EUROPEAN CENTRAL BANK
EUROSYSTEM

Working Paper Series

Simona Delle Chiaie, Laurent Ferrara,
Domenico Giannone

Common factors of commodity prices

No 2112 / November 2017

Abstract

In this paper we extract latent factors from a large cross-section of commodity prices, including fuel and non-fuel commodities. We decompose each commodity price series into a global (or common) component, block-specific components and a purely idiosyncratic shock. 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 in 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.

Non-technical summary

In this paper we analyse the degree of co-movement in international commodity returns by studying a broad range of commodities that are representative of the global market. In doing so, we estimate a dynamic factor model with a block structure to decompose each commodity price series into a global (or common) component, block-specific components related to specific commodity markets and a purely idiosyncratic shock. The distinction between global, block-specific and idiosyncratic components allows for the presence of shocks of different nature, having distinct consequences on the cross-correlation between commodity prices.

We find that there is a single global factor driving the bulk of commodity price fluctuations. 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. It is also strongly related to measures of economic activity, suggesting a close link with demand factors. This is further corroborated by the fact that the global factor has homogenous effects on all markets and hence limited effects on relative prices. Since the start of the new millennium, the relevance of the global factor has increased, especially for oil. We compute model-based historical decompositions of commodity price changes and we find that the global factor explains a large 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. By contrast, block components explain most of the fluctuations in commodity prices during episodes conventionally associated with supply or other commodity-specific shocks.

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.

1 Introduction

Primary commodities, in the form of raw or partially processed goods, have been traditionally case examples of traded goods across borders and account for a significant share of international trade. Despite the secular decline in commodity-intensive sectors in advanced economies, primary commodities continue to have a central role in transportation, manufacturing processes, and in the food supply. The fast economic expansion of emerging market economies, in particular China, has also contributed to a rapid increase in the demand for industrial commodities since the beginning of the new millennium.

Most primary commodities are bought and sold around the globe in well-organized markets where physical and derivative trading take place. Like stock exchanges, commodity exchanges feature institutional and regulatory frameworks that ensure valuable protection to commodities traders and a high level of market liquidity. This is reflected by the fact that purely financial transactions currently outpace transactions in which physical delivery actually occurs.

As far as financial asset returns are concerned, it has been long recognized that they are characterised by a high degree of co-movement (see, for a recent survey, Connor and Korajczyk (2010)). This feature is at the heart of the asset pricing theory and implies that a few underlying factors explain the bulk of the fluctuations in asset returns. Recently, Miranda-Agrippino and Rey (2015) also find evidence of international co-movement in the returns of a large panel of risky assets. A few research studies, as well as more informal narratives, indicate that there are also commonalities in international commodity prices. The presence of strong co-movement among prices of a broad range of seemingly-unrelated commodities might seem puzzling, given that there are many specific factors affecting supply and demand in each market. Pindyck and Rotemberg (1990) described the phenomenon as "excess co-movement" among commodity prices.

In this paper we analyse the degree of co-movement in international commodity returns by studying a broad range of commodities that are representative of the global market. In doing so, we estimate a dynamic factor model with a block structure to decompose each commodity price series into a global (or common) component, block-specific components related to specific commodity markets and a purely idiosyncratic shock. The distinction between global, local and idiosyncratic components allows for the presence of shocks of different nature, having distinct consequences on the cross-correlation between commodity prices.

We find that there is a single global factor driving the bulk of commodity price fluctuations. 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. It is also strongly related to measures of economic activity, suggesting a close link with demand factors. This is further corroborated by the fact that the global factor has homogenous effects on all markets and hence limited effects on relative prices. Since the start of the new millennium, the relevance of the global factor has increased, especially for oil.

We compute model-based historical decompositions of commodity price changes and we find that the global factor accounts for a larger fraction of commodity price fluctuations in episodes typically associated with changes in global economic activity, such as the world economic expansion that started around 2003 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 other commodity-specific shocks. 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 oil-specific demand shocks and exogenous shifts in supply were a more important determinant of the price of oil before the 2000s.

In order to verify the robustness of the modelling strategy, we complement the in-sample analysis by performing an out-of-sample validation of the model. Overall, we find that our factor model performs well in forecasting commodity prices and aggregate indices of commodities, in particular at short horizons. In particular, the predictive performance of the global factor is higher for the group of commodities for which the historical variance explained by the global factor is larger, such as food and metals. These results are in line with other studies – focusing on the oil market – which have also found evidence that proxies of global demand have predictive power for oil prices (see, Baumeister and Kilian (2012)). Similarly, changes in commodity prices indices, in particular industrial raw materials, have been proved to improve the forecast of the price of oil, as these indices are more likely to capture shifts in the global demand for industrial commodities (see, Alquist, Kilian and Vigfusson (2013)). In this respect, our factor-based forecast is a refinement of these 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.

Our paper is not the first that studies the co-movement in commodity prices (see, e.g. Alquist and Coibion (2014), Byrne et al. (2011), West and Wong (2014), Chen et al. (2014)). The focus of this earlier literature is different from ours for two reasons. First, these papers have studied commodity prices in levels instead of the returns, focusing on the co-movement at low frequencies. Second, they do not analyse simultaneously all commodity markets and look at selected groups of commodities.

The importance of global demand has been extensively documented in the context of the oil market (see, e.g. Barsky and Kilian (2002), Kilian (2009), Peersman and Van Robays (2009), Bodenstein, Guerrieri and Kilian (2012), Lippi and Nobili (2012) and Aastveit, Bjorland and Thorsrud (2015)), our paper shows that the association with global economic activity is even stronger when looking at the common factors underlying all the commodities.

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

2 Empirical analysis

2.1 Data

Our dataset includes the spot prices of 52 internationally traded commodities from different categories: food, beverages, agricultural raw materials, metals and fuel commodities. The source of our data is the IMF primary commodity price database. The composition of the IMF dataset has been designed to be representative of the world economy, hence it includes the most relevant commodities in terms of trade values. We use monthly averages of daily prices for a sample from January 1980 to December 2015.[1] Since the estimation of the model described in the next 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. The full list of series used in the analysis and their descriptions are reported in Table 2. 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 constructing the indices are the commodity trade values compared to the total world trade as reported in the UN Comtrade database.

An interesting feature of the dataset that we will exploit later in the empirical analysis is that indices are built to reflect different levels of aggregation; the dataset 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. The non-fuel block can then be broken down into two other group indices, food and beverages 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 an index of crude oil prices.

It is important to note that the number of variables included in each block is not representative of the respective shares. For instance, as shown in Table 1, 60 percent of the overall index is represented by 6 energy commodities. The largest block is the Food and Beverages block which includes 28 series but its share in the overall index is less than 20 percent. The composition of the data and the presence of strong local correlation, are important for the

[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)).

estimation of the common factors. Simple methods, such as principal components, tend to give more weight to categories in the panel that are over-represented. Under these conditions, factors can be poorly estimated (Boivin and Ng (2006)) and the number of factors can be mis-specified (Luciani, (2014)). To mitigate this problem, in practice, one might consider to carefully select commodities before estimation in order to minimize the correlation between idiosyncratic components (Alquist and Coibion (2015)). Alternatively, as done in this paper, one can explicitly model the local correlation in specifying the factor model to alleviate those concerns. This is explained in the next section.

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:

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

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

The idiosyncratic errors are modeled with a block factor structure which 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 $\varepsilon_{it} \sim i.i.d \mathcal{N}\left(0, \sigma_i^2\right)$.

Notice that we have assumed the block factors to be uncorrelated. This implies that while commodities in the same market can be correlated due complementarities or common technology shocks, commodity-specific shocks cannot spillover to other commodity markets. The assumption is consistent with the evidence that the pass-through from shocks to the price of crude oil to other commodities is limited (Baumeister and Kilian (2014)).

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. This is a reasonable initialisation since principal components have been proved to be an asymptotically consistent estimator of the true common factors when the cross-section dimension is large (see, Forni et al. (2000), Stock and Watson (2002a, 2002b) and Bai (2003)). Once we have estimated parameters, the second step consists in updating the estimate of the common factors by using the Kalman smoother. If we stop here, we get 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, for example, in Reis and Watson (2010) and Luciani (2015).

2.3 How many factors?

We begin our analysis by estimating common and 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.

We determine the number of blocks to include in the model by following the structure of our database. In macroeconometrics, data are typically organised either by country, sectoral origin or economic concept, therefore 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 estimation methods, Kose, Otrok, and Whiteman (2003) study the sources of the international business cycle by extracting world, country and regional components; Banbura, Giannone and Reichlin (2011) use blocks of 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 which in our case is based on different categories of commodities (see, Table 1). As a result, 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).

To determine the optimal number of common factors from the observed data, 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. In order to do so, we use 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. Nevertheless, 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 $C_{nT}^{*2} = min\left\{\sqrt{T}, (n/(log(n))\right\}$. 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, T is the number of sample observations, $F_{(r)}$ 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 statistics in Table 3 selects the model with one common factor since this provides the smallest value of the IC statistics. From here on, we will refer to this single common factor in commodity prices as the global factor.

2.4 Empirical results

The Global Factor

The global factor estimated over the full sample is shown in Figure 1 along with the IMF global index of commodity prices. The latter is a linear combination of commodity prices with weights given by trade values. While cross-sectional averages, such as the IMF index, 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. As Figure 1 illustrates, the global factor and the IMF broad index of commodity prices resemble each other,[3] but their second-order properties appear different. A visual inspection of the two series suggests that while the broad index of commodity prices is characterised by swift fluctuations, for instance those associated with the oil price shocks in the early 1990s, the global factor is a smoother and more persistent series.

Particular attention should clearly be paid to what the global factor captures. A natural conjecture is that the global factor, being a pervasive shock affecting a large cross-section of

[2] The information criterion has been recently applied to the quasi-maximum likelihood estimator by Coroneo et al (2016).

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

commodity prices, might capture shifts in the demand for commodities associated with the global business cycle. In fact, as the global economy expands, so does demand for a broad group of commodities, directly via the impact on industrial commodities and indirectly via general equilibrium effects. Barsky and Kilian (2002) argued that broad-based variations in commodity prices are consistent with the evidence of shift in demand driven by macroeconomic conditions. If so, one would expect the global factor to have homogenous effects on all commodity markets and therefore, limited effects on relative prices. This is confirmed by the evidence in Figure 2 that shows that the factor loadings associated with the global factor are mostly positive. Since stronger economic activity is associated with higher commodity prices, it is not surprising that the global factor is also strongly correlated with indicators of global real economic activity. To illustrate this, Figure 3 shows the global factor along with the Kilian's (2009) index of economic activity. This measure, based on percentage changes of dry cargo ocean freight rates, has been developed to capture shifts in the demand for industrial commodities associated with periods of high and low real economic activity. In order to make the comparison meaningful, the global factor is expressed here in year-on-year growth rates. As Figure 3 shows, the two indices are positively correlated and follow the major expansion and contraction phases in the international business cycle over the period considered with the largest declines following recession periods. For example, both measures capture the fast macroeconomic expansion that characterised the world economy and, in particular some emerging market economies, since the earlier 2000s. Moreover, both the global factor and the Kilian's index declined in the second half of 2014, suggesting a weakening in global economic activity which is then reflected in the subsequent decline in the price of oil. Likewise, Figure 4 shows that the global factor is also strongly correlated with monthly indicators of industrial production. As noted by Kilian (2017), the common factors or broad-based indices of commodity prices are actually leading indicators with respect to global industrial production, which makes the global factor a suitable real-time indicator of to estimate aggregate demand pressures in structural models.

To gauge the extent to which the global factor is related to fluctuations in oil prices, 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 only mildly positive over the full sample but the correlation between the two series has increased substantially since the last decade. Second, both 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 and 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 changes in the supply of oil.

As a robustness check, we estimate the model using real commodity prices (i.e. deflated by the US CPI). The resulting global factor (as shown in Figure 6) does not appear to be particularly sensitive to this transformation. However, this might reflect the fact that our sample does not include high inflation periods, such as the Great inflation of the 70s.

Sources of commodity price fluctuations

In this section we study the relative importance of global and block-specific factors in explaining commodity price fluctuations. For expository purposes, we also compute a model-based variance decomposition for the commodity indices in our dataset.[4] Table 4 and 5 report in the first column the share of the variance of commodity prices and indices that is explained by the global factor. The remaning columns show the share of the variance explained by block-specific factors and the purely idiosyncratic component. The global factor explains more than two-third of the variations of the index of non-fuel commodities. This stems from the fact that a large fraction of the variance of food commodities and metals is captured by the global factor. In particular, the global factor explains almost half of the variations in soybean and soybean oil, 40 percent of sunflower oil and about one-third of copper and palm oil price variations. The strong common component in the price of these commodities and in particular copper, explain why researchers had used changes in the price of copper or broad index of non-fuel commodity price in effort of isolating global demand components (Kilian and Lewis (2011) and Hamilton (2014)). The bulk of the fluctuations in beverages, agricultural raw materials and fuel prices is instead mostly captured by block-specific factors. Nevertheless, the global factor explains about 20 percent of oil price fluctuations on average over the sample considered. Given the large weight attributed to oil prices in the IMF index, the fuel-specific factor explains most of the variance of the overall IMF index of commodity prices. One-third of its fluctuations are instead driven by the global factor.

To check the robustness of these results, we include a second global factor in the model. The results shown in Figure 7 and 8 indicate that the second common factor explains a very small share of the variance of commodity prices on average. This fraction is small enough to reinforce the evidence provided by the IC statistic about the presence of a single global factor.

Sub-sample analysis

The analysis over the full sample might mask some important changes that might have happened in the commodity markets, in particular, since the start of the commodity price boom in mid-2003. In Figure 9, we report 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

[4] Let y_t be an m-dimensional 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.

increase in commodity prices. The subsample analysis confirms that the global factor explains an important fraction of the variation of non-fuel commodities before the 2000s while it has little explanatory power for oil and other fuel-commodities. Indeed, block-specific and idiosyncratic components account for the whole variation in oil prices in the first sub-sample. This evidence suggests that commodity-specific 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 VAR analysis in Kilian and Murphy (2014) supports this interpretation showing that key historical events in the oil market over this period, 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. The model-based variance decomposition estimated over the second sub-sample indicates that the importance of the global factor has increased since 2003. The increase is remarkable for oil and metals for which the share of the variance explained by the global factor raised to 40 and 60 percent, respectively. As a result, the share of the variance of the IMF index that can be attributed to the global factor has also increased from less than 10 percent to 60 percent in the period starting from 2003.

Historical decompositions of commodity price changes

As sustained changes in the global factor tend to be indicative of aggregate demand pressures, our factor-structure approach, although it is not structural in nature, allows to disentagle commodity price fluctuations that are driven by demand shifts associated to the global business cycle from those that are commodity-specific, such as supply-driven fluctuations. Commodity-specific shocks are unlikely to spill over all other commodities and their effects are likely to be confined to their specific market or to markets of commodities in their category. In our model their effects will not show up in the common component but on the idiosycratic and block-specific factors. As discussed in previous sections, it is important to note that a clear advantage of the block structure is that there is no need to carefully select the commodities that enter the factor model as practiced in Alquist and Coibion (2014) and discussed in Kilian (2017). In order to keep the panel balanced and avoid that some categories are over represented, which could bias the estimates of the factors toward some markets, Alquist ad Coibion (2014) extract a common factor from a restricted group of commodities that are supposedly unrelated. Rather than selecting the variables before estimation, our approach uses a block structure to mitigate this issue.

In addition, local factors can be confused with global variations in absence of a block structure. To illustrate this, we compare the estimated 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 modelled. Figure 10 shows that the first common factors of the two models 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 third factor of M_2 is highly correlated with the Fuel block in M_1 and, to some extent, with the Food and Beverages sub-block.

In what follows we review a few key historical episodes of commodity price variations in our sample through the lenses of our model. Although the model allows the analysis of a large panel of commodity prices, we focus on an arbitrary small number of commodities with large trading volume for reasons of space. An important event in the commodity market was the run-up in commodity prices from 2003 to mid-2008. There is a widespread agreement that the fast economic expansion that characterized emerging market economies, and in particular China, caused the surge in commodity prices (see, Hamilton (2009), Kilian and Hicks (2013) and Aarsveit et al. (2014)). Figure 11 presents the historical decomposition of Brent crude oil, copper, nickel and corn, showing cumulative changes at each point in time 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 since 2003 while the fuel-specific component had a smaller role. This result is consistent with estimates from empirical models of the global oil market, which attributed the bulk of the cumulative increase in the price of oil to global demand shocks (Kilian (2009), Baumeister and Peersman (2013), Kilian and Murphy (2014). Fuel-specific components were instead important to explain the increase in the price of oil in the early 2000s, in line with previous estimates. As the remaining panels of Figure 11 indicate, the global factor is by far the most important determinant of the surge in non-fuel commodity prices since 2003, suggesting that commodity prices responded to the same economic fundamentals.

A different view expressed by some observers and by a few studies in the financial literature (Tang and Xiong (2012)), has associated the across-the-board surge in commodity prices in 2003-2008 with the growing participation of financial speculators in commodity markets at the beginning of the 2000s. A large body of research, however, has provided compelling evidence that financial speculation did not have an effect on commodity prices (Kilian and Murphy (2014), Kilian and Lee (2014), Juvenal and Petrella (2014)). For a survey of this literature, the reader is referred to Fattouh, Kilian and Mahadeva (2013).

Figure 12 looks at four historical episodes in the oil market. The first two events refer to the oil price fall that followed the collapse of the OPEC cartel in late 1985 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 driven by factors specific to the oil market and unrelated to changes in macroeconomic conditions. The historical decomposition shows that, in both episodes, fuel-specific factors were the main underlying sources of oil price changes that occurred before the 2000s, while the global factor had clearly no role (Figure 12, panel 1 and 2). This is consistent with evidence from structural models of the oil market that showed that shifts in the inventory demand and in the supply of oil were the most important determinant of the oil price in both episodes. Differently from these earlier events, the deep fall in the price of oil which started in mid-2008 as a result of the contraction in world economic activity is mostly explained by the global factor, as shown in the third panel of Figure 12. There is also evidence that oil-specific component exerted further downward pressures on the price of oil since the end of 2008. Finally, the last panel of Figure 12 investigates the oil price fall that started in the second half of 2014. In a first initial assessment, Baumeister and Kilian (2016) find that global demand was the main cause of the oil price decline from June to December 2014. We find that while the global factor explains most of the initial oil price fall, cumulated changes in the fuel-specific component explained most of the variations since the end of 2014. Thus, the model attributes about one-third of the oil price fall from June 2014 to December 2015 to the global factor. The increasing relevance of fuel-specific components since the end of 2014 coincides with the decision of OPEC in November 2014 to hold production unchanged in order to put downward pressures on prices. The empirical findings in Baffes et al. (2015) and Groen and Russo (2015) appear to be consistent with ours.

3 Predictive content of the global factor

A growing empirical literature has used factor models estimated on panels of commodity prices for forecasting purposes. 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 empirical studies have instead investigated whether macroeconomic and financial data have predictive power for commodity prices (see, e.g. Chen, Rogoff and Rossi (2012), Groen and Pesenti (2011)). Focusing on the price of oil, a strand of the literature has found 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 perform an out-of-sample validation of the model to verify the robustness of the modelling strategy. To this end, starting from January 2001, we estimate the model on a rolling window of 20 years of past data and 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 aggregate commodity indices are computed as averages of the individual commodity price forecasts, weighted using their trade weights.[5] After com-

[5]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. 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.

puting 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 calculate the average loss difference 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, is based on the difference between the mean squared forecast error (MSFE) of the candidate model and the benchmark, smoothed over time with a centered rolling window of fixed size. 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 and indices, especially at short horizons. At $h = 1$, the model outperforms the benchmark with gains in accuracy that range from 18% for the non-fuel index to 12% for the fuel-index (Table 6). The forecasts of disaggregated commodity prices in Table 7 indicate that the model provides the largest accuracy gains for food and metals (for instance, 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 predictive performance deteriorates progressively over longer horizons and at $h = 12$, we cannot reject the hypothesis of equal predictive performance between the model and the benchmark. Finally, we find that the predictive ability of the model has changed over time. The evolution of the rolling relative MSFE in Figure 13 indicates that the predictability of oil and other energy commodities increased markedly in the second half of the 2000s. Indeed, from 2007 to 2011, the MSFE of the factor model improved substantially compared to the benchmark. However, given the high level of volatility, the test cannot reject the null of equal predictive accuracy. The finding of a greater predictive performance during the Great Recession is consistent with previous results showing that, for macroeconomic and financial variables, downturn periods are characterised by an increased co-movement (see, e.g. D'Agostino and Giannone (2012)).

4 Concluding remarks

We studied the co-movement in international commodity returns by analyzing a broad range of commodities, which are representative of the global market. Results indicate that the clear co-movement is not only as strong as already documented by Pindyck and Rotemberg's (1990) but also strengthening since the 2000s. Contrary to earlier studies, we find that the co-movement is neither excessive nor puzzling, as it is driven by a pervasive factor that is strongly related to measures of global economic activity, suggesting a close link with demand determinants.

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 alternative, 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 (2014), Do Oil Price Increases Cause Higher Food Prices?, Economic Policy 80, 691-747.

[14] 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.

[15] Baumeister C. and G. Peersman (2013), The Role of Time-Varying Price Elasticities in Accounting for Volatility Changes in the Crude Oil Market", Journal of Applied Econometrics 28(7): 1087-1109.

[16] Bodenstein M., Erceg C. and L. Guerrieri (2011), Oil shocks and External Adjustment, Journal of International Economics 83(2), 168-184.

[17] Bodenstein M., Guerrieri L. and L. Kilian (2012), Monetary Policy Responses to Oil Price Fluctuations, IMF Economic Review 60 (4), 470-504.

[18] Boivin J., and S. Ng (2006), Are More Data Always Better For Factor Analysis? Journal of Econometrics 127, 169-194.

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

[20] 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.

[21] 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.

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

[23] Connor G. and R. Korajczyk (2010), Factor Models of Asset Returns, Encyclopedia of Quantitative Finance, edited by Rama Cont, ed. Chicester: Wiley.

[24] Coroneo, L. D. Giannone and M. Modugno (2016), Unspanned Macroeconomic Factors in the Yields Curve, Journal of Business and Economic Statistics 34, 472-485.

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

[26] 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.

[27] 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

[28] 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.

[29] 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.

[30] 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.

[31] 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.

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

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

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

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

[36] Gospodinov, N. and S. Ng (2013), Commodity prices, Convenience yields and Inflation. Review of Economics and Statistics 95, 206–219.

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

[38] 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.

[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.

[40] 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.

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

[42] Kilian, L. and Zhou X. (2017), Modeling Fluctuations in the Global Demand for Commodities, manuscript, University of Michigan.

[43] 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), 1216-1239, September.

[44] Hamilton J. (2009), Causes and Consequences of the Oil Shock of 2007-08. Brookings Papers on Economic Activity, Spring, 215-283.

[45] Hamilton J. (2014), Oil prices as an Indicator of Global Demand Conditions: http://econbrowser.com/archives/2014/12/oil-prices-as-an-indicator-of-global-economic-conditions

[46] Lippi F. and A. Nobili (2012), Oil and the Macroeconomy: A Quantitative Structural Analysis, Journal of European Economic Association, vol. 10(5), 1059-1083.

[47] Luciani, M. (2014), Forecasting with Approximate Dynamic Factor Models: The Role of Non-Pervasive Shocks, International Journal of Forecasting 30, 20-29.

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

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

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

[51] Pindyck R.S. and J. Rotemberg (1990), The Excess Co-Movement of Commodity Prices, The Economic Journal 100(403), 1173-1189.

[52] Peersman G. and I. Van Robays (2009), Oil and the Euro Area Economy, Economic Policy, vol 24, 603-651.

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

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

[55] 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.

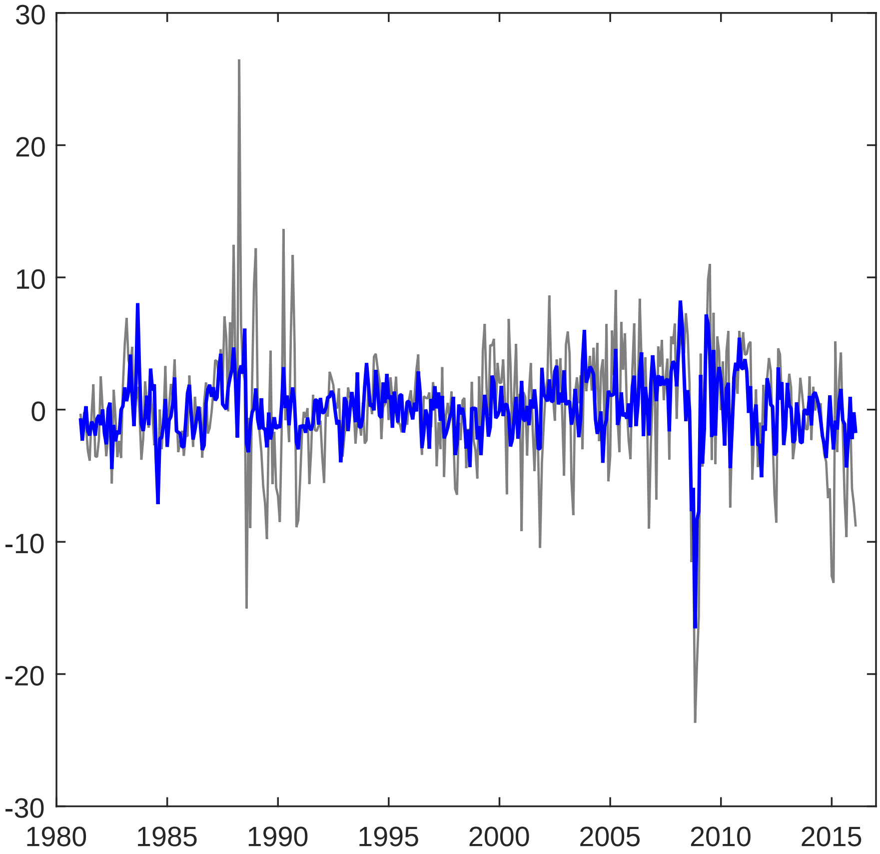
[56] Stock J. H. and M. W. Watson (2002b), Macroeconomic Forecasting Using Diffusion Indexes, Journal of Business and Economics Statistics 20, 147-162.

[57] 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.

[59] White, H. (1982), Maximum Likelihood Estimation of Misspecified Models, Econometrica 50, 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

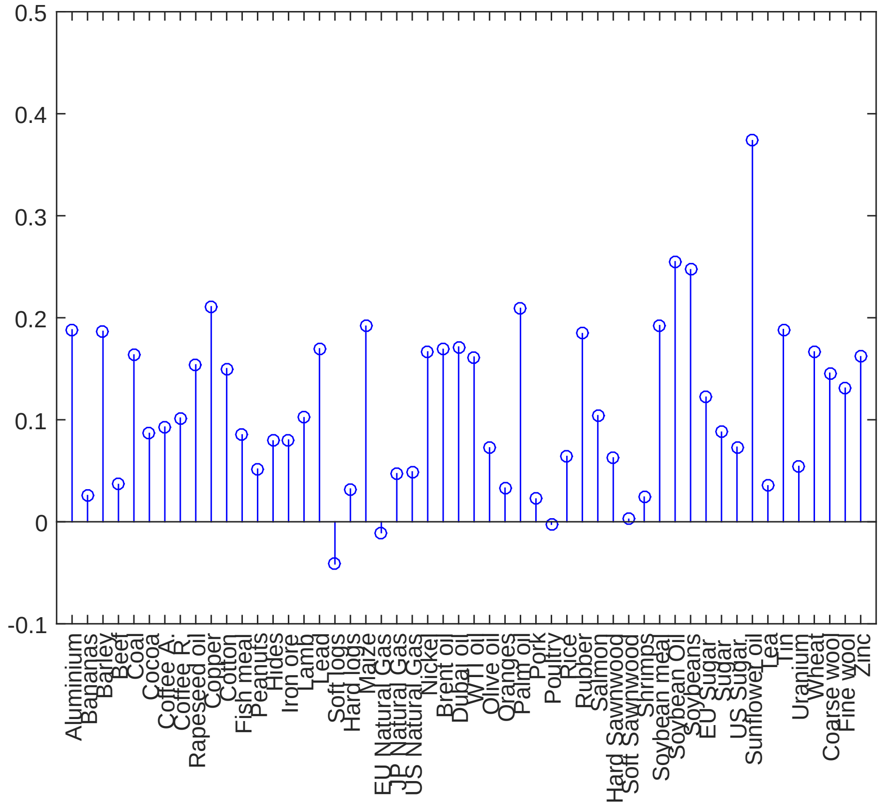
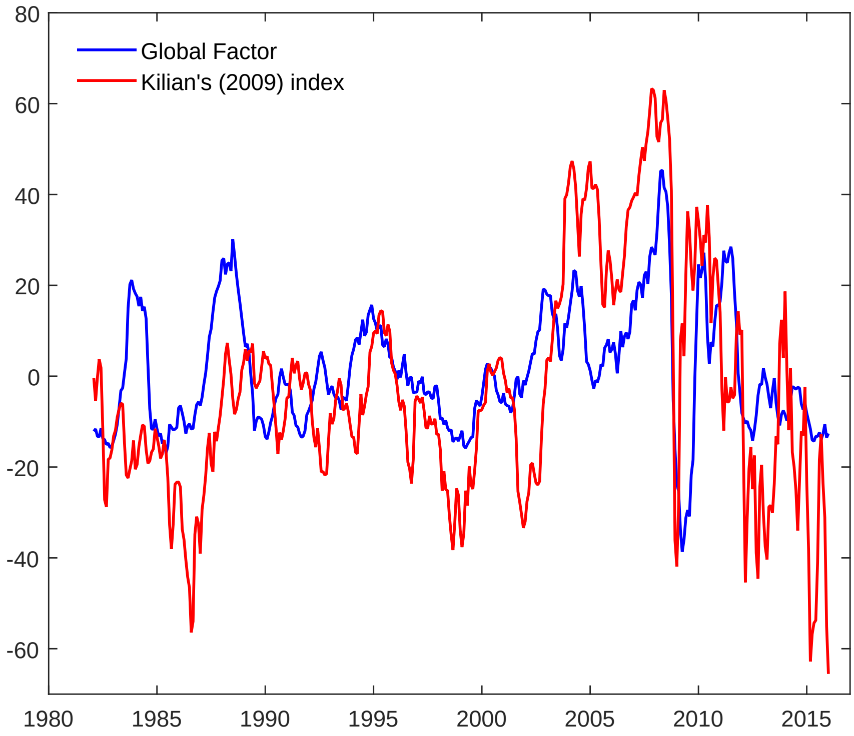
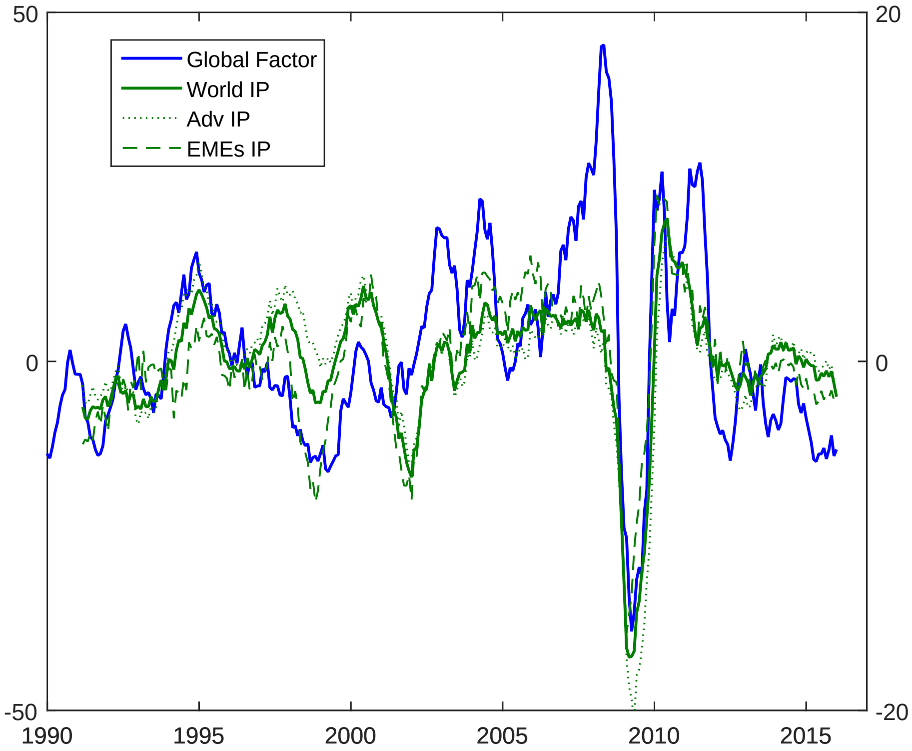


Figure 3: The global factor and the Kilian's index of economic activity



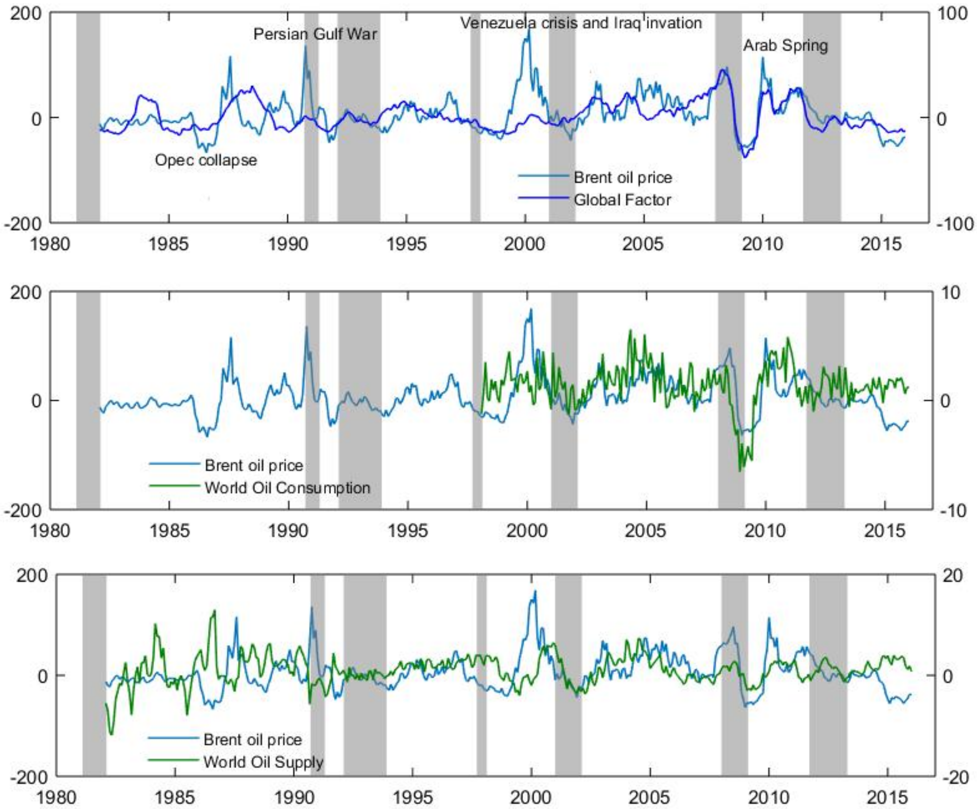
Note: The figure plots the global factor (blue line), extracted from real commodity prices and expressed in year-on-year growth rates, and the Kilian's (2009) index of real economic activity (red line).

Figure 4: The global factor and industrial production indices



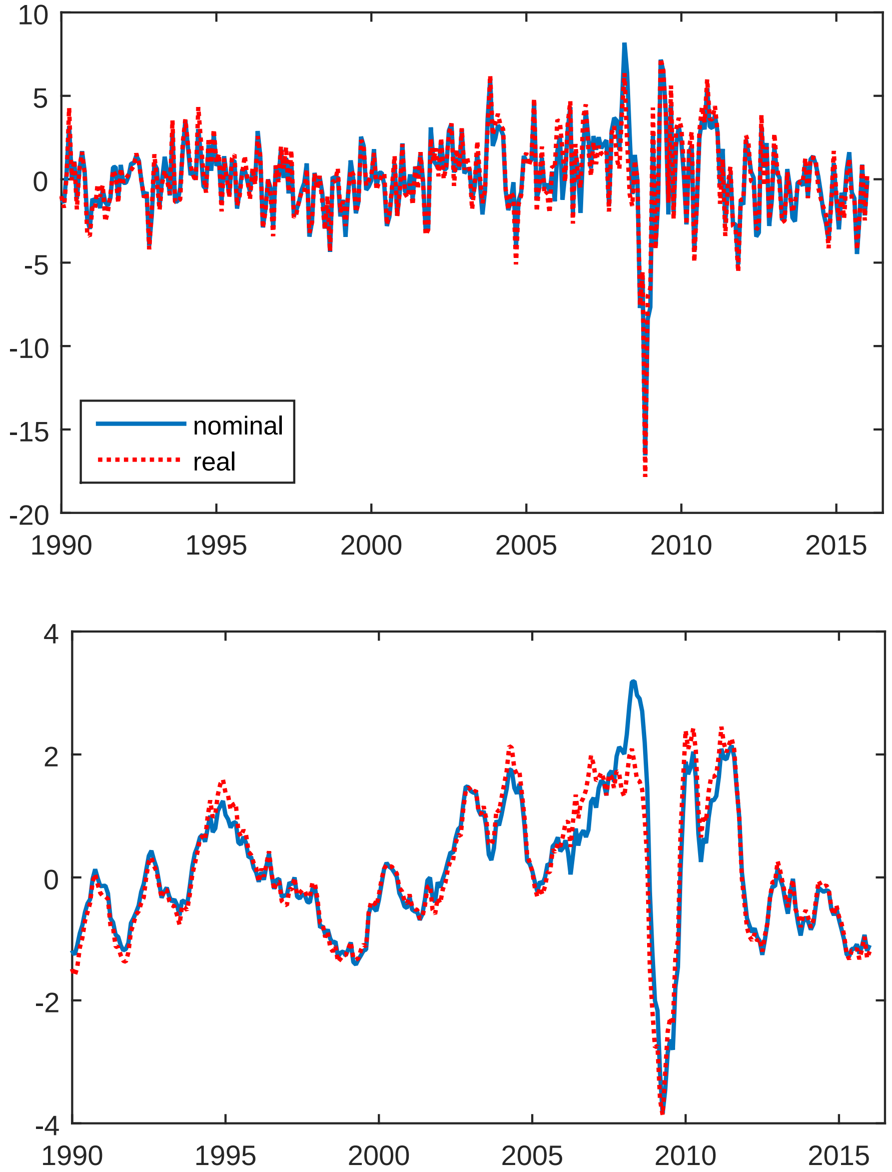
Note: The figure plots the estimated global factor (blue line) extracted from real commodity prices and measures of industrial production in selected areas as provided by the CPB Netherlands Bureau for Economic Policy Analysis. All variables are expressed in year-on-year growth rates.

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. In particular, 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 6: The global factor extracted from real commodity prices



Note: The upper panel of the figure reports the estimated global factor extracted from real commodity prices (US CPI deflated). The lower panel reports the nominal and the real global factor expressed in year-on-year growth rates.

Figure 7: Variance explained by the first two global factors

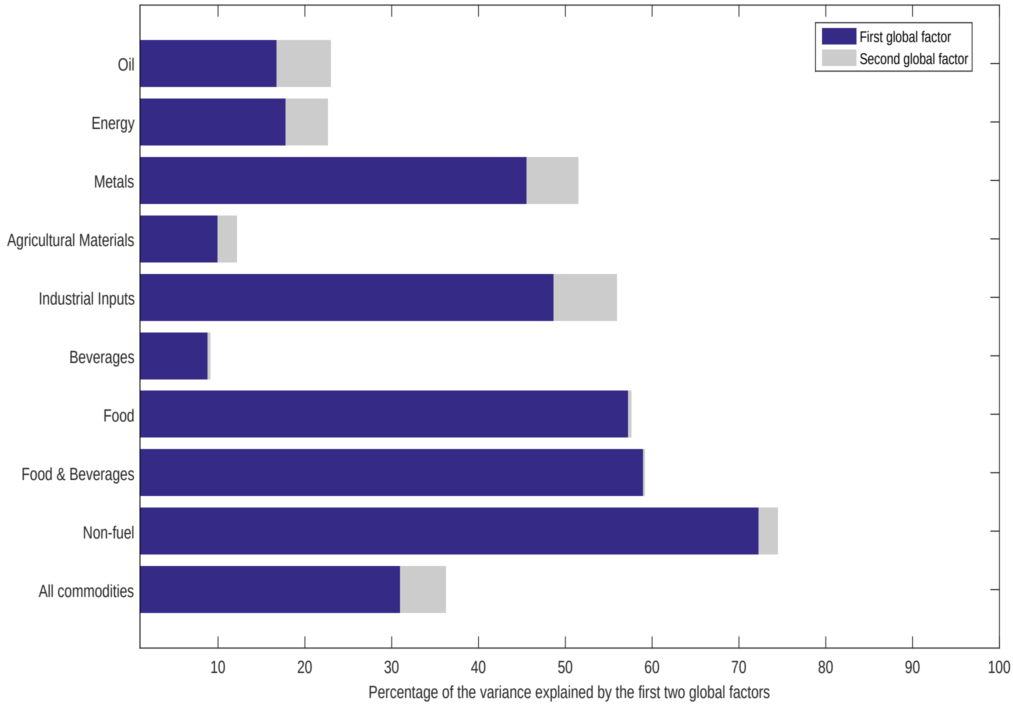


Figure 8: Variance explained by the first two global factors

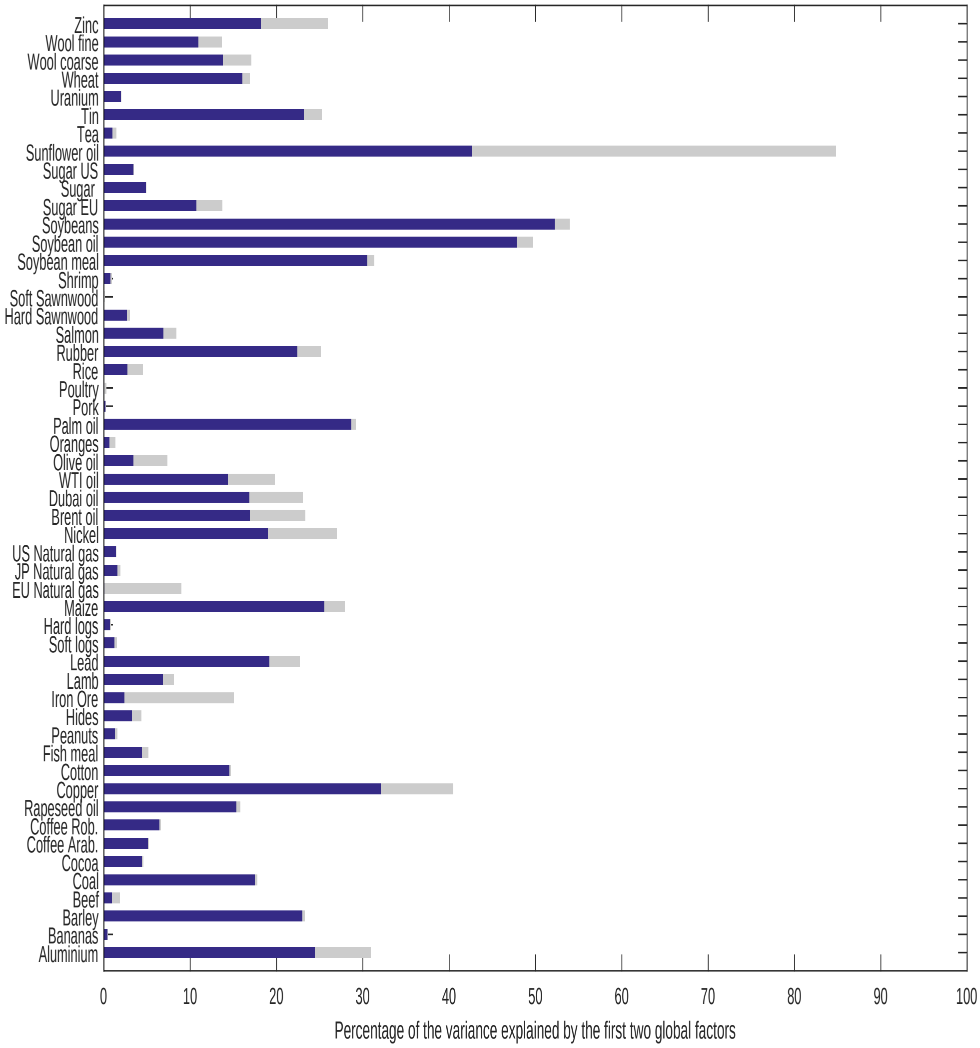
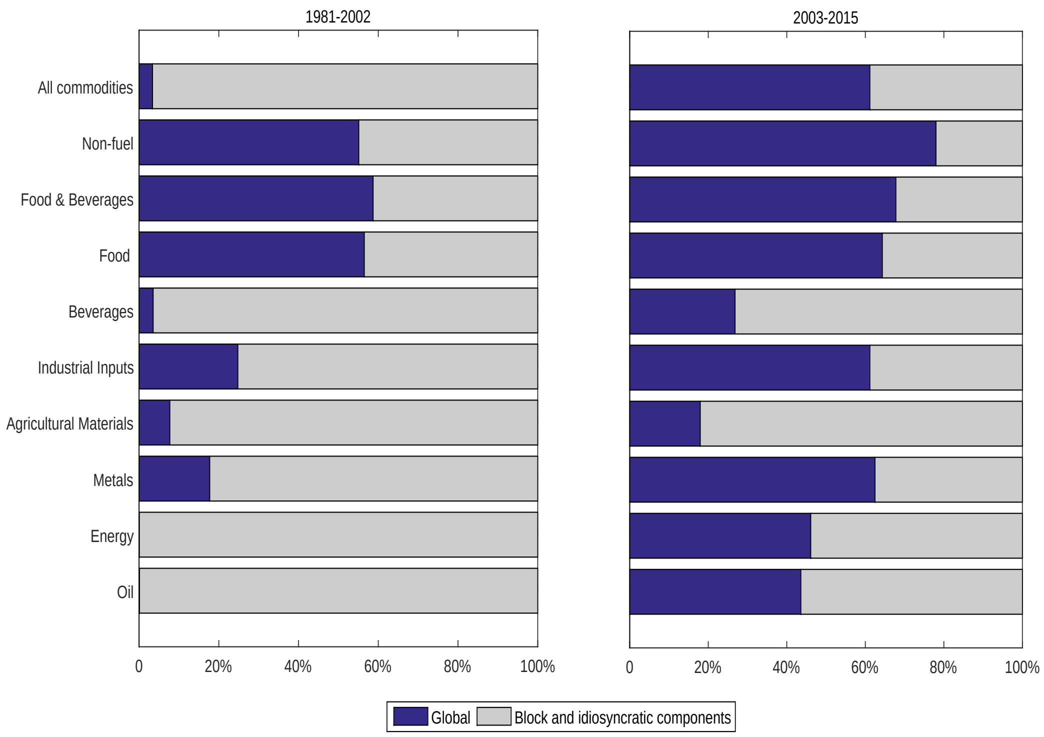
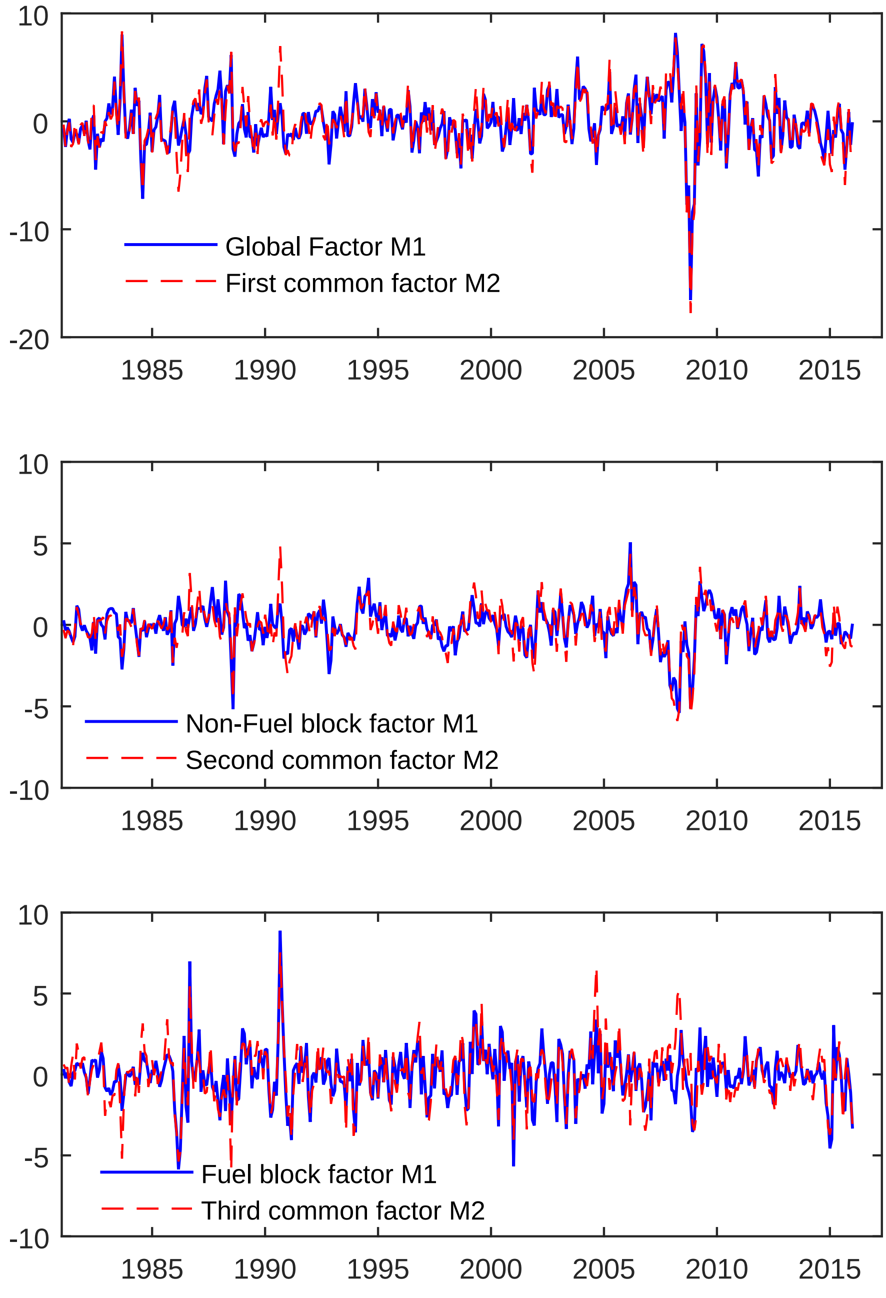


Figure 9: Variance decomposition: sub-sample analysis



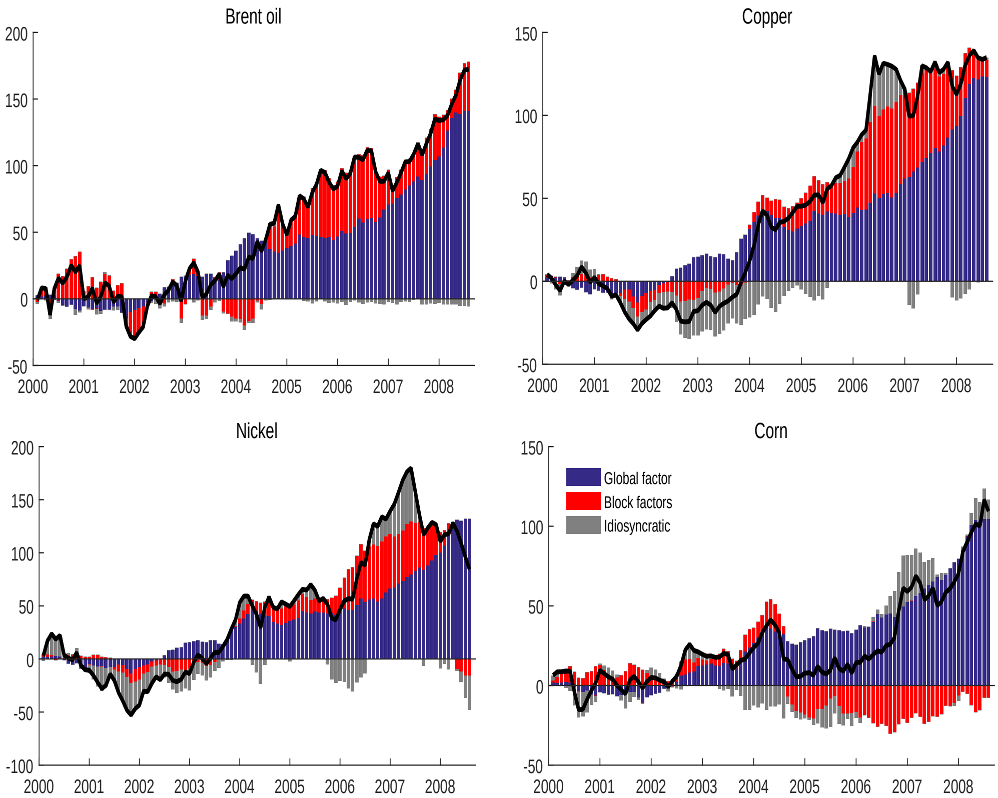
Note: The figure reports the variance decomposition of commodity price indices over two sub-samples. The share of the variance explained by the global factor is captured by the blue bar while the Grey bar is the percentage of the variance explained by block-specific and idiosyncratic components. The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

Figure 10: Block factors and weak common factors



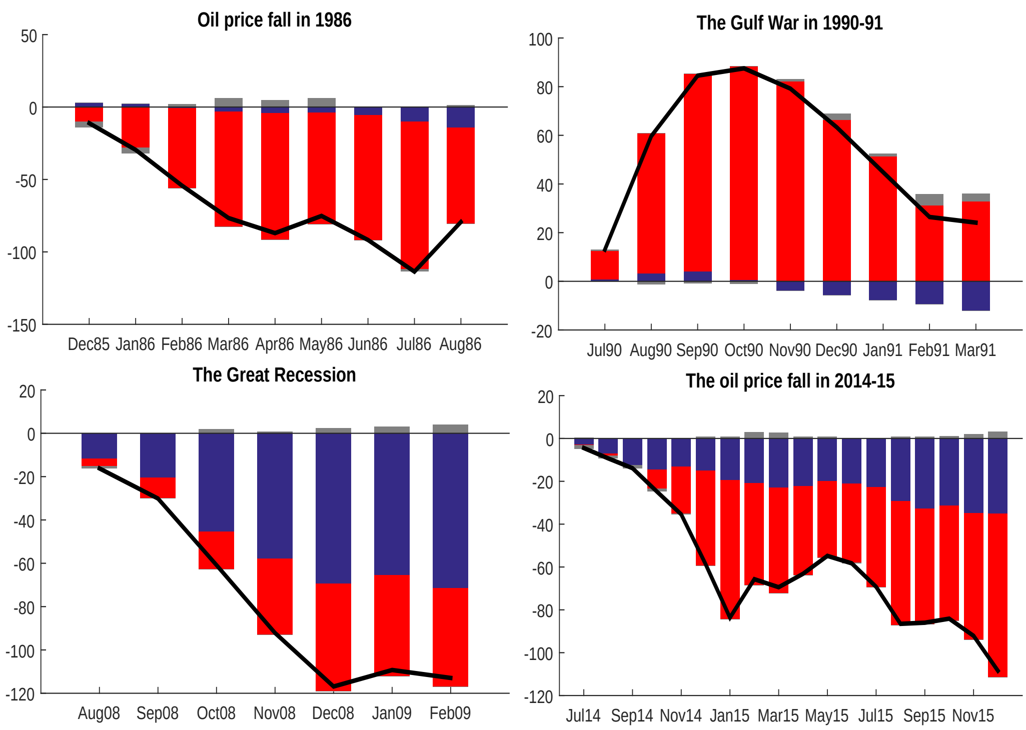
Note: The figure compares the the estimated global and block factors (Non-Fuel and Fuel) of our benchmark specification (M_1) with the first three common factors extracted from a factor model without block structure (M_2).

Figure 11: Historical decompositions of commodity prices



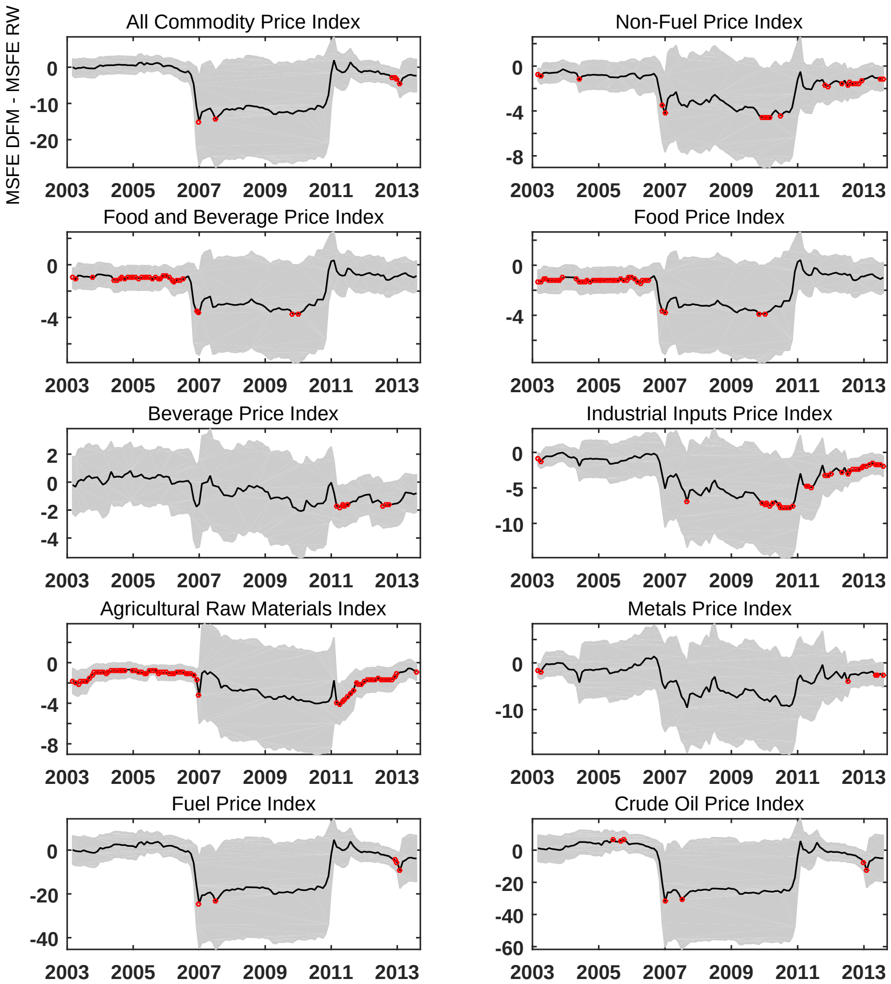
Note: The figure reports the historical decomposition of the price for a selected group of energy, metal and food commodities, showing the cumulative effects at each point in time of global (blue), block-specific (red) and idiosyncratic (Grey) shocks from January 2000 to July 2008.

Figure 12: 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 13: 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 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.

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. A ratio smaller than 1 indicates that the factor model forecasts are on average more accurate. (*), (**) 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. A ratio smaller than 1 indicates that the factor model forecasts are on average more accurate. (*), (**) 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.

Acknowledgements

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, Michele Lenza, 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, the Bank of England-Banque de France Workshop on International Macroeconomics 2015 and the 23rd Annual Conference of the Society for Computational Economics 2017 for their useful comments.

Simona Delle Chiaie (corresponding author)
European Central Bank, Frankfurt am Main, Germany; email: simona.dellechiaie@ecb.europa.eu

Laurent Ferrara
Banque de France, Paris, France; email: laurent.ferrara@banque-france.fr

Domenico Giannone
Federal Reserve Bank of New York, New York, United States, Centre for Economic Policy Research (CEPR), London, United Kingdom; email: domenico.giannone@ny.frb.org

Postal address	60640 Frankfurt am Main, Germany
Telephone	+49 69 1344 0
Website	www.ecb.europa.eu

This paper can be downloaded without charge from www.ecb.europa.eu, from the Social Science Research Network electronic library or from RePEc: Research Papers in Economics. Information on all of the papers published in the ECB Working Paper Series can be found on the ECB's website.

ISSN	1725-2806 (pdf)	DOI	10.2866/116302 (pdf)
ISBN	978-92-899-3026-0 (pdf)	EU catalogue No	QB-AR-17-124-EN-N (pdf)