Common Factors of Commodity Prices*

Simona Delle Chiaie[†], Laurent Ferrara[‡], Domenico Giannone[§]

3rd March 2017

Abstract

In this paper, we extract common factors from a large cross-section of commodity prices. We find that the bulk of the fluctuations in commodity prices is well summarised by a single global factor. This global factor is closely related to fluctuations of global economic activity and its importance in explaining commodity price variations has increased since the 2000s, especially for oil prices.

Keywords: commodity prices, dynamic factor models, forecasting.
JEL Classifications: C51, C53, Q02.

*The views expressed in this paper are those of the authors and do not necessarily reflect those of the Banque de France, the European Central Bank, the Eurosystem, the Federal Reverse Bank of New York or the Federal Reserve System. We are grateful to Christiane Baumeister, Daniele Caratelli, Luca Guerrieri, Raju Huidrom, Lutz Kilian, Evgenia Passari and Robert Vigfusson for helpful comments and suggestions. We would also like to thank seminar participants at the World Bank, at the International Symposium of Forecasting 2013 in Seoul, at the Vienna Workshop "High-dimensional Time Series in Macroeconomics and Finance 2015", at the 2nd Annual Conference of the International Association for Applied Econometrics 2015 and at the 1st Bank of England-Banque de France Workshop on International Macroeconomics 2015 for their useful comments.

[†]Corresponding author: European Central Bank, Sonnemannstrasse 20, 60314 Frankfurt am Main (e-mail address: simona.dellechiaie@ecb.europa.eu).

[‡]Banque de France.

[§]Federal Reserve Bank of New York and CEPR.

Non-technical summary

In the past decade an influential body of research has shown the importance of modelling both supply and demand shocks in order to analyse commodity price fluctuations. Notably, the empirical research that has focussed on oil price fluctuations, has illustrated the importance of identifying shocks to the demand for industrial commodities, driven by changes in global real activity, as opposed to supply shocks and other oil-market specific shocks. In fact, since oil prices, like other commodity prices, respond endogenously to world economic conditions, global demand shocks trigger different dynamic effects on macroeconomic aggregates compared to exogenous oil supply shocks. Since then, global demand shocks have become prominent in the scholarly literature on commodity prices. In fact, the fast macroeconomic expansion of China and other emerging market economies caused an unprecedented increase in the demand for all industrial commodities and resulted in the across-the-board surge in commodity prices observed since 2003. Interestingly, the boom and the burst that followed the Great Recession also revealed that the co-movement among a panel of energy and non-energy commodity prices increases in response to changes in global demand conditions.

This paper uses a dynamic factor model to study the co-movement in commodity price data and shows that this model is able to provide important insights regarding the underlying forces of commodity price variations. To this aim, the paper extracts latent factors from a large panel of both energy and non-energy commodity prices. We use monthly data from the IMF primary commodity database over the sample 1980 Jan-2015 Dec. In particular, we decompose each commodity price series into a global (or common) component, block-specific components related to specific categories of commodity prices and a purely idiosyncratic shock. Our empirical findings show that a single global factor drives the bulk of commodity price fluctuations. This global factor is persistent and closely related to changes in global economic activity. While the global factor has become more important in explaining commodity price changes since the 2000s, its role has increased especially for the price of oil.

Model-based historical decompositions of commodity price changes also show that the global factor explains a larger fraction of commodity price fluctuations during episodes typically associated with changes in global demand conditions, such as the world economic expansion that started around 2003 and the steep contraction during the Great Recession. For oil prices, historical decompositions based on the estimated model also match conventional narratives on the genesis of oil price fluctuations. We found that fuel-specific factors are the main underlying sources of oil price changes that occurred before the 2000s, such as the collapse of OPEC in 1986 and the Persian Gulf War of 1990-1991. As regards the decline in the price of oil that started in the second half of 2014, we find that the global factor explains about one-third of the oil price fall from June 2014 to December 2015 but the largest fraction of the decline is captured by fuel-specific components.

We perform an out-of-sample validation of the model. We find that the factor model

performs well in forecasting commodity prices, in particular at short horizons. The predictive component of the global factor seems to be stronger for the group of commodities for which the variance explained by the global factor is larger, such as food and metals. We also find that for some commodities, in particular crude oil, the predictive content of the global factor has increased during the Great Recession.

1 Introduction

In the past decade an influential body of research has shown the importance of modelling both supply and demand shocks in order to analyse commodity price fluctuations. Building on insights in Barsky and Kilian (2002), Kilian (2009) illustrated the importance of identifying shocks to the demand for industrial commodities, driven by changes in global real activity, as opposed to supply shocks and other oil-market specific shocks. In fact, since oil prices, like other commodity prices, respond endogenously to world economic conditions, global demand shocks trigger different dynamic effects on macroeconomic aggregates compared to exogenous oil supply shocks.

Since then, global demand shocks have become prominent in the scholarly literature on commodity prices (see, e.g. Aastveit, Bjørnland and Thorsrud (2014), Kilian and Murphy (2014) and Baumeister and Kilian (2016)). As also noticed by Hamilton (2009) and documented in Kilian and Hicks (2013), the fast macroeconomic expansion of China and other emerging market economies caused an unprecedented increase in the demand for all industrial commodities and resulted in the across-the-board surge in commodity prices observed since 2003. Interestingly, the boom and the burst that followed the Great Recession also revealed that the co-movement among a panel of energy and non-energy commodity prices increases in response to changes in global demand conditions. Barsky and Kilian (2002) and Juvenal and Petrella (2014) show that the conditional correlation between oil prices and the prices of other commodities rises in face of global demand shocks. If those shocks are more persistent than transitory idiosyncratic shocks, such as supply shifts, then the theory suggests that lower interest rates can further increase price correlation by reducing the cost of inventory holdings (Gruber and Vigfusson (2016)).

In this paper, we propose a dynamic factor model to study the commonality in commodity price data and we show that this model is able to provide important insights regarding the underlying forces of commodity price variations. Our empirical strategy consists in extracting latent factors from a large panel of commodity prices. In doing so, we decompose each commodity price series into a global (or common) component, block-specific components related to specific categories of commodity prices and a purely idiosyncratic shock. The distinction between global, block-specific and idiosyncratic components reflects the underlying idea that different commodity price shocks have distinct consequences on the cross-correlation among commodity prices. Global demand shocks stemming from unexpected changes in global economic activity are likely to affect all or most of commodity prices by shifting the demand curve in most of these markets. Those shocks are thus likely to be pervasive and move prices in the same direction. By contrast, shocks that are commodity-specific, such as supply shocks, are unlikely to represent a common source of variation for all commodity prices since they have a more limited propagation across all markets and capture only sectoral dynamics.

The empirical evidence in this paper shows that the global factor estimated from a large panel of commodity prices is closely related to changes in global economic activity. We find that there is a single global factor driving the bulk of commodity price fluctuations and, by moving prices in the same direction, the global factor has limited effects on relative prices. The global factor is persistent and follows the major expansion and contraction phases in the international business cycle with the largest declines following recession periods. Notably, since the relevance of the global factor has increased since the start of the new millennium, not only the global factor has become more important in explaining commodity price changes but oil prices have become more correlated with other non-fuel commodities because of common global forces.

Some recent empirical studies have also analysed the covariation in commodity prices by extracting common factors and related them to indicators of real economic activity and financial variables, such as the world industrial production and the US exchange rate (see, e.g. Byrne et al. (2011), Chen et al. (2014) and West and Wong (2014), Alquist and Coibion (2014)). In contrast to that literature, one important finding of our paper is that there is only one single factor in commodity prices and this is closely related to global demand conditions. Since these earlier studies did not take into account the presence of local correlation that are not pervasive enough to be treated as common factors, they tended to confuse block variations with truly common ones.

We compute the model-based historical decompositions of commodity price changes and we find that the global factor explains a larger fraction of commodity price fluctuations during episodes typically associated with changes in global demand conditions, such as the world economic expansion that started around 2003 and the steep contraction during the Great Recession. For oil prices, we found that fuel-specific factors are the main underlying sources of oil price changes that occurred before the 2000s, such as the collapse of OPEC in 1986 and the Persian Gulf War of 1990-1991. The structural analyses in the earlier works of Kilian and Murphy (2014) and Kilian and Lee (2014) support this result, showing that exogenous shifts in supply and other oil-specific demand shocks were a more important determinant of the price of oil before the 2000s.

In order to verify the robustness of the modelling strategy, we perform an out-of-sample validation of the model. We find that the factor model performs well in forecasting commodity prices and indices of commodity prices, in particular at short horizons. The predictive component of the global factor seems to be stronger for the group of commodities for which the variance explained by the global factor is larger, such as food and metals. Other studies, in particular focusing on forecasts of the price of oil, have also found evidence that proxies of global demand have predictive power for commodity prices (see, Baumeister and Kilian (2012)). Similarly, changes in index of commodity prices, in particular industrial raw materials, have been proved to improve the forecast of the price of oil, as those price changes are more likely to capture shifts in the global demand for industrial commodities (see, Alquist, Kilian and Vigfusson (2013)). In this respect, the factor-based forecast is a refinement of this earlier approaches. The out-of-sample exercise also shows that for some commodities, in particular crude oil, the predictive content of the global factor has increased during the Great Recession.

The rest of this paper proceeds as follows. Section 2 presents the empirical analysis, the global factors and analyses the sources of commodity price fluctuations. Section 3 studies the predictability and the local forecasting performance of the model. Finally, Section 4 concludes.

2 Empirical analysis

2.1 Data

The data x_t used in the estimation include the spot prices of 52 commodities from different categories, including food, beverages, agricultural raw materials, metals and fuel commodities. Commodity prices come from the IMF primary commodity price database and they cover the period from January 1980 to December 2015[1]. Since the estimation of the model described in the previous section requires covariance-stationary variables, all prices have been taken in log differences. The data have been further standardised to have a zero sample mean and a unit sample variance. It is worth noting that the IMF primary commodity price database also includes 10 price indices and sub-indices, representing the major commodity sectors, constructed as weighted averages of individual commodity prices. The weights used for computing the indices depend on the commodity trade value compared to the total world trade as reported in the UN Comtrade database and they are updated about every 5 years. An interesting feature of these indices is that they are built to reflect different levels of aggregation; the data set is thus characterised by a block structure summarised in Table 1. The first or global index is constructed as a weighted average of all commodity prices in the dataset, hence it represents a broad index of commodity prices. This, in turn, is divided into two main block indices which are constructed using non-fuel and fuel commodity price series, respectively. As shown in Table 1, the energy block represents about 60 percent of the overall index. The non-fuel block can then be broken down into two other group indices, edibles and industrial inputs which are in turn divided into five other subcategories representing food, beverages, agricultural raw materials and metals. The fuel block index contains only one subcategory, represented by the index of crude oil prices. The full list of series used in the analysis and their descriptions is reported in Table 2.

[1] A few series are not available from the beginning of the sample but start only in the 1990s. Maximum likelihood estimates can be adopted to deal with missing data (see, Banbura and Modugno (2014)).

2.2 Model and estimation

The model used here is an approximate dynamic factor model for large cross-sections. This model provides a parsimonious representation of the dynamic co-variation among a set of random variables. Consider an n-dimensional vector of commodity returns $x_t = (x_{1t}, ..., x_{nt})'$ with mean zero. Under the assumption that x_t has a factor representation, each series x_{it} is the sum of two unobservable components, a common component - capturing the bulk of cross-sectional co-movements - and an idiosyncratic component reflecting specific shocks or measurement errors:

$$x_{it} = \lambda_i f_t + e_{it} \tag{1}$$

where $f_t = (f_{1t,...,}f_{rt})^{'}$ is an r-dimensional vector of common pervasive factors affecting all commodities; $\lambda_i = (\lambda_{i1,...}\lambda_{ir})^{'}$ is a vector of factor loadings where each of the element in λ_i measures the effect of the common factors to commodity i; e_{it} is the idiosyncratic component which is assumed to be non-pervasive and weakly correlated across commodities. The common factors f_t and the idiosyncratic component e_{it} are uncorrelated at all leads and lags. Note that if λ_i is similar across commodities, then f_t has a limited impact on relative prices. We model the common factors as following an autoregressive process of finite-order:

$$A\left(L\right) f_t = u_t \tag{2}$$

where $A\left(L\right)= I - A_1 L - \ldots A_p L^p$ an $(r \times r)$ filter of finite length p with roots outside the unit circle, and u_t is a Gaussian white noise, $u_t \sim i.i.d\mathcal{N}\left(0, I_r\right)$.

We further assume that the model has a block structure that represents a parsimoniuos way to model the local correlation among idiosyncratic components. This implies decomposing e_{it} into factors that are specific to groups or blocks of commodities and a purely idiosyncratic component:

$$e_{it} = \sum_{j=1}^{K} \gamma_{ij} g_{jt} + v_{it} \tag{3}$$

$$\gamma_{ij} = \begin{cases} \neq 0 & if\, i\, \in j \\ 0 & otherwise \end{cases}$$

where g_{jt} is (an r_b-dimensional vector) of block-factors; γ_{ij} are block-factor loadings and v_{it} is the purely idiosyncratic disturbance. The block factors g_{jt} and the purely idiosyncratic component v_{it} are assumed to follow an autoregressive process of finite-order:

$$g_{jt} = \phi_j g_{jt-1} + w_{jt} \tag{4}$$

$$v_{it} = \rho_i v_{it-1} + \varepsilon_{it} \tag{5}$$

with $w_{jt} \sim i.i.d\mathcal{N}(0,1)$ and $\varepsilon_{it} \sim i.i.d\mathcal{N}\left(0,\sigma_i^2\right)$.

Principal components are obtained as a special case of our estimates, under the following assumptions:

$$\begin{cases} \gamma_{ij} = 0, \forall i, j \\ \rho_i = 0, \forall i \\ \sigma_i^2 = \bar{\sigma}, \forall i \end{cases}$$

Maximum likelihood estimation is implemented using the Expectation Maximization (EM) algorithm as in Doz et al. (2012). The algorithm consists of two steps. In the first step (M-step), the algorithm is initialised by computing principal components and the model parameters are estimated by OLS regression treating the principal components as if they were the true common factors. Since these estimates have been proved to be an asymptotically consistent estimator of the true common factors (see, Forni et al. (2000), Stock and Watson (2002a, 2002b) and Bai (2003)), the initialisation is good when the cross-section dimension is large. In the second step, given the estimated parameters, an updated estimate of the common factors is obtained using the Kalman smoother. If we stop at the second step, we have the two-step estimates of the common factors studied by Doz et al. (2011). Maximum likelihood is obtained by iterating the two steps until convergence, taking at each step into account the uncertainty related to the fact that factors are estimated.

It is worth noting that in order to keep the number of parameters limited, we have assumed a parsimonious parameterisation of the idiosyncratic dynamics. However, Doz et al. (2012) have shown that under the approximate factor structure (i.e. pervasive factors and limited cross-sectional correlation among idiosyncratic components), maximum-likelihood estimates of the model are robust to misspecification of the cross-sectional and time series correlation of the idiosyncratic components. Moreover, the estimates have been shown to be robust also to non gaussianity. In this respect, the estimator is a quasi-maximum likelihood estimator in the sense of White (1982).

A growing body of research has applied the quasi-maximum likelihood estimator to extract common factors from large cross-sections for a variety of empirical applications. For instance, this method has become a popular tool for now-casting (see, for surveys, Banbura, Giannone and Reichlin (2011), Banbura, Giannone, Modugno and Reichlin (2013) and recently, Luciani (2014)). Banbura, Giannone and Lenza (2015) applied this approach to perform conditional forecasts and scenario analyses; Brave and Butters (2011) constructed a high-frequency indicator of national financial conditions published by the Federal Reserve Bank of Chicago. This method has also been used for structural analyses, as done, example, in Reis and Watson (2010) and Luciani (2015).

2.3 How many factors and blocks?

We begin our analysis by estimating the common factor (the global factor, henceforth) and the other block-specific components using likelihood-based methods described in the previous section. Note that the factor estimates are computed using the commodity returns in x_t which excludes the higher level aggregates represented by the commodity indices. In this way, we avoid introducing - by construction - collinearity in the panel data since commodity indices are linear combinations of commodity prices. First, we determine the number of blocks to include in the model by following the structure of our database. Since typically, in macroeconometrics, data are organised either by country, sectoral origin or economic concept, the empirical literature on factor models has mostly looked at the composition of the data set to have guidance on the extraction of the blocks. For instance, Forni and Reichlin (2001) distinguish between European and national components to study the potential degree of output stabilization deriving from federal policies; using Bayesian methods, Kose, Otrok, and Whiteman (2003) study the sources of the international business cycle by estimating world, country and regional components; Banbura, Giannone and Reichlin (2011) use blocks derived from nominal and real variables for the purpose of nowcasting real economic activity; Miranda-Agrippino and Rey (2015) decompose fluctuations in risky assets into global, regional and asset-specific components. Like these papers, we extract local factors that reflect the composition of the panel data that in our case, as shown in Table 1, is based on different categories of commodities. Thus, we extract two main block factors (fuel and non-fuel), two sub-block factors (food and beverages and industrial inputs) and finally, five group factors (food, beverages, agricultural raw materials, metals and oil). Second, we determine the optimal number of global factors from the observed data. In doing so, we take into account the trade-off between the goodness-of-fit and the loss in parsimony that arises from increasing the number of factors using a modified version of the information criterion in Bai and Ng (2002). These authors derive a penalty function to select the optimal number of factors in approximate factor models when factors are estimated by principal components. The statistical approach of Bai and Ng (2002) can be extended to any consistent estimator of the factors provided that the penalty function is derived from the correct convergence rate. For the quasi-maximum likelihood estimator used in this paper, Doz et al. (2012) show that the convergence rate for the factor estimates is given by $C_{NT}^{*2} = min\{\surd T, (N/(logN))\}$. Hence, a modified version of the Bai and Ng (2002) information criterion (IC) is given by:

$$IC^{*}(r) = log(V(r, F_{(r)})) + rg(N,T),\ g(N,T) = ((log(C_{NT}^{*2}))/(C_{NT}^{*2}))$$

where r is the number of common factors, $F_{(r)}$ denotes the estimated factors, $V(r, F_{(r)})$ is the sum of squared idiosyncratic components divided by NT and finally, $g(N,T)$ represents the penalty function for over-fitting.[2] For our panel data, the figures in Table 3 select the model with only one global factor since this provides the smallest value of the IC statistics.

2.4 Empirical results

The Global Factor

Figure 1 shows the global factor estimated over the full sample. The figure plots the global factor along with the IMF global index of commodity prices. The main motivation for comparing the two indices is that one key feature of the estimated global factor, which might turn out to be useful in practical applications such as forecasting, is that it filters out the idiosyncratic noise in the data to a greater extent than simple averages of commodity prices. In fact, a first and simplest way of extracting a global factor and thus, smoothing out noisy fluctuations in the data is by cross-sectional averaging. Commodity indices, such as the IMF index shown in Figure 1, are linear combinations of commodity prices with weights given by trade values. While cross-sectional averages tend to approximate well the global factor in case of limited cross-correlation among idiosyncratic disturbances (see, for instance Forni and Reichlin, (1998)), in practical applications, simple averages may have a substantial component of noise arising from the idiosyncratic component. Figure 1 illustrates this point. The global factor and the IMF broad index of commodity prices clearly resemble each other and they are indeed highly correlated,[3] however a key difference between the two series lies in their second-order properties. The magnitude of the standard deviation of the overall index of commodity prices is almost twice the standard deviation of the global factor, indicating a smaller variability of the global factor relative to the IMF broad index of commodity prices. A visual inspection of the two series also confirms that while the broad index of commodity prices is characterised by ample fluctuations, for instance those associated with the oil price shocks in the early 1990s, these fluctuations appear to be more limited in the global component as the latter smooths out the idiosyncratic noise as well as local dynamics that do not propagate globally, resulting in a smoother and more persistent series. Particular attention should clearly be paid to what the global factor captures. Recent studies on commodity prices based on principal components have related the estimated common factors to macroeconomic and financial data, such as the real interest rate, the US exchange rate and industrial production (see, e.g. Byrne et al. (2011), Chen et al. (2014) and West and Wong (2014)) by means of simple correlations. Alternatively, studies in the financial literature, in particular Tang and Xiong (2012), have related the co-movement among commodity prices to the growing partici-

[2] The information criterion has been recently applied to the quasi-maximum likelihood estimator by Coroneo et al (2014).

[3] The correlation coefficient between the two series is 0.63.

pation of financial speculators in commodity markets. However, a large body of research, for instance Fattouh et al (2012), Kilian and Murphy (2014), Kilian and Lee (2014), Juvenal and Petrella (2014), among others, did not find substantive evidence supporting this argument.

A natural conjecture is that the global factor, being a pervasive shock affecting a large cross-section of commodity prices, might capture movements in global economic activity. A booming or slumping global economy clearly affects the demand for a broad group of commodities moving the prices in the same direction. This interpretation also finds its theoretical underpinning in Alquist and Coibion (2014) general equilibrium model of global business cycle with commodity prices. The theoretical model predicts that commodity prices can well be represented by a factor structure. The first factor, in particular, captures aggregate shocks that move prices in the same direction and, in the absence of important income effects on the common input into the production of commodities, can be interpreted as a global demand shock. As shown in Figure 2, consistently with the theory, the factor loadings associated with our estimated global factor are mostly positive and are larger for food and metal commodities. Second, the global factor tends to co-move with real indicators of global economic activity. To show this, Figure 3 plots the global factor along with the index of world real economic activity proposed by Kilian (2009). This measure, based on dry cargo ocean freight rates, has been explicitly developed to capture shifts in the demand for industrial commodities associated with periods of high and low real economic activity (see, for details, Kilian (2009)). In order to make the comparison meaningful, the global factor here is extracted from the growth rate of real commodity prices and is expressed in year-on-year growth rates. As a matter of fact, in a standard microeconomic model of demand and supply, the commodity price in the vertical axis is expressed in real terms. Hence, it would be more appropriate to use real commodity prices since the aim is to capture the changes in the real price induced by shifts of the demand curve along the supply curve. However, we have estimated the model both using real and nominal prices and none of the results of the paper appeared to be sensitive to this transformation. This might reflect the fact that our sample does not include high inflation periods, such as the Great inflation of the 70s. Figure 3 also plots vertical bars that correspond to periods of widespread decline in world economic activity. Notably, we include the early 1980s and 1990s recessions, the Asian financial crisis of 1997–1998, the recession that followed the bursting of the dot-com bubble in the 2000s, the Great Recession and, finally, the latest euro area recession starting in the third quarter of 2011. Figure 3 shows that the global factor is positively correlated with the Kilian's index of economic activity. The figure shows clear evidence that both measures follow the major expansion and contraction phases in the international business cycle over the period considered with the largest declines following recession periods. The fast macroeconomic expansion that characterised the world economy from around 2001, and in particular, some emerging market economies, is also captured by the two series. Looking at the most recent period, we observe that both measures indicate a weakening in the global economy, which is consistent with the observed decline in commodity prices. Likewise, Figure 4 compares the global factor with monthly indicators of industrial production for three selected areas (world, advanced and emerging economies) as provided by the CPB Netherlands Bureau for Economic Policy Analysis. Yet, the correlation between these series is clearly positive and very high (0.75). Finally, Figure 5 shows the global factor with the growth rate of the price of Brent crude oil together with estimates of global demand and supply of oil. Three observations can be made. First, the correlation between the global factor and oil prices is positive but the co-movement between the two series has increased in the last decade. Second, the global factor and the price of oil are positively correlated with measures of world consumption of oil. Third, the spikes in the price of oil that coincided with some exogenous events in the oil market such as the Persian Gulf War, the Venezuela crisis which was followed by the Iraq invasion in 2003 are not associated with similar variations in the global factor. Rather, these appear to be associated with important negative shifts in the supply of oil.

Sources of commodity price fluctuations

In this section, we study the relative importance of global and block-specific components in explaining the variance of the commodity price changes. Notice that a model-based variance decomposition can also be derived for the commodity indices in our data set. Let y_t be an m-dimensional vector the vector of commodity indices and W be a given $(m \times n)$ matrix of weights used to compute the indices, then the variance-covariance matrix of y_t is $\Sigma_y = W\Lambda\Sigma_f\Lambda\prime W\prime + W\Sigma_e W\prime$ where Σ_f and Σ_e are the variance-covariance matrices of the factors and idiosyncratic components, respectively. Tables 4-5 present the variance decomposition for all the disaggregated commodity prices and for the commodity indices estimated over the full sample period. The global factor explains an important fraction of the variations, especially for non-fuel commodities, such as food and metals. For instance, this is the case for soybeans and soybean oil (for which the global factor accounts for about 50 percent of the price variation) and to a lesser extent, copper, aluminum, palm and sunflower oil, maize and barley. As a result, almost 70 percent of the variations in the index of non-fuel commodities is attributed to the global factor. Beverages and agricultural raw materials seem to be less correlated with other commodity prices over the full sample, thus their variances are mostly attributed to block-specific components while the global factor captures about 10 percent of the variance. The bulk of the fluctuations in oil prices is also captured by fuel-specific shocks, although the global factor accounts for a non-negligible fraction of the variance (about 20 percent). Overall, a third of the variance of the IMF overall index of commodity prices is attributed to the global factor while the remainder is due to the fuel-specific factor. To check the robustness of these results, we include a second global factor in the model and we calculate the variance explained by the first two global factors. We report the results in Figures 6 and

7. Figure 6 shows the variance decomposition for the indices, while Figure 7 gives the results for the disaggregated commodity returns. Both figures show that the variance explained by the first global factor is robust to the inclusion of a second global factor. Moreover, the latter captures only a small or tiny fraction of the variance. For instance, looking at the commodity indices, the fraction of the variance explained by the second global factor ranges between 0.2 and 7.3 percent. This finding supports the previous evidence provided by the IC statistic. The presence of a single global factor is interesting and it contrasts with previous empirical studies, for instance, Byrne et al (2011), Chen et al. (2014) and West and Wong (2014) among others, who have found than two or more common factors provide a good representation of the data. Both the theory and the empirical analysis in Alquist and Coibion (2014) also predict the presence of two orthogonal common factors. The first one captures aggregate shocks that move price in the same direction while the second factor instead constists of a mix of shocks that directly affect the supply or demand curves. However, However, because of the orthogonality assumption, local comovement originating from common disturbances such as sectoral technology shocks or interdependencies in commodity markets are disregarded. In these previous works, the presence of two common factors in panel of commodity prices is likely to stem from the absence of the block-factor structure since non-pervasive block variations are captured by weak common factors. In order to illustrate this, we compare the global and block factors of our benchmark specification (M_1) with three common factors extracted from a factor model where the local correlation among idiosyncratic components (M_2) is not modeled. By computing linear projections of the common factors of M_2 on M_1 factors, one would find that the first common factors are very much alike, the second common factor in M_2 is akin to the first block component (Non-Fuel) in M_1, while the last factor shows high correlation with the Fuel block and, to a lesser extent with the Food and Beverages sub-block. Guided by these results, Figure (5) shows visually the correlations between the common factors in M_2 and selected block factors of the benchmark specification.

Sub-sample analysis

The analysis over the full sample might mask some important changes that might have happened at the beginning of the 2000s, in particular, since the start of the commodity price boom in mid-2003. In Figure 8, we show the model-based variance decomposition for all the commodity price indices over two sub-samples. We use 2003 as a break date in line with the observed increase in commodity prices. We find that the global factor does not have any explanatory power for oil and other fuel-commodities before the 2000s while it explains a good portion of the variation of non-fuel commodities. Block-specific and idiosyncratic components thus account for the whole variation in oil prices in the first sub-sample. While this result might appear surprising at first sight, it just unveils that supply or other oil-specific demand shocks were, on average, a more important determinant of the price of oil than global demand shocks in the first part of the sample. The structural analysis in Kilian and Murphy (2014) supports this interpretation and it shows how many key historical events in the oil market over these two decades, such as the collapse of the OPEC cartel in 1986, the Gulf war in 1990-91 and the Venezuela crisis in 2002, mostly reflected shocks to the speculative demand of oil together with supply shifts. Interestingly, the sub-sample analysis shows that not only the explanatory power of the global factor has risen substantially since the 2000s for all commodity prices but its role has increased markedly for both the price of oil and the price of industrial inputs, such as metals. In particular, we find that the global factor explains about 60 percent of the variability in the price of metals since the 2000s and more than 40 percent of the variation of oil prices. As a result, the variance of the overall index of commodity prices that can be attributed to the global factor has increased from less than 10 percent in the period preceding the 2000s, to 60 percent afterwards. The historical decompositions in the next section will provide further evidence on the increasing role of the global factor in capturing commodity price movements in the new millennium.

Historical decompositions of commodity price changes

In this section, we present the decomposition of commodity price fluctuations in some key historical episodes. Two important events in the commodity market are, for instance, the world macroeconomic expansion that started around 2001 which coincided with a surge in commodity prices and the Great Recession that resulted in a dramatic across-the-board declines in commodity prices. For reasons of space, Figure 9 reports the historical decomposition of the price for a subset of commodities. In particular, we report the historical decomposition for crude oil, copper, nickel and corn. The figure shows the cumulative effects at each point in time of global and block-specific factors from January 2000 to July 2008. The decomposition of the price of oil in the upper panel of the chart indicates that the cumulative effect of shifts in the global factor largely explains the oil price surge in the 2000s while the fuel-specific component has a more limited explanatory power. The latter component seems to be more important in explaining the increase in the price of oil in the early 2000s. These findings are consistent with previous results in the literature documenting how the increase in the price of oil in the early 2000s was driven by factors specific to the demand of oil while the rise since early 2002 was primarily due to a booming world economy (see, e.g Kilian (2009), Kilian and Murphy (2014)). Historical decompositions of the remaining commodity prices also point to the global factor as the main source of the price surge.

Figure 10 focuses on oil prices and provides further evidence on the ability of the factor-based decomposition to match conventional narratives on the sources of oil price fluctuations. In particular, we look at four historical episodes of large oil price variations. The first two events in Figure 10 refer to the oil price fall that followed the collapse of the OPEC cartel in 1986 and the oil price spike that occurred in response to the Iraqi invasion of Kuwait in

1990. These can be viewed as examples of price variations that are specific to the oil market and unrelated to changes in macroeconomic conditions. The historical decomposition of the data shows that, in both episodes, the main factors underlying the variations in the price of oil are fuel-specific, while the global factor had clearly no role (Figure 10, panel 1 and 2). The third episode, in Figure 10, corresponds to the dramatic fall in the price of oil resulting from the contraction in world real activity which is explained by a large extent by the global factor. However, shocks to the oil-specific component also exerted further downward pressures on the price in 2008. Remarkably, these results are consistent with previous studies based on structural vector autoregressions for the global oil market, such as Kilian and Murphy (2014) and Kilian and Lee (2014) who find evidence of increased inventory demand in the first months of 2008. Finally, the last panel of Figure 10 investigates the underlying causes of the decline in the price of oil starting in the second half of 2014. Baumeister and Kilian (2016) have provided a first quantitative analysis, based on a vector autoregressive model of the global oil market, of the oil price drop between June 2014 and December 2014. These authors find that a slowndown in global demand for oil was the main cause of the earlier oil price decline in addition to a mix of shocks to actual and/or expected global oil supplies prior to July 2014 and a shift in oil price expectations in July 2014. Kilian (2017) also provides a structural analysis of the oil price fall and investigates the extent to which the US shale oil boom played a role in the oil price decline. Our work extends these earlier works and analyses the underlying causes of the decline in the price of oil for the entire episodes up to December 2015. The historical decomposition shows that the global factor had an important role in the oil price decline as it explains about one-third of the oil price fall from June 2014 to December 2015. While the global factor explains most of the initial oil price fall, differently from the earlier analysis of Baumeister and Kilian (2015), a large positive shock to the fuel-specific component also took place at the end of 2014 and coincides with the OPEC announcement in November 2014. The weakening of the global economy exerted further downward pressures on the price of oil, however the slowdown in global economic activity appears to be more pronounced in the second half of 2015. Other recent studies based on correlations of oil price changes with equity prices or a broader array of financial variables to quantify the drivers of oil price movements provide similar results (see, for instance, Baffes et al. (2015) and Groen and Russo (2015)).

3 Predictive content of the global factor

The increased level of co-movement in commodity prices has also sparked a growing empirical literature that uses factor models estimated on a panel of commodity prices for forecasting purposes. For instance, common factors have been used to forecast commodity prices themselves (see, e.g., West and Wong (2014) and Poncela et al. (2015)) or other macro-variables such as inflation (Gospodinov and Ng (2013)). Other studies have also investigated whether macroeconomic and financial data may have predictive power for commodity prices. For instance, Chen, Rogoff and Rossi (2012) found that the US exchange rate helps predict commodity prices. Groen and Pesenti (2011) use a large data set containing a variety of indicators of supply and demand conditions. While the inclusion of these variables does not appear to yield to large gains in the predictive accuracy, other studies, in particular focusing on forecasts of the price of oil, have found strong evidence that proxies of global demand have predictive power for commodity prices (see, Baumeister and Kilian (2012)). Similarly, changes in the spot price of industrial raw materials have been proved to improve the forecast of the price of oil, as those price changes are more likely to capture shifts in the global demand for industrial commodities (see, Alquist, Kilian and Vigfusson (2013)).

In this section, we verify the robustness of the modelling strategy and we perform an out-of-sample validation of the model. In this respect, we investigate whether the global factor has predictive power for commodity prices by conducting a real-time forecasting exercise. To this end, we take the perspective of a researcher who every month starting from January 2001, would have estimated the model using 20 years of past data, and used the estimates to compute out-of-sample forecasts of commodity prices each month from February 2001 to December 2015. The h-step ahead forecasts for individual commodity prices are iterated from the state-space representation using the Kalman filter while forecasts for the aggregate commodity indices are computed as weighted averages of the individual commodity price forecasts using the associated trade weights.[4] After computing the sequence of the out-of-sample forecast error loss differences between the model and a naive benchmark (i.e. a constant growth model), we assume that the researcher computes both the average loss difference over the entire evaluation sample as well as rolling average losses along the lines of Giacomini and Rossi's (2010) fluctuation test. This test, which is useful to study the forecasting performance of a model in a unstable environment, implies to compute the difference between the mean squared forecast error (MSFE) of the factor model and the benchmark, smoothed over time with a centered rolling window of fixed size. We select a window of 4 years. The statistical significance of the relative performance of the model against the benchmark is then tested at each point in time using the Diebold and Mariano's (1995) test of equal predictive accuracy.

The main results of the out-of-sample forecasting exercise can be summarised as follows. First, we observe that the model performs well in predicting commodity prices at shorter horizons. Table 6 shows the relative MSFE of the model against the benchmark at two different horizons. A ratio smaller than 1 indicates that the factor model forecasts are on average more accurate. At $h = 1$, the model outperforms the benchmark with gains in accuracy that range

[4]For each series, the variable that is predicted is: $X^h_{i,t+h} = 100 \times ln(X_{i,t+h}/X_{i,t})$. The model is parameterised as in the previous sections, i.e. it is based on the assumption of a single global factor, one (block) factor for each group and category of commodity prices and one lag in the factor VAR. As a robustness check, the forecasting results are also provided for a model specification with two global factors.

from 18% for the non-fuel index to 12% for oil. Looking at disaggregated commodity prices in Table 7, the predictive component appears to be stronger for the group of commodities such as metals and food for which the common component is more important (copper (19%), rice (19%), poultry (46%), cotton (17%) and aluminium (12%)). However, at $h = 1$, the reduction in MSFE is also marked for oil prices for which gains range between 9% and 12%. Second, the gains fade out progressively over longer horizons and at $h = 12$, we cannot reject the hypothesis of equal predictive performance. Finally, we find that the predictive ability of the model has changed over time. The evolution of the rolling relative MSFE in Figure 11 indicates that the predictability of oil and other energy commodities increased markedly in the second half of the 2000s. Figure 11 shows that from the intensification of the financial crisis to 2011, the MSFE of the factor model improved substantially compared to the benchmark, even though, given the high level of volatility, the MSFE difference is not consistently statistically significant. By contrast, for non-fuel commodities, the factor model forecasts seem to be consistently more accurate than the benchmark. Like fuel-commodities, we observe an increase in the predictive accuracy during the Great Recession even though the test cannot reject the null of equal predictive accuracy. This finding appears to be consistent with previous results in the literature showing that, for macroeconomic and financial variables, downturn periods are characterised by an increasing co-movement (see, e.g. D'Agostino and Giannone (2012)).

4 Concluding remarks

In this paper, we show that a single global factor is a dominant source of fluctuations for a broad set of commodity prices. This factor is persistent and correlated with global economic activity. The factor moves commodity prices in the same direction and has limited effects on relative prices. Historical decompositions of major commodity price changes indicate that the global factor accounts for a larger fraction of commodity price movements in episodes that coincided with changes in global activity, such as the world economic expansion that started around 2002 and the steep contraction during the Great Recession. By contrast, block components explain most of the fluctuations in commodity prices during episodes conventionally associated with supply or any other commodity-specific shocks. The model performs well in predicting commodity prices in real-time confirming that the high level of commonality is not an artifact due to over-fitting but is a genuine feature of the data that remains in the out-of-sample assessment, in spite of potential structural changes.

References

[1] Aastveit, Bjørnland and Thorsrud (2014), What drives Oil Prices? Emerging versus Developed Economies, Journal of Applied Econometrics 30, 1013-1028.

[2] Alquist R. and O. Coibion (2014), Commodity Price Comovement and Global Economic Activity, NBER WP 20003.

[3] Alquist R., L. Kilian and R.J. Vigfusson (2013), Forecasting the Price of Oil, Handbook of Economic Forecasting, Volume 2A, Elsevier.

[4] Bai, J. (2003), Inferential Theory for Factor Models of Large Dimensions, Econometrica 71, 135-171.

[5] Bai J. and S. Ng (2002), Determining the Number of Factors in Approximate Factor Models, Econometrica, 70, 191-221.

[6] Baffes J, M.A. Kose, F. Ohnsorge and M. Stocker (2015), The Great Plunge in Oil Prices: Causes, Consequences, and Policy Responses, World Group Policy Research Note.

[7] Bańbura M, D. Giannone and M. Lenza (2015), Conditional Forecasts and Scenario Analysis with Vector Autoregressions for Large Cross-Sections, International Journal of Forecasting 31(3), 739-756.

[8] Bańbura M., D. Giannone and L. Reichlin (2011), Nowcasting, The Oxford Handbook of Economic Forecasting.

[9] Bańbura M, D. Giannone, M. Modugno and L. Reichlin (2013), Now-casting and the real-time data flow, Handbook of Economic Forecasting 2, 195-237.

[10] Bańbura, M. and Modugno, M., (2014), Maximum likelihood estimation of factor models on data sets with arbitrary pattern of missing data, Journal of Applied Econometrics, 29(1): 199-160.

[11] Barsky R. and L. Kilian (2002), Do we really know that oil caused the Great Stagflation? A monetary alternativee, NBER Macroeconomics Annual 2001, 16, 137-183.

[12] Baumeister C. and L. Kilian (2012), Real-Time Forecasts of the Real Price of Oil, Journal of Business and Economic Statistics, 30(2), 326-336.

[13] Baumeister C. and L. Kilian (2016), Understanding the decline in the price of oil since June 2014, Journal of the Association of Environmental and resource economists, 3(1), 131-158.

[14] Brave S. and R.A. Butters (2011), Monitoring Financial Stability: A Financial Conditions index approach, Economic Perspectives Q(I), 22-43.

[15] Byrne J.P, G. Fazio and N. Fiess (2013), Primary Commodity Prices: Co-movements, Common Factors and Fundamentals, Journal of Development Economics 101: 16-26.

[16] Chen, Y.C., K.S. Rogoff and B. Rossi (2010), Can Exchange Rates Forecast Commodity Prices? The Quarterly Journal of Economics, 125(3): 1145-1194.

[17] Chen, S.L., J.D. Jackson, H. Kim and P. Resiandini (2014), What Drives Commodity Prices? American Journal of Agricultural Economics 96, 1455-1468.

[18] Coroneo, L. D. Giannone and M. Modugno (2015), Unspanned Macroeconomic Factors in the Yields Curve, Journal of Business and Economic Statistics, forthcoming.

[19] Diebold F.X. and R.S. Mariano (1995), Comparing Predictive Accuracy, Journal of Business and Economic Statistics 13, 253-263.

[20] Doz C., D. Giannone and L. Reichlin (2011), A Two Step Estimator for Large Approximate Dynamic Factor Models Based on Kalman Filtering", Journal of Econometrics 164 (2011), 188-205.

[21] Doz C., D. Giannone and L. Reichlin (2012), A Quasi-Maximum Likelihood Approach for Large, Approximate Dynamic Factor Models, The Review of Economics and Statistics, 94(4): 1014-1024

[22] D'Agostino A. and D. Giannone (2012), Comparing Alternative Predictors Based on Large-Panel Factor Models, Oxford Bulletin of Economics and Statistics, 74(2), 306-326.

[23] Fattouh, B., L. Kilian and L. Mahadeva (2012). The Role of Speculation in Oil Markets: What Have we Learned so Far? Energy Journal, 34(3), 385-394.

[24] Forni, M., M. Hallin, M. Lippi and L. Reichlin (2000), The Generalized Dynamic Factor Model: Identification and Estimation, The Review of Economic and Statistics 82, 540-554.

[25] Forni, M. & Reichlin, L. (1998), Let's Get Real: A Factor Analytical Approach to Disaggregated Business Cycle Dynamics, Review of Economic Studies, Wiley Blackwell, vol. 65(3), pages 453-73.

[26] Forni M. and L. Reichlin (2001), Federal policies and local economies: Europe and the us. European Economic Review, 45(1):109- 134, January 2001.

[27] Giacomini R. and B. Rossi (2010), Forecast Comparisons in Unstable Environments, Journal of Applied Econometrics, 25, 595-610.

[28] Groen, Jan J.J., Pesenti, Paolo A., (2010), Commodity Prices, Commodity Currencies, and Global Economic Developments. NBER Working Paper 15743.

[29] Groen, J. and P. Russo (2015). Is Cheaper Oil Good News or Bad News for the US Economy? Liberty Street Economics, June 2015.

[30] Gruber, J. W. and R. Vigfusson (2016), Interest Rates and The Volatility Correlation of Commodity Prices", Macroeconomic Dynamics 2016, 1-20.

[31] Gospodinov, N. and S. Ng (2013),Commodity prices, convenience yields and inflation. Rev. Econ. Stat. 95, 206–219.

[32] Hamilton, J. D. (2009), Causes and Consequences of the Oil Shock of 2007-08, Brookings Papers on Economic Activity, Spring 2009.

[33] Juvenal L. and I. Petrella (2014), Speculation in the Oil Market. Journal of Econometrics, 30(4), 621-649.

[34] Kilian, L. (2009), Not All Oil Price Shocks Are Alike: Disentangling Demand and Supply Shocks in the Crude Oil Market, American Economic Review, 99(3), June 2009, 1053-1069.

[35] Kilian, L. (2017), The impact of the fracking boom on Arab oil producers, Energy Journal, 38(6), 137-160.

[36] Kilian, L. and B. Hicks (2013), Did Unexpectedly Strong Economic Growth Cause the Oil Price Shock of 2003-2008?, Journal of Forecasting, 32(5), 385-394.

[31] Kilian, L and D.P. Murphy (2014), The Role of Inventories and Speculative Trading in the Global Market for Crude Oil, Journal of Applied Econometrics, 29(3), April 2014, 454-478.

[38] Kilian, L. and T.K. Lee (2014). Quantifying teh Speculative Components in the Real Price of Oil: The Role of Global Oil Inventories, Journal of International Money and Finance, 42, 71-87.

[39] Kose M. A, C. Otrok, and C.H. Whiteman (2003). International Business Cycles: World, Region, and Country-Specific Factors, American Economic Review, vol. 93(4), pages 1216-1239, September.

[40] Luciani, M. (2015), Monetary Policy and the Housing Market. A Structural Factor Analysis, Journal of Applied Econometrics 30(2), 199-218.

[41] Luciani M. and L. Ricci (2014), Nowcasting Norway, International Journal of Central Banking 10(4), 215-248.

[42] Miranda-Agrippino S. and H. Rey (2015), World Asset Markets and the Global Financial Cycle, NBER WP No. 21722.

[43] Poncela, P., E. Senra and L.P. Sierra (2015). The predictive content of co-movement in non-energy commodity price changes, mimeo.

[44] Reis R. and M.W. Watson (2010), Relative Goods' Prices, Pure Inflation and the Phillips Correlation, American Economic Journal: Macroeconomics 2(3), 128-157.

[45] Stock J. H. and M. W. Watson (2002a), Forecasting using Principal Components from a Large Number of Predictors, Journal of American Statistical Association 97, 147-162.

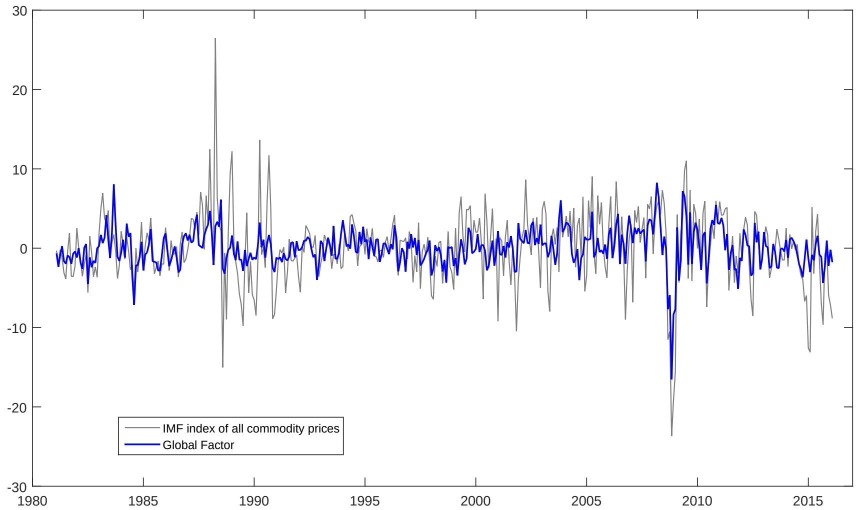
[46] Stock J. H. and M. W. Watson (2002b), Macroeconomic Forecasting Using Diffusion Indexes, Journal of Business and Economics Statistics 20, 147-162.

[47] Tang, K. and W. Xiong (2012). Index Investment and Financialization of Commodities, Financial Analysts Journal 68, 54-74.

[39] West K.D. and K.F. Wong (2014), A factor model for co-movements of commodity prices, Journal of International Money and Finance 42: 289-309.

[49] White, H. (1982), Maximum Likelihood Estimation of Misspecified Models, Econometrica 50(1982): 1-25.

Figure 1: The global factor



Note: The top panel of the figure shows the estimated global factor (blue line) and the IMF overall index of commodity prices (grey line).

Figure 2: Factor loadings associated with the global factor

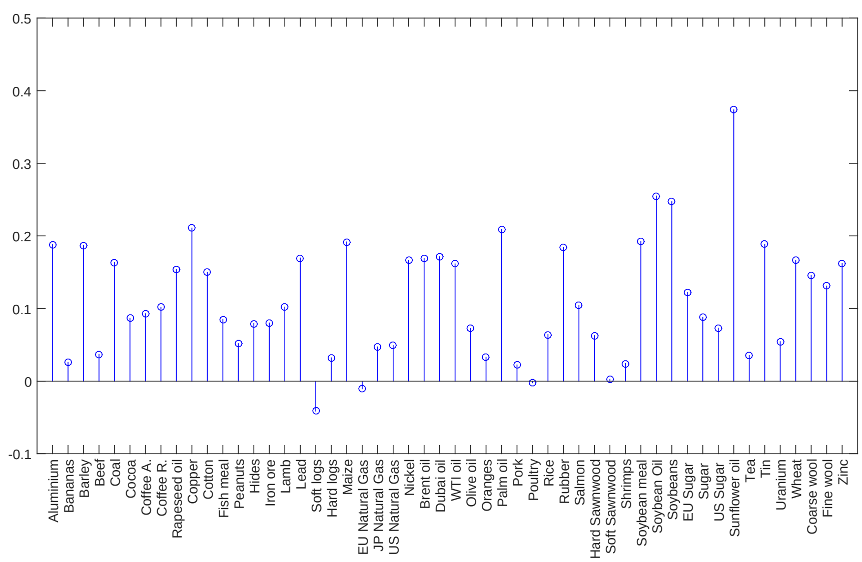
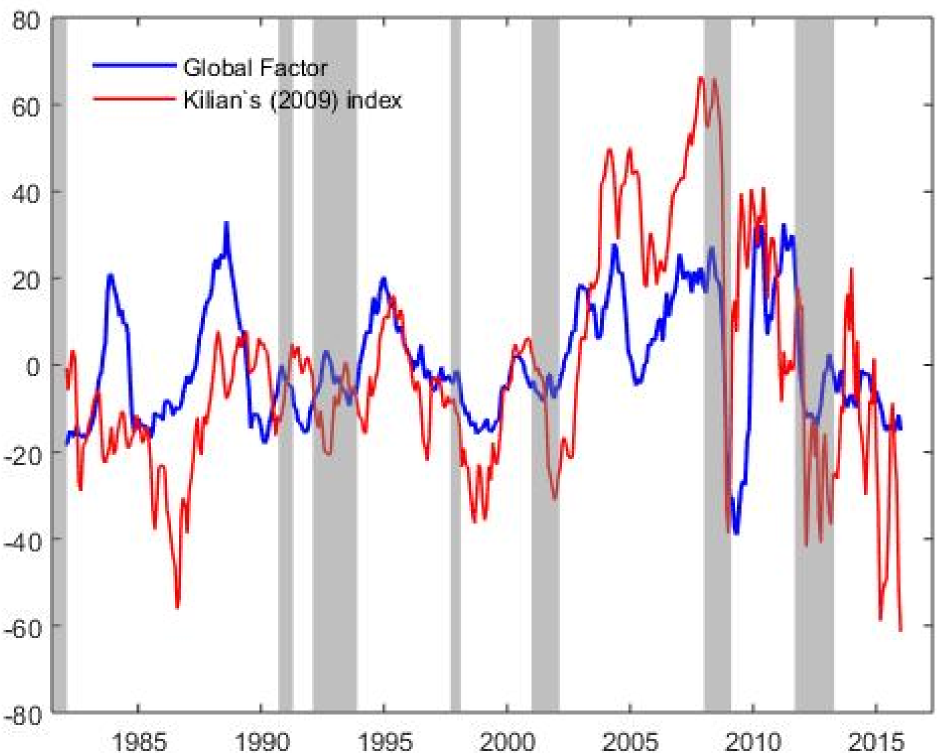


Table 1: Structure of the database

Global	Blocks	Sub-blocks	Groups	N. Series
All commodities (PALLFNF) 100.0				52
	Non-Fuel (PNFUEL) 36.9			45
		Food & Beverages (FFOBEV) 18.5		28
			Food (PFOOD) 16.7	24
			Beverages (PBEV) 1.8	4
		Industrial Inputs (PINDU) 18.4		17
			Agricultural Raw Materials (PAGR) 7.7	9
			Metals (PMET) 10.7	8
	Energy (PNRG) 63.1			7
			Oil (POILAPSP) 53.6	3

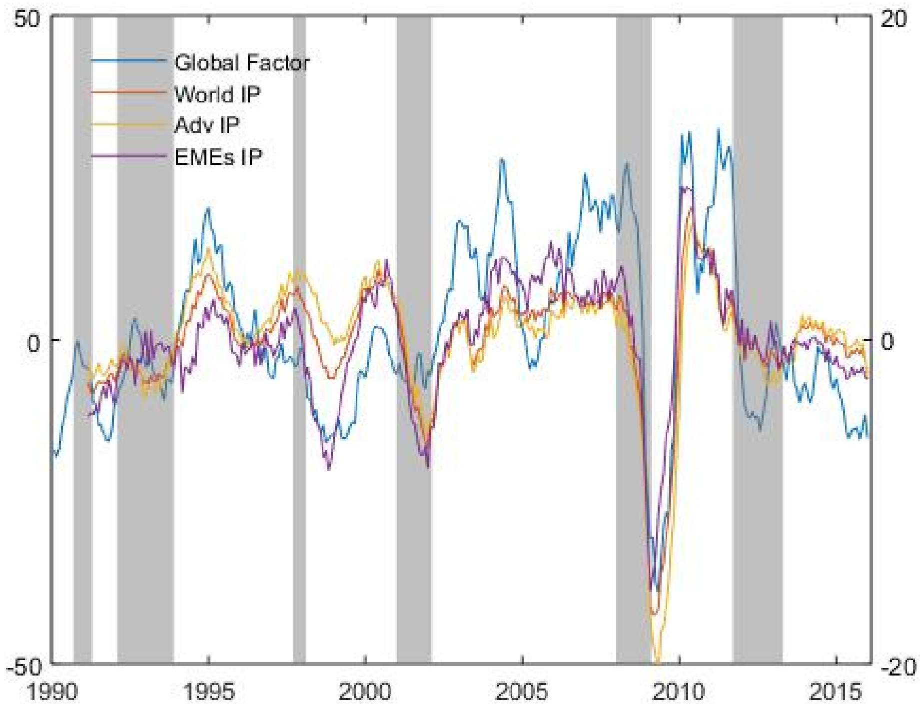
Note: The data set includes one main index for all commodity prices and 9 sub-indices representing different levels of aggregation. The weights reported in the table represent the share of each sub-index in the overall index of commodity prices.

Figure 3: The global factor and the Kilian's index of economic activity



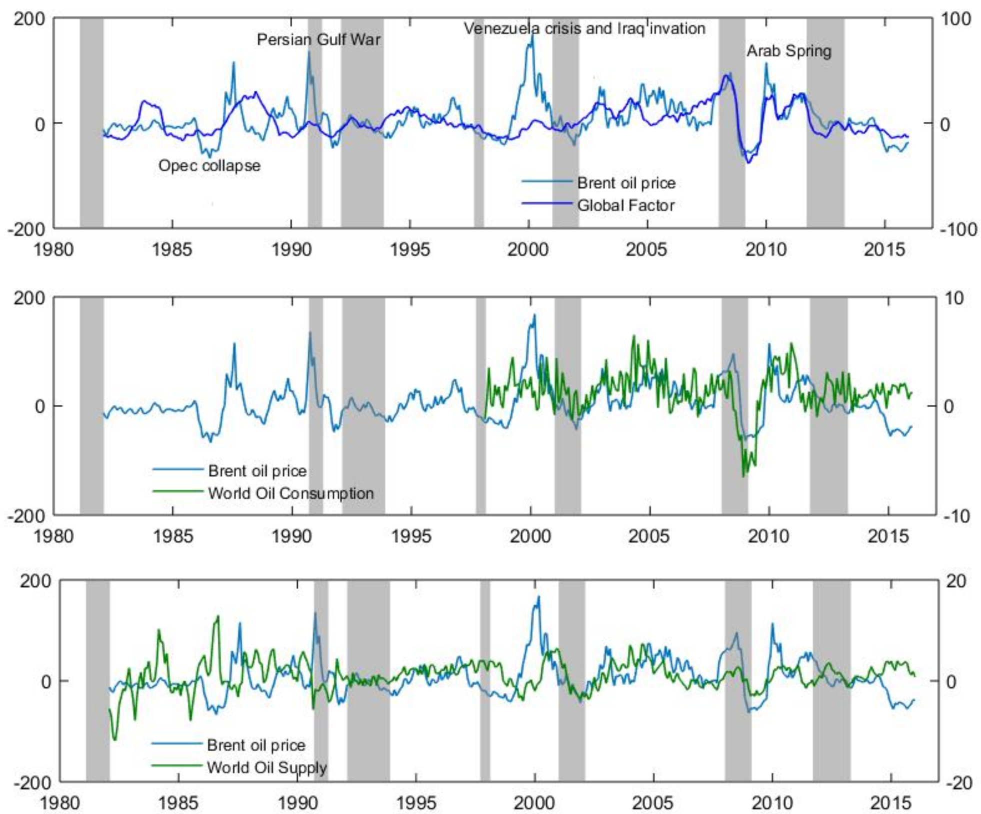
Note: The figure plots the global factor (blue line), extracted from real commodity prices and expressed in year-on-year growth rates, and the Kilian's (2009) index of real economic activity (red line). The vertical bars represent periods of widespread economic slowdown as described in the text.

Figure 4: The global factor and industrial production indices



Note: The figure plots the estimated global factor (blue line) extracted from real commodity prices and measures of industrial production in selected areas as provided by the CPB Netherlands Bureau for Economic Policy Analysis. All variables are expressed in year-on-year growth rates. The vertical bars represent periods of widespread economic slowdown as described in the text.

Figure 5: Oil and the Global Factor



Note: All variables are expressed in year-on-year growth rates. Estimates for the world oil consumption are taken from Short-Term Energy Outlook of the Energy Information Administration (EIA) while the world crude oil production is taken from the Monthly Energy Review of EIA. The vertical bars represent periods of widespread economic slowdown as described in the text.

Figure 6: Variance explained by the first two global factors

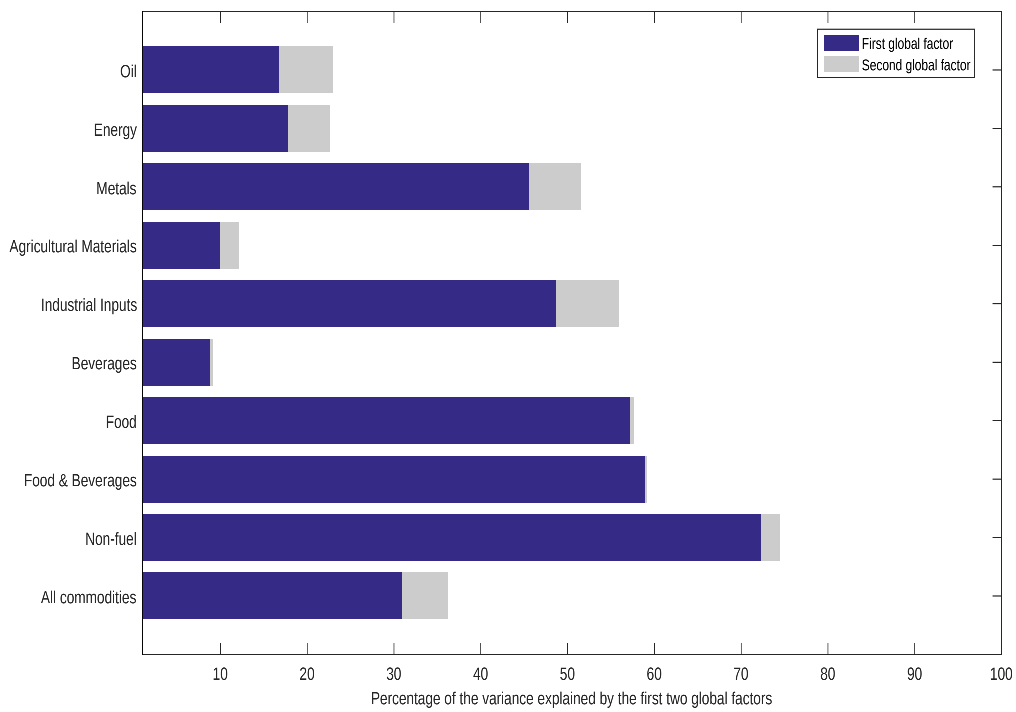


Figure 7: Variance explained by the first two global factors

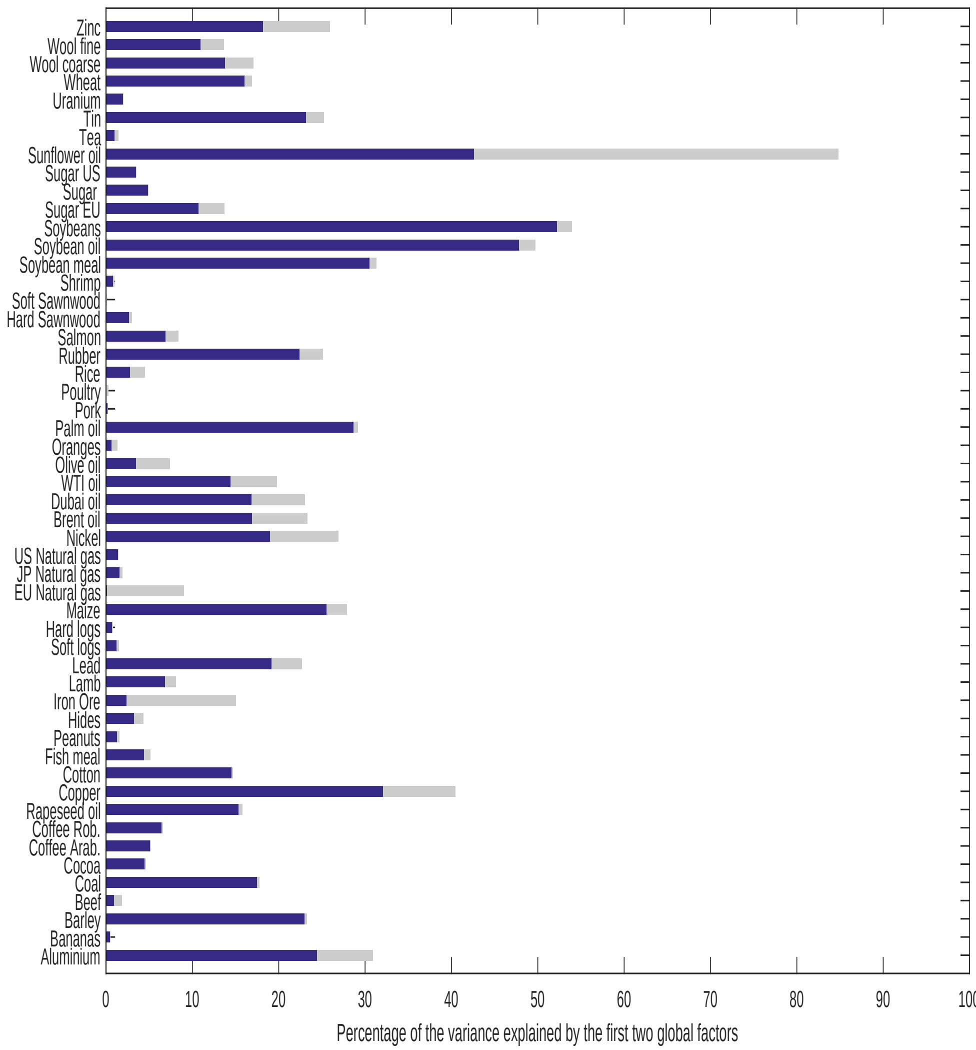


Figure 8: Block factors and weak common factors

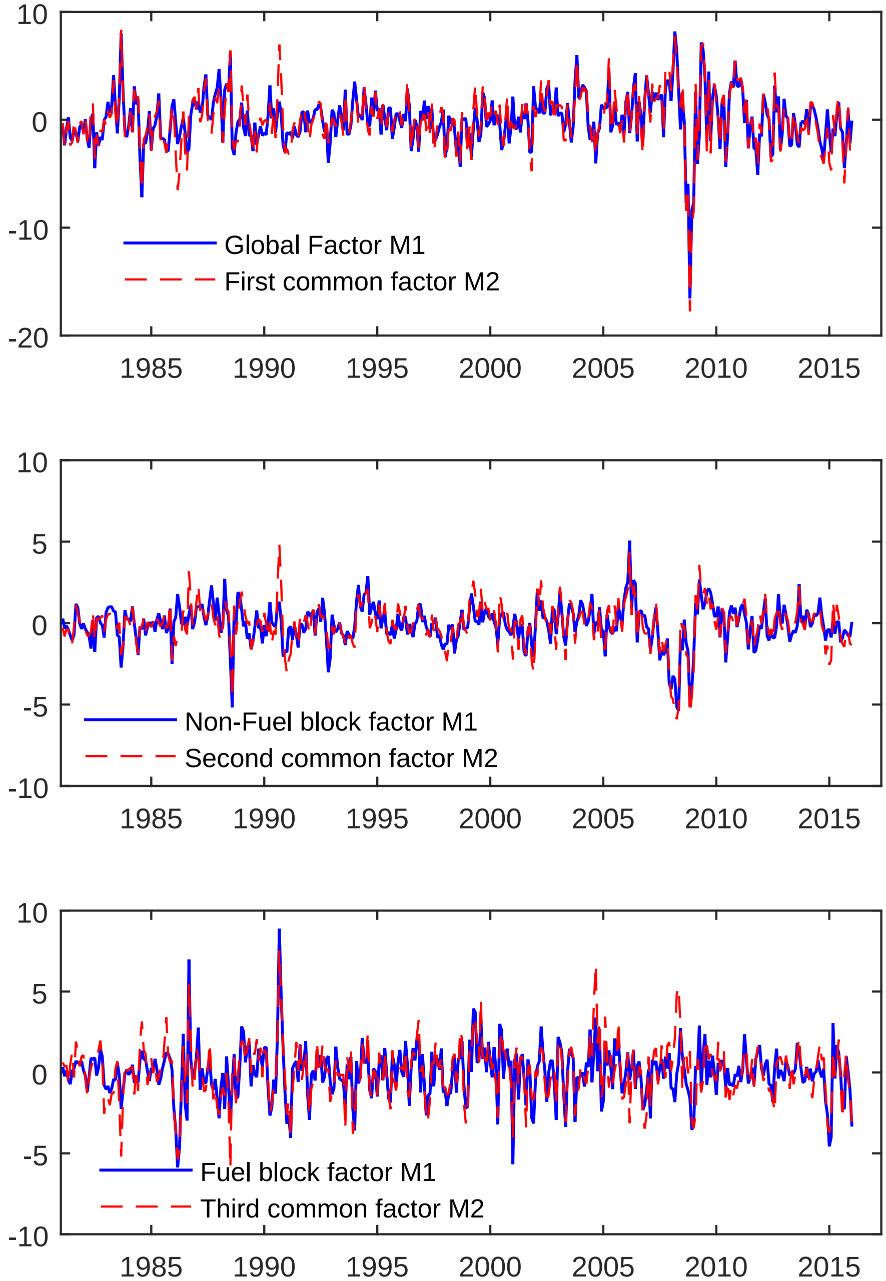
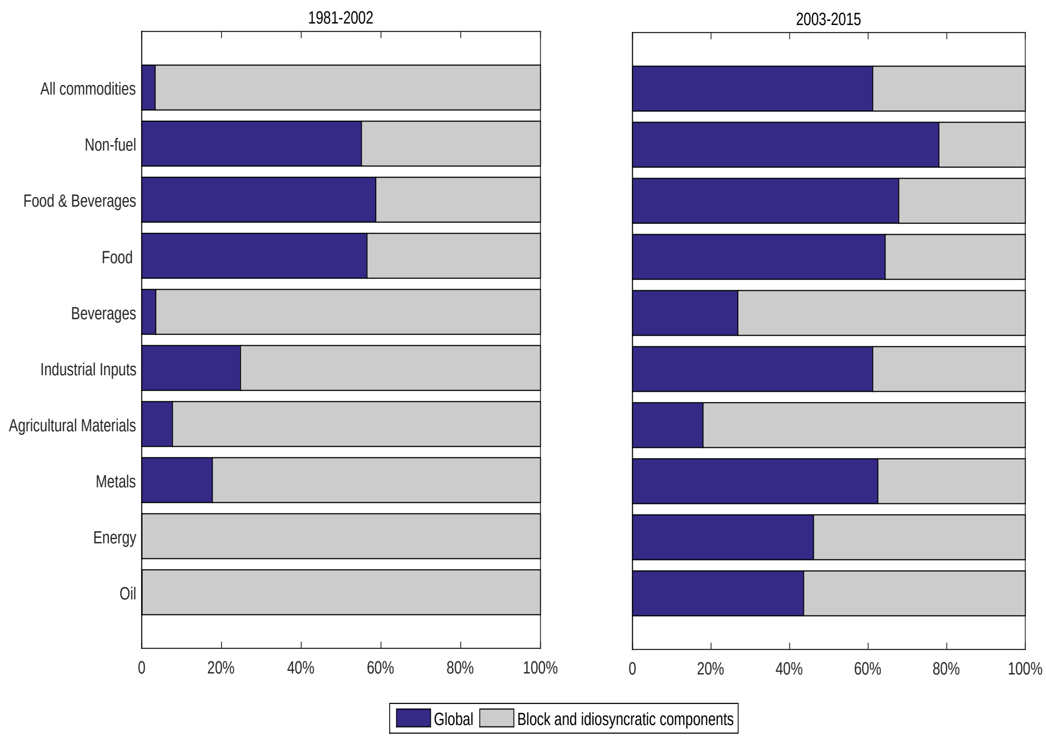
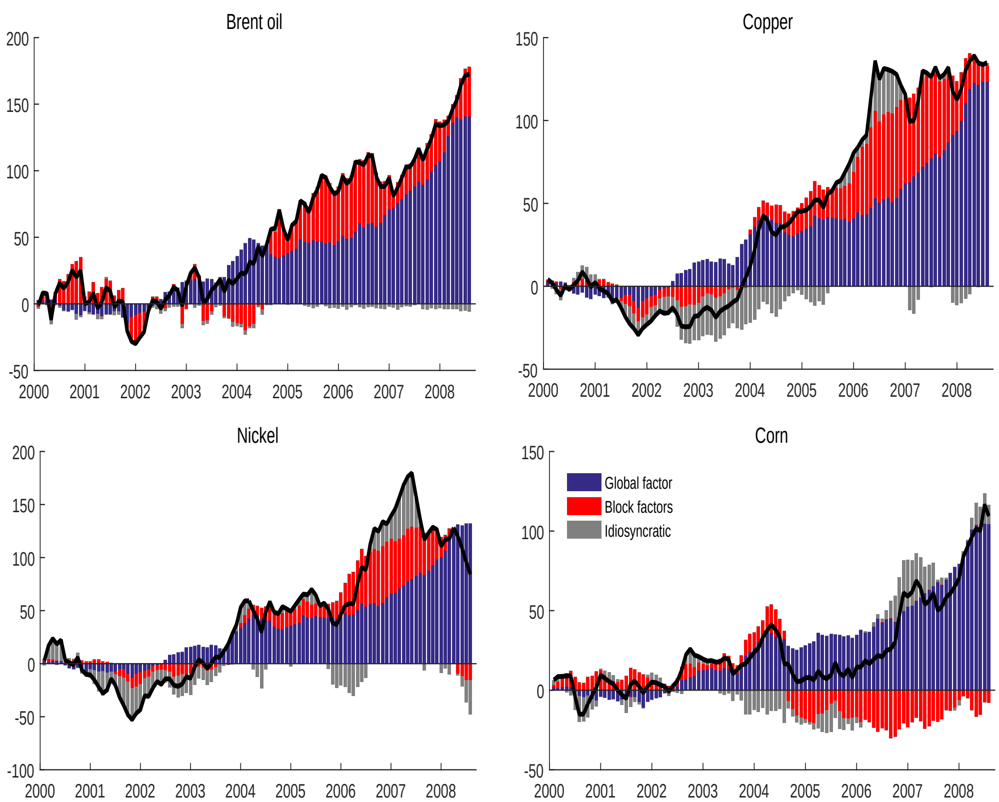


Figure 9: Variance decomposition: sub-sample analysis



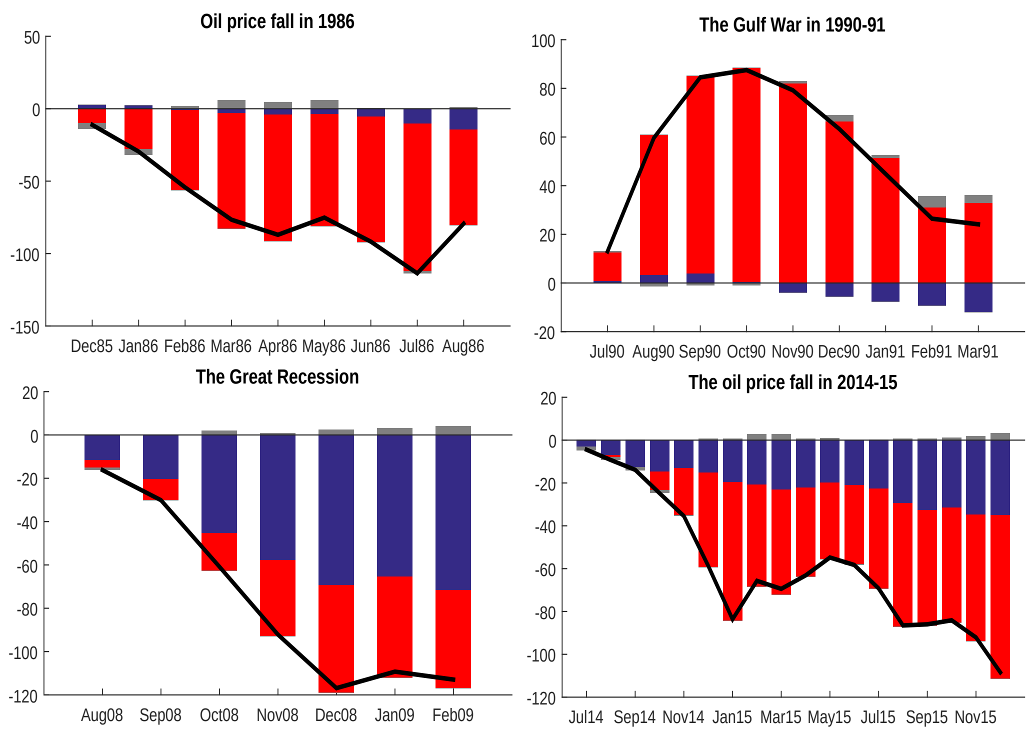
Note: The figure reports the variance decomposition of commodity price indices over two sub-samples. The share of the variance explained by the global factor is captured by the blue bar while the grey bar is the percentage of the variance explained by block-specific and idiosyncratic components. The first subsample goes from Jan. 1981 to Dec. 2002 while the second goes from Jan. 2003 to Dec. 2015.

Figure 10: Historical decompositions of commodity prices: The price boom of the 2000s



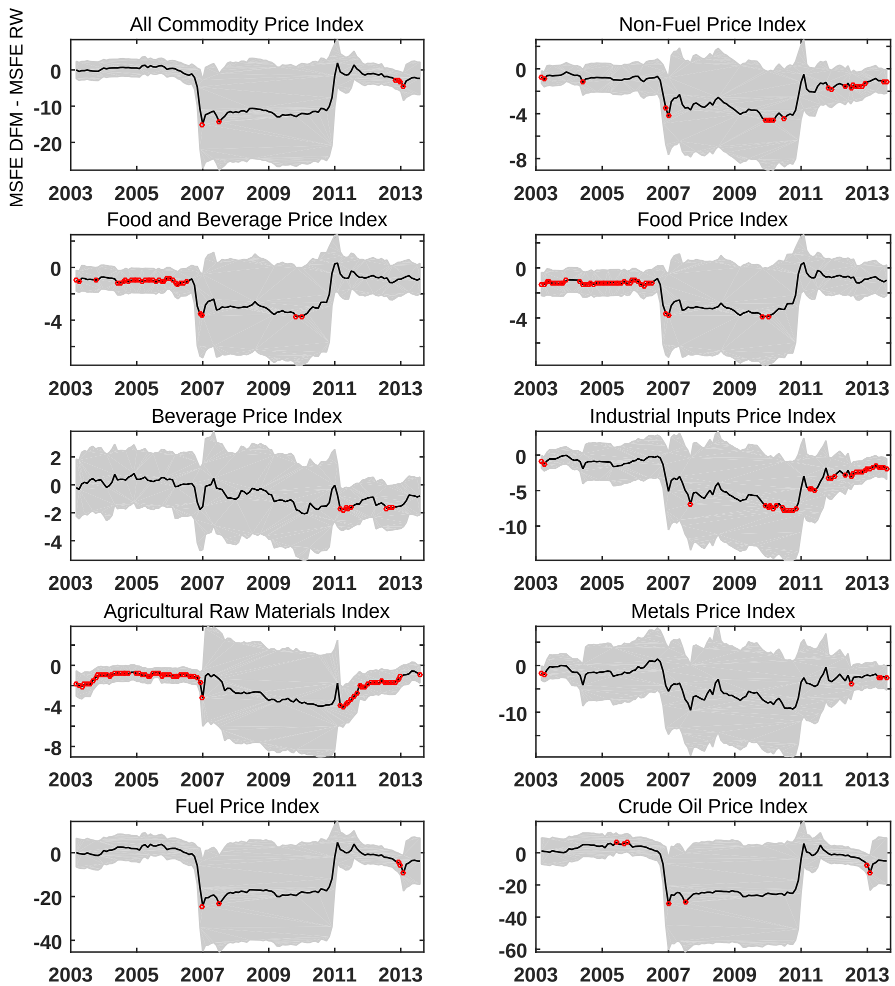
Note: The figure reports the historical decomposition of the price for a selected group of energy, metal and food commodities, showing the cumulative effects at each point in time of global (blue), block-specific (red) and idiosyncratic (grey) shocks from January 2000 to July 2008.

Figure 11: Historical decompositions of the price of oil in selected episodes



Note: The figure presents the historical decomposition of the price of oil, showing the cumulative effects at each point in time of global (blue), block-specific (red) and idiosyncratic shocks (grey) during four historical episodes of large oil price variations.

Figure 12: Time-varying predictability



Note: The figure shows the difference between the MSFE of the factor model and the MSFE of the benchmark, smoothed over time with a centered rolling window spanning 4 years. A negative number indicates that the factor model has a higher predictive accuracy than the benchmark. The 90% confidence bands are represented by the shaded area and are derived from testing the null hypothesis of equal predictive accuracy at each point in time. Red circles indicate rejection of the hypothesis of equal predictive accuracy at the 5% level.

Table 2: Data description

Mnemonic	Unit	description	IMF global commodity index 2002-2004 weights
PALLFNF	Index	All Commodity Price Index, 2005 = 100, includes both Fuel and Non-Fuel	100.0
PNFUEL	Index	Non-Fuel Price Index, 2005 = 100, Food and Beverages and Industrial Inputs	36.9
PFANDB	Index	Food and Beverage Price Index, 2005 = 100, Food and Beverage	18.5
PFOOD	Index	Food Price Index, 2005 = 100, Cereal, Vegetable Oils, Meat, Seafood, Sugar, Bananas, Oranges	16.7
PBEVE	Index	Beverage Price Index, 2005 = 100, Coffee, Tea, Cocoa	1.8
PINDU	Index	Industrial Inputs Price Index, 2005 = 100, Agricultural Raw Materials and Metals	18.4
PRAWM	Index	Agricultural Raw Materials Index, 2005 = 100, Timber, Cotton, Wool, Rubber, Hides	7.7
PMETA	Index	Metals Price Index, 2005 = 100, Copper, Aluminium, Iron Ore, Tin, Nickel, Zinc, Lead,Uranium	10.7
PNRG	Index	Fuel (Energy) Index, 2005 = 100, Crude oil , Natural Gas, Coal	63.1
POILAPSP	Index	Crude Oil (petroleum), Price index, 2005 = 100, simple average of Brent, WTI, Dubai Fateh	53.6
PALUM	USD	Aluminium, 99.5% minimum purity, LME spot price, CIF UK ports, USD per metric ton	3.9
PBANSOP	USD	Bananas, Central American and Ecuador, FOB U.S. Ports, USD per metric ton	0.4
PBARL	USD	Barley, Canadian no.1 Western Barley, spot price, USD per metric ton	0.3
PBEEF	USD	Beef, Australian and New Zealand 85% lean fores, CIF U.S. import price, US cents per pound	1.4
PCOALAU	USD	Coal, Australian thermal coal, 12,000- btu/pound, FOB Newcastle/Port Kembla, USD(metric ton)	2.6
PCOCO	USD	Cocoa beans, Int. Cocoa Org. cash price, CIF US and European ports, USD per metric ton	0.7
PCOFFOTM	USD	Coffee, Other Mild Arabicas, Int. Coffee Org. NY cash price, US cents per pound	0.5
PCOFFROB	USD	Coffee, Robusta, Int. Coffee Org. NY cash price, US cents per pound	0.4
PROIL	USD	Rapeseed oil, crude, fob Rotterdam, USD per metric ton	0.3
PCOPP	USD	Copper, grade A cathode, LME spot price, CIF European ports, USD per metric ton	2.8
PCOTTIND	USD	Cotton,Outlook 'A Index', Middling 1-3/32 inch staple, CIF Liverpool, US cents per pound	0.7
PFISH	USD	Fishmeal, Peru Fish meal/pellets 65% protein, CIF, USD per metric ton	0.2
PGNUTS	USD	Groundnuts (peanuts), 40/50 , cif Argentina, USD per metric ton	0.2
PHIDE	USD	Hides, Heavy native steers, over 53 pounds, wholesale price, US, Chicago, US cents per pound	2.6
PIORECR	USD	China import Iron Ore Fines 62% FE spot (CFR Tianjin port), US dollars per metric ton	1.3
PLAMB	USD	Lamb, frozen carcass Smithfield London, US cents per pound	0.3
PLEAD	USD	Lead, 99.97% pure, LME spot price, CIF European Ports, USD per metric ton	0.2
PLOGORE	USD	Soft Logs, Average Export price from the U.S. for Douglas Fir, USD per cubic meter	0.4
PLOGSK	USD	Hard Logs, Best quality Malaysian meranti, import price Japan, USD per cubic meter	0.4
PMAIZMT	USD	Maize (corn), U.S. No.2 Yellow, FOB Gulf of Mexico, U.S. price, USD per metric ton	1.0

PNGASEU	USD	Natural Gas, Russian Natural Gas border price in Germany, USD per thousands of cubic meters of gas	3.2
PNGASJP	USD	Natural Gas, Indonesian Liquefied Natural Gas in Japan, USD per cubic meter of liquid	1.9
PNGASUS	USD	Natural Gas, Henry Hub terminal in Louisiana, USD per thousands of cubic meters of gas	1.9
PNICK	USD	Nickel, melting grade, LME spot price, CIF European ports, USD per metric ton	1.1
POILBRE	USD	Crude Oil (petroleum), Dated Brent, light blend 38 API, fob U.K., USD per barrel	17.9
POILDUB	USD	Oil; Dubai, medium, Fateh 32 API, fob Dubai Crude Oil, Dubai Fateh Fateh 32 API, USD per barrel	17.9
POILWTI	USD	Crude Oil (petroleum), West Texas Intermediate 40 API, Midland Texas, USD per barrel	17.9
POLVOIL	USD	Olive Oil, extra virgin less than 1% free fatty acid, ex-tanker price U.K., USD per metric ton	0.3
PORANG	USD	Oranges, miscellaneous oranges CIF French import price, USD per metric ton	0.5
PPOIL	USD	Palm oil, Malaysia Palm Oil Futures (first contract forward) 4-5 percent FFA, USD per metric ton	0.7
PPORK	USD	Swine (pork), 51-52% lean Hogs, U.S. price, US cents per pound	1.1
PPOULT	USD	Poultry (chicken), Whole bird price, Ready-to-cook, whole, iced, Georgia, US cents per pound	0.9
PRICENPQ	USD	Rice, 5 percent broken milled white rice, Thailand nominal price quote, USD per metric ton	0.6
PRUBB	USD	Rubber, Singapore Commodity Exchange, No. 3 Rubber Smoked Sheets, 1st contract, US cents-pound	0.5
PSALM	USD	Fish (salmon), Farm Bred Norwegian Salmon, export price, USD per kilogram	2.5
PSAWMAL	USD	Hard Sawnwood, Dark Red Meranti, select and better quality, C&F U.K port, USD per cubic meter	0.8
PSAWORE	USD	Soft Sawnwood, average export price of Douglas Fir, U.S. Price, USD per cubic meter	1.8
PSHRI	USD	Shrimp, No.1 shell-on headless, 26-30 count per pound, Mexican origin, NY port, US cents-pound	0.7
PSMEA	USD	Soybean Meal, Chicago Soybean Meal Futures Minimum 48 percent protein, USD per metric ton	0.8
PSOIL	USD	Soybean Oil, Chicago Soybean Oil Futures exchange approved grades, USD per metric ton	0.4
PSOYB	USD	Soybeans, U.S. soybeans, Chicago Soybean futures contract No. 2 yellow and par, USD per metric ton	1.2
PSUGAEEC	USD	Sugar, European import price, CIF Europe, US cents per pound	0.2
PSUGAISA	USD	Sugar, Free Market, CSCE contract n.11, US cents a pound	0.6
PSUGAUSA	USD	Sugar, U.S. import price, contract no.14 nearest futures position, US cents a pound	0.1
PSUNO	USD	Sunflower oil, Sunflower Oil, US export price from Gulf of Mexico, USD per metric ton	0.2
PTEA	USD	Tea, Mombasa, Kenya, Auction Price, US cents per kg, From July 1998, Best Pekoe Fannings	0.3
PTIN	USD	Tin, standard grade, LME spot price, USD per metric ton	0.2
PURAN	USD	Uranium, NUEXCO, Restricted Price, Nuexco exchange spot, USD per pound	0.5
PWHEAMT	USD	Wheat, No.1 Hard Red Winter, ordinary protein, FOB Gulf of Mexico, USD per metric ton	1.7
PWOOLC	USD	Wool, coarse, 23 micron, Australian Wool Exchange spot quote, US cents per kilogram	0.3
PWOOLF	USD	Wool, fine, 19 micron, Australian Wool Exchange spot quote, US cents per kilogram	0.2
PZINC	USD	Zinc, high grade 98 percent pure, USD per metric ton	0.6

Table 3: Model Selection

	Number of global factors				
	$r = 1$	$r = 2$	$r = 3$	$r = 4$	$r = 5$
IC^*	11.77	11.92	12.07	12.26	12.41
$log(V)$	11.57	11.53	11.48	11.47	11.43

Table 4: Variance decomposition of commodity indices

Indices	***Global***	***Non Fuel***	***Food-Bev.***	***Food***	***Bev.***	***Ind. Inputs***	***Agric. Raw Mat.***	***Metals***	***Fuel***	***Oil***	***Idiosyncratic***
All Commodities	34.1	0.2	0.1	0	0	0	0	0.2	62.7	0	2.6
Non-Fuel	68.6	3.1	1.7	0.1	0.5	0	0.8	3.7	0	0	21.3
Food and Beverages	58.0	0	6.7	0.4	1.8	0	0	0	0	0	33.2
Food	55.4	0.1	8.6	0.4	0	0	0	0	0	0	35.6
Beverages	9.3	1.0	1.0	0	50.4	0	0	0	0	0	38.4
Industrial Inputs	48.4	7.9	0	0	0	0.7	2.0	9.2	0	0	31.8
Agricultural Raw Materials	10.5	1.4	0	0	0	31.2	8.3	0	0	0	48.6
Metals	43.2	7.5	0	0	0	5.2	0	13.8	0	0	30.3
Energy	19.9	0	0	0	0	0	0	0	78.1	0	2.0
Oil	17.7	0	0	0	0	0	0	0	81.2	0	1.1

Note: The table reports the model-based variance decomposition of commodity indices estimated over the sample January 1981 - December 2015.

Table 5: Variance decomposition of commodity prices

Commodity prices	***Global***	***Non-Fuel***	***Food-Bev.***	***Food***	***Bev.***	***Ind. Inputs***	***Agric. Raw Mat.***	***Metals***	***Fuel***	***Oil***	***Idiosyncratic***
Aluminium	23.6	7.4	-	-	-	3.4	-	0.5	-	-	65.1
Bananas	0.4	0.3	1	1.2	-	-	-	-	-	-	96.9
Barley	21.6	0.1	8	0.01	-	-	-	-	-	-	70.0
Beef	0.8	1	0.5	0.4	-	-	-	-	-	-	97.7
Coal	16.7	-	-	-	-	-	-	-	0.5	-	82.8
Cocoa	4.4	0.3	2.4	-	4.0	-	-	-	-	-	88.9
Coffee Arabica	5.2	0.7	0.1	-	67.3	-	-	-	-	-	26.7
Coffee Robusta	6.4	0.5	0.1	-	70.1	-	-	-	-	-	22.9
Rapeseed Oil	15.2	0.2	0.01	0.04	-	-	-	-	-	-	84.5
Copper	30.8	9.4	-	-	-	1.4	-	6.2	-	-	52.2
Cotton	13.7	0.4	-	-	-	0.2	2.1	-	-	-	83.6
Fish meal	4.3	1	4.1	2.7	-	-	-	-	-	-	87.9
Peanuts	1.1	0.8	20.3	8.0	-	-	-	-	-	-	69.8
Hides	3.7	0.6	-	-	-	15.2	32.2	-	-	-	48.3
Iron ore	2.6	11.1	-	-	-	3.3	-	77.0	-	-	6.0
Lamb	6.7	1.1	9.3	0.2	-	-	-	-	-	-	82.7
Lead	18.9	3.9	-	-	-	1.3	-	1.7	-	-	74.3
Soft logs	1.1	0	-	-	-	3.3	1.1	-	-	-	94.2
Hard logs	0.7	0	-	-	-	54.1	4.3	-	-	-	40.9
Maize	23.0	0.9	21	0.4	-	-	-	-	-	-	55.0
EU Natural gas	0.0	-	-	-	-	-	-	-	5.7	-	94.3
JP Natural gas	1.3	-	-	-	-	-	-	-	0.4	-	98.3
US Natural gas	1.6	-	-	-	-	-	-	-	1.0	-	97.4
Nickel	19.2	10.1	-	-	-	1.8	-	0.8	-	-	68.1
Brent oil	17.0	-	-	-	-	-	-	-	78.1	0.4	4.5
Dubai oil	17.6	-	-	-	-	-	-	-	77.0	2.0	3.3
WTI oil	16.1	-	-	-	-	-	-	-	77.2	5	1.3
Olive oil	2.7	2.8	10.1	0.4	-	-	-	-	-	-	84.0
Oranges	0.6	0	1	0.8	-	-	-	-	-	-	97.6
Palm oil	28.7	0.3	2.3	0.05	-	-	-	-	-	-	68.6
Pork	0.4	0.01	0.2	0.0	-	-	-	-	-	-	99.4
Poultry	0.0	1	3	0.9	-	-	-	-	-	-	95.8
Rice	2.7	1.6	3.4	0.02	-	-	-	-	-	-	92.4
Rubber	22.1	2.8	-	-	-	0.01	0.7	-	-	-	74.4
Salmon	6.0	1.5	2.8	0.06	-	-	-	-	-	-	89.7
Hard Sawnwood	2.7	0.1	-	-	-	47.1	3.4	-	-	-	46.7
Soft Sawnwood	0.0	0.0	-	-	-	3.1	0.9	-	-	-	96.0
Shrimp	0.4	0.0	0.1	23.8	-	-	-	-	-	-	75.8
Soybean meal	28.6	0.3	38	0.1	-	-	-	-	-	-	33.2
Soybean oil	45.9	1.0	13	0.0	-	-	-	-	-	-	39.6
Soybeans	48.7	0.8	38	0.2	-	-	-	-	-	-	12.2
EU sugar	10.1	2.7	8.4	0.2	-	-	-	-	-	-	78.5
Sugar	4.6	0.0	1	0.9	-	-	-	-	-	-	93.8
US sugar	3.1	0	0.8	1.3	-	-	-	-	-	-	94.8
Sunflower oil	40.1	43	9.9	0.1	-	-	-	-	-	-	6.5
Tea	0.9	0.4	0.3	-	1	-	-	-	-	-	97.6
Tin	21.9	2.0	-	-	-	0.4	-	2.0	-	-	73.7
Uranium	1.8	0.0	-	-	-	0.01	-	1.1	-	-	97.1
Wheat	17.1	0.2	10.1	0.5	-	-	-	-	-	-	72.0
Coarse wool	13.0	4.1	-	-	-	1.0	2.1	-	-	-	79.9
Fine wool	10.4	3	-	-	-	0.0	4.7	-	-	-	81.9
Zinc	18.6	10.7	-	-	-	3.1	-	3.0	-	-	64.6

Note: The table reports the model-based variance decomposition of commodity prices estimated over the sample January 1981 - December 2015.

Table 6: Out-of-sample forecasting performance

	h=1			*h=12*		
Indices	***RMSE Benchmark***	***Relative MSE r=1***	***r=2***	***RMSE Benchmark***	***Relative MSE r=1***	***r=2***
All commodities	5.22	0.84*	0.85*	24.98	1.05	1.06
Non-fuel	3.06	0.82**	0.84*	15.21	1.11	1.11
Food and Beverages	3.19	0.85**	0.85**	14.27	1.09**	1.09**
Food	3.30	0.85**	0.85**	14.64	1.11***	1.11***
Beverages	4.19	0.97	0.96	17.67	1.04	1.04
Industrial Inputs	3.92	0.83**	0.86	20.65	1.01	1.01
Agricultural Raw Materials	3.12	0.81**	0.81**	14.71	0.98*	0.98
Metals	5.05	0.88	0.95	26.15	1.04	1.04
Energy	7.33	0.88	0.89	32.57	1.05	1.05
Oil	8.55	0.88	0.89	35.68	1.01	1.01

Note: The table shows the root mean forecast error (RMSE) of a benchmark model, i.e. a constant growth model and the MSE of the candidate forecasting model relative to the benchmark. (*), (**) and (***) indicate rejection of the null of equal predictive accuracy at the 10%, 5% and 1% level based on the Diebold and Mariano (1995) statistic. The model estimation is rolling using a fixed window of 20 years and the estimation starts in 2001:1. The evaluation period goes from 2001:2 to 2015:12. As robustness check, the table also displays the relative MSE for a model specification with two global factors.

Table 7: Out-of-sample forecasting performance

	h=1			*h=12*		
Commodity prices	***RMSE Benchmark***	***Relative MSE r=1***	***r=2***	***RMSE Benchmark***	***Relative MSE r=1***	***r=2***
Aluminium	5.14	0.89	0.88	22.10	1.03	1.03
Bananas	11.61	0.99	1.00	22.69	1.00	1.00
Barley	6.49	0.89 *	0.90	27.69	1.01	1.01
Beef	4.52	0.94 *	0.94 *	16.08	1.00	1.00
Coal	7.06	0.85 *	0.85 *	37.41	1.02	1.02
Cocoa	6.05	0.98	0.98	23.34	1.04	1.04
Coffee Arabica	6.51	0.99	0.99	28.24	0.98	0.98
Coffee Robusta	5.91	0.98	0.98	26.51	0.97**	0.97**
Rapeseed Oil	5.75	0.84 **	0.85 **	27.12	0.98	0.98
Copper	7.05	0.82 *	0.81 *	32.63	1.05	1.05
Cotton	6.32	0.83 **	0.83 **	32.30	0.98 *	0.98 *
Fish meal	4.88	0.88 ***	0.87 ***	22.02	1.04	1.05
Peanuts	4.91	1.09	1.10	27.16	1.06	1.06
Hides	6.87	0.94	0.95	25.77	1.04	1.04
Iron ore	8.77	1.88 ***	3.78 ***	35.50	3.12 ***	3.12 ***
Lamb	3.44	0.88 **	0.87 **	17.87	1.00	1.00
Lead	7.99	0.96	0.96	36.88	1.03	1.03
Soft logs	6.64	0.86 **	0.86 **	13.24	1.02	1.02
Hard logs	3.15	0.88 ***	0.88 ***	14.86	1.00	1.00
Maize	6.30	0.96	0.97	27.03	1.02	1.03
EU Natural gas	6.41	1.26	1.38 **	34.63	1.44	1.47
JP Natural gas	7.13	0.98	0.98	29.33	1.00	1.00
US Natural gas	13.22	1.00	1.00	44.65	1.01	1.01
Nickel	8.99	0.92	0.91	42.67	1.05	1.05
Brent oil	8.97	0.90	0.91	36.18	1.00	1.01
Dubai oil	8.48	0.87	0.88	35.38	1.01	1.01
WTI oil	8.78	0.89	0.90	36.07	1.01	1.02
Olive oil	4.19	0.90 **	0.90 **	18.33	1.03	1.03
Oranges	12.05	0.96	0.97	23.08	1.13 ***	1.13 ***
Palm oil	7.83	0.92	0.92	32.21	1.04	1.04
Pork	9.14	0.98	0.98	23.38	0.99	0.99
Poultry	1.27	0.54 ***	0.54 ***	6.32	0.96	0.96
Rice	5.92	0.81	0.81	25.86	1.01	1.01
Rubber	8.27	0.89 *	0.90 *	37.20	1.00	1.00
Salmon	7.03	0.92	0.93	22.26	1.04	1.04
Hard Sawnwood	2.12	0.99	1.00	8.13	1.05**	1.05 *
Soft Sawnwood	5.72	0.91	0.91	10.04	0.96**	0.96 **
Shrimp	5.00	0.98	0.99	20.98	1.00	1.00
Soybean meal	7.10	0.93	0.92 *	25.04	1.05	1.05
Soybean oil	6.06	0.91	0.91	27.67	1.02	1.02
Soybeans	6.51	0.91 *	0.90 *	26.61	1.03	1.03
EU sugar	2.15	0.86	0.86 *	9.67	1.06*	1.06 *
Sugar	7.71	0.97	0.97	30.31	1.04	1.04
US sugar	3.53	0.88 **	0.88 **	17.71	0.97	0.97
Sunflower oil	9.31	0.88	0.88	38.76	1.19	1.19
Tea	7.08	0.99	0.99	19.23	1.05	1.05
Tin	6.73	0.88**	0.88 **	33.83	1.00	1.00
Uranium	6.55	0.90	0.90	38.87	0.98	0.98
Wheat	7.41	0.96	0.96	29.83	1.05**	1.05 *
Coarse wool	5.80	0.91*	0.91 *	28.39	1.03	1.03
Fine wool	6.06	0.91**	0.91 **	25.82	1.04	1.04
Zinc	6.84	0.92	0.92	36.65	1.01	1.01

Note: The table shows the root mean forecast error (RMSE) of a benchmark model, i.e. a constant growth model and the MSE of the candidate forecasting model relative to the benchmark. (*), (**) and (***) indicate rejection of the null of equal predictive accuracy at the 10%, 5% and 1% level based on the Diebold and Mariano (1995) statistic. The model estimation is rolling using a fixed window of 20 years and the estimation starts in 2001:1. The evaluation period goes from 2001:2 to 2015:12. As robustness check, the table also displays the relative MSE for a model specification with two global factors.