling and forecasting commodity prices in several respects. On the one hand, we present predictive ability results from the most comprehensive set of commodity forecasting models entertained hitherto with respect to different performance measures (both loss-based and profit-based indicators). In addition to homoskedastic and heteroskedastic univariate time series models, we include vector AR (VAR) and vector error correction (VEC) specifications aimed at exploiting the relationship between commodity price forecasting and macroeconomic, financial and climatic variables. On the other hand, we also assess explicitly the issue of specification uncertainty by assessing the potential improvements in predictive ability that can be gained from forecast averaging methods. Similar forecast combination techniques (albeit with less model types than those in this contribution) have been proposed in the literature on exchange rate forecasting (Costantini et al., 2016) and coffee price forecasting (Crespo Cuaresma et al., 2018). The analysis expands the work in Crespo Cuaresma et al. (2018) and is applied to wheat, soybeans and corn spot prices.[1] We also analyse the forecasting performance of our battery of commodity price models with respect to different states of the economy and examine the responses of agricultural prices to their determinants. Our results indicate that in addition to market fundamentals and macroeconomic variables, especially international competitiveness, as measured by real effective exchange rate (REER) (and, to a lesser extent, financial variables) appear important to improve out-of-sample predictive ability. VEC specifications tend to be promising individual model structures when it comes to predicting commodity prices, whereas forecast combination methods do not appear to improve predictive ability significantly.

Our results concerning the response of agricultural commodity prices to their determinants highlight the role played by real exchange rate dynamics and market fundamentals, as well as by global macroeconomic and financial shocks as captured by US industrial production and stock market fluctuations. We find scattered

[1]The analysis in Drachal (2019) is also applied to the prices of these particular agricultural commodities.

evidence for differences in out-of-sample predictive performance across states of the business cycle.

The structure of the paper is as follows. Section 2 presents the econometric framework of the model selection and averaging exercise and describes the specifications and methods employed in the analysis. It also presents the different predictive accuracy measures used in the comparison across model specifications. Section 3 presents the results for the different commodities and assesses potential non-linearities in the dynamics along the business cycle. Section 4 highlights the responses of agricultural prices to their determinants implied by the best models in our pool of specifications, and Section 5 concludes the study.

2 | THE ECONOMETRIC FRAMEWORK AND MEASURING PREDICTIVE ACCURACY OF COMMODITY PRICE MODELS

In this section, we provide a description of the models employed to analyse the predictability of agricultural commodity prices, as well as their reaction to selected determinants. We include univariate and multivariate model structures where the corresponding commodity price is assumed to depend on its own past values and, in the case of the multivariate models, on past values of other potential determinants.

2.1 | Models

Within the class of linear univariate time series models, we consider AR models, where the price of commodity m (in our case, alternatively wheat, soybeans or corn) is assumed to depend on its own k lags and a random white noise shocks. Denoting $P_{m,\,t}$ as the price of commodity m at time t, this implies that

$$P_{m,t} = \phi_0 + \sum_{i=1}^{k} \phi_i P_{m,t-i} + \epsilon_t, \epsilon_t \sim \mathbf{IID}(\mathbf{0}, \sigma), \tag{1}$$

where the error term ϵ_t is assumed to be white noise. Alternatively, AR models in first differences are also considered as potential model structures for the commodity price variable, leading to specifications, which are given by

$$\Delta P_{m,t} = \theta_0 + \sum_{i=1}^{k} \theta_i \Delta P_{m,t-i} + \rho_t, \qquad \rho_t \sim \mathbf{IID}(\mathbf{0}, \sigma_\rho), \tag{2}$$

with ρ_t being a white noise error.

In addition to univariate models with homoskedastic shocks, such as those given by Equations (1) and (2), we also entertain specifications whose dynamics are driven by heteroskedastic disturbances in the form of AR conditional heteroskedastic (ARCH) and generalized AR conditional heteroskedastic (GARCH) errors. This implies that the variance of the shock in the respective model is assumed to change over time following the specification

$$\sigma_t^2 = \nu_0 + \sum_{h=1}^{q} \psi_h \sigma_{t-h}^2 + \sum_{g=1}^{q} \psi_g \epsilon_{t-g}^2, \tag{3}$$

where ϵ_t is the corresponding error term of the specification.

Linear multivariate time series specifications are also entertained in our analysis. In such specifications, we consider $P_{m,\,t}$ as an element of the vector x_t, which includes other fundamental, macroeconomic, financial or climatic variables. The vector x_t is assumed to depend on its past values and on a multivariate random shock, so that

$$x_t = \Psi_0 + \sum_{l=1}^{p} \Psi_l x_{t-l} + \varepsilon_t, \qquad \varepsilon_t \sim \mathbf{IID}(\mathbf{0}, \Sigma_\varepsilon), \tag{4}$$

where Ψ_l for $l = 1,...,p$ are matrices of coefficients and Ψ_0 is a vector of intercept terms. Instead of assuming a relationship in levels, it can be assumed that the linear linkage is among first differences of the variables, so that the corresponding model would be given by

$$\Delta x_t = \chi_0 + \sum_{l=1}^{p} \chi_l \Delta x_{t-l} + \mu_t, \qquad \mu_t \sim \mathbf{IID}(\mathbf{0}, \Sigma_\mu). \tag{5}$$

Alternatively, if the elements of x_t are integrated of order one and linked by a cointegration relationship, the corresponding VEC representation is given by

$$\Delta x_t = \delta_0 + \alpha\beta' x_{t-1} + \sum_{l=1}^{p} \delta_l \Delta x_{t-l} + u_t, \qquad u_t \sim \mathbf{IID}(\mathbf{0}, \Sigma_u). \tag{6}$$

A summary of all models used in this study can be found in Table A2.

2.2 | Forecasts, forecast averaging methods and forecast performance measures

Specification uncertainty is addressed by using model averaging methods based on both frequentist and Bayesian methods. For this purpose, we employ a large number of forecast combination techniques that have been put forward in Costantini et al. (2016) and Crespo Cuaresma et al. (2018). Thus, combinations of out-of-sample predictions of individual specifications within these classes of univariate and multivariate models are considered in addition to the forecasts of each individual model. Pooled forecasts, $\hat{P}_{m,c,t+h|t}$, take the form of the linear combination of the predictions generated by individual specifications

$$\hat{P}_{m,c,t+h|t} = w^h_{m,c,0t} + \sum_{i=1}^{M} w^h_{m,c,it} \hat{P}_{m,i,t+h|t}, \tag{7}$$

where $\hat{P}_{m,i,t+h|t}$ is the forecast of the price of commodity m, using model i ($i = 1,...,M$, with M being the number of forecasts from individual models) for time $t+h$ conditional on the information available at time t, c is the combination method and the weights are given by $\{w^h_{m,c,it}\}^M_{i=0}$. Each one of the forecast averaging methods used in this study employs different weights. Table A3 presents the definition of the weights corresponding to each one of the methods entertained.[2] The forecast averaging methods employed to aggregate out-of-sample predictions from the individual specifications entertained in our application are described in more detail in Costantini et al. (2016) and Crespo Cuaresma et al. (2018).

We evaluate the forecasts of commodity spot prices using performance measures based on both loss minimization and profit maximization. The former build upon continuous predictive error measures, whereas the latter rely on conceptual settings using trading strategies in the commodity market, which build upon the predictions of the individual models and their combinations. We consider traditional loss-based measures such as mean-square error (MSE), mean absolute error (MAE) and profit-based measures such as directional accuracy (DA), directional value (DV), the returns from ‘buy low, sell high’ trading strategy generated by our forecasts and a risk-adjusted performance measure given by the Sharpe ratio. All forecast performance measures under consideration are listed in Table 1. To simplify the notation, we omit the index m, which describes the commodity, and the index c, which describes the model or the forecast combination method used to obtain the prediction.

In addition, we also entertain composite forecasts based on the relative performance of predictions from all models over certain out-of-sample periods. In particular, for this technique at each time point t, we choose the model or forecast combination method (and thus also the forecast for time point $t + h$) with the best performance over a certain time window ending at time point t. A summary of the composite forecasts employed, which build upon to the last k observations, can be found in Table 2.

3 | EMPIRICAL RESULTS: PREDICTING COMMODITY PRICES

The lag length of all multivariate model specifications under consideration is selected using the Akaike information criterion (AIC) for potential lag lengths ranging from one to six lags. For the VEC models, selection of the lag length and the number of cointegration relationships is carried out simultaneously using the AIC. Because VAR models are known to forecast poorly due to overfitting, we also estimate subset-VAR specifications, where individual parameters of the VAR specification are set equal to zero using recursive t testing. The same specifications were used for all the commodities analysed in this study. Our dataset on monthly observations of all commodity prices and their potential predictors spans the period from January 1980 to December 2016. The choice of variables included in our models strikes a balance between data availability and covering determinants related to the

TABLE 1 Predictive performance measures

Performance measure	Description
	Loss-based performance measures
Mean-square error (MSE)	$\frac{1}{T_3-T_2+1}\sum_{j=0}^{T_3-T_2}\left(\hat{P}_{T_2+j\|T_2+j-h}-P_{T_2+j}\right)^2$
Mean absolute error (MAE)	$\frac{1}{T_3-T_2+1}\sum_{j=0}^{T_3-T_2}\left\|\hat{P}_{T_2+j\|T_2+j-h}-P_{T_2+j}\right\|$
	Profit-based performance measures
Directional accuracy (DA)	$100\sum_{j=0}^{T_3-T_2}\frac{DA_{T_2+j,h}}{T_3-T_2+1}$, where
	$DA_{t,h}=I\left(\text{sgn}(P_t-P_{t-h})=\text{sgn}(\hat{P}_{t\|t-h}-P_{t-h})\right)$
Directional value (DV)	$100\frac{\sum_{j=0}^{T_3-T_2}\left\|\hat{P}_{T_2+j\|T_2+j-h}-P_{T_2+j-h}\right\|DA_{T_2+j,h}}{\sum_{j=0}^{T_3-T_2}\left\|P_{T_2+j}-P_{T_2+j-h}\right\|}$
Total return (return)	$\frac{1}{h}\sum_{j=0}^{h-1}\Pi_{i=0}^{n_j}\left(R_{t+j+ih,h}+1\right)-1$, where
of the trading strategy	
over period [t, $t+n$], $n \geq h$	$R_{t,h}=1-\frac{P_t}{P_{t-h}}$, if $\hat{P}_{t\|t-h}<P_{t-h}$
	$R_{t,h}=\frac{P_t}{P_{t-h}}-1$, if $\hat{P}_{t\|t-h}>P_{t-h}$
	n_j, $j=1,...,h-1$, is the largest integer such that $t+j+n_jh\leq n$
Sharpe ratio	Return per unit of standard deviation implied by the trading strategy

Note: $I(\cdot)$ is the indicator function, T_2 is the beginning of the out-of-sample period and T_3 is the end of the data sample. The trading strategy under consideration is simple ‘buy low, sell high’ trading strategy, that is, buying the commodity if its price is forecast to rise and selling it when its price is forecast to fall.

TABLE 2 Composite forecasts based on the last k time periods

Composite forecasts with respect to	Description
Minimum MSE	$\hat{P}_{c_{kth}^{MSE},t+h\|t}$, where $c_{kth}^{MSE} = \text{argmin}_c \sum_{j=t-k+1}^{t} \left(\hat{P}_{c,j\|j-h} - P_j\right)^2$
Minimum MAE	$\hat{P}_{c_{kth}^{MAE},t+h\|t}$, where $c_{kth}^{MAE} = \text{argmin}_c \sum_{j=t-k+1}^{t} \left\|\hat{P}_{c,j\|j-h} - P_j\right\|$
Maximum DA	$\hat{P}_{c_{kth}^{DA},t+h\|t}$, where $c_{kth}^{DA} = \text{argmax}_c \sum_{j=t-k+1}^{t} DA_{cjh}$
	where $DA_{cjh} = I\left(\text{sgn}(P_j - P_{j-h}) = \text{sgn}(\hat{P}_{c,j\|j-h} - P_{j-h})\right)$
Maximum DV	$\hat{P}_{c_{kth}^{DV},t+h\|t}$, where $c_{kth}^{DV} = \text{argmax}_c \sum_{j=t-k+1}^{t} \left\|P_j - P_{j-h}\right\| DA_{cjh}$
Maximum return	$\hat{P}_{c_{kth}^{return},t+h\|t}$, where $c_{kth}^{return} = \text{argmax}_c \sum_{j=t-k+1}^{t} R_{c,j,h}$

Abbreviations: DA, directional accuracy; DV, directional value; MAE, mean absolute error; MSE, mean-square error.

categories (i) market fundamentals, (ii) macroeconomic factors, (iii) financial variables and (iv) determinants related to climatic variability. The covariates correspond to predictors used in the literature aimed at modelling and forecasting agricultural commodity prices (see, e.g., the battery of variables employed in Drachal, 2019). For the category of market fundamentals, they include total production figures at the global level and for the most important producers for each one of the grains, yields and stock-to-use ratio. The group of macroeconomic variables includes global gross domestic product (GDP) and industrial production indices, as well as business climate indices and the REER. The category of financial variables contains information on stock market indices and interest rate spreads, whereas the group of climatic variables is given by indices of temperature anomalies and air pressure fluctuations. These variables cover the most important drivers highlighted in the literature on agricultural price prediction (see Drachal, 2019; Rezitis & Sassi, 2013, for discussions on the choice of variables to proxy the channels of price determination in agricultural markets).[3] All of the variables included in our models have nonstationary features. We entertain VAR specifications in levels, which can be interpreted as the VAR representation of VEC models, as well as VAR models with first-differenced covariates, so all of our multivariate models are balanced in terms of order of integration.

The beginning of the holdout forecasting sample for individual models used in order to obtain weights based on predictive accuracy is chosen to be January 2000 (T_1). The beginning of the actual out-of-sample forecasting period is January 2005 (T_2), and the end of the data sample is December 2016 (T_3). We consider rolling-window estimation for our analysis: we keep the size of the estimation sample constant (equal to 240 observations, i.e., 20 years) and move forward the sample by 1 month, as we reestimate the model parameters.[4] 'Best' models are chosen on the basis of the individual forecast performance of all single models and all combinations of variables under consideration. Forecasts of all forecast averaging methods are calculated from forecasts of these single models with respect to all combination of variables. In addition, we also present the results of the Diebold–Mariano test of equal forecasting accuracy against the benchmark random walk model (Diebold & Mariano, 1995).

3.1 | Wheat price

After corn and rice, wheat is the crop with the highest production worldwide. Its starch content is about 70% and thus can be converted to ethanol. In Europe, wheat is currently the main starch crop for ethanol production. The following variables are employed to summarize information in each one of the categories proposed (fundamental, macroeconomic, financial and climatic): (i) we use as *fundamental variables* wheat production for world, y_w^{world}, wheat production for European Union (EU), y_w^{EU}, wheat production for the United States, y_w^{US}, wheat yield for world, $yield_w^{world}$, wheat yield for EU, $yield_w^{EU}$, wheat yield for the United States, $yield_w^{US}$, and world stock-to-use ratio for wheat,[5] $stock_w^{world}$; (ii) as *macroeconomic variables*, we use world output, y^{world}, output for the EU, y^{EU}, output for the United States, y^{US}, a leading indicator for Germany, li^{EU}, a leading indicator for the United States, li^{US}, and the REER, with respect to the USD (as the USD

[3]Fundamental variables (production, yield and stock-to-use ratio) are available on annual basis. We used linear interpolation to obtain information at the monthly frequency. All macroeconomic variables used are logged, with the exception of interest rate spreads.

[4]The sources for all variables used are given in Appendix C1.

[5]Ending stocks over consumption. Inventory stock levels play an important role in commodity pricing as the market-clearing price is determined not only by current production and consumption but also by changes in inventory holdings. Existing stocks are thus a fundamental source of stability in commodity markets.

is the main currency in which the wheat and most commodities are traded); (iii) as *financial variables*, we employ the stock market index for the EU, $stock^{EU}$, the stock market index for the United States, $stock^{US}$, the S&P Goldman Sachs commodity index, GSCI, the interest rate spread for the EU, $spread^{EU}$, and the interest rate spread for the United States, $spread^{US}$ (the difference between 10-year interest rate on government bonds and 3-month interbank rates); and (iv) as *climatic variables*, we use sea surface temperature anomalies (the index measuring deviations between the sea surface temperatures in the El Niño region 3.4 and its historical average), SSTA, and Southern Oscillation Index anomalies (the index capturing fluctuations in air pressure occurring between the western and eastern tropical Pacific during El Niño and La Niña episodes), SOI. As the EU and the United States are the main producers of wheat, we include fundamental, macroeconomic and financial variables related to these two world regions.[6]

We start by performing the forecast analysis for each of these groups of variables separately. For all combinations of variables within a thematic group, all models/methods are evaluated on the basis of both the loss and profit-based measures introduced in Section 2. Tables 3 and 4 summarize the best models and corresponding results for each individual group and forecast horizons of 1, 3, 6, 9 and 12 months. Comparing the results over all forecasting horizons used in the analysis, specifications based on fundamental market variables for wheat tend to present the best predictive performance for loss measures at the 1-month-ahead forecasting horizon. Specifications based on macroeconomic variables, on the other hand, lead to the best performance in terms of profit measures, achieving the highest outcome for the DA and DV statistics at 1-year-ahead forecasts and for the return at 1-month-ahead predictions.

The importance of macroeconomic variables in terms of improving the predictive content of long-term forecasts in multivariate time series models for wheat prices is evident when assessing the results of the forecasting exercise for long predictive horizons presented in Table 4. For the 12-month-ahead forecast horizon, models including macroeconomic variables beat their counterparts over all loss and profit forecast accuracy measures, whereas the same is true for all measures but the MSE in the case of a 9-month-ahead forecasting horizon. Specifications with macroeconomic covariates dominate other models in terms of loss measured by MAE for all forecasting horizons beyond 1 month, and at the horizon of 1 month, the models with macroeconomic variables perform best in terms of profit measures. The particular specifications with the best performance for models of this group tend to include REER and output variables.

Besides the set of models including macroeconomic information, only specifications with predictors related to market fundamentals present a good record of forecasting power for wheat prices. These models beat those of other variable groups in terms of MSE for all horizons with the exception of the furthest one (1 year ahead) and perform particularly well at horizons of 3 and 6 months ahead. For the groups of financial and climatic variables, univariate specifications tend to present the best performance, thus providing evidence that the predictive content of variables in these two groups is very limited.

Making use of specifications based on combinations of the variables, which present best performance as predictors, we perform the complete forecasting exercise for wheat prices, where we mix variables across categories. The results of the forecasting comparison are presented in Table 5. The good performance of VEC specifications can already be observed in the group-specific results in Tables 3 and 4. The same VEC model that delivers the best results in the group results tends to achieve the best performance when considering variables from all categories. This specification combines information of wheat production and yields, as well as the stock-to-use ratio and the REER, which is part of all best models over all forecast horizons and all performance measures. The results of the Diebold–Mariano test indicate that the vast majority of these models significantly outperform the random walk model in terms of predictive ability.

3.2 | Soybean price

The soybean has become one of the most important beans globally as it provides oil and protein around the world. With applications as diverse as vegetable oil, animal feed and foodstuffs, soybeans have become staples in countries far from its original roots in Eastern Asia. Fifty-five per cent of the world's soybean production occurs in the Americas. We again employ four thematic groups of variables: fundamental, macroeconomic, financial and climatic variables. Financial and climatic variables remain the same as in the case of the wheat, whereas as *fundamental variables*, we use soybean production for world, y_s^{world}, soybean production for the United States, y_s^{US}, soybean production for Brazil, y_s^{BR}, soybean yield for world, $yield_s^{world}$, soybean yield for the

[6]We also conducted forecasting exercises including futures prices as a variable among the set of explanatory covariates. However, even when they were included in the best models, the forecast performance of these models did not improve in the long run. The results based on these additional models are available from the authors upon request.

TABLE 3 Summary of forecast performance of best models for wheat over different variable groups (fundamental, macroeconomic, financial and climatic) for 1-, 3- and 6-month horizons

	MAE	MSE	DA	DV	Return	Sharpe ratio
1-month horizon						
Fundamental	**7.800**** VAR(2) y_w^{US} y_w^{EU} $stock_w^{world}$	**154.244** s-VAR(2) y_w^{US} y_w^{EU} $yield_w^{EU}$	**66.667***** DVAR(1) $yield_w^{US}$ $stock_w^{world}$	78.795** VEC(1,2) y_w^{US} y_w^{EU} $yield_w^{world}$ $yield_w^{US}$ $yield_w^{EU}$	31.523 MAE whole	0.433 MAE whole
Macroeconomic	7.962** VEC(1,1) y^{EU} REER	157.739** VEC(2,1) y^{US} REER	**66.667***** DVAR(1) y^{US} y^{world} li^{EU} REER	**81.221***** VEC(2,1) y^{US} REER	**33.696***** VEC(2,1) y^{US} REER	**0.464***** VEC(2,1) y^{EU} REER
Financial	8.115*** DVAR(1) $stock^{US}$ GSCI	162.302*** ARCH(3,3)	**66.667***** s-DVAR(1) $stock^{US}$	78.873*** s-DVAR(1) $stock^{EU}$ $stock^{US}$ $spread^{US}$	31.154*** s-DVAR(1) $stock^{US}$	0.427*** s-DVAR(1) $stock^{US}$
Climatic	8.146 GARCH(2,2)	161.309** whole	64.583*** median	78.577*** whole	29.801*** VEC(2,2) SOI SSTA	0.408** VEC(2,2) SOI SSTA
3-month horizon						
Fundamental	19.083 VEC(1,1) y_w^{EU} $stock_w^{world}$	**776.139** DVAR(6) y_w^{world} y_w^{US} $yield_w^{world}$ $yield_w^{US}$	**66.667**** VEC(1,1) y_w^{EU} $stock_w^{world}$	**77.159** DVAR(1) y_w^{EU} $yield_w^{world}$ $yield_w^{US}$ $stock_w^{world}$	**22.196***** VEC(1,2) y_w^{world} y_w^{EU} $yield_w^{world}$ $yield_w^{EU}$ $stock_w^{world}$	**0.437**** VEC(1,2) y_w^{world} y_w^{EU} $yield_w^{world}$ $yield_w^{EU}$ $stock_w^{world}$
Macroeconomic	**18.733*** VEC(1,1) y^{EU} REER	839.468 VEC(2,1) y^{US} REER	**66.667***** VEC(2,1) y^{US} REER	74.838** VEC(2,1) y^{US} REER	21.952** VEC(2,1) y^{US} REER	0.432** VEC(2,1) y^{US} REER
Financial	19.755*** DARCH(2,6)	857.303*** DARCH(2,6)	61.702 last 3 months	71.729*** last 3 months	15.681*** BDVAR(1) $stock^{US}$ $spread^{EU}$	0.302*** BDVAR(1) $stock^{US}$ $spread^{EU}$
Climatic	19.252 EHR	878.954 DARCH(2,3)	65.957** DMSFE	73.531 DARCH(2,3)	14.053** DARCH(2,3)	0.337 DARCH(2,3)

(Continues)

TABLE 3 (Continued)

	MAE	MSE	DA	DV	Return	Sharpe ratio
6-month horizon						
Fundamental	27.025***	**1495.214**	**75.694*****	**82.655*****	**21.920*****	**0.602****
	VEC(1,2)	VEC(6,1)	VEC(1,2)	VEC(1,2)	VEC(1,2)	VEC(1,2)
	y_w^{US}	y_w^{US}	y_w^{world}	y_w^{world}	y_w^{world}	y_w^{world}
	$yield_w^{world}$	$yield_w^{US}$	y_w^{EU}	y_w^{EU}	y_w^{EU}	y_w^{EU}
	$yield_w^{US}$	$yield_w^{EU}$	$yield_w^{world}$	$yield_w^{world}$	$yield_w^{world}$	$yield_w^{world}$
	$yield_w^{EU}$		$yield_w^{EU}$	$yield_w^{EU}$	$yield_w^{EU}$	$yield_w^{EU}$
			$stock_w^{world}$	$stock_w^{world}$	$stock_w^{world}$	$stock_w^{world}$
Macroeconomic	**26.355*****	1501.842**	74.306***	82.126***	21.489***	0.587*
	VEC(1,1)	VEC(2,1)	VEC(1,1)	VEC(2,1)	VEC(2,1)	VEC(2,1)
	y^{EU}	y^{US}	y^{EU}	y^{US}	y^{US}	y^{US}
	REER	REER	REER	REER	REER	REER
Financial	29.739**	1698.854	66.667*	67.554	13.739**	0.348*
	DARCH(2,6)	DGARCH(4,2)	DGARCH(2,2)	DVAR(1)	last 6 months, MAE	last 6 months, MAE
				$stock^{EU}$		
				$stock^{US}$		
				GSCI		
				$spread^{US}$		
Climatic	29.450*	1736.581	66.667**	71.303*	15.325	0.393
	DGARCH(2,2)	DARCH(2,3)	DGARCH(2,2)	last month	last 3 months, MSE	last 3 months, MSE

Note: Bold figures in black indicate the best performance among all groups but within certain forecast horizon, and bold figures in red indicate the best performance among all groups and forecast horizons. See Table A2 for the abbreviation of the models.
*Indicates rejection of the null hypothesis of equal forecasting accuracy at 10%.
**Indicates rejection of the null hypothesis of equal forecasting accuracy at 5%.
***Indicates rejection of the null hypothesis of equal forecasting accuracy at 1%.

United States, $yield_s^{US}$, soybean yield for Brazil, $yield_s^{BR}$, and world stock-to-use ratio for soybean, $stock_s^{world}$. Finally, as *macroeconomic variables*, we use world output, y^{world}, output for the United States, y^{US}, output for Brazil, y^{BR}, output for the EU, y^{EU}, a leading indicator for Germany, li^{EU}, a leading indicator for the United States, li^{US}, and the REER, with respect to the USD. As the United States is the biggest producer of soybeans (with 117,208,000 metric tons in February 2017), followed by Brazil (with 104,000,000 metric tons in February 2017), we have focused on the fundamental variables for these two countries as well as the world.

Tables 6 and 7 show the results of the forecast analysis for soybean prices by variable category group. Models based on macroeconomic variables dominate in terms of predictive performance for practically all loss and profit measures and for all forecasting horizons. Within this category of models, a single specification (a VEC model including output variables for the United States and Brazil as well as the REER) appears as the most successful tool for predicting soybean prices for 3, 6 and 9 months ahead. Macroeconomic leading indicator variables also appear in the specification of the best models for 1-month-ahead predictions. Models based on fundamental variables perform poorly for predictive horizons beyond 1 month ahead, with univariate models such as the random walk or AR specifications in first differences beating multivariate specifications in terms of loss measures. The specifications with the best performance among those using market fundamentals as covariates tend to include the global stock-to-use ratio as a variable. Models based on financial or climatic variables tend to perform poorly compared with those of the other two categories. Univariate specifications tend to beat the best multivariate specifications based on loss measures of predictive performance.

When evaluating forecast performance based on combination of variables that give the best performance as predictors, Table 8 shows that the VEC model was systematically picked as the best model in

TABLE 4 Summary of forecast performance of best models for wheat over different variable groups (fundamental, macroeconomic, financial and climatic) for 9- and 12-month horizons

	MAE	MSE	DA	DV	Return	Sharpe ratio
9-month horizon						
Fundamental	35.750^{**} VEC(1,1) y_w^{EU} $stock_w^{world}$	**2537.179** VEC(6,1) y_w^{US} $yield_w^{US}$ $yield_w^{EU}$	71.528^{***} VEC(1,2) y_w^{world} y_w^{EU} $yield_w^{world}$ $yield_w^{EU}$ $stock_w^{world}$	78.586^{***} VEC(1,2) y_w^{world} y_w^{EU} $yield_w^{world}$ $yield_w^{EU}$ $stock_w^{world}$	17.655^{***} VEC(1,2) y_w^{world} y_w^{EU} $yield_w^{world}$ $yield_w^{EU}$ $stock_w^{world}$	0.512 VEC(1,2) y_w^{world} y_w^{EU} $yield_w^{world}$ $yield_w^{EU}$ $stock_w^{world}$
Macroeconomic	**35.577***** VEC(2,1) y^{US} REER	2616.299^{***} VEC(2,1) y^{US} REER	**74.306***** VEC(2,1) y^{US} REER	**84.095***** VEC(2,1) y^{US} REER	**20.784***** VEC(2,1) y^{US} REER	**0.634** VEC(2,1) y^{US} REER
Financial	39.574^{**} DARCH(2,6)	3076.385 DGARCH(4,2)	63.889^{*} DGARCH(2,2)	66.941 DGARCH(2,2)	11.846^{**} BDVAR(4) $stock^{US}$ GSCI $spread^{EU}$ $spread^{US}$	0.324 BDVAR(4) $stock^{US}$ GSCI $spread^{EU}$ $spread^{US}$
Climatic	39.309 VAR(2) SOI	3104.981 DARCH(2,3)	63.889^{*} DGARCH(2,2)	66.941 DGARCH(2,2)	12.030 VEC(1,1) SOI	0.329 VEC(1,1) SOI
12-month horizon						
Fundamental	43.733^{***} VEC(1,1) $yield_w^{US}$ $yield_w^{EU}$ $stock_w^{world}$	3524.100 VEC(6,1) y_w^{US} $yield_w^{US}$ $yield_w^{EU}$	72.222^{**} VEC(1,2) y_w^{EU} $yield_w^{US}$ $yield_w^{EU}$ $stock_w^{world}$	75.010^{**} VEC(1,2) y_w^{world} y_w^{EU} $yield_w^{world}$ $yield_w^{EU}$ $stock_w^{world}$	14.297^{**} VEC(1,2) y_w^{world} y_w^{EU} $yield_w^{world}$ $yield_w^{EU}$ $stock_w^{world}$	0.471 VEC(1,2) y_w^{world} y_w^{EU} $yield_w^{world}$ $yield_w^{EU}$ $stock_w^{world}$
Macroeconomic	**43.182***** VEC(2,1) y^{US} REER	**3407.905***** VEC(2,1) y^{US} REER	**77.778***** VEC(2,1) y^{US} REER	**84.183***** VEC(2,1) y^{US} REER	**19.292***** VEC(2,1) y^{US} REER	**0.703***** VEC(2,1) y^{US} REER
Financial	48.109^{***} DARCH(2,6)	4102.830 DARCH(2,6)	64.583^{*} DGARCH(2,2)	65.985 DGARCH(2,2)	9.265 VEC(4,1) $stock^{EU}$ $stock^{US}$ $spread^{US}$	0.287 VEC(4,1) $stock^{EU}$ $stock^{US}$ $spread^{US}$
Climatic	47.001 s-VAR(2) SSTA	4102.830 DARCH(2,6)	64.583 DGARCH(2,2)	65.985 DGARCH(2,2)	8.039 DGARCH(2,2)	0.246 DGARCH(2,2)

Note: Bold figures in black indicate the best performance among all groups but within certain forecast horizon, and bold figures in red indicate the best performance among all groups and forecast horizons. See Table A2 for the abbreviation of the models.

*Indicates rejection of the null hypothesis of equal forecasting accuracy at 10%.

**Indicates rejection of the null hypothesis of equal forecasting accuracy at 5%.

***Indicates rejection of the null hypothesis of equal forecasting accuracy at 1%.

TABLE 5 Summary of forecast performance of best models for wheat over variables with highest predictive power

Forecast horizon	MAE	MSE	DA	DV	Return	Sharpe ratio
1-month	**7.331^{**}**	**128.674^{**}**	70.833^{***}	84.351^{***}	**37.015^{***}**	0.514^{***}
	s-VAR(2)	VAR(3)	VEC(1,1)	s-VAR(2)	s-VAR(2)	s-VAR(2)
	y_w^{US}	y_w^{world}	y^{US}	y_w^{world}	y_w^{world}	y_w^{world}
	y_w^{EU}	y_w^{US}	REER	$stock_w^{world}$	$stock_w^{world}$	$stock_w^{world}$
	$yield_w^{world}$	y_w^{EU}	$stock^{US}$	REER	REER	REER
	$stock_w^{world}$	$yield_w^{world}$	SSTA	$stock^{US}$	$stock^{US}$	$stock^{US}$
	y^{US}	$stock_w^{world}$		SSTA	SSTA	SSTA
	REER	y^{US}				
	$stock^{US}$	REER				
	SSTA	$stock^{US}$				
3-month	17.930^{**}	702.687^{*}	74.306^{***}	83.032^{***}	28.086^{***}	0.574^{***}
	VAR(2)	VAR(2)	s-VAR(3)	s-VAR(3)	s-VAR(3)	s-VAR(3)
	y_w^{EU}	y_w^{world}	y_w^{world}	y_w^{world}	y_w^{world}	y_w^{world}
	$stock_w^{world}$	$yield_w^{world}$	$yield_w^{world}$	$yield_w^{world}$	$yield_w^{world}$	$yield_w^{world}$
	REER	y^{EU}	$stock_w^{world}$	$stock_w^{world}$	$stock_w^{world}$	$stock_w^{world}$
	$stock^{US}$	REER	y^{EU}	y^{EU}	y^{EU}	y^{EU}
		$stock^{US}$	REER	REER	REER	REER
6-month	22.839^{***}	1099.885^{***}	82.639^{***}	**92.076^{***}**	26.808^{***}	**0.799^{***}**
	VEC(1,2)	VEC(2,2)	VEC(2,3)	s-VAR(3)	VAR(2)	VAR(2)
	y_w^{US}	$yield_w^{world}$	$yield_w^{world}$	$yield_w^{world}$	y_w^{world}	y_w^{world}
	$yield_w^{world}$	y^{US}	y^{EU}	$stock_w^{world}$	$yield_w^{world}$	$yield_w^{world}$
	y^{EU}	REER	REER	y^{EU}	$stock_w^{world}$	$stock_w^{world}$
	y^{US}	$stock^{US}$	$stock^{EU}$	REER	REER	REER
	REER	$spread^{US}$	GSCI		$stock^{EU}$	$stock^{EU}$
	$stock^{US}$		$spread^{US}$			
	GSCI					
9-month	30.679^{***}	2090.37^{**}	81.944^{***}	88.849^{***}	23.242^{***}	0.744
	VEC(2,1)	VEC(2,1)	VEC(1,2)	VEC(4,1)	VEC(1,2)	VEC(1,2)
	$yield_w^{EU}$	y_w^{US}	y_w^{EU}	y_w^{EU}	y_w^{EU}	y_w^{EU}
	y^{US}	y_w^{EU}	$yield_w^{EU}$	y^{US}	$yield^{EU}$	$yield^{EU}$
	REER	y^{US}	y^{US}	REER	y^{US}	y^{US}
	$stock^{US}$	REER	REER	$spread^{US}$	REER	REER
	$spread^{US}$	$spread^{US}$	$stock^{US}$		$stock^{US}$	$stock^{US}$
			$spread^{US}$		$spread^{US}$	$spread^{US}$
12-month	35.394^{***}	2491.006^{***}	**83.333^{***}**	90.397^{***}	20.739^{***}	0.786^{***}
	VEC(1,3)	VEC(1,3)	VEC(2,1)	VEC(2,2)	VEC(2,2)	VEC(2,2)
	y_w^{US}	y_w^{US}	$yield_w^{EU}$	$yield_w^{EU}$	$yield_w^{EU}$	$yield_w^{EU}$
	$yield_w^{US}$	$yield_w^{US}$	y^{US}	$stock_w^{world}$	$stock_w^{world}$	$stock_w^{world}$
	$yield_w^{EU}$	$yield_w^{EU}$	REER	REER	REER	REER
	y^{US}	y^{US}	$stock^{US}$	$stock^{US}$	$stock^{US}$	$stock^{US}$
	REER	REER	$spread^{US}$	$spread^{US}$	$spread^{US}$	$spread^{US}$

TABLE 5 (Continued)

Forecast horizon	MAE	MSE	DA	DV	Return	Sharpe ratio
	$stock^{US}$ $spread^{US}$	$stock^{US}$ $spread^{US}$				

Note: Bold figures indicate the best performance among all forecast horizons. See Table A2 for the abbreviation of the models.
*Indicates rejection of the null hypothesis of equal forecasting accuracy at 10%.
**Indicates rejection of the null hypothesis of equal forecasting accuracy at 5%.
***Indicates rejection of the null hypothesis of equal forecasting accuracy at 1%.

TABLE 6 Summary of forecast performance of best models for soybeans over different variable groups (fundamental, macroeconomic, financial and climatic) for 1-, 3- and 6-month horizons

	MAE	MSE	DA	DV	Return	Sharpe ratio
1-month horizon						
Fundamental	17.664* BDVAR(1) $yield_s^{world}$ $yield_s^{US}$	659.854* DVAR(1) y_s^{BR} $yield_s^{US}$	60.417* DVAR(1) y_s^{BR}	69.735** VEC(1,1) $yield_s^{US}$	22.522** VEC(1,1) $yield_s^{US}$	0.323** VEC(1,1) $yield_s^{US}$
Macroeconomic	**17.384*** s-DVAR(1) y^{EU} li^{EU} REER	**644.562** s-DVAR(1) y^{US} li^{EU} li^{US} REER	**62.500**** VEC(4,2) y^{US} li^{EU} li^{US}	**76.025**** DVAR(4) y^{US} li^{US} REER	**26.927**** DVAR(4) y^{US} li^{US} REER	**0.388** DVAR(4) y^{US} li^{US} REER
Financial	17.690* BDVAR(1) $spread^{EU}$	666.307 s-DAR(1)	61.111* BDVAR(1) $stock^{EU}$ $stock^{US}$ GSCI $spread^{EU}$	69.318* DVAR(1) $stock^{US}$ GSCI $spread^{EU}$	21.478* DVAR(1) $stock^{US}$ GSCI $spread^{EU}$	0.308* DVAR(1) $stock^{US}$ GSCI $spread^{EU}$
Climatic	17.693 s-DAR(1)	666.307 s-DAR(1)	57.639 DVAR(1) SSTA	64.987 DVAR(1) SSTA	16.843 DGARCH(4,2)	0.242* DGARCH(4,2)
3-month horizon						
Fundamental	39.170 RW	2942.875 s-DAR(1)	60.417 RW y_s^{world}	63.948 DVAR(1) y_s^{world} $yield_s^{world}$ $stock_s^{world}$	11.382 DVAR(1) y_s^{world} $yield_s^{world}$ $stock_s^{world}$	0.249 DVAR(1) $yield_s^{world}$ $stock_s^{world}$
Macroeconomic	**38.618** VEC(1,1) y^{BR} REER	**2816.018** VEC(1,2) y^{EU} y^{US} y^{BR} REER	**63.889** VEC(1,2) y^{BR} li^{US} REER	**68.493** VEC(1,2) y^{US} y^{BR} REER	**12.863** VEC(1,2) y^{US} y^{BR} REER	**0.282** VEC(1,2) y^{US} y^{BR} REER

(Continues)

TABLE 6 (Continued)

	MAE	MSE	DA	DV	Return	Sharpe ratio
Financial	39.170	2942.875	60.417	60.572	9.435	0.206
	RW	s-DAR(1)	RW	last 3 months	BDVAR(1)	BDVAR(1)
					$stock^{EU}$	$stock^{EU}$
					GSCI	GSCI
					$spread^{EU}$	$spread^{EU}$
Climatic	39.170	2942.875	60.417	56.108	6.647	0.145
	RW	s-DAR(1)	RW	BDVAR(1)	RW	RW
				SOI		
6-month horizon						
Fundamental	62.063	6371.589	61.806	62.692	8.209	0.232
	s-DAR(1)	s-DAR(1)	s-DVAR(2)	VEC(4,1)	s-DVAR(2)	s-DVAR(2)
			y_s^{world}	y_s^{world}	y_s^{world}	y_s^{world}
			$yield_s^{US}$	y_s^{BR}	$yield_s^{US}$	$yield_s^{US}$
				$yield_s^{US}$		
Macroeconomic	**60.545**	**6072.505**	**67.361**	**75.749**	**14.835**	**0.440**
	VEC(1,1)	VEC(1,2)	VEC(1,2)	VEC(1,2)	VEC(1,2)	VEC(1,2)
	y^{US}	y^{US}	y^{US}	y^{US}	y^{US}	y^{US}
	y^{BR}	y^{BR}	y^{BR}	y^{BR}	y^{BR}	y^{BR}
		REER	REER	REER	REER	REER
Financial	62.045	6371.589	57.639	61.496	9.201	0.261
	s-DVAR(1)	s-DAR(1)	DVAR(6)	DVAR(6)	DVAR(1)	DVAR(1)
	$stock^{EU}$		$spread^{US}$	$spread^{US}$	$stock^{EU}$	$stock^{EU}$
	GSCI				GSCI	GSCI
	$spread^{EU}$				$spread^{EU}$	$spread^{EU}$
	$spread^{US}$					
Climatic	62.063	6371.589	55.556 56.894	6.891	0.194	
	s-DAR(1)	s-DAR(1)	s-DVAR(1)	s-DVAR(1)	RW	RW
			SOI	SOI		

Note: Bold figures in black indicate the best performance among all groups but within certain forecast horizon, and bold figures in red indicate the best performance among all groups and forecast horizons. See Table A2 for the abbreviation of the models.
*Indicates rejection of the null hypothesis of equal forecasting accuracy at 10%.
**Indicates rejection of the null hypothesis of equal forecasting accuracy at 5%.
***Indicates rejection of the null hypothesis of equal forecasting accuracy at 1%.

longer forecast horizons. The variable appearing most often among the best models is the REER, which is chosen in all models as the variable with the best predictive power. Financial variables are not chosen in long forecast horizons. Note that the best forecasting accuracy and return are achieved for the 1-month-ahead forecast horizon, whereas the highest DA, DV and Sharpe ratio were achieved for 12-month-ahead forecasts. For both forecast horizons, 1 and 12 months ahead, the 'best' models significantly outperform the random walk model, whereas for other horizons (3, 6 and 9 months ahead), the improvement with respect to the random walk is only marginal. This contrasts the forecast performance of the wheat price where the random walk is significantly outperformed by the 'best' models in terms of predictive ability.

3.3 | Corn price

Corn is the crop with the highest production worldwide. Its starch content is about 70% and thus can be converted

TABLE 7 Summary of forecast performance of best models for soybeans over different variable groups (fundamental, macroeconomic, financial and climatic) for 9- and 12-month horizons

	MAE	MSE	DA	DV	Return	Sharpe ratio
9-month horizon						
Fundamental	76.380 RW	9638.842 RW	64.583 s-DVAR(2) y_s^{world} $yield_s^{US}$	66.209* DVAR(4) $yield_s^{US}$ $stock_s^{world}$	9.843 DVAR(4) $yield_s^{US}$ $stock_s^{world}$	0.331 DVAR(4) $yield_s^{US}$ $stock_s^{world}$
Macroeconomic	**72.724** VEC(1,2) y^{US} y^{BR} REER	**8585.876** VEC(1,2) y^{US} y^{BR} REER	**70.833*** VEC(1,2) y^{US} y^{BR} REER	**82.777**** VEC(1,2) y^{US} y^{BR} REER	**15.600**** VEC(1,2) y^{US} y^{BR} REER	**0.570**** VEC(1,2) y^{US} y^{BR} REER
Financial	76.380 RW	9638.842 RW	57.778 last 12 months	57.316 last 9 months	6.659 DVAR(1) $stock^{EU}$ GSCI $spread^{EU}$	0.218 DVAR(1) $stock^{EU}$ GSCI $spread^{EU}$
Climatic	76.380 RW	9638.842 RW	57.778 last 12 months	54.530 BDVAR(1) SOI	5.811 last 9 months, DA	0.190 last 9 months, DA
12-month horizon						
Fundamental	86.620 RW	10,906.1 RW	65.278*** DVAR(4) $yield_s^{US}$ $stock_s^{world}$	72.551*** DVAR(4) $yield_s^{US}$ $stock_s^{world}$	10.995** DVAR(4) $yield_s^{US}$ $stock_s^{world}$	0.445** DVAR(4) $yield_s^{US}$ $stock_s^{world}$
Macroeconomic	**76.329** VEC(1,1) y^{US} y^{BR} REER	**9195.996**** VEC(6,1) y^{US} y^{BR} REER	**69.444***** VEC(1,2) y^{US} y^{BR} REER	**80.363***** VEC(1,2) y^{US} y^{BR} REER	**12.422**** VEC(1,2) y^{US} y^{BR} REER	**0.518** VEC(1,2) y^{US} y^{BR} REER
Financial	82.620 RW	10,906.1 RW	56.061 last 6 months	58.940* last 6 months	6.082* BDVAR(1) $stock^{EU}$ $spread^{EU}$	0.231* BDVAR(1) $stock^{EU}$ $spread^{EU}$
Climatic	82.620 RW	10,906.1 RW	56.061 last 6 months	58.385 last 6 months	5.676 DVAR(1) SOI	0.215 DVAR(1) SOI

Note: Bold figures in black indicate the best performance among all groups but within certain forecast horizon, and bold figures in red indicate the best performance among all groups and forecast horizons. See Table A2 for the abbreviation of the models.
*Indicates rejection of the null hypothesis of equal forecasting accuracy at 10%.
**Indicates rejection of the null hypothesis of equal forecasting accuracy at 5%.
***Indicates rejection of the null hypothesis of equal forecasting accuracy at 1%.

TABLE 8 Summary of forecast performance of best models for soybeans over variables with highest predictive power

Forecast horizon	MAE	MSE	DA	DV	Return	Sharpe ratio
1-month	**17.095*** VEC(1,1) y_s^{BR} $yield_s^{US}$ y^{US} REER $spread^{EU}$	**633.045*** DVAR(1) y_s^{BR} $yield_s^{US}$ y^{US} li^{EU} li^{US} REER $spread^{EU}$	65.278*** DVAR(3) y^{US} li^{EU} li^{US} REER $stock^{US}$ $spread^{EU}$	78.027*** DVAR(3) y^{US} li^{EU} li^{US} REER $stock^{US}$ $spread^{EU}$	**29.924**** DVAR(3) y^{US} li^{EU} li^{US} REER $stock^{US}$ $spread^{EU}$	0.433*** DVAR(3) y^{US} li^{EU} li^{US} REER $stock^{US}$ $spread^{EU}$
3-month	37.902 VEC(1,2) y_s^{world} y^{US} REER GSCI SOI	2740.142 VEC(1,2) y_s^{world} y^{US} REER GSCI SOI	66.667** VEC(1,2) $yield_s^{US}$ y^{BR} REER $stock^{EU}$	70.004 VEC(1,2) $yield_s^{US}$ y^{US} y^{BR} REER	15.763 DVAR(1) $yield_s^{US}$ $stock^{world_s}$ y^{BR} REER GSCI $spread^{EU}$ SOI	0.349 DVAR(1) $yield_s^{US}$ $stock^{world_s}$ y^{BR} REER GSCI $spread^{EU}$ SOI
6-month	59.150 VEC(1,2) $yield_s^{US}$ y^{BR} REER $stock^{EU}$	5731.902 VEC(1,2) $yield_s^{world}$ y^{US} REER GSCI SOI	67.361 VAR(5) y_s^{world} $yield_s^{US}$ y^{BR} REER	75.749 VEC(1,2) y^{US} y^{BR} REER	14.835 VEC(1,2) y^{US} y^{BR} REER	0.439 VEC(1,2) y^{US} y^{BR} REER
9-month	70.974 VEC(1,2) $yield_s^{US}$ y^{US} y^{BR} REER	8585.876 VEC(1,2) y^{US} y^{BR} REER	70.833* VEC(1,2) y^{US} y^{BR} REER	82.777*** VEC(1,2) y^{US} y^{BR} REER	15.600** VEC(1,2) y^{US} y^{BR} REER	0.570** VEC(1,2) y^{US} y^{BR} REER
12-month	75.716 VEC(1,2) $yield_s^{US}$ y^{US} y^{BR} REER SOI	9039.717** VEC(1,2) $yield_s^{US}$ y^{US} y^{BR} REER	**74.306***** VEC(1,2) $yield_s^{US}$ y^{US} y^{BR} REER	**86.452***** VEC(1,2) $yield_s^{US}$ y^{US} y^{BR} REER	14.845*** VEC(1,2) $yield_s^{US}$ y^{US} y^{BR} REER	**0.658** VEC(1,2) $yield_s^{US}$ y^{US} y^{BR} REER

Note: Bold figures indicate the best performance among all forecast horizons. See Table A2 for the abbreviation of the models.

*Indicates rejection of the null hypothesis of equal forecasting accuracy at 10%.

**Indicates rejection of the null hypothesis of equal forecasting accuracy at 5%.

***Indicates rejection of the null hypothesis of equal forecasting accuracy at 1%.

to ethanol. Corn is the feedstock for more than 90% of ethanol production in the United States due to its abundance and low price. For corn prices, we again employ the four thematic groups of variables: fundamental, macroeconomic, financial and climatic variables. Macroeconomic, financial and climatic variables remain the same as in the case of the wheat: as *fundamental variables*, we use corn production for the whole world, y_c^{world}, corn production for the United States, y_c^{US}, corn production for China, y_c^{CH}, corn yield for world, $yield_c^{world}$, corn yield for the United States, $yield_c^{US}$, corn yield for China, $yield_c^{CH}$, and world stock-to-use ratio for corn, $stock_c^{world}$. As the United States is the biggest producer of corn (with 384,778,000 metric tons in February 2017), followed by China (with 219,554,000 metric tons in February 2017), a focus on the fundamental variables for these two countries appears justified.

The results for predictive ability differences across the different groups are shown in Tables 9 and 10. Specifications containing market fundamental variables achieve the best predictive performance over practically all the forecast accuracy measures. The only exceptions can be found for DA at horizons of 3 and 6 months ahead, where models based on macroeconomic variables overperform the rest of the specifications. All fundamental variables (output, yield and stock-to-use ratio) are employed in the models with best predictive performance. Models based on financial and climatic variables perform poorly, and the random walk and other univariate specifications tend to present the best performance in these groups.

When evaluating forecast performance based on combination of variables that give the best performance as predictors (see Table 11), again, the VEC specification appears as the best model in a majority of cases. The most used variables among the best performing models are stock-to-use variable and yield, which are chosen in nearly all models as the variables with highest predictive power. Financial and climatic variables are picked in Horizons 6, 9 and 12. The interest rate spread was the financial variable, which was most often chosen, and the Southern Oscillation Index anomalies is chosen among the climatic variables. The most used macroeconomic variable is the REER. Note that the best forecasting accuracy and return are achieved for the 1-month-ahead forecast horizon, whereas the highest DA, DV and Sharpe ratio are achieved for 12-month-ahead forecast horizons. As in the case of wheat price forecasts, the 'best' models for the corn price significantly outperform the random walk model (with the exception of the performance based on the loss measures for the 1- and 3-month forecast horizons).

3.4 | Asymmetric effects along the business cycle?

Several theoretical arguments imply that the predictive power of econometric models for commodity prices may depend on the particular phase of the business cycle in which they are performed (Gargano & Timmermann, 2014). We evaluate the statistical significance of differences in predictive power in expansions relative to recessions using regression models for differences in forecast performance measures (M_{th}) between the benchmark model (the random walk, *RW*) and the best model in our battery of specifications with a recession indicator (D_t) as the explanatory variable. We thus estimate the regression model

$$M_{RW,t,h} - M_{best,t,h} = c_0 + c_1 D_t + \varepsilon_t, \tag{8}$$

where the performance measure M_{th} is alternatively the absolute error $AE_{th} = \left|\hat{P}_{t|t-h} - P_t\right|$; square forecast error $SE_{th} = \left(\hat{P}_{t|t-h} - P_t\right)^2$; DA measure $DA_{th} = I(\text{sgn}(P_t - P_{t-h}) = \text{sgn}(\hat{P}_{t|t-h} - P_{t-h}))$, where $I(\cdot)$ is the indicator function; DV measure $DV_{th} = |P_t - P_{t-h}| DA_{th}$; the return R_{th} implied by the 'buy low, sell high' trading strategy and the Sharpe ratio implied by this return. The dummy variable D_t represents periods of recession when $D_t = 1$ and expansion $D_t = 0$ and is calculated on the basis of the turning points of the growth cycle as captured by the corresponding Organisation for Economic Co-operation and Development (OECD) composite leading indicator. We perform the analysis based alternatively on recessions and expansions for the euro area (EU), United States, OECD, Brazil (in case of the soybean) and China (in case of the corn). The recession period is identified as the period between a peak and a trough. Positive and significant values of c_1 for loss measures (AE and SE) suggest that the best model is more accurate relative to the benchmark during recessions than during expansions, whereas negative and significant values of c_1 for loss measures suggest that the best model is more accurate (relative to the benchmark) during expansions than during recessions. Regarding the profit measures (DA, DV, return and SR), negative and significant values of c_1 suggest that the best model is more accurate relative to the benchmark during recessions than during expansions, whereas positive and significant values of c_1 suggest that the best model is more accurate during expansions than during recessions.

Table 12 presents the results of the analysis by forecasting horizon and recession indicator. The figures in Table 12 without an asterisk show the forecasting horizons at which the best model performs significantly better than the benchmark in expansions, whereas figures

TABLE 9 Summary of forecast performance of best models for corn over different variable groups (fundamental, macroeconomic, financial and climatic) for 1-, 3- and 6-month horizons

	MAE	MSE	DA	DV	Return	Sharpe ratio
1-month horizon						
Fundamental	**9.518** VEC(6,1) y_c^{world} y_c^{US} $yield_c^{CH}$	**184.722** VEC(4,2) y_c^{world} y_c^{US} y_c^{CH} $yield_c^{CH}$ $stock_c^{world}$	**61.111**** DVAR(6) y_c^{US} $yield_c^{world}$ $yield_c^{US}$	**69.609**** DVAR(1) $yield_c^{US}$ $stock_c^{world}$	**23.544*** VEC(1,2) y_c^{CH} $yield_c^{world}$ $yield_c^{CH}$ $stock_c^{world}$	**0.272*** VEC(1,2) y_c^{CH} $yield_c^{world}$ $yield_c^{CH}$ $stock_c^{world}$
Macroeconomic	9.522 RW	193.527 BDVAR(4) y^{US} y^{world} li^{US} REER	**61.111*** VEC(1,2) y^{EU} y^{US} li^{EU} li^{US} REER	68.166 VEC(1,2) y^{EU} y^{US} li^{EU} li^{US} REER	21.728 VEC(1,2) y^{EU} y^{US} li^{EU} li^{US} REER	0.252 VEC(1,2) y^{EU} y^{US} li^{EU} li^{US} REER
Financial	9.522 RW	197.459 BDAR(1) $stock^{US}$	55.556 DVAR(1) $stock^{EU}$	63.301 DVAR(4) $stock^{US}$ $spread^{US}$	16.041 DVAR(4) $stock^{US}$ $spread^{US}$	0.188 DVAR(4) $stock^{US}$ $spread^{US}$
Climatic	9.522 RW	198.019 last month	56.250 BDVAR(2) SOI SSTA	62.220 last 9 months	13.731 DVAR(1) SSTA	0.161 DVAR(1) SSTA
3-month horizon						
Fundamental	**18.435** VEC(4,2) y_c^{world} y_c^{US} y_c^{CH} $yield_c^{CH}$ $stock_c^{world}$	**623.051*** VEC(4,2) y_c^{world} y_c^{US} y_c^{CH} $yield_c^{CH}$ $stock_c^{world}$	66.664** VEC(4,2) y_c^{world} y_c^{US} $stock_c^{world}$	**78.470***** DVAR(5) y_c^{US} y_c^{CH} $yield_c^{world}$ $yield_c^{US}$ $yield_c^{CH}$ $stock_c^{world}$	**22.734**** VEC(4,2) y_c^{US} $yield_c^{US}$ $yield_c^{CH}$ $stock_c^{world}$	**0.418**** VEC(4,2) y_c^{US} $yield_c^{US}$ $yield_c^{CH}$ $stock_c^{world}$
Macroeconomic	18.985 VEC(1,2) y^{EU} y^{US} li^{EU} li^{US} REER	731.485 VEC(1,1) y^{US} li^{EU} REER	**69.444***** VAR(2) y^{EU} li^{EU} REER	70.888 VEC(1,1) y^{US} li^{EU} REER	18.246 VEC(1,1) y^{US} li^{EU} REER	0.331 VEC(1,1) y^{US} li^{EU} REER

TABLE 9 (Continued)

	MAE	MSE	DA	DV	Return	Sharpe ratio
Financial	19.497	764.358	57.639	63.162	13.043	0.235
	RW	s-DAR(1)	VAR(5)	DVAR(1)	last 9 months, MAE	last 9 months, MAE
			$stock^{US}$	GSCI		
			GSCI			
			$spread^{US}$			
Climatic	19.497	764.358	56.944	62.104	11.682	0.210
	RW	s-DAR(1)	ARCH(2,4)	VEC(1,1)	whole, MAE	whole, MAE
				SOI		
				SSTA		
6-month horizon						
Fundamental	**27.150***	**1445.155***	68.750*	**81.570***	**21.524****	**0.518****
	VEC(1,2)	VEC(4,2)	VAR(5)	VEC(5,2)	DVAR(4)	DVAR(4)
	y_c^{US}	y_c^{world}	y_c^{US}	y_c^{US}	y_c^{world}	y_c^{world}
	y_c^{CH}	y_c^{US}	$yield_c^{US}$	y_c^{CH}	y_c^{CH}	y_c^{CH}
	$yield_c^{world}$	y_c^{CH}	$stock_c^{world}$	$yield_c^{world}$	$yield_c^{world}$	$yield_c^{world}$
	$yield_c^{CH}$	$yield_c^{CH}$		$yield_c^{US}$	$yield_c^{US}$	$yield_c^{US}$
	$stock_c^{world}$	$stock_c^{world}$		$yield_c^{CH}$	$yield_c^{CH}$	$yield_c^{CH}$
Macroeconomic	29.076	1710.724	**70.139****	77.560	19.517	0.461**
	VEC(1,1)	VEC(2,1)	VEC(2,1)	VEC(2,1)	VEC(2,1)	VEC(2,1)
	y^{US}	y^{US}	y^{US}	y^{US}	y^{US}	y^{US}
	li^{EU}	REER	REER	REER	REER	REER
	REER					
Financial	30.672	1861.807	59.722	58.457	7.917	0.177
	RW	s-DAR(1)	VAR(2)	DVAR(4)	RW	RW
			$stock^{US}$	$spread^{US}$		
			GSCI			
Climatic	30.672	1861.807	57.639	58.393	7.917	0.177
	RW	s-DAR(1)	VAR(3)	whole	RW	RW
			SSTA			

Note: Bold figures in black indicate the best performance among all groups but within certain forecast horizon, and bold figures in red indicate the best performance among all groups and forecast horizons. See Table A2 for the abbreviation of the models.
*Indicates rejection of the null hypothesis of equal forecasting accuracy at 10%.
**Indicates rejection of the null hypothesis of equal forecasting accuracy at 5%.
***Indicates rejection of the null hypothesis of equal forecasting accuracy at 1%.

with asterisk refer to significantly better performance in recessions.

We can infer from our results that the best models for wheat outperform the benchmark model in recessions for both loss- and profit-based measures and for all three recession dummies. The asymmetry of the forecast performance along the business cycle is the most pronounced for wheat (with respect to the recession dummies for the EU, the United States and OECD) when compared with soybean and corn. The only exception is for the DA, return and the Sharpe ratio profit measures for the euro area recession dummy variable, when the best models for wheat outperform the benchmark model in expansions.

Results are more mixed for soybean. The most systematic results are for business cycles based on the indicator for Brazil with profit-based measures, for which best models outperform benchmark models in recessions (for 6- to 12-month forecast horizons for DA measure and for 3- to 9-month forecast horizons for

TABLE 10 Summary of forecast performance of best models for corn over different variable groups (fundamental, macroeconomic, financial and climatic) for 9- and 12-month horizons

	MAE	MSE	DA	DV	Return	Sharpe ratio
9-month horizon						
Fundamental	**31.792*** VAR(5) y_c^{US} $yield_c^{US}$ $stock_c^{world}$	**2130.211** VAR(6) y_c^{CH} $yield_c^{world}$ $yield_c^{US}$ $stock_c^{world}$	**78.472***** VAR(2) $yield_c^{US}$ $yield_c^{CH}$ $stock_c^{world}$	**84.070***** DVAR(5) y_c^{world} y_c^{CH} $yield_c^{world}$ $yield_c^{US}$ $yield_c^{CH}$ $stock_c^{world}$	**20.044**** DVAR(4) y_c^{world} $yield_c^{world}$ $yield_c^{US}$	**0.588**** DVAR(4) y_c^{world} $yield_c^{world}$ $yield_c^{US}$
Macroeconomic	37.335 VEC(1,1) y^{US} li^{EU} REER	2551.644 VEC(2,1) y^{US} REER	69.444*** VEC(2,1) y^{US} REER	74.461 VEC(2,1) y^{US} REER	15.170 VEC(2,1) y^{US} REER	0.418* VEC(2,1) y^{US} REER
Financial	39.080 RW	2785.99 RW	64.583 VAR(2) $stock^{US}$ GSCI	57.923 s-DAR(1)	8.614 RW	0.226 RW
Climatic	38.807 FMA-aic	2785.99 RW	61.806 VAR(3) SSTA	58.254 whole	8.614 RW	0.226 RW
12-month horizon						
Fundamental	**35.870**** VAR(6) y_c^{CH} $yield_c^{world}$ $yield_c^{US}$ $stock_c^{world}$	**2613.324*** VAR(6) y_c^{CH} $yield_c^{world}$ $yield_c^{US}$ $stock_c^{world}$	**79.167***** VEC(4,2) y_c^{US} y_c^{CH} $yield_c^{US}$ $yield_c^{CH}$ $stock_c^{world}$	**86.137***** VEC(4,2) y_c^{US} y_c^{CH} $yield_c^{US}$ $yield_c^{CH}$ $stock_c^{world}$	**19.329***** VEC(4,2) y_c^{US} y_c^{CH} $yield_c^{US}$ $yield_c^{CH}$ $stock_c^{world}$	**0.704***** VEC(4,2) y_c^{US} y_c^{CH} $yield_c^{US}$ $yield_c^{CH}$ $stock_c^{world}$
Macroeconomic	43.778 VEC(1,1) y^{world} li^{EU} REER	3174.461 VEC(1,1) y^{world} li^{EU} REER	65.278 VEC(1,1) y^{world} li^{EU} REER	75.032 VEC(1,1) y^{world} li^{EU} REER	13.433 VEC(1,1) y^{world} li^{EU} REER	0.436 VEC(1,1) y^{world} li^{EU} REER
Financial	45.229 RW	3562.25 RW	60.417 VAR(2) GSCI	56.008 RW	8.404 RW	0.258 RW
Climatic	45.229 RW	3562.25 RW	59.028 ARCH(2,4)	80.431 whole	8.404 RW	0.258 RW

Note: Bold figures in black indicate the best performance among all groups but within certain forecast horizon, and bold figures in red indicate the best performance among all groups and forecast horizons. See Table A2 for the abbreviation of the models.

*Indicates rejection of the null hypothesis of equal forecasting accuracy at 10%.

***Indicates rejection of the null hypothesis of equal forecasting accuracy at 5%.

***Indicates rejection of the null hypothesis of equal forecasting accuracy at 1%.

TABLE 11 Summary of forecast performance of best models for corn over variables with highest predictive power

Forecast horizon	MAE	MSE	DA	DV	Return	Sharpe ratio
1-month	**9.505** VEC(4,2) y_c^{world} y_c^{US} y_c^{CH} $yield_c^{CH}$ $stock_c^{world}$ REER	**180.601** VEC(4,3) y_c^{world} y_c^{US} y_c^{CH} $yield_c^{CH}$ $stock_c^{world}$ y^{US} REER	61.806* VEC(1,1) $yield_c^{world}$ y^{EU} REER	73.650** VEC(1,3) y_c^{world} y_c^{CH} $yield_c^{CH}$ $stock_c^{world}$ y^{US} li^{EU}	**28.938**** VEC(1,3) y_c^{world} y_c^{CH} $yield_c^{CH}$ $stock_c^{world}$ y^{US} li^{EU}	0.334** VEC(1,3) y_c^{world} y_c^{CH} $yield_c^{CH}$ $stock_c^{world}$ y^{US} li^{EU}
3-month	18.149 VEC(1,1) y_c^{world} y_c^{CH} $yield_c^{CH}$ $stock_c^{world}$ REER	613.232 VEC(6,2) y_c^{world} y_c^{CH} $yield_c^{world}$ $yield_c^{US}$ $yield_c^{CH}$ $stock_c^{world}$ li^{EU}	71.528*** VAR(2) y_c^{CH} $yield_c^{world}$ $yield_c^{CH}$ $stock_c^{world}$ li^{EU} REER	80.985** VAR(2) y_c^{CH} $yield_c^{world}$ $yield_c^{CH}$ $stock_c^{world}$ li^{EU} REER	26.250*** VEC(1,2) $yield_c^{US}$ $yield_c^{CH}$ $stock_c^{world}$ li^{EU} REER GSCI	0.491*** VEC(1,2) $yield_c^{US}$ $yield_c^{CH}$ $stock_c^{world}$ li^{EU} REER GSCI
6-month	25.822** VEC(1,2) y_c^{CH} $yield_c^{world}$ $yield_c^{CH}$ $stock_c^{world}$ REER $stock^{US}$ $spread^{US}$	1241.291** VAR(2) y_c^{CH} $yield_c^{US}$ $yield_c^{CH}$ $stock_c^{world}$ REER $stock^{US}$ $spread^{US}$ SOI	76.389*** s-VAR(2) $yield_c^{world}$ $yield_c^{US}$ $yield_c^{CH}$ $stock_c^{world}$ REER $stock^{US}$ SOI	84.256** VEC(5,2) $yield_c^{world}$ $yield_c^{US}$ REER $stock^{US}$ SOI	23.513** VEC(4,1) y_c^{CH} $yield_c^{US}$ $stock_c^{world}$ REER $spread^{US}$	0.578*** VEC(4,1) y_c^{CH} $yield_c^{US}$ $stock_c^{world}$ REER $spread^{US}$
9-month	28.901** VAR(5) y_c^{CH} $yield_c^{US}$ $stock_c^{world}$ y^{US} $spread^{US}$	1696.506*** VEC(4,3) y_c^{US} y_c^{CH} $yield_c^{US}$ $yield_c^{CH}$ $stock_c^{world}$ y^{US} $spread^{US}$ SOI	81.250*** VEC(1,2) $yield_c^{US}$ $stock_c^{world}$ y^{US} REER $stock^{US}$ $spread^{US}$ SOI	88.729*** VEC(4,1) $yield_c^{US}$ $stock_c^{world}$ y^{US} REER $spread^{US}$	23.412*** VEC(4,1) $yield_c^{US}$ $stock_c^{world}$ y^{US} REER $spread^{US}$	0.730*** VEC(4,1) $yield_c^{US}$ $stock_c^{world}$ y^{US} REER $spread^{US}$
12-month	30.973*** VEC(4,3) y_c^{US}	1774.714*** VEC(4,3) y_c^{US}	**88.194***** VEC(4,3) y_c^{US}	**92.452***** VEC(4,3) y_c^{US}	21.657*** VEC(4,3) y_c^{US}	**0.844***** VEC(4,3) y_c^{US}

(Continues)

TABLE 11 (Continued)

Forecast horizon	MAE	MSE	DA	DV	Return	Sharpe ratio
	y_c^{CH}	y_c^{CH}	y_c^{CH}	y_c^{CH}	y_c^{CH}	y_c^{CH}
	$yield_c^{US}$	$yield_c^{US}$	$yield_c^{US}$	$yield_c^{US}$	$yield_c^{US}$	$yield_c^{US}$
	$yield_c^{CH}$	$yield_c^{CH}$	$yield_c^{CH}$	$yield_c^{CH}$	$yield_c^{CH}$	$yield_c^{CH}$
	$stock_c^{world}$	$stock_c^{world}$	$stock_c^{world}$	$stock_c^{world}$	$stock_c^{world}$	$stock_c^{world}$
	REER	REER	REER	REER	REER	REER
	$spread^{US}$	$spread^{US}$	$spread^{US}$	$spread^{US}$	$spread^{US}$	$spread^{US}$
	SOI	SOI	SOI	SOI	SOI	SOI

Note: Bold figures indicate the best performance among all forecast horizons. See Table A2 for the abbreviation of the models.
*Indicates rejection of the null hypothesis of equal forecasting accuracy at 10%.
**Indicates rejection of the null hypothesis of equal forecasting accuracy at 5%.
***Indicates rejection of the null hypothesis of equal forecasting accuracy at 1%.

TABLE 12 Forecast horizons when the best models significantly outperform the benchmark model (RW) in either expansion or recession times

Wheat	Euro area	The United States	OECD	
AE	6*	6*	6*	
SE	6*, 9*	6*, 9*, 12*	1*, 6*, 9*	
DA	9	6*	6*, 9*	
DV	1*, 3*, 6*, 9*	1*, 6*, 9*	1*, 6*, 9*	
Return	1*, 12	6*, 9*	6*, 9*	
Sharpe ratio	1*, 12	6*	6*	
Soybean	**Euro area**	**The United States**	**OECD**	**Brazil**
AE	6	6	6	3
SE	6*	6*	3*, 6*	6
DA				6*, 9*, 12*
DV	1*		12*	3*, 6*, 9*
Return	1*	12*	12*	3*, 6*, 9*
Sharpe ratio	1*			3*, 6*, 9*
Corn	**Euro area**	**The United States**	**OECD**	**China**
AE	9*, 12	12	12	
SE		6, 9, 12	12	
DA	12	9		12*
DV				
Return				
Sharpe ratio		12	12	

Note: No asterisk values indicate horizons when best models outperform the benchmark model in expansion times, whereas asterisk values indicate horizons when best models outperform the benchmark model in recession times.

DV, return and the Sharpe ratio). The best models outperform the benchmark in expansions for loss measures such as for the AE for 3-month forecast horizon and for the MS measure for 6-month forecast horizon. The AE loss measure of the benchmark is outperformed also in expansions for the euro area, the United States and OECD for the 6-month-ahead horizon, whereas the SE of the benchmark is outperformed in recessions (for the EU, the United States and OECD and 6-month-ahead horizon and 3-month horizon for OECD). The rest of the results are scattered and not very systematic.

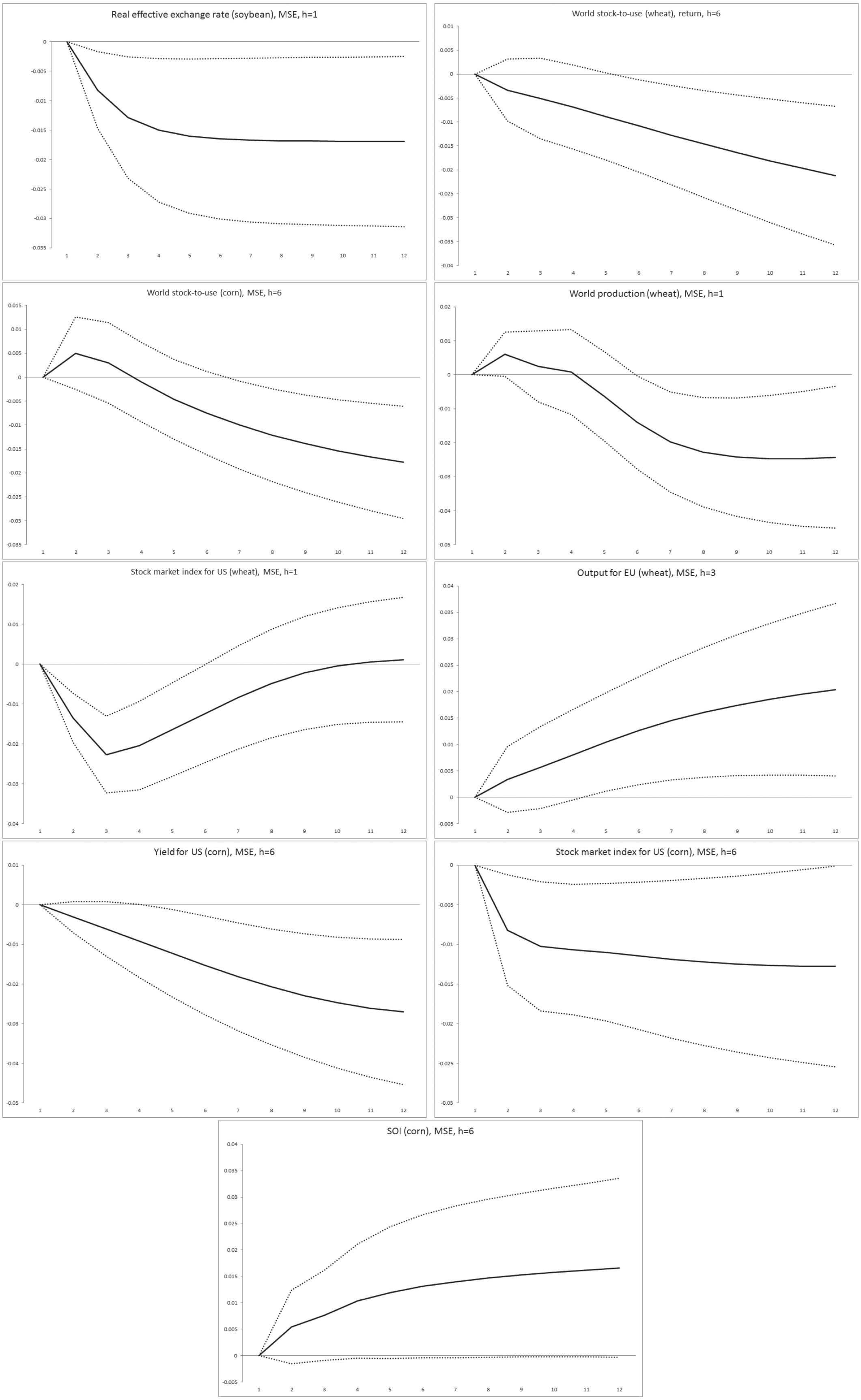

FIGURE 1 Impulse response analysis. EU, European Union; MSE, mean-square error

In contrast to the case of wheat, the best models for corn that significantly outperform the benchmark model appear for expansion periods for longer forecast horizons (9 and 12 months ahead) for both loss- and profit-based measures (DA or Sharpe ratio) and for the EU, the United States and OECD recession dummies.[7] However, these occurrences are much less frequent than in the case of wheat. Regarding indicators for China, there is only one case when the best model outperforms the benchmark model, namely, for the DA performance measure, for 12-month forecast horizon in recession periods.

4 | QUANTIFYING THE RESPONSES OF AGRICULTURAL COMMODITY PRICES TO THEIR DETERMINANTS

In this section, we present the main results of the impulse response analysis for the best models for wheat, soybean and corn presented in Tables 5,8 and 11. For the impulse response analysis, we use Cholesky decomposition to identify the structural shocks in the respective multivariate time series model[8] and 12-period (months) impulse response functions. We analyse the responses of the price of wheat, soybean and corn to one standard deviation shock in all of the variables included in the corresponding best models, with the results presented in Figure 1.

The most robust result across commodities is the negative price response of one standard deviation shock in the REER of the USD against a currency basket. An appreciation of the USD with respect to the currency basket[9] lowers the price of all three commodities, which is denominated in USD. The fact that the REER is included in every best forecasting model for wheat and soybean and in the majority of the best models for corn highlights the importance of changes in the overall competitiveness in international markets as a determinant of price changes in agricultural commodities.

The most robust outcome regarding the impulse responses of the model for wheat prices across all horizons and performance measures (in addition to the one described for the REER) is the negative price response to positive shocks in the world and the US wheat production (as predicted by theory), as well as in the world stock-to-use ratio. In addition, we observe also a positive price response to shocks in US industrial production and the negative price response to the stock market index for the United States, where the price reaction occurs in the short run.

For soybean, the most salient reactions captured by the impulse response analysis are the negative price response to positive shocks in the yield in the United States and the positive price response to output increases in the United States and in the output for Brazil, which can be interpreted as reactions to increased demand.

Finally, for corn prices, the analysis highlights a long-run negative price response to shock in the world stock-to-use ratio following a price increase, a negative price response to increases in the yield in the United States and the positive price response to corn production in China. This last result may be related to price support policies for corn implemented in China during the period under scrutiny in our analysis.

5 | CONCLUSIONS

We present a comprehensive modelling framework to perform out-of-sample prediction of commodity prices for forecasting horizons from 1 to 12 months ahead. The commodities are wheat (748.24 million tons of global production for 2016/2017), soybeans (337.85 million tons of global production for 2016/2017) and corn (1040.21 million tons of global production for 2016/2017). Our method relies on assessing explicitly model uncertainty by incorporating not only single univariate and multivariate specifications, but also forecast combinations that rely on both Bayesian and frequentist weighting schemes and also on composite forecasts based on predictive ability over certain period. We incorporate information related to market fundamentals, global macroeconomic and financial developments, and climatic covariates as potential predictors of commodity prices. We measure the predictive ability of different models, forecast combination methods and composite forecasts making use of a battery of performance measures that build on predictive loss and on the potential return from 'buy low, sell high' trading strategy based on the point forecasts.

When comparing the DA and DV measures across all four commodities, our results indicate that our models for wheat prices (followed by those for corn prices) have the best performance for forecast horizons of 1, 3, 6 and 9 months ahead, whereas corn performs the best for 12-month forecast horizon. The common feature for all commodities is that the performance based on the loss

[7]One exception is for the absolute error measure, for 9-month forecast horizon with respect to the euro area recession dummy, where the best model overperforms the benchmark during recessions.

[8]The results presented are robust to the ordering of variables in the system, as well as to the use of different identification methods.

[9]The currencies of all main producers of wheat, soybean and corn (BRL, CNH, EUR and USD) are part of the basket.

measures and the return is the best for 1-month-ahead horizon, whereas the performance for the remaining profit-based measures is the best for higher forecast horizons.

The role of fundamental variables as out-of-sample predictors of commodity price dynamics is different across the particular agricultural products considered. The stock-to-use ratio plays the most important role for corn and also to a certain extent for wheat. Financial variables such as the interest rate spread and the stock market index for the United States appear to contain out-of-sample predictive ability for corn and wheat at longer forecast horizons, whereas neither these nor any other financial variables appear in the best forecasting models for soybeans at longer forecast horizons (but they do in the shorter forecast horizons). VEC specifications tend to be the most promising individual model structures when it comes to predicting the prices of commodities entertained in this study, and forecast combination methods do not appear to provide significant improvements in predictive ability.

We find mixed evidence concerning the differential predictive ability of the models entertained across states of the business cycle, and the examination of impulse response functions highlights significant responses of agricultural commodities to macroeconomic, financial and fundamental variables, with a particularly robust response to shocks in international competitiveness, as measured by changes in the real exchange rate.

ACKNOWLEDGEMENTS

The authors would like to thank an anonymous referee for very helpful comments and Rudolf Neubauer for research assistance. Funding from the European Union's Horizon 2020 research and innovation programme (H2020 Food) under Grant Agreement No. 633692 'Metrics, Models and Foresight for European SUStainable Food And Nutrition Security' (SUSFANS) is gratefully acknowledged.

DATA AVAILABILITY STATEMENT

The data that support the findings of this study are available from the corresponding author upon reasonable request.

ORCID

Jesus Crespo Cuaresma https://orcid.org/0000-0003-3244-6560

REFERENCES

Ahumada, H., & Cornejo, M. (2015). Explaining commodity prices by a cointegrated time series-cross section model. *Empirical Economics*, *48*(4), 1667–1690.

Ahumada, H., & Cornejo, M. (2016). Forecasting food prices: The case of corn, soybeans and wheat. *International Journal of Forecasting*, *32*(3), 838–848.

Bernard, J.-T., Khalaf, L., Kichian, M., & McMahon, S. (2008). Forecasting commodity prices: GARCH, jumps, and mean reversion. *Journal of Forecasting*, *27*(4), 279–291.

Brooks, C., & Prokopczuk, M. (2013). The dynamics of commodity prices. *Quantitative Finance*, *13*(4), 527–542.

Chen, Y.-C., Rogoff, K. S., & Rossi, B. (2010). Can exchange rates forecast commodity prices? *The Quarterly Journal of Economics*, *125*(3), 1145–1194.

Costantini, M., Crespo Cuaresma, J., & Hlouskova, J. (2016). Forecasting errors, directional accuracy and profitability of currency trading: The case of EUR/USD exchange rate. *Journal of Forecasting*, *35*(7), 652–668.

Crespo Cuaresma, J., Hlouskova, J., & Obersteiner, M. (2018). Fundamentals, speculation or macroeconomic conditions? Modelling and forecasting Arabica coffee prices. *European Review of Agricultural Economics*, *45*(4), 583–615.

Deaton, A., & Laroque, G. (1992). On the behaviour of commodity prices. *The Review of Economic Studies*, *59*(1), 1–23.

Deaton, A., & Laroque, G. (2003). A model of commodity prices after Sir Arthur Lewis. *Journal of Development Economics*, *71*(2), 289–310.

Degiannakis, S., Filis, G., Klein, T., & Walther, T. (2020). Forecasting realized volatility of agricultural commodities. *International Journal of Forecasting*. https://doi.org/10.1016/j.ijforecast.2019.08.011

Diebold, F. X., & Mariano, R. S. (1995). Comparing predictive accuracy. *Journal of Business & Economic Statistics*, *13*(3), 253–263. https://doi.org/10.1080/07350015.1995.10524599

Drachal, K. (2019). Analysis of agricultural commodities prices with new Bayesian model combination schemes. *Sustainability*, *11*(19), 5305.

Gargano, A., & Timmermann, A. (2014). Forecasting commodity price indexes using macroeconomic and financial predictors. *International Journal of Forecasting*, *30*(3), 825–843.

Husain, M. A. M., & Bowman, C. (2004). *Forecasting commodity prices: Futures versus judgment*: International Monetary Fund Working Paper.

Just, R. E., & Rausser, G. C. (1981). Commodity price forecasting with large-scale econometric models and the futures market. *American Journal of Agricultural Economics*, *63*(2), 197–208.

Kohzadi, N., Boyd, M. S., Kermanshahi, B., & Kaastra, I. (1996). A comparison of artificial neural network and time series models for forecasting commodity prices. *Neurocomputing*, *10*(2), 169–181.Financial Applications, Part I.

Ramirez, O. A., & Fadiga, M. (2003). Forecasting agricultural commodity prices with asymmetric-error GARCH models. *Journal of Agricultural and Resource Economics*, *28*(1), 71–85.

Rezitis, A. N., & Sassi, M. (2013). Commodity food prices: Review and empirics. *Economics Research International*, *2013*, 1–15.

AUTHOR BIOGRAPHIES

Jesus Crespo Cuaresma is a Professor of Economics at the Vienna University of Economics and Business (WU), as well as Director of Economic Analysis at the Wittgenstein Centre for Demography and Global Human Capital (WIC) and research scholar at the International Institute of Applied Systems Analysis (IIASA). His research interests are in the fields of applied econometrics, macroeconomics, economic growth, human capital and economic policy.

Jaroslava Hlouskova is a senior researcher at the Institute for Advanced Studies (research unit Macroeconomics and Business Cycles), Vienna, and holds a PhD in Mathematics from the Comenius University in Bratislava. Her research interests are in the field of forecasting, quantitative finance, and behavioural economics and finance.

Michael Obersteiner is Director of the Environmental Change Institute at Oxford University. He holds a PhD in Forest Economics from the University of Natural Resources and Life Sciences in Vienna. His main research interest lies in the (re)combination of science fields ranging from biology, earth system science, engineering, economics and finance.

How to cite this article: Crespo Cuaresma, J., Hlouskova, J., & Obersteiner, M. (2021). Agricultural commodity price dynamics and their determinants: A comprehensive econometric approach. *Journal of Forecasting*, *40*(7), 1245–1273. https://doi.org/10.1002/for.2768

APPENDIX A: DATA SOURCES AND ABBREVIATIONS

TABLE A1 Variable description

Variable	Abbreviation	Description	Source	Start date
Wheat spot price		All wheat (US season average), dollars per metric ton	USDA	1980:1
Soybean spot price		Soybeans (US), c.i.f. Rotterdam, dollars per metric ton	USDA, World Bank commodity price data (pink sheet)	1980:1
Corn spot price		Maize (US), no. 2, yellow, f.o.b. US Gulf ports, dollars per metric ton	USDA, World Bank commodity price data (pink sheet)	1980:1
Fundamental variables				
Wheat production for world	y_w^{world}	Million metric tons	USDA	1980:1
Wheat production for EU	y_w^{EU}	Million metric tons	FAO	1980:1
Wheat production for the United States	y_w^{US}	Million metric tons	USDA	1980:1
Soybean production for world	y_s^{world}	Million metric tons	FAO	1980:1
Soybean production for the United States	y_s^{US}	Million metric tons	FAO	1980:1
Soybean production for Brazil	y_s^{BR}	Million metric tons	FAO	1980:1
Corn production for world	y_c^{world}	Million metric tons	FAO	1980:1
Corn production for the United States	y_c^{US}	Million metric tons	FAO	1980:1
Corn production for China	y_c^{CH}	Million metric tons	FAO	1980:1
Wheat yield for world	$yield_w^{world}$	Metric tons per hectare	USDA	1980:1
Wheat yield for EU	$yield_w^{EU}$	Metric tons per hectare	FAO	1980:1
Wheat yield for the United States	$yield_w^{US}$	Metric tons per hectare	USDA	1980:1
Soybean yield for world	$yield_s^{world}$	Metric tons per hectare	FAO	1980:1
Soybean yield for the United States	$yield_s^{US}$	Metric tons per hectare	FAO	1980:1
Soybean yield for Brazil	$yield_s^{BR}$	Metric tons per hectare	FAO	1980:1
Corn yield for world	$yield_c^{world}$	Metric tons per hectare	FAO	1980:1
Corn yield for the United States	$yield_c^{US}$	Metric tons per hectare	FAO	1980:1
Corn yield for China	$yield_c^{CH}$	Metric tons per hectare	FAO	1980:1
Stock-to-use ratio for world and wheat	$stock_w^{world}$	Ending stocks over consumption	USDA FAS	1980:1
Stock-to-use ratio for world and soybeans	$stock_s^{world}$	Ending stocks over consumption	USDA FAS	1980:1

(Continues)

TABLE A1 (Continued)

Variable	Abbreviation	Description	Source	Start date
Stock-to-use ratio for world and corn	$stock_c^{world}$	Ending stocks over consumption	USDA FAS	1980:1
Macroeconomic variables				
World output	y^{world}	GDP (market prices 2010 USD), index	World Bank WDI	1980:1
		Data for 2016 were extrapolated using the World Bank forecasts		
		Indexed 2000:1 = 100. Seasonally adjusted.		
Output for EU	y^{EU}	Industrial production index for eurozone.	Datastream: EKESIMANG	1980:1
		Indexed 2000:1 = 100. Seasonally adjusted.		
Output for the United States	y^{US}	Industrial production index for the US	Datastream: USIPTOT.G	1980:1
		Indexed 2000:1 = 100. Seasonally adjusted.		
Output for Brazil	y^{US}	Industrial production index for Brazil	Datastream: BRIPTOT.G	1985:1
		Indexed 2000:1 = 100. Seasonally adjusted.		
Leading indicator for Germany	li^{EU}	Ifo: business climate index	Datastream: BDIFOIDXE	1980:1
Leading indicator for the United States	li^{US}	ISM: manufacturing index	Datastream: USCNFBUSQ	1980:1
Real effective exchange rate	REER	With respect to the USD (with respect to Brazilian real for coffee)	Bloomberg	1980:1
Financial variables				
Stock market index for EU	$stock^{EU}$	Index covers at least 80% of the market capitalization in the EMU	Datastream: TOTMKEM	1980:1
Stock market index for the United States	$stock^{US}$	Index covers at least 80% of the market capitalization in the United States	Datastream: TOTMKUS	1980:1
S&P Goldman Sachs Commodity Index	GSCI	A composite index of commodity sector returns representing an	Datastream: GSCITOT	1980:1
		unleveraged, long-only investment in commodity futures that is broadly		
		diversified across a spectrum of commodities		
Interest rate spread for EU	$spread^{EU}$	Difference between 10-year interest rate on government bonds	Datastream: EMBRYLD, EIBOR3M	1981:1
		and 3-month interbank rates		
Interest rate spread for the United States	$spread^{US}$	Difference between 10-year interest rate on government bonds	Datastream: BMUS10Y, BBUSD3M	1981:1
		and 3-month interbank rates		

TABLE A1 (Continued)

Variable	Abbreviation	Description	Source	Start date
Climatic variables				
Sea surface temperature anomalies	SSTA	The index measuring deviations between the sea surface temperatures in	NCDC, NOAA	1982:1
		the El Niño region 3.4 and its historical average		
Southern Oscillation Index anomalies	SOI	The index capturing fluctuations in air pressure occurring between	NOAA	1980:1
		the western and eastern tropical Pacific during El Niño and La Niña episodes		

Abbreviations: EMU, Economic and Monetary Union; EU, European Union; FAO, Food and Agriculture Organization of the United Nations; FAS, Foreign Agricultural Service; GDP, gross domestic product; NCDC, National Climatic Data Center; NOAA, National Oceanic and Atmospheric Administration; USDA, United States Department of Agriculture; WDI, World Development Indicators.

TABLE A2 Models and combination methods

Abbreviations	Description
	Individual models
RW	Random walk (benchmark model)
AR(p)	Autoregression in levels with p lags
DAR(p)	Autoregression in first differences with p lags
s-AR(p)	Subset autoregression in levels with p lags
s-DAR(p)	Subset autoregression in first differences with p lags
ARCH(p, q)	Autoregression conditional heteroskedasticity in levels with p lags in mean equation
	and q lags in variance equation
DARCH(p, q)	Autoregression conditional heteroskedasticity in first differences with p lags in mean equation
	and q lags in variance equation
GARCH(p, q)	Generalized autoregression conditional heteroskedasticity in levels with p lags in mean equation
	and q lags in variance equation
DGARCH(p, q)	Generalized autoregression conditional heteroskedasticity in first differences with p lags in mean equation
	and q lags in variance equation
VAR(p)	Vector autoregression in levels with p lags
DVAR(p)	Vector autoregression in first differences with p lags
VEC(c,p)	Vector error correction model with c cointegration relationships and p lags
s-VAR(p)	Subset vector autoregression in levels with p lags
s-DVAR(p)	Subset vector autoregression in first differences with p lags
BVAR(p)	Bayesian vector autoregression in levels with p lags
BDVAR(p)	Bayesian vector autoregression in first differences with p lags
	Forecast combination methods
mean	Forecasting combination based on mean of individual predictions
tmean	Forecasting combination based on trimmed mean of individual predictions
median	Forecasting combination based on median of individual predictions
OLS	Forecasting combination based on pooling using OLS
PC	Forecasting combination based on principal components
DMSFE	Forecasting combination based on discounted mean-square forecast errors
HR	Forecasting combination based on hit rates
EHR	Forecasting combination based on exponential of hit rates
EEDF	Forecasting combination based on the economic evaluation of directional forecasts
BMA	Forecasting combination based on Bayesian model averaging weights using the predictive likelihood
FMA-aic	Forecasting combination based on AIC weights
FMA-bic	Forecasting combination based on BIC weights
FMA-hq	Forecasting combination based on Hannan–Quinn weights

Abbreviations: AIC, Akaike information criterion; BIC, Bayesian information criterion; OLS, ordinary least squares.

TABLE A3 Weights of forecast combination methods

Method	Weights, $w^h_{m,it}$
Mean	$\frac{1}{k}$
Trimmed mean	$\frac{1}{k-2}$,
	where the smallest and largest forecasts are discarded
OLS	Coefficients from regressing actual values on forecasted values
PC	Coefficients from regressing actual values on factors
DMSFE	$\sum_{s=T_{1-1+h}}^{t} \theta^{T-h-s}\left(P_{m,s+h}-\hat{P}_{m,i,s+h\|s}\right)^2$,
	where $\theta = 0.95$ is a discount factor
HR	$\frac{\sum_{j=T_{1+h-1}}^{t} DA_{m,i,jh}}{\sum_{c=1}^{M}\left(\sum_{j=T_{1+h-1}}^{t} DA_{m,c,jh}\right)}$,
	where $DA_{m,c,jh} = I\left(\mathrm{sgn}(P_{m,j}-P_{m,j-h}) = \mathrm{sgn}(\hat{P}_{m,c,j\|j-h}-P_{m,j-h})\right)$
	and $I(\cdot)$ is the indicator function
EHR	$\frac{\exp\left(\sum_{j=T_{1+h-1}}^{t}(DA_{m,i,jh}-1)\right)}{\sum_{c=1}^{M}\exp\left(\sum_{j=T_{1+h-1}}^{t}(DA_{m,c,jh}-1)\right)}$
EEDF	$\frac{\sum_{j=T_{1+h-1}}^{t} DV_{m,i,jh}}{\sum_{c=1}^{M}\left(\sum_{j=T_{1+h-1}}^{t} DV_{m,c,jh}\right)}$,
	where $DV_{m,c,th} = \|P_{m,t}-P_{m,t-h}\| DA_{m,c,th}$
P	$(M_i)\sum_{l=1}^{k} P(\mathbf{S}_{T_{1+h-1:t}}\|M_l)P(M_l)$
BMA	$\frac{(t-T_{1-h+2})^{\frac{p_1-p_i}{2}}\left(\frac{\sum_{j=T_{1+h-1}}^{t} SE_{m,1,jh}}{\sum_{j=T_{1+h-1}}^{t} SE_{m,i,jh}}\right)^{\frac{t-T_{1-h+2}}{2}}}{\sum_{c=1}^{M}(t-T_{1-h+2})^{\frac{p_1-p_l}{2}}\left(\frac{\sum_{j=T_{1+h-1}}^{t} SE_{m,1,jh}}{\sum_{j=T_{1+h-1}}^{t} SE_{m,c,jh}}\right)^{\frac{t-T_{1-h+2}}{2}}}$,
	where $SE_{m,c,th} = \left(\hat{P}_{m,c,t\|t-h}-P_{m,t}\right)^2$
FMA	$\frac{\exp\left(-\frac{1}{2}IC_{it}\right)}{\sum_{c=1}^{M}\exp\left(-\frac{1}{2}IC_{ct}\right)}$,
	where IC_{ct} is the information criterion of model c
	and t is the last time point of the data over which are models estimated