Australian National University

NATURAM PRIMUM COGNOSCERE RERUM

Crawford School of Public Policy

CAMA

Centre for Applied Macroeconomic Analysis

Common factors and the dynamics of cereal prices. A forecasting perspective

Marek Kwas
SGH Warsaw School of Economics, Collegium of Economic Analysis, Warsaw, Poland

Alessia Paccagnini
University College Dublin, Michael Smurfit Graduate Business School
Centre for Applied Macroeconomic Analysis, ANU

Michal Rubaszek
SGH Warsaw School of Economics, Collegium of Economic Analysis, Warsaw, Poland

Abstract

This article investigates what determines the price dynamics of the main cereals: barley, maize, rice and wheat. Using an extensive dataset of monthly time series covering the years 1980 - 2019, we extract four different common factors explaining the dynamics of commodity prices, exchange rates, financial and macroeconomic indicators. Next, we examine whether these factors are useful in explaining the movements of cereal prices. We show that models incorporating all four factors outperform significantly the naive random walk model in out-of-sample forecasting competition, especially for longer horizons. However, they have only marginally better performance than a simpler model based on the commodity factor alone.

Keywords

Cereal prices, Forecasting, Factor models, Autoregressive models.

JEL Classification

C32, C53, Q11

Address for correspondence:

(E) cama.admin@anu.edu.au

Common factors and the dynamics of cereal prices.
A forecasting perspective.*

Marek Kwas[a], Alessia Paccagnini[b], and Michał Rubaszek[a]

[a]*SGH Warsaw School of Economics, Collegium of Economic Analysis, Warsaw, Poland*
[b]*University College Dublin, Michael Smurfit Graduate Business School & CAMA*

Abstract

This article investigates what determines the price dynamics of the main cereals: barley, maize, rice and wheat. Using an extensive dataset of monthly time series covering the years 1980 - 2019, we extract four different common factors explaining the dynamics of commodity prices, exchange rates, financial and macroeconomic indicators. Next, we examine whether these factors are useful in explaining the movements of cereal prices. We show that models incorporating all four factors outperform significantly the naive random walk model in out-of-sample forecasting competition, especially for longer horizons. However, they have only marginally better performance than a simpler model based on the commodity factor alone.

Keywords: Cereal prices; Forecasting; Factor models; Autoregressive models.

JEL classification: C32; C53; Q11.

1 Introduction

Understanding the dynamics of cereal prices is valuable for many economic agents. For the agriculture sector, the sale of cereals is the main source of revenue. For many other industries, cereals are an important input in their production processes. For countries where cereals constitute a significant fraction of export, their prices influence many key macroeconomic variables, for instance, the current account balance, terms of trade, and exchange rates. On top of that, cereal prices exert a substantial impact on food security, hence poverty, across the globe (Gutierrez et al., 2014). For the above reasons, understanding the dynamics of cereal prices and examining the feasibility of econometric models to forecast their dynamics is essential to improve economic decisions, especially in emergency periods such as the recent COVID-19 pandemic.

For that reason, some authors investigate the sources and characteristics of cereal price dynamics. For instance, Jumah and Kunst (2008) indicate that real prices of barley and wheat are affected by seasonality. Baffes and Haniotis (2010) document the existence of a link between food and energy prices, emphasizing the role of fertilizer and transport costs. Gutierrez et al. (2014) apply a global vector autoregressive model to analyze the dynamics of the wheat market. Haile et al. (2015) point out, relying on a dynamic panel data analysis, how reducing agriculture commodity price volatility can increase food supply around the world, particularly in developing countries. They also provide evidence of a strong connection between energy and financial markets related to volatility, which can destabilize prices of agriculture commodities, as earlier stated in Tadesse et al. (2014). Kyriazi et al. (2019) point to an irregular cyclical component in agricultural price dynamics, which depends on productivity and the business cycle. They also indicate that adaptive learning techniques can be exploited in forecasting their prices. Recently, Karali et al. (2020), relying on an instrumental variable estimation, show how fundamental supply and market news plays a relevant role in explaining cereal futures price movements. These are just a few examples of contributions that use macroeconomic and financial variables, as well as exchange rates and other commodity prices, as potential determinants of cereal prices. However, the debate on what determines the dynamics of cereal prices and whether these prices are forecastable is still ongoing.

In this article, we join this debate and contribute to two margins. First, we explain the dynamics of real prices of main cereals – barley, maize, rice, and wheat – using factor models. For that purpose, we collect hundreds of monthly time-series from the period of Jan. 1980 – Aug. 2019 and extract four first common factors describing the dynamics of the commodity market, the foreign exchange market, financial markets, and the real economic activity sector. Second, we verify whether these four factors are leading indicators for cereal prices by running an out-of-sample forecasting competition. Our main finding is that vector autoregressive (VAR) models incorporating all four factors outperform significantly the naive random walk model in terms of point and density forecasts accuracy. We also show that the quality of forecasts from VAR models is higher than those from models based on direct forecast regressions. Finally, we present evidence that among the four analyzed common factors, the predictive content of the commodity factor is the highest.

Our study relates to a large literature that investigates drivers for a broader group of commodity prices within a factor model framework. This literature exploits recent advancements in data processing focusing on exploiting information content of large datasets on macroeconomic variables, financial indicators as well as volumes and prices of commodities. In most studies, factors are extracted with the principal components (PC) method and have only loose economic interpretation. For instance, Zagaglia (2010) extracts factors from a large dataset to show that the first PC is strongly correlated with the crude oil price index, whereas the other factors describe stock volumes of oil-related products or are related to a broader set of finan-

cial variables. Similarly, Lombardi et al. (2010) investigate interlinkages between commodity prices (metals and foods) and macroeconomic variables by estimating a factor augmented vector autoregression (FAVAR) introduced to the literature by Bernanke et al. (2005). Their main finding is that exchange rates and industrial production affect non-energy commodity prices. A notable exception is the study of Alquist et al. (2019), who justify the factor structure within a micro-founded structural model.

In terms of applications, factors are used to identify the effects of aggregate shocks on the commodity markets (Byrne et al., 2013; Aastveit et al., 2015; Juvenal and Petrella, 2015) or to forecast commodity prices. As regards the latter, the dynamic factor model (DFM), introduced to the literature by Stock and Watson (2005), and FAVAR are the most popular frameworks (Lombardi et al., 2010; Zagaglia, 2010; Juvenal and Petrella, 2015; Kagraoka, 2016; Bunn et al., 2017). In a few studies, however, the information content of factors is exploited within a single equation framework (Groen and Pesenti, 2011; Cummins et al., 2016). The forecasting studies focus predominantly on the predictive content of commodity factors (CF). West and Wong (2014) show that CF are good predictors for energy commodities, poor predictors for industrial metals, with agricultural products falling somewhere in between. Chen et al. (2014) indicate that CF can be used in forecasting exchange rates. Delle Chiaie et al. (2017) present that the DFM framework with CF performs well in commodity price forecasting at short horizons, also for food prices. The literature is less explicit about the predictive content of other factors, extracted from exchange rates, real sector or financial variables. We, therefore, contribute by presenting a comprehensive study that uses the methodology based on a wider range of common factors applied to forecasting real prices of agriculture products.

The rest of the paper is organized as follows. Sections 2 describes the data with a preliminary statistical analysis. Section 3 discusses econometric models applied in our study. Section 4 presents in-sample evidence, whereas Section 5 reports the results of the forecasting contest. The last Section 6 concludes.

2 Data

Our focus is on four main cereals – barley, maize, rice, and wheat – throughout Jan 1980 – Sep 2019. Their prices are taken from the World Bank Commodity Price database, where they also form the cereal price index with weights 4.1% (barley), 40.7% (maize), 30.1% (rice) and 25.2% (wheat). To express commodity prices in real terms, we use the US consumer price index, which is taken from the St. Louis Fed FRED database. Finally, the real time series are seasonally adjusted with X-13 ARIMA-SEATS metod.

The values of real prices (expressed in USD from Aug 2019) for four analyzed cereals are presented in Figure 1. It shows that these time series are characterized by a common long-term trend. In particular, at the beginning of the sample, they are relatively high, then they gradually decline, to jump sharply before the Great Financial Crisis and decrease afterwards. On the contrary, the upper panel of Table 1 indicates that their short-term dynamics are more heterogeneous. Even though all series display a similar degree of volatility, there are notable differences in higher moments of their distributions. For instance, growth rates of rice prices feature positive skewness and significantly higher kurtosis than the remaining three series. Looking at Figure 1, we notice that relatively low volatility periods of rice prices are separated by sudden spikes, unobserved in the other series with more uniform volatilities. Additionally, for all series, we observe moderate autocorrelation of their growth rates, the highest for rice and the lowest for barley.

Let us now focus on times series for common factors, which are extracted as the first principal components from four datasets:

$f_{\mathbf{com}}$: commodity factor based on 33 real commodity prices from the World Bank Commodity Price database, all series deflated by the US CPI and seasonally adjusted.

$f_{\mathbf{er}}$: exchange rate factor based on 16 real exchange rates against the USD calculated using data from the OECD Main Economic Indicator database.

$f_{\mathbf{fin}}$: financial factor based on 35 financial time series extracted from the FRED database.

$f_{\mathbf{mac}}$: macroeconomic factor based on 19 real activity indicators taken from the FRED database.

Figure 2 presents the dynamics of these four factors. It well illustrates that the characteristics of f_{com} and f_{er} differ from the characteristics of the remaining two factors. In particular, it is visible that f_{com} and f_{er} are more persistent than f_{fin} and f_{mac}. This is quantified in the lower panel of Table 1, which shows that autocorrelations of commodity and exchange rate factors is much higher than those for the financial and real sector factors. The panel also shows that the financial factor is characterized by the highest kurtosis and is highly negatively skewed.

3   Competing models

We use two popular model classes, which allow exploring the predictive content of the four factors (f_{kt} for $k \in \{\mathrm{com}, \mathrm{er}, \mathrm{fin}, \mathrm{mac}\}$) for the four time series of real cereal prices (q_{it} for $i \in \{\mathrm{barley}, \mathrm{maize}, \mathrm{rice}, \mathrm{wheat}\}$).

The first model is based on the single equation "direct forecast" (DF) regressions:

$$\Delta_h q_{it} = \alpha_h + \beta_h q_{i,t-h} + \gamma_h f_{k,t-h} + \epsilon_t, \tag{1}$$

where $\Delta_h q_{it} = q_{it} - q_{i,t-h}$ is the (logarithmic) percentage change in the real price of cereal i over h-period horizon. We also consider a model, in which the price is regressed on all four factors:

$$\Delta_h q_{it} = \alpha_h + \beta_h q_{i,t-h} + \Gamma_h F_{t-h} + \epsilon_t,$$

where $F_t = (f_{\mathrm{com},t}, f_{\mathrm{er},t}, f_{\mathrm{fin},t}, f_{\mathrm{mac},t})'$ is a vector containing the values of all four factors.

The second model is a vector autoregression (VAR) with a lag of length 2:

$$y_t = A_0 + A_1 y_{t-1} + A_2 y_{t-2} + D\epsilon_t, \quad \epsilon_t \sim \mathcal{N}(0, I), \tag{2}$$

where y_t is the vector of n endogenous variables, A_0 is an $n \times 1$ vector of constant terms, A_1, A_2 are $n \times n$ matrices of coefficients and ϵ_t is the $n \times 1$ vector of residuals. As regards the endogenous vector y_t, we consider 16 bivariate models (with $n = 2$) where $y_t = (q_{it}, f_{kt})'$, as well as 4 models with $n = 5$ and y_t containing the real prices of a given cereal and all four factors $y_t = (q_{it}, f_{\mathrm{com},t}, f_{\mathrm{er},t}, f_{\mathrm{fin},t}, f_{\mathrm{mac},t})'$.

Finally, for the benchmark we take the random walk (RW) model commonly used in the forecasting literature, for which:

$$q_{i,t+1} = q_{it} + \epsilon_t. \tag{3}$$

This model provides naïve "no-change" forecasts.

4   In-sample analysis

We start by analyzing the relationship between cereal prices and factors using the full sample of data. To verify whether factors are leading indicators of cereal prices, we compute correlation coefficients of h period growth rates of cereal prices $\Delta_h q_{it}$ against lagged factor values $f_{k,t-h}$, i.e.

the two variables from equation (1). Figure 3 indicates that the commodity factor f_{com} leads to subsequent changes in cereal prices, where the strength of correlation is increasing with h. This suggests the existence of predictive content in f_{com}. For the remaining factors, correlations are insignificant or relatively weak. In the case of f_{er} there is a negative correlation with maize and rice prices, but a positive one with wheat prices for selected horizons h. In turn, f_{fin} weakly leads maize, rice and wheat prices. Barley prices, which are correlated neither with f_{er} nor with f_{fin}, exhibit moderate negative correlation with f_{mac}. Overall, cross-correlations indicate that commodity factor f_{com} contains some predictive power for all cereal prices, whereas the remaining three factors can potentially provide some auxiliary information.

We continue by estimating DF regressions, described by equation (1), using all observations from the sample. This essentially expands the previous analysis of correlations by using also information on the level of the dependent variable ($q_{i,t-h}$). We observe that, after controlling for ($q_{i,t-h}$), f_{com} exhibits the strongest relationship with all cereal price series (but barley) in terms of the absolute values of γ_h as well as their statistical significance. The estimates for γ_h are also significant in regressions with the exchange rate factor f_{er}, apart from wheat price models. As regards the remaining factors, f_{fin} and f_{mac}, coefficients γ_h are often small and insignificant. We thus observe that, after controlling for the price level, the commodity f_{com} and exchange rate f_{er} factors seem to be the best leading indicators for cereal prices.

We complete the in-sample analysis by conducting Granger causality tests using VAR models described in equation (2). The test statistics, which are reported in Table 3, point to strong Granger causality from the commodity factor to real prices of barley, rice, and maize. The Table also shows that the real sector factor f_{mac} Granger causes barley prices and the exchange rate factor f_{er} Granger causes rice prices. Unexpectedly, no causality is identified between factors and wheat prices. In general, we conclude that among the four investigated factors, the commodity factor seems to be the best leading indicator for subsequent changes in cereal prices. In the following section, we shall verify whether this in-sample finding is confirmed within an out-of-sample forecasting contest.

5 Forecasting contest

We continue our investigation by conducting an out-of-sample competition among the models described in Section 3. We use the recursive forecast evaluation scheme, which means that forecasting models are estimated on expanding windows. It can be noted that we also update the times series for factors recursively so that their values do not depend on the full sample but on history up to the moment at which the forecast is formulated. We calculate forecasts for horizons h ranging from one month to five years ahead. The first set of forecasts is generated for the period Jan 1995 – Dec 1997, with models estimated on the period Jan 1980 – Dec 1994, whereas the last set of forecasts is generated for the period Sep 2019 – Aug 2022, with models estimated on the period Jan 1980 – Aug 2019. As a result, for a given forecast horizon h, we assess the quality of forecasts using $297 - h$ forecast errors.

In what follows, we evaluate the accuracy of point, directional and density forecasts. For point forecasts, we assess the accuracy with the mean forecast error (MFE) and the root mean square forecast error (RMSFE). For directional forecasts, we calculate the fraction of correct sign predictions as well as the correlation of forecasts and realizations. As regards density forecasts, we use log predictive scores (LPS) and continuous ranked probability scores (CRPS). For the relative accuracy measures (RMSFE, LPS, and CRPS) we assess the performance of a given model by comparing it to the benchmark random walk model using Diebold and Mariano (1995) and Amisano and Giacomini (2007) tests.

5.1 Point forecasts

We begin the out-of-sample analysis by calculating the MFE for cereal prices (Table 4). The slightly negative and insignificant values of MFE for the random walk benchmark indicate that in the evaluation sample the real prices of all cereals experienced a gradual decline. In fact, nominal prices of the four cereals increased over the period from Dec. 1994 to Sep. 2019, where the scale of the price change amounted to 32.8% (barley), 53.4% (maize), 63.9% (rice) and 28.3% (wheat). At the same time, the overall US CPI index grew up slightly more, i.e. by 70.9%. Looking at the performance of other models, a general finding is that they deliver unbiased forecasts or tend to overpredict future movements of all four analyzed price series, which can be explained by their already mentioned gradual downward trend in the forecast evaluation sample.

We continue the evaluation of point forecasts by comparing the second moments of forecast errors using the most popular measure of ex-post forecast quality – RMSFE. In Table 5, we report the levels of the RMSFE for the benchmark naïve forecast form the RW model, whereas the remaining numbers are the ratios relative the benchmark RMSFE levels. Thus, values below unity indicate that a given method outperforms the naïve forecast. Moreover, we test whether a given method dominates the benchmark using the Diebold and Mariano (1995) test. Several interesting findings become evident. First, at longer horizons, all models deliver more accurate forecasts than the benchmark. We interpret this as evidence that any autoregressive structure helps explain the dynamics of real cereal prices, also in the out-of-sample setting. Second, Table 5 illustrates that, for shorter horizons, VAR models tend to deliver more accurate forecasts than the benchmark, which is not the case for the DF models. The third finding is that the models based on four factors ($4f$) perform best, but they are closely followed by the models based on the commodity factor alone (f_{com}). This means that the higher complexity of the $4f$ model is only mildly remunerated by the accuracy of forecasts it delivers.

5.2 Directional forecasts

We proceed by comparing the realized ($q_{i,t+h} - q_{it}$) and forecasted ($q_{i,t+h|t} - q_{it}$) changes in real cereal prices. In Table 6, we report the correlations between both variables for different horizons h. It can be noted that the values in the table can be interpreted as the fit of the Mincer-Zarnowitz type of regression of forecast efficiency:

$$q_{i,t+h} - q_{it} = \alpha + \beta(q_{i,t+h|t} - q_{it}) + \eta_{i,t+h},$$

so that negative or low values indicate that a given method does not pass the efficiency test. We observe that all correlation coefficients for all models and cereals are positive. Moreover, in most cases, their values are increasing with the forecast horizon. Table 6 also shows that the performance of VAR models is comparable or slightly better than their DF counterparts. Finally, the models with four factors or with the commodity factor alone tend to dominate the remaining three specifications, i.e. based on f_{er}, f_{fin} and f_{mac}, which is especially visible for maize and rice prices.

The information about the correlation of forecasted and realized changes is complemented by the percentages of correctly forecasted signs of change in Table 7. They are accompanied by the results of the goodness-of-fit χ^2 test of Pesaran and Timmermann (1992), which enables to check if they are significantly different from 50%. The table shows that in most cases the share of forecasts with correct signs is well above 50%. Moreover, for longer horizons, this fraction becomes significantly higher than for a random toss of a coin. Once again, the quality of forecasts from VAR models with f_{mac} and with all four factors is comparable or better than that from the remaining specifications.

5.3 Density forecasts

We finalize the out-of-sample investigation with the assessment of density forecasts. The aim is to check to what extent the analyzed forecasts provide a realistic description of actual uncertainty. To quantify whether density forecasts from one model are better or worse than the RW benchmark, we compare their quality by using the two most popular scoring rules (Jordan et al., 2019): LPS and CRPS. For the DF models, we calculate the dispersion of forecasts with the standard analytical formula for the forecast standard error in linear regression, which takes into account both random and estimation error. For the VAR models, we follow the method proposed by Adolfson et al. (2007) and assume that for each horizon, forecasts are normally distributed, where the two moments of this distribution can be approximated using draws from the predictive density.

Table 8 presents the average LPS differences of a given model in comparison to the RW, so that positive values indicate that the model outperforms the benchmark. We test whether the values are significantly above zero with the one-tailed Amisano and Giacomini (2007) test. It shows that DF models deliver density forecasts of comparable (barley), superior (maize) or inferior quality (rice and wheat) in comparison to the RW. However, for all cereals density forecasts from DF models are less accurate than those from the VAR models. As regards VAR models, for maize and rice they deliver forecasts of higher quality than those from the RW model, whereas for the remaining two cereals they are comparable to the benchmark. Among VAR models, the version based on f_{er} is very competitive for barley and wheat, whereas VAR models with f_{com} and with all four factors are best performing for maize and rice.

Finally, Table 9 reports the ratios of the CRPS from a given model in comparison to the RW benchmark, so that the values below unity indicate that the method outperforms the naïve forecast. The results are very similar to those given by RMSFE, and point to the outstanding quality of forecasts from VAR models with commodity factor f_{com} alone or those including all four factors.

6 Conclusions

This study presents the analysis of the dynamics of real prices for four main cereals: barley, maize, rice, and wheat. Using monthly data from the period 1980:1–2019:8 we have recursively extracted four common factors from commodity, exchange rate, financial and macroeconomic time-series databases. Next, we have explored to what extent these factors are leading real cereal price developments. For that purpose, we tested two specifications: direct forecast models as well as vector autoregressions. For both specifications, we analyzed the predictive power in-sample as well as out-of-sample.

The broad picture that emerges from our analysis is as follows. The VAR models including real prices of a given cereal and all four factors perform best among all analyzed methods, being followed closely by a simpler bivariate VAR models based on the commodity factor alone. Forecasts from these models turned out to be highly correlated with realizations (true future values) and characterized by the lowest RMSFE and CRPS as well as the highest LPS. Our interpretation of the results is that the predictive content of the commodity factor is the highest among the four analyzed factors. Moreover, boosting the complexity of the model by incorporating the additional three factors helps little in terms of forecasting accuracy.

References

Aastveit KA, Bjornland H, Thorsrud LA. 2015. What drives oil prices? Emerging versus developed economies. *Journal of Applied Econometrics* **30**: 1013–1028.

Adolfson M, Linde J, Villani M. 2007. Forecasting performance of an open economy DSGE model. *Econometric Reviews* **26**: 289–328.

Alquist R, Bhattarai S, Coibion O. 2019. Commodity-price comovement and global economic activity. *Journal of Monetary Economics* : in Press.

Amisano G, Giacomini R. 2007. Comparing density forecasts via weighted likelihood ratio tests. *Journal of Business & Economic Statistics* **25**: 177–190.

Baffes J, Haniotis T. 2010. Placing the 2006/08 commodity price boom into perspective. Policy Research Working Paper 5371, World Bank.

Bernanke B, Boivin J, Eliasz P. 2005. Measuring the effects of monetary policy: A factor-augmented vector autoregressive (FAVAR) approach. *The Quarterly Journal of Economics* **120**: 387–422.

Bunn D, Chevallier J, Pen YL, Sevi B. 2017. Fundamental and financial influences on the cco-movement of oil and gas prices. *The Energy Journal* **38**: 201–228.

Byrne JP, Fazio G, Fiess N. 2013. Primary commodity prices: Co-movements, common factors and fundamentals. *Journal of Development Economics* **101**: 16–26.

Chen SL, Jackson JD, Kim H, Resiandini P. 2014. What drives commodity prices? *American Journal of Agricultural Economics* **96**: 1455–1468.

Cummins M, Dowling M, Kearney F. 2016. Oil market modelling: A comparative analysis of fundamental and latent factor approaches. *International Review of Financial Analysis* **46**: 211–218.

Delle Chiaie S, Ferrara L, Giannone D. 2017. Common factors of commodity prices. Working Paper Series 2112, European Central Bank.

Diebold FX, Mariano RS. 1995. Comparing predictive accuracy. *Journal of Business & Economic Statistics* **13**: 253–263.

Groen JJJ, Pesenti PA. 2011. Commodity prices, commodity currencies, and global economic developments. In *Commodity Prices and Markets, East Asia Seminar on Economics, Volume 20*, NBER Chapters. National Bureau of Economic Research, Inc, 15–42.

Gutierrez L, Piras F, Paolo Roggero P. 2014. A global vector autoregression model for the analysis of wheat export prices. *American Journal of Agricultural Economics* **97**: 1494–1511.

Haile MG, Kalkuhl M, von Braun J. 2015. Worldwide acreage and yield response to international price change and volatility: A dynamic panel data analysis for wheat, rice, corn, and soybeans. *American Journal of Agricultural Economics* **98**: 172–190.

Jordan A, Kruger F, Lerch S. 2019. Evaluating probabilistic forecasts with scoringRules. *Journal of Statistical Software* **90**: 1–37.

Jumah A, Kunst RM. 2008. Seasonal prediction of European cereal prices: good forecasts using bad models? *Journal of Forecasting* **27**: 391–406.

Juvenal L, Petrella I. 2015. Speculation in the oil market. *Journal of Applied Econometrics* **30**: 621–649.

Kagraoka Y. 2016. Common dynamic factors in driving commodity prices: Implications of a generalized dynamic factor model. *Economic Modelling* **52**: 609 – 617.

Karali B, Irwin SH, Isengildina-Massa O. 2020. Supply fundamentals and grain futures price movements. *American Journal of Agricultural Economics* **In Press**.

Kyriazi F, Thomakos DD, Guerard JB. 2019. Adaptive learning forecasting, with applications in forecasting agricultural prices. *International Journal of Forecasting* **35**: 1356 – 1369.

Lombardi M, Osbat C, Schnatz B. 2010. Global commodity cycles and linkages: A FAVAR approach. *Empirical Economics* **43**.

Pesaran MH, Timmermann A. 1992. A simple nonparametric test of predictive performance. *Journal of Business & Economic Statistics* **10**: 561–565.

Stock JH, Watson MW. 2005. Implications of dynamic factor models for VAR analysis. Working Paper 11467, National Bureau of Economic Research.

Tadesse G, Algieri B, Kalkuhl M, von Braun J. 2014. Drivers and triggers of international food price spikes and volatility. *Food Policy* **47**: 117 – 128.

West KD, Wong KF. 2014. A factor model for co-movements of commodity prices. *Journal of International Money and Finance* **42**: 289–309.

Zagaglia P. 2010. Macroeconomic factors and oil futures prices: A data-rich model. *Energy Economics* **32**: 409–417.

Table 1: Descriptive statistics

	Mean	SD	Min.	Max.	Skew.	Kurt.	AR(1)
	Log growth rates of real cereal prices						
BARLEY	0.00	0.06	-0.30	0.32	0.43	3.59	0.13
MAIZE	0.00	0.06	-0.25	0.29	-0.11	3.14	0.26
RICE	0.00	0.05	-0.23	0.44	1.57	12.47	0.37
WHEAT	0.00	0.07	-0.26	0.26	-0.05	1.73	0.18
	Levels of common factors						
f_{com}	0.00	0.73	-2.16	1.48	-0.39	0.11	0.99
f_{er}	0.00	0.86	-1.54	2.61	0.62	-0.31	0.99
f_{fin}	0.00	0.46	-4.10	1.32	-2.72	18.64	0.44
f_{mac}	0.00	0.68	-2.29	3.62	0.75	3.39	0.22

Notes: AR(1) stands for the first-order autoregressive coefficients.

Table 2: In-sample direct forecast regression estimates

h	f_{com}		f_{er}		f_{fin}		f_{mac}	
	β_h	γ_h	β_h	γ_h	β_h	γ_h	β_h	γ_h
	BARLEY							
1	-0.05•	-0.01°	-0.04•	-0.01*	-0.03•	0.00	-0.03•	-0.01°
6	-0.29•	-0.04•	-0.25•	-0.04•	-0.22•	-0.03*	-0.21•	-0.02°
12	-0.55•	-0.05*	-0.52•	-0.05•	-0.46•	0.00	-0.45•	-0.03°
24	-0.98•	-0.05*	-0.97•	-0.08•	-0.88•	0.00	-0.89•	0.04*
36	-1.09•	-0.04	-1.11•	-0.09•	-1.01•	0.02	-1.02•	0.04*
	MAIZE							
1	-0.10•	-0.04•	-0.02*	-0.01°	-0.02*	-0.01	-0.02*	-0.01
6	-0.63•	-0.23•	-0.17•	-0.03•	-0.17•	-0.05•	-0.16•	-0.01
12	-0.93•	-0.30•	-0.32•	-0.04•	-0.31•	-0.04°	-0.30•	-0.01
24	-1.32•	-0.37•	-0.57•	-0.06•	-0.55•	-0.01	-0.55•	0.03
36	-1.57•	-0.43•	-0.69•	-0.08•	-0.67•	-0.03	-0.67•	0.07•
	RICE							
1	-0.07•	-0.03•	-0.03•	-0.01•	-0.02•	0.00	-0.02*	0.00
6	-0.46•	-0.17•	-0.20•	-0.06•	-0.15•	-0.01	-0.15•	0.00
12	-0.69•	-0.23•	-0.35•	-0.08•	-0.28•	0.00	-0.28•	0.00
24	-0.91•	-0.20•	-0.63•	-0.09•	-0.56•	-0.01	-0.56•	0.01
36	-0.95•	-0.13•	-0.75•	-0.06•	-0.71•	0.02	-0.71•	0.03°
	WHEAT							
1	-0.06•	-0.02°	-0.03•	0.00	-0.03•	-0.01	-0.03•	-0.01
6	-0.48•	-0.14•	-0.21•	-0.01	-0.20•	-0.03°	-0.20•	-0.01
12	-0.71•	-0.18•	-0.36•	-0.03*	-0.34•	0.00	-0.34•	-0.02°
24	-0.93•	-0.17•	-0.58•	-0.01	-0.57•	0.01	-0.58•	0.03°
36	-0.82•	-0.09•	-0.62•	0.00	-0.62•	-0.01	-0.62•	0.04*

Note: Coefficients β_h and γ_h of regressions $\Delta_h q_{it} = \alpha_h + \beta_h q_{i,t-h} + \gamma_h f_{k,t-h} + \epsilon_{it}$. Symbols •, *and °denote the 1%, 5% and 10% significance levels.

Table 3: Granger causality

	f_{com}	f_{er}	f_{fin}	f_{mac}	$4f$
→BARLEY	11.77•	1.55	0.72	3.25*	4.11•
→MAIZE	2.91°	1.00	1.16	1.62	1.90°
→RICE	5.48•	2.51°	0.12	1.81	2.08*
→WHEAT	1.20	0.57	0.88	1.33	1.16

Note: Tests whether factors (levels) Granger-cause cereal prices (growth rates). Test statistics values are associated with symbols •, * and ° denoting rejections of the nulls of noncausality at 1%, 5% and 10% significance levels.

Table 4: Mean forecasts error (MFE)

		Direct Forecast					Vector Autoregression				
h	RW	f_{com}	f_{er}	f_{fin}	f_{mac}	$4f$	f_{com}	f_{er}	f_{fin}	f_{mac}	$4f$
						BARLEY					
1	0.00	0.01	0.00	0.00	0.00	0.01*	0.00	0.00	0.00	0.00	0.01*
6	0.00	0.04	0.01	0.01	0.01	0.05*	0.03	0.01	0.01	0.01	0.04*
12	-0.01	0.06°	0.02	0.01	0.01	0.08*	0.05	0.00	0.00	0.00	0.06°
24	-0.03	0.07	0.02	0.03	0.02	0.08°	0.05	0.00	-0.01	0.00	0.06
36	-0.05	0.06	0.02	0.03	0.03	0.06	0.06	0.00	-0.01	-0.01	0.06
						MAIZE					
1	0.00	0.00	0.00	0.00	0.00	0.00	-0.01°	0.00	0.00	0.00	0.00
6	0.00	-0.04°	-0.01	-0.01	-0.01	-0.03	-0.03°	-0.01	-0.01	-0.01	-0.03
12	-0.01	-0.06*	-0.02	-0.02	-0.02	-0.05°	-0.05°	-0.02	-0.02	-0.02	-0.05°
24	-0.03	-0.09*	-0.04	-0.04	-0.05	-0.09*	-0.06°	-0.05	-0.05	-0.05	-0.06
36	-0.04	-0.08*	-0.05	-0.04	-0.05	-0.10•	-0.05	-0.06	-0.06	-0.06	-0.04
						RICE					
1	0.00	0.00	0.00	0.00	0.00	0.00	0.00°	0.00	0.00	0.00	0.00
6	0.00	-0.03°	-0.01	-0.01	-0.01	-0.03°	-0.03*	-0.01	-0.01	-0.01	-0.03*
12	-0.01	-0.06*	-0.02	-0.02	-0.02	-0.07•	-0.05*	-0.03	-0.02	-0.02	-0.06•
24	-0.02	-0.10*	-0.05	-0.04	-0.04	-0.12•	-0.06°	-0.06°	-0.04	-0.03	-0.08*
36	-0.03	-0.11*	-0.07	-0.06	-0.06	-0.15•	-0.06	-0.09*	-0.04	-0.04	-0.08°
						WHEAT					
1	0.00	-0.01	0.00	0.00	0.00	0.00	-0.01*	0.00	0.00	0.00	-0.01
6	0.00	-0.05*	-0.02	-0.02	-0.02	-0.05*	-0.05*	-0.02	-0.02	-0.02	-0.04°
12	-0.01	-0.07*	-0.03	-0.03	-0.03	-0.07*	-0.07*	-0.04	-0.04	-0.04	-0.06*
24	-0.03	-0.11•	-0.07	-0.07	-0.07°	-0.12•	-0.09*	-0.07°	-0.07°	-0.07°	-0.08*
36	-0.04	-0.08°	-0.06	-0.05	-0.06	-0.09*	-0.08*	-0.09*	-0.09*	-0.09*	-0.08°

Notes: The errors are calculated for log real prices. MFE values represent average percent difference between realisations and forecasts, where positive and negative values stand for under- and overprediction, respectively. Symbols •, *and °denote the 1%, 5% and 10% significance levels for the unbiasedness null, where the long-run variance is calculated with the Newey-West method.

Table 5: Root mean square forecasts error (RMSFE)

h	RW	Direct Forecast					Vector Autoregression				
		f_{com}	f_{er}	f_{fin}	f_{mac}	$4f$	f_{com}	f_{er}	f_{fin}	f_{mac}	$4f$
BARLEY											
1	0.05	1.04	1.00	1.01	1.00	1.05	1.01	0.98°	0.98	0.98	1.02
6	0.17	1.07	0.98	0.98	0.99	1.10	1.02	0.97	0.98	0.98	1.05
12	0.25	1.04	0.95	0.96	0.96	1.08	0.99	0.94	0.93	0.93	1.02
24	0.35	0.86*	0.81*	0.84*	0.84*	0.85*	0.82•	0.80•	0.82•	0.82•	0.83•
36	0.40	0.79•	0.75•	0.78•	0.77•	0.77•	0.77•	0.73•	0.77•	0.77•	0.75•
MAIZE											
1	0.06	0.98	1.00	1.00	1.00	0.98	0.95*	0.96*	0.96*	0.96°	0.94*
6	0.18	0.91	0.97*	0.96	0.98	0.91	0.89*	0.97	0.97	0.97	0.90°
12	0.24	0.89°	0.95°	0.96	0.97	0.90	0.87*	0.95°	0.95	0.95	0.87°
24	0.33	0.80*	0.90*	0.92°	0.91°	0.82*	0.78•	0.90*	0.91*	0.91°	0.79•
36	0.38	0.74•	0.86•	0.88*	0.86*	0.74•	0.74•	0.85•	0.88*	0.88*	0.74•
RICE											
1	0.05	0.98	0.98*	1.00	1.00	0.98	0.91	0.93°	0.94°	0.94°	0.93
6	0.16	0.92	0.94*	1.00	0.99	0.95	0.92	0.96	1.01	1.01	0.94
12	0.21	0.89	0.94°	1.01	1.00	0.85	0.86°	0.93°	1.00	1.00	0.86
24	0.30	1.00	0.97	1.02	1.02	0.94	0.84•	0.90*	0.96	0.96	0.83•
36	0.36	1.00	0.98	0.99	0.99	1.02	0.89•	0.91*	0.94°	0.94°	0.89•
WHEAT											
1	0.07	1.00	1.00	0.99	1.00	0.99	0.99	0.98°	0.98°	0.98	0.98
6	0.18	1.03	0.98	0.98	0.99	1.03	0.99	0.98	0.98	0.97	0.98
12	0.25	0.98	0.97	0.97	0.97	0.98	0.93	0.96	0.96	0.96	0.94
24	0.33	0.96	0.96	0.94	0.94	0.97	0.88*	0.93°	0.93°	0.93°	0.90*
36	0.36	0.89•	0.89•	0.87•	0.87•	0.91•	0.86•	0.89*	0.90*	0.90*	0.86•

Notes: For RW the figures describe the values of RMSFE, whereas for the remaining methods the ratio of RMSFE in comparison to the RMSFE from RW, so that numbers below 1 denote that a given method outperforms the random walk model. Symbols •, *and °denote the 1%, 5% and 10% significance levels of the one-sided Diebold-Mariano test with the alternative that a given model performs better than RW.

Table 6: Correlation of forecasts and realization

h	Direct Forecast					Vector Autoregression				
	f_{com}	f_{er}	f_{fin}	f_{mac}	$4f$	f_{com}	f_{er}	f_{fin}	f_{mac}	$4f$
BARLEY										
1	0.03	0.06	0.06	0.07	0.05	0.21	0.22	0.20	0.21	0.21
6	0.17	0.21	0.23	0.22	0.16	0.25	0.25	0.25	0.25	0.24
12	0.36	0.37	0.37	0.37	0.32	0.34	0.36	0.36	0.37	0.32
24	0.61	0.61	0.59	0.59	0.61	0.60	0.61	0.58	0.58	0.59
36	0.63	0.66	0.63	0.63	0.65	0.68	0.70	0.66	0.66	0.69
MAIZE										
1	0.19	0.09	0.08	0.06	0.20	0.34	0.27	0.28	0.27	0.34
6	0.46	0.26	0.27	0.21	0.45	0.49	0.23	0.25	0.25	0.47
12	0.52	0.31	0.29	0.28	0.51	0.54	0.34	0.31	0.33	0.52
24	0.65	0.44	0.42	0.42	0.64	0.67	0.47	0.43	0.44	0.65
36	0.71	0.55	0.48	0.51	0.73	0.71	0.58	0.50	0.50	0.71
RICE										
1	0.22	0.21	0.00	0.02	0.25	0.43	0.37	0.35	0.35	0.41
6	0.46	0.38	0.14	0.15	0.51	0.46	0.29	0.13	0.13	0.51
12	0.53	0.38	0.17	0.18	0.65	0.58	0.40	0.20	0.20	0.64
24	0.41	0.34	0.28	0.28	0.54	0.58	0.48	0.30	0.30	0.63
36	0.40	0.37	0.36	0.37	0.44	0.49	0.47	0.36	0.36	0.51
WHEAT										
1	0.08	0.07	0.11	0.09	0.14	0.21	0.19	0.21	0.19	0.23
6	0.28	0.21	0.23	0.21	0.27	0.32	0.22	0.24	0.26	0.32
12	0.38	0.27	0.29	0.29	0.36	0.46	0.31	0.32	0.33	0.44
24	0.46	0.38	0.39	0.40	0.46	0.56	0.43	0.43	0.43	0.51
36	0.50	0.48	0.50	0.52	0.48	0.57	0.50	0.48	0.48	0.56

Note: The figures represent the correlation of forecast and realized changes in real cereal prices.

Table 7: Fraction of correct sign predictions

h	Direct Forecast					Vector Autoregression				
	f_{com}	f_{er}	f_{fin}	f_{mac}	$4f$	f_{com}	f_{er}	f_{fin}	f_{mac}	$4f$
BARLEY										
1	0.51	0.48	0.49	0.49	0.50	0.54	0.52	0.49	0.52	0.54
6	0.52	0.51	0.54	0.55°	0.51	0.55	0.52	0.55°	0.54	0.53
12	0.56*	0.57*	0.58•	0.57*	0.57*	0.59•	0.55	0.59•	0.57*	0.55
24	0.74•	0.74•	0.73•	0.73•	0.72•	0.70•	0.72•	0.73•	0.72•	0.69•
36	0.81•	0.82•	0.79•	0.79•	0.80•	0.77•	0.82•	0.79•	0.78•	0.80•
MAIZE										
1	0.57*	0.56°	0.51	0.50	0.58•	0.59•	0.62•	0.59•	0.61•	0.59•
6	0.61•	0.60•	0.56*	0.53	0.63•	0.61•	0.53	0.56*	0.54	0.61•
12	0.49	0.57*	0.51	0.51	0.49	0.52	0.55	0.52	0.55	0.51
24	0.61•	0.60•	0.58•	0.59•	0.64•	0.59•	0.60•	0.57*	0.58*	0.54
36	0.61•	0.60•	0.55°	0.56*	0.65•	0.61•	0.62•	0.56*	0.58•	0.62•
RICE										
1	0.53	0.58•	0.58•	0.53	0.57*	0.61•	0.63•	0.64•	0.65•	0.60•
6	0.62•	0.56*	0.61•	0.61•	0.63•	0.62•	0.56*	0.60•	0.61•	0.64•
12	0.68•	0.64•	0.65•	0.67•	0.71•	0.69•	0.58•	0.66•	0.65•	0.70•
24	0.69•	0.71•	0.72•	0.74•	0.72•	0.75•	0.61•	0.69•	0.70•	0.78•
36	0.66•	0.73•	0.72•	0.73•	0.64•	0.70•	0.63•	0.71•	0.70•	0.71•
WHEAT										
1	0.49	0.51	0.53	0.53	0.52	0.52	0.54	0.56*	0.55	0.54
6	0.52	0.58•	0.58•	0.58*	0.54	0.53	0.57*	0.55°	0.57*	0.52
12	0.58•	0.63•	0.62•	0.62•	0.57*	0.59•	0.63•	0.63•	0.61•	0.57*
24	0.68•	0.65•	0.65•	0.66•	0.69•	0.66•	0.66•	0.64•	0.64•	0.64•
36	0.71•	0.72•	0.73•	0.73•	0.71•	0.75•	0.74•	0.73•	0.73•	0.73•

Note: The figures represent the fraction of forecasts that correctly predict the sign of the change in the real price of a given cereal. Symbols •, * and ° denote the rejection of the null of the goodness-of-fit χ^2 test (Pesaran and Timmermann, 1992), stating that the fraction of correct sign forecast is 50%, at the 1%, 5% and 10% significance levels.

Table 8: Log Predictive Scores (LPS)

| h	RW	Direct Forecast					Vector Autoregression				
		f_{com}	f_{er}	f_{fin}	f_{mac}	$4f$	f_{com}	f_{er}	f_{fin}	f_{mac}	$4f$
BARLEY											
1	1.45	-0.02	0.00	0.00	0.00	-0.02	0.01	0.02*	0.02°	0.02	0.01
6	0.34	-0.10	0.02	0.02	0.02	-0.18	-0.02	0.04	0.04	0.04	-0.06
12	-0.03	-0.17	0.03	0.01	0.02	-0.32	-0.06	0.06	0.07	0.07	-0.12
24	-0.41	-0.03	0.10	0.04	0.03	-0.07	0.08	0.18*	0.15°	0.16*	0.07
36	-0.51	-0.01	0.09	0.06	0.07	0.02	0.14	0.29•	0.22•	0.23•	0.21*
MAIZE											
1	1.39	0.01	0.00	0.00	0.00	0.02	0.05°	0.04*	0.04*	0.04°	0.06*
6	0.17	0.16	0.12*	0.13*	0.11°	0.16	0.20*	0.12*	0.12°	0.12°	0.19°
12	-0.07	0.12	0.08°	0.07	0.06	0.10	0.16°	0.10°	0.09°	0.10°	0.15
24	-0.35	0.21°	0.05	0.03	0.04	0.18°	0.26•	0.12°	0.11°	0.12°	0.25•
36	-0.46	0.24*	0.01	0.00	0.02	0.22*	0.28•	0.15*	0.12°	0.12°	0.29•
RICE											
1	1.53	0.02°	0.01•	0.00	-0.01	0.02	0.11*	0.09•	0.08•	0.08*	0.10*
6	0.35	0.15	0.10	0.04	0.05	0.10	0.16	0.09	0.04	0.04	0.15
12	0.10	0.14	0.08°	-0.01	0.00	0.16	0.20°	0.08	0.01	0.01	0.23°
24	-0.25	-0.35	-0.34	-0.42	-0.44	-0.23	0.19•	0.10	0.03	0.03	0.23•
36	-0.42	-0.79	-0.78	-0.78	-0.81	-0.79	0.07	0.06	0.03	0.03	0.08°
WHEAT											
1	1.25	-0.01	0.00	0.01	0.00	0.01	0.00	0.02	0.02	0.02	0.02
6	0.24	-0.10	0.03	0.04	0.03	-0.10	-0.02	0.04	0.05	0.05	0.00
12	-0.06	-0.23	-0.09	-0.08	-0.09	-0.23	-0.06	0.03	0.04	0.04	-0.04
24	-0.33	-0.37	-0.30	-0.27	-0.26	-0.38	-0.12	0.01	0.01	0.01	-0.12
36	-0.41	-0.68	-0.52	-0.49	-0.48	-0.69	-0.19	-0.01	-0.04	-0.04	-0.13

Notes: For RW the figures describe the average values of the LPS, whereas for the remaining models the average differences in comparison to the RW benchmark. Positive values indicate that density forecasts from a given model are more accurate than those from the benchmark. The values can be interpreted as a percentage gain in the fit of forecasts to realizations in comparison to the benchmark. Symbols •, * and ° denote, respectively, the 1%, 5% and 10% significance levels of the one-tailed Amisano and Giacomini (2007) test, where the long-run variance is calculated with the Newey-West method.

Table 9: Continuous Ranked Probability Scores (CRPS)

| h	RW	Direct Forecast					Vector Autoregression				
		f_{com}	f_{er}	f_{fin}	f_{mac}	$4f$	f_{com}	f_{er}	f_{fin}	f_{mac}	$4f$
						BARLEY					
1	0.03	1.04	1.00	1.01	1.01	1.05	1.01	0.98°	0.99	0.98	1.02
6	0.09	1.11	1.01	1.02	1.02	1.15	1.05	1.00	1.01	1.01	1.08
12	0.14	1.07	0.98	0.98	0.98	1.13	1.01	0.96	0.95	0.95	1.05
24	0.21	0.87°	0.81*	0.84°	0.83°	0.87	0.82•	0.81•	0.82•	0.82•	0.84*
36	0.23	0.81*	0.77•	0.80*	0.79•	0.80*	0.78•	0.74•	0.78•	0.77•	0.76•
						MAIZE					
1	0.03	0.99	1.00	1.00	1.00	0.99	0.95*	0.97*	0.97°	0.97°	0.95°
6	0.10	0.95	0.97	0.97	0.99	0.94	0.92	0.98	0.98	0.98	0.93
12	0.14	0.94	0.97	0.98	0.99	0.94	0.91	0.96	0.97	0.96	0.91
24	0.19	0.84*	0.93	0.94	0.94	0.85°	0.80•	0.92°	0.92°	0.92	0.81*
36	0.22	0.77•	0.90*	0.92	0.90°	0.78•	0.75•	0.86*	0.90*	0.90°	0.75•
						RICE					
1	0.03	0.99	0.98•	1.00	1.00	1.00	0.95*	0.95•	0.96*	0.96*	0.97
6	0.08	0.99	0.97	1.01	1.01	1.08	0.99	0.99	1.03	1.03	1.05
12	0.11	0.93	0.94°	1.01	1.01	0.91	0.89	0.96	1.02	1.02	0.90
24	0.17	1.04	0.99	1.04	1.05	0.99	0.84•	0.91°	0.96	0.96	0.83*
36	0.21	1.06	1.02	1.03	1.03	1.09	0.88*	0.91*	0.93°	0.93°	0.89*
						WHEAT					
1	0.04	1.01	1.00	1.00	1.00	1.00	1.00	0.99	0.99	0.99	0.99
6	0.10	1.07	0.99	0.98	0.99	1.06	1.02	0.99	0.98	0.98	1.01
12	0.14	1.03	0.99	0.99	0.99	1.04	0.98	0.98	0.97	0.97	0.98
24	0.18	1.00	1.00	0.98	0.98	1.01	0.91	0.94	0.94	0.94	0.93
36	0.20	0.93°	0.93°	0.91*	0.90*	0.95	0.89*	0.91°	0.92°	0.92°	0.90*

Notes: For RW the figures describe the average values of the CRPS, whereas for the remaining models the ratio relative to the RW benchmark. Values below unity indicate that density forecasts from a given model are more accurate than those from the benchmark. Symbols •, * and ° denote, respectively, the 1%, 5% and 10% significance levels of the one-tailed Amisano and Giacomini (2007) test, where the long-run variance is calculated with the Newey-West method.
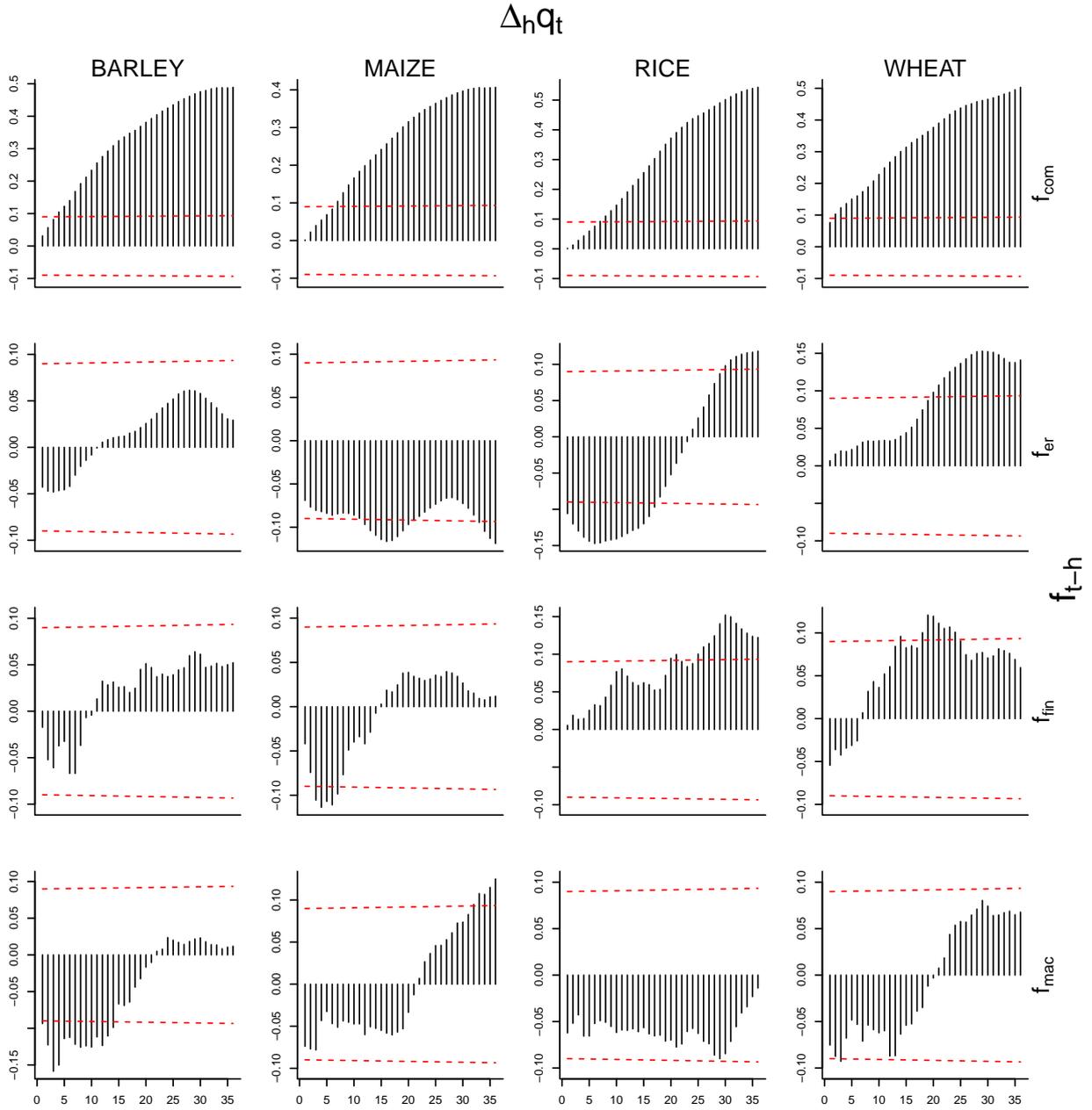
Figure 1: The dynamics cereal prices
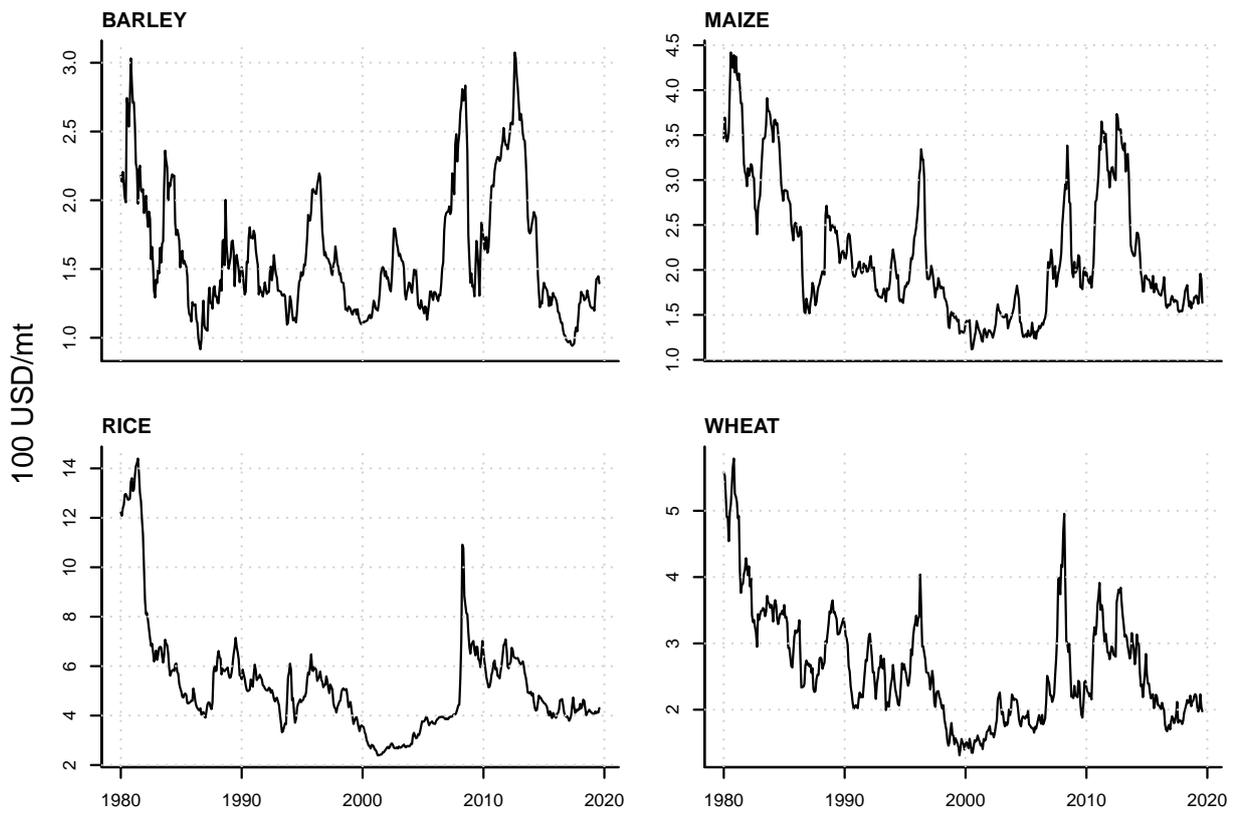
Note: The figure presents the dynamics of prices for four cereals in real prices from last observation (Aug 2019).

Figure 2: The dynamics of factors
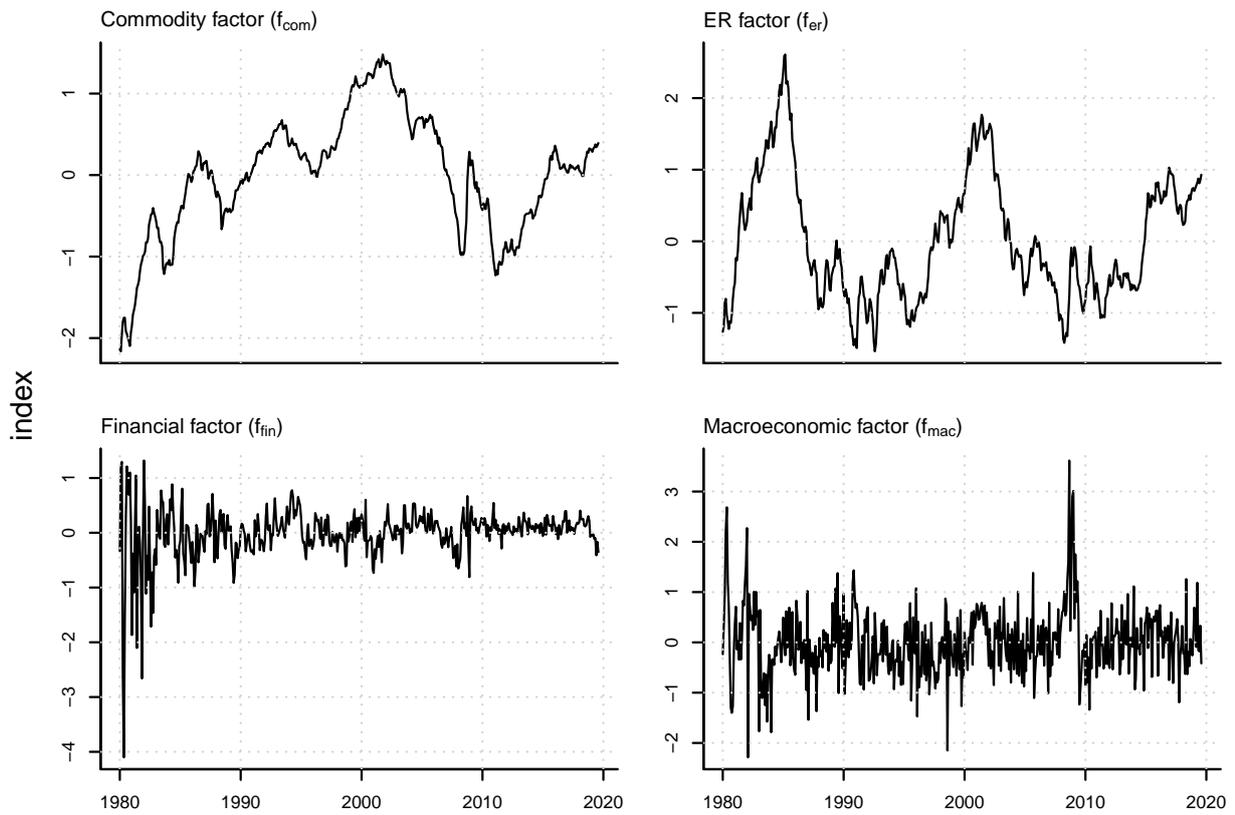
Note: The figure presents the dynamics of four factors: commodity, exchange rates, financial and macroeconomic.

Figure 3: In sample correlations

Note: The figure presents the correlations of growth rates for cereal prices $\Delta_h q_{it}$ and h-lagged values of factors $f_{k,t-h}$ with $h = 1, 2, \ldots, 24$. Dashed lines represent 95% confidence intervals.