The decomposition of agricultural commodity markets volatility between fundamentals and market speculation

Deliverable No. 8.1

SUSFANS DELIVERABLES

Jesus Crespo Cuaresma, Jaroslava Hlouskova, and Michael Obersteiner (IIASA)

Abstract We analyse the role played by market fundamentals, speculation and macroeconomic conditions as empirical determinants of commodity price changes. We combine model averaging techniques to explain historical patterns with an in-depth analysis of out-of-sample predictability of commodity prices using fundamentals as well as macroeconomic and financial variables. Our results indicate that variables related to global macroeconomic and financial developments contain valuable information to explain the historical pattern of coffee price developments, as well as to improve out-of-sample predictions of coffee prices.

Version	Release date	Changed	Status	Distribution
V1	30/09/2016	-	final	

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 633692 to the sum of €5M over 2015 to 2019.

SUSFANS DELIVERABLE

Document information

Project name	SUSFANS
Project title:	Metrics, Models and Foresight for European SUStainable Food And Nutrition Security
Project no	633692
Start date:	April 2015
Report:	D8.1 Fundamentals, Speculation or Macroeconomic Conditions? On the Determinants of Commodity Price Dynamics, with an Application to Arabica Coffee
Work package	WP 8
WP title (acronym):	
WP leader:	IIASA, Jesus Crespo Cuaresma
Period, year:	1, 2016
Responsible Authors:	
Participant acronyms:	IIASA
Dissemination level:	**Public**
Version	V1
Release Date	September 30, 2016
Planned delivery date:	September 30, 2016
Status	final
Distribution	

Dissemination level of this report

Public

ACKNOWLEDGMENT & DISCLAIMER

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 633692.
Neither the European Commission nor any person acting on behalf of the Commission is responsible for how the following information is used. The views expressed in this publication are the sole responsibility of the author and do not necessarily reflect the views of the European Commission.

1 Introduction

There is an emerging consensus that the world is experiencing increasingly volatile food prices leading to considerable impacts on food security. Rapidly rising food prices during the last commodity market super-cycle have sparked riots, contributed to political instability and caused the number of malnourished people to rise to 1 billion. World food markets have experienced greater volatility than at any time since the 1970s, a sign that supply is struggling to keep up with demand or other systemic drivers might be at work. A number of recent studies have attempted to explain the volatility and potential level shift in global food prices. They have identified exacerbating factors such as trade restrictions, financial speculation, currency movements and lower food reserves. They have pointed to growing demand caused by an expanding population and bio-fuels policies, as well as supply constraints caused by rising energy costs, water depletion, extreme weather events, ever increasing competing pressures on land use and under-investment. All of these have played a part, however, little work has been done to quantify and understand the interaction of these drivers. Furthermore, no distinction has been made between long-run pressure variables and short-run shocks.

Changes in aggregate demand driven by growth in emerging economies (in particular China) have been often claimed responsible for trends in commodity prices over the last decades (see Cheung et al., 2007, for example). Speculation with storable commodities which are traded on futures markets has also been put forward as an explanatory factor for commodity price dynamics (see Gilbert, 2006).

Here we focus on coffee - one of the internationally most traded agricultural commodities with an estimated total consumption value of some 174 billion dollars in 2012.[1] Based on the trade statistics data of the International Coffee Organization (ICO) the largest producer and exporter in 2015 is Brazil, followed by Vietnam, Colombia, Indonesia and Ethiopia.[2]

The world coffee market has undergone a significant transformation over the last 50 years. The coffee market was regulated, up until 1989, by a series of International Coffee Agreements which were intended to manage supply and maintain price stability. This system subsequently collapsed, and since 1990 the coffee market has been subject to free market forces. Price levels during the regulated market period (1965 to 1989) were relatively high since both upward and downward trends were corrected through the application of export quotas. The free market period beginning in 1990 had two sub-periods of markedly low price levels: 1990 to 1993 and 1999 to 2004 (see Figure 1). The latter sub-period (known as the *coffee crisis* period) was the longest period of low prices ever recorded with severe negative consequences on the economies of exporting countries. Prices recovered strongly after 2004, reaching a 34-year high in mid-2011. However, there has subsequently been a severe deterioration in prices while costs of coffee production inputs, particularly fertilizers and labor, continued to rise. These price increases were in part

[1]See Pendergrast (2010) which is often referred to as as the 'Bible of coffee'.
[2]http://www.ico.org/prices/po-production.pdf

driven by higher expenditures for pesticides to combat emerging large scale diseases attacking coffee plantations and increasing fertilizer prices both squeezing the margins for labor inputs.

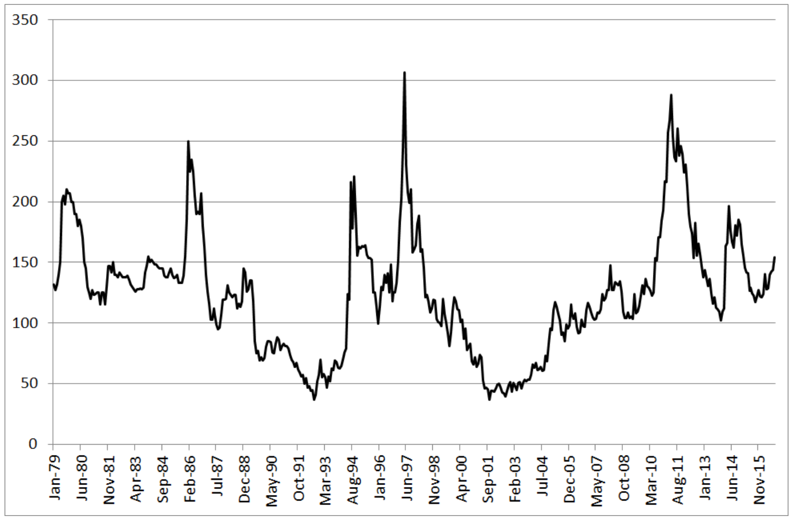

Figure 1: Arabica (Brazilian) coffee price (Cents/lb).

During the regulated market period the highest volatility was recorded in years following severe climate shocks recorded in exporting countries, notably in Brazil in 1975 and 1985. The highest volatility levels are generally recorded for the months of May, June, July and August, since they cover the period of possible frosts in Brazil, thought to fuel speculative activity. During the free market high volatility was recorded in 1994 and 1997 (see Figure 2) where in 1994 a climate shock was recorded in Brazil.

Prices in the futures markets have been significantly more volatile than the yearly price indicator recorded by the International Coffee Organization. Futures prices contain additional information beyond the fundamentals of the market - production, trade, stocks and consumption. While the fundamentals are strongly discussed by the communities of producers and traders the impact of macroeconomic and financial variables is typically underestimated or entirely ignored. Also the academic literature on this topic seems to be void of appropriate methods and analyses. Here we apply a large set of models to coffee market data including fundamental, macroeconomic, financial and climatic explanatory variables with the aim to arrive at a better understanding of drivers' contributions to coffee price phenomena.

In addition to studies which concentrate on assessing the historical determinants of commodity price dynamics, a series of contributions focus on addressing out-of-sample predictability. Numerous existing studies aim at evaluating the predictive power of commodity futures prices for actually realized spot prices, as well as exploiting the information of macroeconomic and financial variables as leading indicators of commodity price dynamics. Husain and Bowman (2004), for instance, analyse 15 commodities and

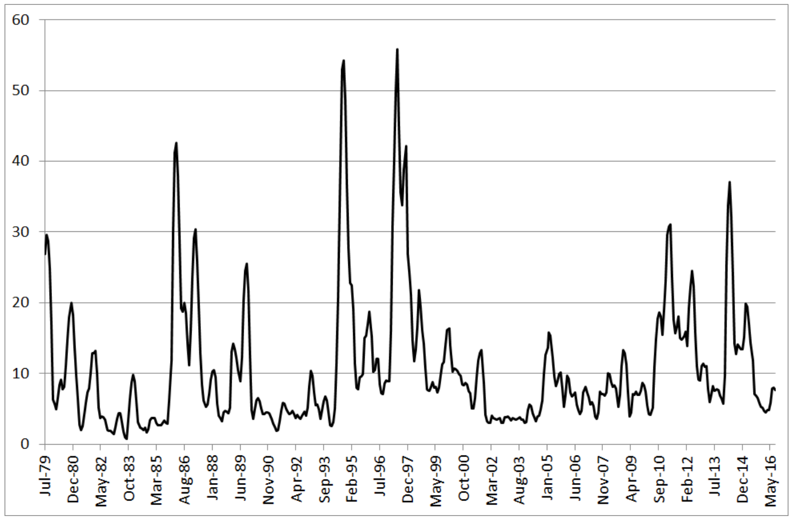

Figure 2: Volatility of Arabica (Brazilian) coffee price (based on 6-months rolling window).

conclude that statistical models based on futures yield better results in terms of predictive ability than those based exclusively on spot price dynamics or on judgement.

In this contribution, we explicitly assess the role that different theoretical driving factors of coffee prices play as predictors of their dynamics. We explicitly address specification uncertainty using forecast averaging methods which aggregate predictions from different models using different techniques to integrate the uncertainty over specification choice. We entertain individual models that contain information about climate, global macroeconomic developments and speculation. Making use of forecast pooling techniques that account for model uncertainty, we are also able to assess quantitatively the differences in predictive ability of competing explanatory factors. Our results indicate that variables related to global macroeconomic and financial developments contain valuable information to explain the historical pattern of coffee price developments, as well as to improve out-of-sample predictions of coffee prices.

2 Explaining historical commodity price dynamics: The case of Arabica coffee

2.1 Model uncertainty and commodity price dynamics: An in-sample analysis

In order to analyse the relative importance of different potential explanatory factors of the dynamics of coffee prices, we start by entertaining models of the class of autoregressive distributed lags. The specifications we consider are of the form

$$\Delta P_{m,t} = \alpha_m + \sum_{k=1}^{q_{m0}} \theta_m \Delta P_{m,t-k} + \sum_{i=1}^{v_m} \sum_{j=1}^{q_{mv}} \phi_{mij} x_{m,i,t-j} + \varepsilon_{mt}, \tag{1}$$

where $\Delta P_{m,t}$ is the (annual) log-change in the price of commodity m, which is assumed to be explained by its own lags and by lags of a set of variables $\{x_{mit}\}_{i=1}^{v_m}$ and a random normally distributed shock, ε_{mt}, assumed to fulfil the standard assumptions of linear regression specifications.

We address uncertainty about the specification of the model (in the sense of covariate inclusion) by using Bayesian model averaging (BMA) techniques to carry out inference in the class of models given by equation (1).[3] The extensive number of candidate variables proposed in the literature as candidates to enter the model presented in equation (1) implies that model uncertainty may lead to flawed inference if it is not explicitly assessed. Instead of basing our inference on a particular selected model, we learn about the drivers of commodity prices using a weighted average of single regressions. In the Bayesian framework, the natural weighting scheme is based on the corresponding *posterior model probabilities* (PMP) of the individual specifications. In particular, if we are interested in performing inference in a quantity χ, which could be a parameter of the model, a combination of parameters or a predicted value of the dependent variable, then the posterior probability over χ can be obtained as

$$p(\chi|y) = \sum_{l=1}^{2^K} p(\chi|M_l, y) p(M_l|y), \tag{2}$$

with $p(\cdot|y)$ denoting posterior distributions (that is, conditional on the data, y) and $p(\cdot|M_l, y)$ denoting posterior distributions conditional on the structure given by model M_l.[4] Assuming that v potential independent variables are available and that up to q lags are allowed to enter the specification, the cardinality of the model space based on equation (1) is given by $K = 2^{(v+1)q}$, which corresponds to the number of models that can be built by combining these covariates and lags in addition to the autoregressive terms.

Bayesian reasoning allows us to write the posterior model probabilities in equation (2) as proportional to the product of the marginal likelihood of the corresponding model, $p(y|M_l)$ and the model prior $p(M_l)$,

$$p(M_l|y) \propto p(y|M_l) p(M_l). \tag{3}$$

[3] See Koop (2003) for an introduction to BMA techniques.

[4] To easy the exposition we drop the commodity index m for now.

Obtaining PMPs implies that prior distributions need to be elicited on the parameters of the models that can be formed by combining the covariates and lags, as well as on the variance of the error term, σ^2. Following the literature on BMA for linear models, improper priors are placed on the intercept $p(\alpha) \propto 1$ and variance $p(\sigma) \propto \sigma^{-1}$, reflecting lack of prior subjective information about these quantities. For the rest of the parameters in a given specification within the class of models described by equation (1), we follow the standard convention in BMA and use Zellner's g prior ((Zellner, 1986)),

$$\phi_{ij} | \left(\sigma^2, M_l, g\right) \ \sim \ N\left(0, \sigma^2 g (X_l' X_l)^{-1}\right), \tag{4}$$

where X_l is the matrix of observations of the independent variables included in model M_l. Characteristic choices of the parameter g are T, the number of observations (unit information prior, UIP), proposed by Kass and Wasserman (1995) and K^2, (the risk inflation criterion, RIC) put forward by Foster and George (1994). Fernández et al. (2001a) propose $g = \max(T, K^2)$ (BRIC prior) after comparing the performance of the UIP and RIC priors in simulated settings.[5]

The prior probabilities assigned to individual models, $p(M_l)$, allow the researcher to include prior beliefs about the relative adequacy of the different specifications nested in the class of models given by 1. Following Ley and Steel (2009), in our application we use a binomial-beta prior for inclusion of a given variable with a prior expected model size of $K/2$ regressors. Such a prior over the model space is uninformative about model size.

2.2 The determinants of coffee price dynamics

In order to assess the nature of the factors affecting changes in Arabica coffee prices, we start by applying BMA using data on variables which we divide into four thematic groups: (i) *fundamental variables* (coffee production in Brazil, y^{BR}_{coffee}, world coffee production, y^{world}_{coffee}), (ii) *macroeconomic variables* (output for Brazil, y^{BR}, output for the EU, y^{EU}, output for the US, y^{US}, leading indicator for Germany, li^{EU}, leading indicator for the US, li^{US}, real effective exchange rate, REER), (iii) *financial variables* (stock market index for the EU, $stock^{EU}$, stock market index for the US, $stock^{US}$, S&P Goldman Sachs commodity index, GSCI) and (iv) *other climatic and meteorological variables* (precipitation, temperature of the area in Brazil where Arabica coffee is grown). We employ monthly data spanning the period from April 1990 until March 2016. The description of the variables and source of the data can be found in Table 1.

[5] Alternatively, a hierarchical structure can be imposed by defining a prior on g, as put forward by Liang et al. (2008), Feldkircher and Zeugner (2009) or Ley and Steel (2012).

Table 1: Variable description

Variable	Description	Source	Start date
Coffee price	Coffee-Brazilian (Arabica), (NY) Cents/lb	Datastream: COFBRAZ	1980:1
Fundamental variables			
Coffee production in Brazil	In thousand 60kg bags	International Coffee Organization	1990:4
Coffee production in the world	In thousand 60kg bags	International Coffee Organization	1990:4
Marcoeconomic variables			
Output for Brazil	Industrial production index. Indexed 2000:1=100. Seasonally adjusted.	Datastream: BRIPTOT.G	1985:1
Output for EU	Industrial production index for Eurozone. Indexed 2000:1=100. Seasonally adjusted.	Datastream: EKESIMANG	1980:1
Output for US	Industrial production index for the US Indexed 2000:1=100. Seasonally adjusted.	Datastream: USIPTOT.G	1980:1
Leading indicator for Germany	IFO: Business climate index	Datastream: BDIFOIDXE	1980:1
Leading indicator for the US	ISM: Manufacturing index	Datastream: USCNFBUSQ	1980:1
Real effective exchange rate	with respect to Brazil	Bloomberg	1980:1
Financial variables			
Stock market index for EU	Index covers at leats 80% of the market capitalization in the EMU	Datastream: TOTMKEM	1980:1
Stock market index for the US	Index covers at leats 80% of the market capitalization in the US	Datastream: TOTMKUS	1980:1
S&P Goldman Sachs commodity index	A composite index of commodity sector returns representing an unleveraged, long-only investment in commodity futures that is broadly diversified across a spectrum of commodities	Datastream: GSCITOT	1980:1
Climatic variables			
Precipitation	Mean precipitation of the Arabica coffee region (mm/pentad)	Google Earth Engine for Brazil (Arabica coffee region)	1981:1
Temperature	Mean value of daytime land surface temperature for each aggregated pixel, in degrees Celsius	Google Earth Engine for Brazil (Arabica coffee region)	1980:1

We apply BMA using the class of models defined by equation (1) and the set of variables described in Table 1, after transforming trending and seasonal variables when necessary by using annual changes. We employ the BRIC prior for the parameters corresponding to the covariates and their lags and the beta-binomial described above to define prior model probabilities. In order to overcome the computational constraints that are given by the large cardinality of the model space, we employ Markov Chain Monte Carlo (MCMC) methods proposed by Kass and Wasserman (1995) to explore the model space. The results presented in this section are based on two million model draws after a burn-in of one million draws.

We start by presenting results that correspond to models without an autoregressive term, that is, based on specifications of the form

$$\Delta P_{m,t} = \alpha_m + \sum_{i=1}^{v_m} \phi_{mi} x_{m,i,t-1} + \varepsilon_{mt}, \tag{5}$$

By abstracting away from modelling the persistence of the commodity price series, the analysis may be polluted by identifying partial correlation structures which are driven by common persistence patterns across variables. However, it serves as a first approach to pinpoint the robustness of partial correlations between our dependent variable and the lagged variables entertained in the analysis. The first column of Table 2 shows the posterior inclusion probability (PIP) of each one of the covariates considered within the set of potential determinants. The PIP is defined as the sum of posterior model probabilities of specifications containing a particular variable, and is routinely interpreted as a measure of the importance of that covariate as a robust determinant of changes in the dependent variable under model uncertainty (see for example Fernández et al., 2001b; Ley and Steel, 2009). Our prior elicitation implies an expected prior inclusion probability of 0.5 for the variables considered, so we will label variables with PIP above 0.5 as *robust covariates*, since the confidence on model inclusion increases after observing the data.

While the results of the BMA exercise which explores models without an autoregressive component unveil robust partial correlations for covariates belonging to the group of fundamental, macroeconomic and financial variables (see first column in Table 2), most of these factors lose their robustness once the persistence of commodity prices is explicitly included in the model. In the second column of Table 2 we present the PIPs of the variables based on entertaining models that include the lagged dependent variable (with a one-month lag) in addition to the rest of the lagged covariates. Modelling the persistence of commodity prices leads to only two variables besides the lag of the dependent variable having PIP above 0.5. These covariates (US leading indicator and the yearly log-change in the US stock market index) reflect global macroeconomic and financial developments. The corresponding commodity market fundamentals (global coffee production or production in Brazil) lose their relevance as determinants of coffee price dynamics under model uncertainty once the persistence of this variable is modelled through a lag.

Table 2: BMA analysis: Posterior inclusion probabilities

	$q_{m0}=0$ $q_{mv}=1$	$q_{m0}=1$ $q_{mv}=1$	$q_{m0}=2$ $q_{mv}=2$
Coffee production in Brazil	**1.00**	0.02	0.01
Coffee production in the world	**0.81**	0.02	0.01
Output for EU	**0.51**	0.02	0.01
Output for US	0.14	0.02	0.23
Output for Brazil	**0.99**	0.03	0.14
Leading indicator, Germany	**0.64**	0.06	0.01
Leading indicator, US	**1.00**	**0.73**	0.01
Real effective exchange rate	**1.00**	0.03	0.01
Stock market index, EU	0.10	0.04	0.01
Stock market index, US	**1.00**	**0.59**	0.07
S&P GS commodity index	0.27	0.03	0.01
Precipitation	0.10	0.02	0.03
Temperature	0.12	0.02	0.01
Lagged dependent variable	–	**1.00**	**1.00**

PIPs based on the class of models defined by equation (1). Bold figures if PIP>0.5
PIPs in column "$q_{m0}=2$ and $q_{mv}=2$" correspond to the lag with maximum PIP.
Dependent variable is the annual log-change in Arabica coffee price.
Results based on two million MCMC replications after one million burn-in draws.

Finally, in the third column of Table 2 we enlarge the number of potential lags of the dependent and independent variables included in the model to two, thus allowing for more complex dynamic relationships between changes in the commodity price and its determinants. For each variable, the figures in Tale 2 present the maximum PIP for the two lags included. Once that more complex autoregressive dynamics are allowed for in the specification, no single covariate besides the lagged dependent variables achieve PIP above the implied prior expectation. Such a result is a reflection of the difficulty of finding individual drivers that are able to explain historical changes in commodity prices.

On the one hand, our results emphasize the role of global macroeconomic developments and financial markets as drivers of coffee prices. However, the large degree of model uncertainty renders this result unrobust once slightly more complex autoregressive dynamics are assumed for commodity prices. Although historical in-sample dynamics may be difficult to assess with the covariates proposed, they may still contain information which is useful for out-of-sample forecasting of commodity price changes. In the following section we provide a comprehensive assessment of the out-of-sample predictive power of fundamental, macroeconomic, financial and climatic variables for commodity prices, applying the analysis to Arabica coffee prices.

3 Out-of-sample predictability analysis

3.1 The econometric setting: Prediction models and forecast averaging techniques

In order to provide a comprehensive analysis of the predictive power of the variables put forward above for Arabica coffee prices, we consider a battery of univariate and multivariate model structures as potential prediction models and perform a systematic comparison of their out-of-sample forecasting power.[6] We consider a large number of univariate and multivariate models as well as forecast combination methods for variables corresponding to the different categories described above (fundamentals, macroeconomic, financial and climatic variables). The particular models and combination methods used for the analysis are presented in Table 3 and a full description of the forecast evaluation exercise can be found in Crespo Cuaresma et al. (2016).

The assessment of predictive ability is based on a series of profit/cost measures. Denoting $\hat{P}_{m,c,t+h|t}$ the forecast of the price of commodity m for time $t+h$ conditional on the information available at time t obtained by model or forecast combination method c, $c = 1, \ldots, M$, the loss measures we evaluate include the standard square forecast error,

$$SE_{m,c,t,h} = \left(\hat{P}_{m,c,t|t-h} - P_{m,t}\right)^2 \tag{6}$$

and the absolute error

$$AE_{m,c,t,h} = \left|\hat{P}_{m,c,t|t-h} - P_{m,t}\right|, \tag{7}$$

which are standard loss measures in assessments of forecasting models for continuous variables.

Denoting T_3 the end of the available in-sample period and T_2 the beginning of the out-of-sample period, the statistics of interest based on these two measures of predictive error are the *mean square error* (MSE) at horizon h

$$MSE_{m,c,h} = \frac{1}{T_3 - T_2 + 1} \sum_{j=0}^{T_3 - T_2} SE_{m,c,T_2+j,h} \tag{8}$$

and the *mean absolute error* (MAE) at horizon h,

$$MAE_{m,c,h} = \frac{1}{T_3 - T_2 + 1} \sum_{j=0}^{T_3 - T_2} AE_{m,c,T_2+j,h}. \tag{9}$$

In our forecast analysis we also use composite forecasts based on the relative performance of particular methods over certain out-of-sample periods. In particular, for this technique at each time point t we

[6] See Crespo Cuaresma et al. (2016) for a detailed description of the comparison exercise.

Table 3: Models and combination methods

Abbreviations	Description
	Individual models
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

choose the model or forecast combination method (and thus also the forecast for time point $t+h$) with the best performance (i.e. minimum MSE and/or MAE) over a certain time window ending at time point t. Namely,

$$\hat{P}^{MSE,l}_{m,t+h|t} = \hat{P}_{m,c^{MSE}_{mlth},t+h|t} \quad \text{where} \quad c^{MSE}_{mlth} = \text{argmin}_c \sum_{j=l}^{t} SE_{m,c,j,h}. \tag{10}$$

Time point l, such that $T_2 \leq l \leq t$, defines the beginning of the window over which the performance is evaluated, i.e., the evaluation window is $[l, t]$ where $l \leq t \leq T_3$. In a similar way

$$\hat{P}^{MAE,l}_{m,t+h|t} = \hat{P}_{m,c^{MAE}_{mlth},t+h|t} \quad \text{where} \quad c^{MAE}_{mlth} = \text{argmin}_c \sum_{j=l}^{t} AE_{m,c,j,h}. \tag{11}$$

In terms of profit measures, we use directional accuracy (*DA*), directional value (*DV*), the returns from a trading strategy generated by our forecasts and a risk adjusted performance measure given by the Sharpe ratio.

The DA measure is given by

$$DA_{m,c,t,h} = I\left(\mathsf{sgn}(P_{m,t} - P_{m,t-h}) = \mathsf{sgn}(\hat{P}_{m,c,t|t-h} - P_{m,t-h})\right) \quad (12)$$

where $I(\cdot)$ is the indicator function. $DA_{m,c,th}$ is thus a binary variable indicating whether the direction of the price change was correctly forecast at horizon h ($DA_{m,c,t,h} = 1$) or not ($DA_{m,c,t,h} = 0$). The economic value of directional forecasts is better captured by assigning to each correctly predicted change its magnitude (see Blaskowitz and Herwartz, 2011). We use the directional value (DV) statistic for this purpose,

$$DV_{m,c,t,h} = |P_{m,t} - P_{m,t-h}| DA_{m,c,t,h}. \quad (13)$$

We entertain composite forecasts based on forecasts from all models and forecast combination methods. 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 largest DA or DV over certain time window ending at time point t. That is,

$$\hat{P}^{DA,l}_{m,t+h|t} = \hat{P}_{m,c^{DA}_{mlth},t+h|t} \qquad \text{where} \qquad c^{DA}_{mlth} = \mathsf{argmax}_c \sum_{j=l}^{t} DA_{m,c,j,h}, \quad (14)$$

where l, $T_2 \leq l \leq t$, defines the beginning of the window over which is the performance evaluated, i.e., the evaluation window is $[l, t]$ where $l \leq t \leq T_3$. In a similar way

$$\hat{P}^{DV,l}_{m,t+h|t} = \hat{P}_{m,c^{DV}_{mlth},t+h|t} \qquad \text{where} \qquad c^{DV}_{mlth} = \mathsf{argmax}_c \sum_{j=l}^{t} DV_{m,c,j,h}. \quad (15)$$

The performance of commodity price forecasts based on their profitability is also evaluated by the returns or Sharpe ratios implied by a simple trading strategy that is based on predictions. Selling/buying signals are based on the difference between the current spot price and the forecast for horizon h. Positive returns are executed as long positions while negative returns are executed as short positions (see for example Gencay, 1998). The (discrete) return of the spot price for commodity m at time t over period h is $r_{m,th} = P_{m,t}/P_{m,t-h} - 1$. If the trading signal implied my model or model combination c for commodity m at time t is given by

$$y_{m,c,t-h,h} = \begin{cases} -1, & \text{for selling signal (forecast downward movement for horizon } h) \\ & \hat{P}_{m,c,t|t-h} < P_{m,t-h}, \\ & \\ 1, & \text{for buying signal (forecast upward movement for horizon } h) \\ & \hat{P}_{m,c,t|t-h} > P_{m,t-h}, \end{cases}$$

then the return of the trading strategy (at time t over period h) implied by model c is $R_{m,c,th} = y_{m,c,t-h,h} r_{m,th}$ for $t = 1, \ldots, n$, and the total return of the trading strategy over n periods, i.e., over interval $[t, t+n]$, implied by model c and with respect to all realized $h-$period returns ($h \leq n$), is given by

$$R_{m,c,h,[t,t+n]} = \frac{1}{h} \sum_{j=0}^{h-1} \Pi_{i=0}^{n_j} \left(R_{m,c,t+j+ih,h} + 1 \right) - 1 \tag{16}$$

where n_j, $j = 1, \ldots, h-1$, is the largest integer such that $t + j + n_j h \leq n$.[7]

As in the previous cases, we create an aggregate/composite forecast with the maximum averaged or realized return – based on forecasts from all models and forecast combination methods. I.e., 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 largest average return over time window $[l, t]$, namely

$$\hat{P}^{TS,l}_{m,t+h|t} = \hat{P}_{m,c^{TS}_{mlth},t+h|t} \quad \text{where} \quad c^{TS}_{mlth} = \text{argmax}_c \sum_{k=l}^{t} R_{m,c,k,h} \tag{17}$$

and the largest total realized return until time point t, namely

$$\hat{P}^{TS}_{m,t+h|t} = \hat{P}_{m,c^{TS}_{mth},t+h|t} \quad \text{where} \quad c^{TS}_{mth} = \text{argmax}_c R_{m,c,h,[1,t]}. \tag{18}$$

We also perform comparisons based on *Sharpe ratios* - the excess return per unit of deviation generated by a trading strategy. In our application we take zero return as a benchmark return in the definition of the Sharpe ratio.

The forecast averaging methods employed use different weights, with some of the schemes using the predictive ability of each one of the specifications to compute them. Starting with the simplest methods, forecast pooling based on the *mean* uses the average of the forecasts of the individual models. The

[7]Note that for $h = 1$ is the total return over $[t, t+n]$ given by $\Pi_{i=0}^{n} \left(R_{m,c,t+i,h} + 1 \right) - 1$. Note in addition that the total return given by equation (16) is the average of all possible $h-$period returns. We decided to proceed this way so as to take into account all $h-$step ahead forecasts.

trimmed mean method uses the same type of weighting after discarding the lowest and highest forecast generated by the set of models considered. The *median* combination method uses the median of the predictions produced by the battery of specifications entertained.

Granger and Ramanathan (1984) propose to use weights based on the parameter estimates obtained from regressing the actual realizations in a hold-out sample on the corresponding forecasts from the individual models. We denote this combination method *OLS*. To avoid potential problems caused by multicolinearity, we also use a similar forecast pooling method based on building OLS weights based on the principal components of the model-specific forecasts instead of the individual set of predictions (*PC*). Stock and Watson (2004) put forward a forecast combination technique based on discount mean square forecast errors (*DMSFE*) which corresponds to using weights in equation (19) which depend inversely on the discounted squared forecast errors obtained in the hold-out sample for each model. Such a discounting scheme implies that the recent predictive performance of the individual models is considered more relevant for this weighting strategy.

We also use a combination method based on the proportion of correctly predicted directions of change in the commodity price by model i (the hit rate, *HR*), as well as a pooling strategy based on the exponential of the hit rate (*EHR*), a method put forward, for instance, in Bacchini et al. (2010). While these methods base the weight of the individual specifications on their ability to predict direction of change, we can also construct weights based on the economic evaluation of directional forecasts (*EEDF*), that is, taking into account the magnitude of the realized change in the commodity price of interest. In this case, the weights are built using the relative performance of the individual models in terms of the variable created by multiplying the absolute change in the commodity price by a variable that takes value one if the direction of change was forecast adequately and zero otherwise.

Bayesian model averaging (*BMA*) techniques provide a framework which can be used to construct weights for pooling forecasts. In the spirit of weighting based on posterior model probabilities, weights for the individual models can be obtained making use of the Laplace approximation of the marginal likelihood of each model evaluated using the out-of-sample forecast errors, as proposed by Kapetanios et al. (2006). While the Laplace approximation of the marginal likelihood relies on the use of the Bayesian Information Criterion (BIC), frequentist approaches also propose the use of the Akaike Information Criterion (*AIC*) or the Hannan-Quinn Information Criterion (HQ) as alternatives to the BIC when building the model averaging weights see (see, for instance, Claeskens et al., 2008).

The pooled forecast methods considered in this analysis build linear combination of the predictions of individual specifications,

$$\hat{P}_{m,c,t+h|t} = w^h_{m,c,0t} + \sum_{i=1}^{F} w^h_{m,c,it} \hat{P}_{m,i,t+h|t}, \tag{19}$$

where c is the combination method, F is the number of individual forecasts and the weights are given by $\{w^h_{m,c,it}\}_{i=0}^F$. Table 4 presents the exact definition of the weights corresponding to each one of the methods entertained.[8]

3.2 Out-of-sample results on Arabica coffee price

We base our comparisons on monthly data spanning the period from April 1990 until March 2016 for Arabica coffee. The beginning of the hold-out forecasting sample for individual models used in order to obtain weights based on predictive accuracy is given by January 2000. The beginning of the actual out-of-sample forecasting sample is January 2005, and the end of the data sample is March 2016. The lag length of all multivariate model specifications under consideration is selected using the AIC criterion for potential lag lengths ranging from 1 to 6 lags. For the VEC models, selection of the lag length and the number of cointegration relationships is carried out simultaneously using the AIC. We also estimate subset-VAR specifications, where individual parameters of the VAR specification are set equal to zero recursively using t-tests.

In a first stage, the forecasting exercise is performed for groups of variables corresponding to each one of the groups, in order to assess the relative performance of each one of the potential types of determinants of commodity price dynamics. Table 5 summarizes these results (from forecast performance analysis) executed for each individual group and forecast horizons of one, three, six, nine and twelve months. The results identify fundamental variables as containing predictive information for short-term movements in coffee prices, while macroeconomic and financial variables appear important in terms of predictability at longer horizons. The variables included in the best performing models at each forecasting horizon are:

- Forecast horizon of one month: coffee production in Brazil and the world, output for the US, real effective exchange rate, stock market indices for the EU and the US, S&P Goldman Sachs commodity index, precipitation and temperature (y^{BR}_{coffee}, y^{world}_{coffee}, y^{US}, REER, $stock^{EU}$, $stock^{US}$, GSCITOT, precipitation, temperature);
- Forecast horizon of three months: coffee production in Brazil and the world, output for the US and Brazil, leading indicators for both Germany and the US, real effective exchange rate, stock market indices for the EU and the US and temperature (y^{BR}_{coffee}, y^{world}_{coffee}, y^{US}, y^{BR}, li^{EU}, li^{US}, REER, $stock^{EU}$, $stock^{US}$, temperature);
- Forecast horizon of six months: coffee production in Brazil and the world, output for the US and Brazil, leading indicators for both Germany and the US, real effective exchange rate, stock market

[8]We use also the median of forecasts, i.e., $\hat{P}_{m,\text{median},t+h|t} = \text{median}\{\hat{P}_{m,c,t+h|t}\}_{c=1}^M$, which can not be expressed by (19).

Table 4: 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 $DA_{m,c,jh} = I\left(\text{sgn}(P_{m,j} - P_{m,j-h}) = \text{sgn}(\hat{P}_{m,c,j\|j-h} - P_{m,j-h}) \right)$ 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 $DV_{m,c,th} = \|P_{m,t} - P_{m,t-h}\| DA_{m,c,th}$
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

indices for the EU and the US and temperature (y^{BR}_{coffee}, y^{world}_{coffee}, y^{US}, y^{BR}, li^{EU}, li^{US}, REER, $stock^{EU}$, $stock^{US}$, temperature);

- Forecast horizon of nine months: coffee production in Brazil and the world, output for Brazil, leading indicators for both Germany and the US and stock market index for the EU (y^{BR}_{coffee}, y^{world}_{coffee}, y^{BR}, li^{EU}, li^{US}, $stock^{EU}$);

- Forecast horizon of twelve months: output for Brazil, leading indicators for both Germany and the US, real effective exchange rate and stock market index for the EU (y^{BR}, li^{EU}, li^{US}, REER, $stock^{EU}$).

Using these variables, we perform the forecasting exercise again for each forecasting horizon, now mixing across groups of potential determinants. The findings of this exercise are summarized in Table 6. The performance results based on one month forecast horizon improve when comparing them to the performance for individual groups for one month ahead forecasts for profit measures (DA, DV, return, Sharpe ratio), while for loss measures they slightly decline. In the case of six months forecast horizon both loss and profit measures outperform the results implied by individual groups. The combination of variables (with highest performance power) from different groups helps to improve the performance for three and six months forecast horizons. In more detail, for three months forecast horizon the improvement in performance is given by the combination of fundamental, macro and financial variables for the MAE; fundamental and macro variables for the MSE; macro and financial variables for directional accuracy and the Sharpe ratio; fundamental, macro and other variables for directional value; and variables from all four groups for the returns implied by the trading strategy. For six months forecast horizon the performance improvement is implied by the combination of: macro and financial variables for loss measures; fundamental and macro variables for directional accuracy and directional value; fundamental macro and other variables for return implied by the 'buy low sell high' trading strategy and fundamental, macro and financial variables for Sharpe ratio. The performance for forecast horizons of nine and twelve months coincides with the best performance (for forecast horizons of nine and twelve months, see Table 5) implied by the group of macro variables and fundamental variables (horizon of nine months) and by the group of macro variables and financial variables (horizon of twelve months). This implies that the larger group of variables which encompasses variables (with best predictive power) of individual groups does not add any additional value to the performance and thus, the best performance is implied by variables that belong into certain (individual) groups, namely by macro variables (horizon of nine months) and macro and financial variables (horizon of twelve months).

The smallest forecast errors, MSE and MAE, were obtained for the one month forecast horizon (similar result as in the groups based analysis, see Table 5), the largest directional accuracy (71.1%) was reached for six months forecast horizon (the largest DA value in the group based analysis was 64.4% for twelve month forecast horizon), the largest directional value (78.6%) and the Sharpe ratio were achieved for the nine months forecast horizon, finally, the biggest improvement was observed in returns for three months forecast horizon from 3.2% achieved in the group based analysis to 9.7%.

Similarly, as in the group based analysis, only VAR and subset VAR models were chosen as the best ones (again forecast combination methods or the multivariate methods in the first differences were never chosen as the best ones). In contrast to the groups-based analysis, univariate models were never among the best models, which suggests the importance of other variables different from the price of the Arabica coffee itself.

4 A variance decomposition exercise

In order to quantify the relative role of the different sets of determinants of commodity prices as driving factors of Arabica coffee price dynamics, we select the largest vector autoregression (VAR) model among those with the best forecasting ability presented in Table 6 and use it to compute a variance decomposition of coffee price dynamics. The model includes coffee production in Brazil as fundamental variable, as well as the leading indicator variable for EU and US, and the real effective exchange rate as macroeconomic covariates. Stock indices for EU and US are incorporated as financial variables and temperature as an (exogenous) climatic covariate.

In order to perform the variance decomposition, we assume a Cholesky ordering that implies that shocks originating in changes of expectations reflect first in the stock indices, then in leading indicators and the real effective exchange rate and only then affect coffee prices and production.[9] Figure 3 shows the variance decomposition for up to 24 months ahead, depicting the proportion of the variance of coffee prices that can be explained by shocks in each one of the different groups of explanatory factors.

The minimal importance of our fundamental variable in terms of explaining historical variation in coffee prices is visible in Figure 3, which shows that the percentage of price variance explained by changes in the production of Arabica coffee in Brazil does not reach more than 2.5% at any prediction horizon. A similar conclusion applies to the group of purely financial variables, whose variation is able to explain a maximum of about 3.5% of the variance of coffee prices in the two-year ahead horizon. In the framework of the VAR model used, most of the variance of coffee prices is explained by idiosyncratic shocks to the price variable, labeled "Persistence" in Figure 3.

These results emphasize the importance of global macroeconomic dynamics as an explanatory factor behind changes in commodity price dynamics. Although financial variables have been shown to add to the in-sample and out-of-sample predictive power of time series models for coffee price dynamics, the variance decomposition exercise reveals that their contribution is quantitatively small once macroeconomic variables are included in the specification.

[9]The results concerning the relative importance of the factors are not qualitatively affected by changes in the ordering.

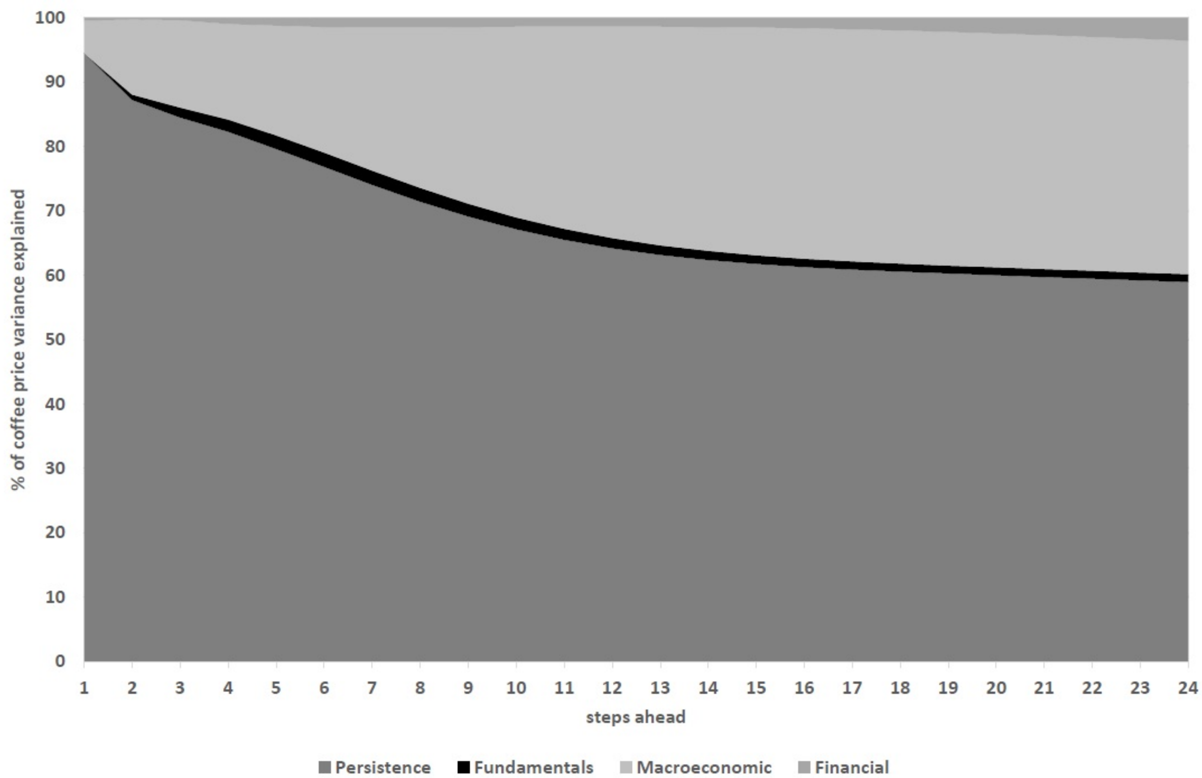

Figure 3: Variance decomposition by variable category

5 Conclusions

As is the case of many other commodities, price trends and volatility is a major concern for stakeholders in the world coffee market. In exporting countries, the price volatility is a source of uncertainty in relation to export earnings and tax revenues, as well as instability in producer incomes many of which are smallholders. Sustained low coffee prices can imply considerable social hardship in many export dependent countries. In importing countries, price volatility makes it difficult for roasters to control processing costs and affects profit margins along the supply chain.

In the free market period since 1990, smallholder farmers in many countries have been more exposed to fluctuations in coffee prices, as the internal regulatory mechanisms in producing countries were predominantly dismantled. These price fluctuations have increased rural poverty as it became difficult for small producers to efficiently plan their resource allocations. As a result, risk management strategies are becoming increasingly recommended to producers in developing countries. However, the scope and applicability of these instruments can vary significantly depending on the nature of the underlying and direct drivers of price trends and volatility. This paper provides a first quantified assessment of a comprehensive set of drivers to price formation.

Price volatility has significantly changed since the dismantling of the coffee agreement. Information of climate shocks are now more readily reflected in futures prices, allowing for more immediate market responses to such exogenous impacts and overall dampening instantaneous price volatility effects as well as their durability. Market response affecting the durability of shocks include the avoidance of price illusion trigged periods of overproduction following marked events of supply shortage as well as the management of strategic stocks.

We conclude that improvements in market information and the application of financial risk management contributed to a decreasing role of market fundamentals driving coffee prices. Our results suggest that macro-economic and financial market variables are more important to understand and predict coffee prices. This has important implications for how individual producers, including smallholder coffee producers and producer countries, should manage the consequences of commodity price risks. Predictive tools such as the ones presented in this paper appear to be key for the implementation of such a risk management systems.

Table 5: Summary of forecast performance of best models for Brazilian Arabica coffee over different variable groups: fundamentals, macro financial and other.

1-month horizon	MAE	MSE	DA	DV	return	Sharpe ratio
Fundamentals	**6.981** last 6 months	**85.716** last month	58.519 s-AR(3)	65.325 last 9 months	**7.619** last 9 months	**0.260** last 9 months
Macroeconomic	7.299 s-VAR(3) REER	91.253 VAR(3) REER	**61.481** s-VAR(4) y^{US}	**66.501** VAR(3) REER	7.093 s-VAR(4) y^{US}	0.251 s-VAR(4) y^{US}
Financial	7.309 AR(3)	92.322 AR(3)	59.259 s-VAR(3) $stock^{EU}$	63.463 last month	4.351 last month	0.195 last month
Climatic	7.223 VAR(2) precipitation	91.646 VAR(2) precipitation	58.519 s-AR(3)	61.829 last 3 months	3.012 DARCH(5,2)	0.153 DARCH(5,2)
3-months horizon	**MAE**	**MSE**	**DA**	**DV**	**return**	**Sharpe ratio**
Fundamentals	**15.469** last month	**436.635** whole	54.074 VAR(2) y^{BR}_{coffee} y^{world}_{coffee}	60.761 DARCH(1,3)	3.173 DA - last month	0.237 DA - last month
Macroeconomic	15.800 RW	447.884 VAR(1) li^{US} REER	**62.963** s-VAR(4) y^{US}	**70.930** s-VAR(4) y^{US} li^{EU} li^{US} REER	**5.344** s-VAR(6) y^{US} y^{BR} li^{EU} REER	**0.294** s-VAR(4) y^{US}
Financial	15.800 RW	458.373 RW	55.556 VAR(1) $stock^{EU}$ $stock^{US}$	56.970 VAR(3) $stock^{US}$	2.290 DA - last month	0.171 DA - last month
Climatic	15.800 RW	458.373 RW	53.788 last 3 months	58.019 last 9 months	2.323 s-VAR(3) temperature	0.154 s-VAR(3) temperature
6-months horizon	**MAE**	**MSE**	**DA**	**DV**	**return**	**Sharpe ratio**
Fundamentals	25.329 RW	1107.249 RW	59.259 VAR(2) y^{BR}_{coffee} y^{world}_{coffee}	61.973 DARCH(1,3)	2.984 DARCH(1,3)	0.306 DARCH(1,3)
Macroeconomic	**24.634** VAR(6) y^{BR}	**988.213** VAR(6) y^{BR} li^{EU}	**62.963** VAR(3) y^{US} y^{BR} li^{EU} li^{US} REER	**76.421** VAR(4) li^{EU} li^{US}	**5.584** VAR(4) li^{EU} li^{US}	**0.404** VAR(4) li^{EU} li^{US}
Financial	25.329 RW	1107.249 RW	60.000 VAR(1) $stock^{EU}$ $stock^{US}$	65.144 s-VAR(3) $stock^{EU}$	2.772 s-VAR(3) $stock^{EU}$	0.268 DARCH(1,5)
Climatic	25.329 RW	1107.249 RW	49.630 s-DAR(2)	59.211 s-AR(3)	1.330 s-VAR(3) temperature	0.108 s-AR(3)
9-months horizon	**MAE**	**MSE**	**DA**	**DV**	**return**	**Sharpe ratio**
Fundamentals	33.863 RW	1895.082 RW	**62.222** DARCH(1,5)	63.867 last month	2.879 DARCH(1,3)	0.357 DARCH(1,3)
Macroeconomic	**30.666** VAR(6) y^{BR}	**1532.005** s-VAR(6) y^{BR} li^{EU} li^{US}	62.222 DARCH(1,5)	**78.619** VAR(4) li^{EU} li^{US}	**5.191** VAR(4) li^{EU} li^{US}	**0.483** VAR(4) li^{EU} li^{US}
Financial	33.863 RW	1895.082 RW	**62.222** DARCH(1,5)	70.301 s-VAR(3) $stock^{EU}$	4.061 s-VAR(3) $stock^{EU}$	0.412 s-VAR(3) $stock^{EU}$
Climatic	33.863 RW	1895.082 RW	50.370 s-AR(3)	60.585 s-AR(3)	1.207 s-AR(3)	0.151 s-AR(3)
12-months horizon	**MAE**	**MSE**	**DA**	**DV**	**return**	**Sharpe ratio**
Fundamentals	41.232 RW	2667.403 RW	62.963 DARCH(1,5)	61.088 DARCH(1,5)	2.324 DARCH(1,5)	0.326 DARCH(1,5)
Macroeconomic	**38.583** VAR(6) y^{BR}	**2124.918** s-VAR(6) y^{BR} li^{EU} li^{US}	63.704 VAR(1) li^{US} REER	**73.689** VAR(4) li^{EU} li^{US}	**3.853** VAR(4) li^{EU} li^{US}	0.441 VAR(4) li^{EU} li^{US}
Financial	41.232 RW	2667.403 RW	**64.444** s-VAR(3) $stock^{EU}$	69.567 s-VAR(3) $stock^{EU}$	3.299 s-VAR(3) $stock^{EU}$	**0.444** s-VAR(3) $stock^{EU}$
Climatic	41.232 RW	2667.403 RW	52.593 s-AR(3)	60.031 s-AR(3)	1.262 s-AR(3)	0.187 s-AR(3)

See Table 3 for the abbreviation of the models. **Bold** figures indicate the best performance among all groups but within certain forecast horizon and **bold** figures indicate the best performance among all groups and forecast horizons.

Table 6: Summary of forecast performance of best models for Brazilian Arabica coffee over variables with highest predictive power.

Forecast horizon	MAE	MSE	DA	DV	return	Sharpe ratio
1-month	**7.062** VAR(2) y^{BR}_{coffee} y^{world}_{coffee} REER precipitation	**86.871** VAR(2) y^{BR}_{coffee} REER precipitation	62.963 VAR(2) y^{world}_{coffee} REER $stock^{EU}$ temperature	69.972 VAR(3) y^{BR}_{coffee} REER	7.682 last 9 months	0.262 last 9 months
3-months	14.497 s-VAR(1) y^{BR}_{coffee} y^{US} y^{BR} li^{EU} $stock^{US}$	405.226 VAR(2) y^{BR}_{coffee} li^{US} REER	65.926 VAR(3) y^{US} li^{EU} REER $stock^{EU}$	76.771 VAR(2) y^{BR}_{coffee} y^{US} li^{EU} li^{US} REER temperature	**9.724** VAR(4) y^{BR}_{coffee} li^{EU} li^{US} REER $stock^{EU}$ $stock^{US}$ temperature	0.400 VAR(3) y^{US} li^{EU} REER $stock^{EU}$
6-months	22.681 s-VAR(2) y^{BR} li^{EU} $stock^{US}$	928.783 VAR(6) y^{BR} li^{EU} $stock^{EU}$	**71.111** VAR(3) y^{world}_{coffee} y^{US} y^{BR} li^{EU}	77.624 VAR(4) y^{BR}_{coffee} li^{EU} li^{US} REER	7.764 VAR(3) y^{world}_{coffee} y^{US} y^{BR} li^{EU} temperature	0.433 VAR(2) y^{world}_{coffee} y^{US} y^{BR} li^{EU} $stock^{EU}$
9-months	30.666 VAR(6) y^{BR}	1532.005 s-VAR(6) y^{BR} li^{EU} li^{US}	62.222 VAR(6) y^{BR}	**78.619** VAR(4) li^{EU} li^{US}	5.191 VAR(4) li^{EU} li^{US}	**0.483** VAR(4) li^{EU} li^{US}
12-months	38.583 VAR(6) y^{BR}	2124.918 s-VAR(6) y^{BR} li^{EU} li^{US}	64.444 s-VAR(3) $stock^{EU}$	73.689 VAR(4) li^{EU} li^{US}	3.853 VAR(4) li^{EU} li^{US}	0.444 s-VAR(3) $stock^{EU}$

See Table 3 for the abbreviation of the models. **Bold** figures indicate the best performance among all forecast horizons.

References

Bacchini, F., Ciammola, A., Iannaccone, R., and Marini, M. (2010). Combining forecasts for a flash estimate of euro area gdp. *Contributi ISTAT*, (3).

Blaskowitz, O. and Herwartz, H. (2011). On economic evaluation of directional forecasts. *International Journal of Forecasting*, 27(4):1058–1065.

Cheung, C., Morin, S., et al. (2007). The impact of emerging asia on commodity prices. *Money Affairs*, 20(2):181–224.

Claeskens, G., Hjort, N. L., et al. (2008). *Model selection and model averaging*, volume 330. Cambridge University Press Cambridge.

Crespo Cuaresma, J., Hlouskova, J., and Obersteine, M. (2016). Forecasting Commodity Prices: A Comprehensive Approach. *SUSFANS Project*, mimeo.

Feldkircher, M. and Zeugner, S. (2009). Benchmark Priors Revisited: On Adaptive Shrinkage and the Supermodel Effect in Bayesian Model Averaging. *IMF Working Paper*, WP/09/202.

Fernández, C., Ley, E., and Steel, M. F. (2001a). Benchmark Priors for Bayesian Model Averaging. *Journal of Econometrics*, 100:381–427.

Fernández, C., Ley, E., and Steel, M. F. (2001b). Model Uncertainty in Cross-Country Growth Regressions. *Journal of Applied Econometrics*, 16:563–576.

Foster, D. P. and George, E. I. (1994). The Risk Inflation Criterion for Multiple Regression. *The Annals of Statistics*, 22:1947–1975.

Gencay, R. (1998). The predictability of security returns with simple technical trading rules. *Journal of Empirical Finance*, 5(4):347–359.

Gilbert, C. (2006). The prices prospects for latin american export commodities. In *Twenty Fifth Meeting of the Latin American Network of Central Banks and Finance Ministries*, pages 19–20.

Granger, C. W. and Ramanathan, R. (1984). Improved methods of combining forecasts. *Journal of Forecasting*, 3(2):197–204.

Husain, M. A. M. and Bowman, C. (2004). *Forecasting Commodity Prices: Futures Versus Judgment*. Number 4-41. International Monetary Fund Working Paper.

Kapetanios, G., Labhard, V., and Price, S. (2006). Forecasting using predictive likelihood model averaging. *Economics Letters*, 91(3):373–379.

Kass, R. and Wasserman, L. (1995). A reference Bayesian test for nested hypotheses and its relationship to the Schwarz criterion. *Journal of the American Statistical Association*, pages 928–934.

Koop, G. (2003). Bayesian Econometrics. *John Wiley & Sons*.

Ley, E. and Steel, M. F. (2009). On the Effect of Prior Assumptions in Bayesian Model Averaging with Applications to Growth Regressions. *Journal of Applied Econometrics*, 24:4:651–674.

Ley, E. and Steel, M. F. (2012). Mixtures of g-priors for bayesian model averaging with economic applications. *Journal of Econometrics*, 171(2):251–266.

Liang, F., Paulo, R., Molina, G., Clyde, M. A., and Berger, J. O. (2008). Mixtures of g Priors for Bayesian Variable Selection. *Journal of the American Statistical Association*, 103:410–423.

Pendergrast, M. (2010). *Uncommon Grounds: The History of Coffee and how it Transferred our World*. Basic Books.

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

Zellner, A. (1986). *Bayesian Inference and Decision Techniques: Essays in Honor of Bruno de Finetti*, chapter On Assessing Prior Distributions and Bayesian Regression Analysis with g-Prior Distributions. North-Holland: Amsterdam.