What Makes Commodity Prices Move Together?
An Answer from a Dynamic Factor Model

Abstract

This paper aims to investigate the common movement of commodity prices. Two alternative hypotheses explaining the origin and nature of this common movement are put forward: the interdependence and the common latent factor hypotheses. This latter is assessed by specifying a DF/FAVAR model whose latent factors move around a zero-mean short-term level and a non-stationary long-run equilibrium level, respectively. Four heterogeneous and mostly unrelated commodities are considered (crude oil, copper, wheat, beef). Using IMF monthly prices over the 1980:1-2016:4 period, a Kalman Filter ML estimation is performed and results suggest that, beside the increasing price volatility, the last decade experienced a significant rise of the long-term equilibrium price. Some implications of this major result are also discussed.

Keywords: Commodity Prices, Dynamic Factor/FAVAR Models, State-space Representation, Kalman Filter

1. Introduction: objectives of the paper

This paper deals with the common movement of commodity prices. In particular, the objective is to identify and estimate a long-term pattern (if any) shared by quite heterogeneous commodities, from agricultural products to minerals. While the existence of seemingly common movements of commodity prices can be easily demonstrated with descriptive statistics, the actual identification and estimation of their nature and, above, all, causes is challenging. High-frequency price movements are quite complex as they show some regularities but, at the same time, also periods where these regularities seem to be contradicted. Therefore, the main empirical difficulty consists in identifying that common stochastic process that is consistent with this price behavior.

Especially in the last two decades, a major research effort has been spent on this issues in order to detect the main features and drivers of commodity price movements (Baffes e Haniotis, 2016). Most of this large body of literature, however, concentrated on price interdependence and, therefore, on the causal relationships among prices. Causal linkage generates price transmission and, then, statistical interdependence. The nature of these causal linkages can be either real or financial (or both) but the empirical implications are eventually the same. However, real linkages can be hardly argued when very different and apparently unrelated commodities are under investigation, like beef and copper, for instance. Nonetheless, they still show strongly common dynamics. Financial arguments can be put forward to explain this fact[1] but while such arguments may be valid for futures markets they seem less likely in the case of spot prices.

With respect to this recent literature, the main hypothesis of the present work is that heterogeneous commodities show common price movements but this is not the consequence of some causal relationship or market interdependence. Their commonality actually comes from the fact that they are all dependent on the same underlying drivers. These drivers can not be actually observed thus they behave as common latent factors.

The paper firstly presents some descriptive evidence on the common movement of a small group of heterogeneous commodity prices and puts forward a battery of tests to assess the magnitude and nature of their statistical interdependence. Then, the common latent factor hypothesis is investigated

[1] In this respect, several works refer to the so-called financialization of commodity markets (or financialization of food in a more specific context) to express the increasing common movements of commodity returns (thus prices) not generated by some underlying supply and demand factors (Bruno et al., 2017).

within a Dynamic Factor framework. In order to admit both short-term price interdependence and the presence of common latent drivers, the stochastic generation process of these prices is modelled as a Factor Augmented VAR. The presence of latent factors requires this model to be specified in the state-space form and the consequent ML estimation is then performed via Kalman Filter.

2. The main regularities of commodity prices: literature review and some descriptive evidence

In any empirical investigation on commodity price dynamics and its drivers, the first and often toughest challenge consists in the proper identification of the price series stochastic properties. Such properties emerge as some combination of short and long-run movements as well as individual and common components (Kim et al., 2003; Schleicher, 2003; Bai, 2013; Martin et al., 2013). Moreover, this combined effect may concern either the price levels or their variances. In fact, it often concerns both (Piot-Lepetit and M'Barek, 2011; Listorti and Esposti, 2012; Esposti and Listorti, 2013).

In order to solve this puzzle, the wide recent literature has outlined some regularities about the underlying stochastic generation process. A first regularity is that commodity prices are expected to behave like *mean-reverting series* (Bobenrieth et al., 2014; Valera and Lee, 2016). This property comes from the fact that these markets depend on a stable and often inelastic demand while, on the supply side, they are strongly limited by natural constraints (i.e., some underlying natural resource stock). As a consequence, while in the short-term market prices may also significantly deviate (often due to storers' and speculators' expectations), in the medium and long-run they will tend to revert back towards the stable supply and demand market-clearing equilibrium. The main consequence of this feature, from the statistical point of view, is that these prices are expected to be stationary series, or integrated of order 0 (Schwartz, 1997; Routledge et al., 2000; Bobenrieth et al., 2014). As a matter of fact, however, statistical tests performed on these series much more often tend to accept the hypothesis of an unit root rather than of stationarity. In practice, commodity prices are expected to behave like I(0) while, in fact, empirical evidence mostly suggest they behave like I(1) series.

This apparent contradiction may have a twofold statistical explanation. On the one hand, commodity prices (especially the agricultural ones), either I(0) or I(1), always show a strong serial correlation. The main implication is that price at a given time *t* is strongly influenced by its lagged values, thus temporary price shocks may persist for a long period of time. This long memory is also called fractional integration and it may motivate why, low power unit roots may tend to designate these prices as I(1) series (Wei and Leuthold, 1998; Esposti and Listorti, 2013). A second possible explanation of evidence suggesting mean reversion and non-stationarity at the same time consists on the fact that the long-term equilibrium value towards which commodity prices tend to revert is not a constant value but it is itself a I(1) stochastic process (for instance, a random walk or a stochastic trend) (Valera and Lee, 2016). This dynamics of the long-term equilibrium price is the consequence of the permanent changes (or structural breaks) of the respective market fundamentals.

The combination of these conflicting properties eventually motivate the highly non-linear and, thus, often surprising and unpredictable behaviour of commodity price series and the consequent difficulties in disentangling the short-term (temporary) movements from the long-run dynamics (Gilbert, 1995). The recurrence of "*spikes, runs and bubbles*", i.e., of relatively short periods of rapid price increase then followed by an as much quick price drop is a typical feature of these markets and it has become even more frequent in the last decade. In some cases, at least part of these abrupt price changes remain also in the longer-term, that is, they turn to be permanent *jumps* or *structural breaks* (Brooks e Prokopczuk, 2013).

A further issue emerged in the last decade is that the abovementioned complex features of price series do not only concern their levels but also their variances or *volatility*. Prices' volatility tends to remain quite stable over long periods then followed by shorter periods of rapid increase (*volatility* *clusters*). Again, this increase often disappears in few months but sometime it may remain for longer periods and also become permanent. This is what allegedly occurred in many commodity markets in the last decade, at least starting from 2007 (Piot-Lepetit and M'Barek, 2011).

The aspect of major interest here, however, is a further regularity of these prices' dynamics. In fact, though this dynamics is complex, they seem to move together or to show a *common movement* (Brooks and Prokopczuk, 2013). The nature and possible explanations of this common movement is the main focus of the present work. Figure 1 juxtaposes the monthly price dynamics of 4 quite different commodities over a long period, i.e. from January 1980 to mid 2016 (see also section 4). Two are agricultural commodities but with quite different production processes, supply chains and uses: wheat and beef. The other two commodities are crude oil, with its wide set of possible uses but mostly related to energy demand, and copper that represents one of the most used material in a large set of manufacturing activities as well as in construction. Although a larger set of commodities could be considered, due to space limitations here we want to focus only on few very relevant and very heterogeneous commodities.

By computing simple statistics (sample averages and standard deviations) over three sub-periods (1980-1994, 1995-2004, 2005-2016) a common behaviour emerges. For all commodities the average price observed in the last decade is significantly (more than 50%) higher than previous periods when it remained quite stable. In fact, for all commodities the observed average price in the last sub-period exceeds the upper 95% confidence bound computed by adding up twice the standard deviation to the average of the previous periods. In other words, for all commodities the average level observed in the last decade can be hardly attributed to the same statistical distribution (i.e., stochastic process) of the previous periods (Grantham, 2011).

Not only the average level remarkably increased, but also the standard deviation has sharply grown during the last decade for all commodities. Therefore, both price level and volatilities strongly increased and at least part of this increase seems to be permanent and generated by a new underlying stochastic process. This change seemingly occurred for all commodities in the same way and at the same time. In Figure 1 this is evident for the 4 commodities here considered but this behaviour is actually shared by most of the about 50 commodities included in the IMF Primary Commodity Price Database with only few exceptions (Esposti, 2016). The explanation of this common behaviour, however, is methodological challenging.

There are two very different reasons why we observe prices' common movement and they have quite different methodological implications, i.e., in terms of how the underlying stochastic process is eventually represented. The first reason is that prices are interdependent: there is some, possibly reciprocal, causation relationship due to some real or financial linkage. For instance, the price increase of one commodity (e.g., crude oil or wheat) causes an increase of production costs of another commodity (e.g., beef). This interdependence takes the form of price transmission, as the shocks experienced by one price in its level or volatility are transmitted to other prices. As a consequence, price series are generated by interdependent stochastic processes and as such should be represented within empirical analysis.

The second reason is that prices are actually independent stochastic processes since there is no relevant causation relationship among them. Nonetheless, they share some common exogenous drivers and whenever these drivers experience a shock, this is transmitted to all commodity prices (Baffes e Haniotis, 2016). Therefore, even though no real or financial linkage among them actually occurs we still observe a common price dynamics. These common drivers might be either observable variables (e.g., the interest rate) or unobserved (i.e. latent) factors. These latter are, in fact, particularly interesting as they can be expression of generalized market fundamentals like a generalized long-term demand growth not compensated by long-term output growth (Gilbert, 1995).[2]

A detailed review of these possible underlying common market fundamentals is well beyond the scope of the present study. In a recently published paper, Le Mouël and Forslund (2017) review more than 200 scenario studies concerning the long-term dynamics of the long-term of agricultural commodity markets up to 2050. Most of these studies share market fundamentals that can be in fact generalized also to non-food commodities: permanent demand growth due to income and population growth and changes of consumer habits (for instance, reduction of wastes) and preferences (for instance dietary changes) due to higher income and urbanization, on the demand side; technological innovations (also related to recycling), climate change, natural resource limitations (for instance, limitations in land use extension and changes) and restrictions imposed by policies in this respect, on the supply side. Some of these common market fundamentals are actually mostly qualitative and, in fact, unobservable. Others are observable also at the global level (for instance, population and income growth) but not at the same high frequency at which commodity prices are observed and are here investigated (Esposti, 2016). For these reasons, these market fundamentals can not be entered as quantitative variables in the present analysis but are regarded as components of a one or more common latent factors.

In practice, the actual commodity price dynamics is likely a combination of these two stochastic processes as some form of price transmission can not be excluded in the short term and, at the same time, common latent drivers might operate especially in the longer run. Although the stochastic process taking both these two forms of common movement (interdependence and common drivers) into account represents a major methodological issue in terms of specification, identification and estimation, particularly when the common drivers are latent factors.

3. The modelling approach: state-space representation and the FAVAR model

From the discussion above it follows that the empirical investigation of prices' interdependence or common movement necessarily concerns balanced panel dataset where *N* commodity price series are observed over *T* periods (years/months/weeks). Firstly assume that the generic i-th price observed at the generic time *t* ($p_{i,t}$) follows an autoregressive (AR) process:

$$(1) \qquad p_{i,t} = \alpha_i + \sum_{s=1}^{s=S<T} \alpha_s p_{i,t-s} + \varepsilon_{it}, \forall i \in N, \forall t, s \in T \ ,$$

where α_i is a constant term (drift), α_s are autocorrelation coefficients and $\pounds_{it}$ is a disturbance term assumed to be i.i.d. Once these AR processes are estimated, usual tests can be performed on such estimates in order to assess the underlying stochastic processes and, in particular, their order of integration and the presence of conditional heteroskedatisticy (ARCH). However, if there is some common movement across these prices, this representation of the underlying stochastic process is incomplete.

The presence of such common movement could be expressed by correlation across estimated error terms, that is, $E\left(\pounds_{it,} \pounds_{jt}\right) \neq 0, \forall i, j \in N, \forall t \in T$ (Chen et al., 2014). No correlation evidently provides a strong evidence against the common movement hypothesis. However, cross-price correlation does not tell anything about the form and, above all, the causes of this common movement. As mentioned, the first hypothesis is that common movement is actually the consequence of price interdependence, or price transmission. Therefore, the stochastic process generating any single price should be rather represented as:

[2] "Were demand [for food] to being to outpace supply, the inevitable consequence would be an increase in [food] prices" (Godfray, 2015, p. 201),

(2) $$p_{i,t} = \alpha_i + \sum_{s=1}^{s=S<T} \sum_{i,j=1}^{N} \alpha_{js} p_{j,t-s} + \varepsilon_{it}, \forall i,j \in N, \forall t,s \in T$$

As prices are reciprocally interdependent, the actual stochastic process generating price series has to be represented in a vector form, i.e., as a VAR process:

(3) $\mathbf{p}_t = \Phi(L)\mathbf{p}_{t-1} + \boldsymbol{\varepsilon}_t$

where $\mathbf{p}_t$ is the Nx1 vector of commodity prices and $\boldsymbol{\varepsilon}_t$ the Nx1 vector of the i.i.d. disturbance terms.

The second hypothesis is that of common drivers. Within representation in (3), they behave as exogenous possibly lagged variables affecting price dynamics in addition to their interdependence. Assume there exist *M* common drivers. Any single price formation can be represented as follows:

(3) $$p_{i,t} = \alpha_i + \sum_{i,j=1}^{N} \sum_{s=1}^{s=S<T} \alpha_{js} p_{j,t-s} + \sum_{k=1}^{M} \sum_{s=0}^{s=S<T} \delta_{iks} z_{kt-s} + \varepsilon_{it}, \forall i,j \in N, \forall t,s \in T, \forall k \in M$$

where z_{kt} indicates the value of the k-th generic common driver at time t and δ_{iks} are coefficients expressing the lonkage between p_i and z_k. In a more compact matrix form, (3) can be written as VAR model with exogenous variables (VARX):

(4) $\mathbf{p}_t = \Phi(L)\mathbf{p}_{t-1} + \Theta(L)\mathbf{z}_t + \mathbf{G}\boldsymbol{\varepsilon}_t$

where $\mathbf{z}$ indicates the vector of common factors. The NxN $\mathbf{G}$ matrix is usually assumed to be an identity matrix. However, if a more complex interdependence has to be modelled and, in particular, interdependence also occurs in volatilities (*volatility spillovers*), $\mathbf{G}$ can be specified in order to admit and model such kind of interdependence. MGARCH models are aimed to provide such representation though they concentrates only on volatility transmission. A combination of the two (VAR-MGARCH models) can be also specified though their consistent estimation may be challenging (Carnero and Eratalay, 2014; Maekawa and Setiawan, 2014; Ohashi and Okimoto, 2016).

Unfortunately, these common factors might be unobservable. Not only for the lack of data (at least at the same high frequency at which we observe commodity prices), but also because these factors are actually the outcome of complex and multiple phenomena that can be hardly expressed by single observed variables.[3] Not only they can not be observed, but also their stochastic processes are unknown and must be somehow assumed *ex ante*.

One possible way to deal with the presence of latent factors is to treat them as state variables and represent the stochastic process expressing price formation in a state-space form (Schwartz, 1997). Following previous works, let's assume two latent state variables expressing two complementary market fundamentals or market imbalances (Gilbert, 1995). The first, z_1, represents the long-term equilibrium level towards which prices tend to revert. The second, z_2, represents the short-run deviations from this equilibrium. Here we assume that z_{1t} follows a Brownian motion process and, in particular, a random walk, while z_{2i} follows a zero-mean Ornstein-Uhlenbeck process, i.e., a stationary AR process without a drift (Schwartz and Smith 2000; Sørensen 2002).

According to this representation, commodity price dynamics presents short-run deviations mostly resulting from temporary changes of supply (e.g., extreme weather conditions). At the same time, in the longer run prices tend to revert to their long-term mean which is, however, not constant and it behaves like an I(1) process due to persistent shocks in the market underlying fundamentals (i.e., in supply and demand). As a consequence, price themselves eventually behave as I(1) since one of their drivers is a random walk (Valera and Lee, 2016).

[3] For instance, market fundamentals, such as long-term demand (or supply) growth or excess, may be easy concepts but they can be hardly measured with a single variable (Gilbert, 1995).

Model (4) thus combines the price interdependence and the common drivers hypotheses. Its state-space representation consists, as usual, by a *transition* (or *state*) *equation* and a *measurement* (or *observation*) *equation.* The transition equation describes the stochastic dynamics of the (unobserved) state variables. (Schwartz and Smith, 2000; Sørensen, 2002). The measurement equation relates observables (i.e. commodity prices) among them and to the state variables according to (4). In compact matrix notation, the state-space representation of the model here adopted is thus the following:

(5)
$$\mathbf{Z}_t = \mathbf{A}\mathbf{Z}_{t-1} + \mathbf{C}\nu_t$$
$$\mathbf{Y}_t = \sum_s \mathbf{B}_s \mathbf{Y}_{t-s} + \mathbf{D}\mathbf{Z}_t + \mathbf{G}\varepsilon_t$$

where **A**, $\mathbf{B}_S$ and **D** are matrices of unknown coefficients to be estimated while **C** and **G** are coefficient matrices allowing for a more complex specification of disturbance terms' cross-correlation, including the presence of a MGARCH structure in the measurement equation.
As in large part of this literature (Gilbert, 1995; Esposti and Listorti, 2013), the logarithm of commodity prices is actually considered and all estimated models are thus specified in a log-linear form. Though more than 50 different commodity prices are available and the main advantage of the FAVAR approach more clearly emerges when many observed series are included (Bernanke et al, 2005), here the approach is applied to only 4 commodities: crude oil, copper, wheat, beef. This reduces the computationally complexity implied by the adopted estimation approach and also facilitates the interpretation of results. At the same time, this small group of variables should still represent the whole set of different production processes and uses, i.e. the wide heterogeneity, across these commodity markets.
According to the discussion above, model (5) takes the following extensive form:

Transition/State equation:

(6a)
$$\begin{pmatrix} z_{1t} \\ z_{2t} \end{pmatrix} = \begin{pmatrix} c_L \\ 0 \end{pmatrix} + \begin{pmatrix} 1 & 0 \\ 0 & \delta_1 \end{pmatrix}\begin{pmatrix} z_{1t-1} \\ z_{2t-1} \end{pmatrix} + \begin{pmatrix} 0 & 0 \\ 0 & \delta_2 \end{pmatrix}\begin{pmatrix} z_{1t-2} \\ z_{2t-2} \end{pmatrix} + \underset{(2x2)}{\mathbf{I}}\begin{pmatrix} \nu_{1t} \\ \nu_{12} \end{pmatrix}$$

Measurement/Observation equation:

(6b)
$$\begin{pmatrix} cr_t \\ cp_t \\ wh_t \\ be_t \end{pmatrix} = \begin{pmatrix} c_{cr} \\ c_{cp} \\ c_{wh} \\ c_{be} \end{pmatrix} + \begin{pmatrix} \gamma_{1cr} & \gamma_{2cr} \\ \gamma_{1cp} & \gamma_{2cp} \\ \gamma_{1wh} & \gamma_{2wh} \\ \gamma_{1be} & \gamma_{2be} \end{pmatrix}\begin{pmatrix} z_{1t} \\ z_{2t} \end{pmatrix} + \begin{pmatrix} \alpha_{11} & \alpha_{12} & \alpha_{13} & \alpha_{14} \\ \alpha_{21} & \alpha_{22} & \alpha_{23} & \alpha_{24} \\ \alpha_{31} & \alpha_{32} & \alpha_{33} & \alpha_{34} \\ \alpha_{41} & \alpha_{42} & \alpha_{43} & \alpha_{44} \end{pmatrix}\begin{pmatrix} cr_{t-1} \\ cp_{t-1} \\ wh_{t-1} \\ be_{t-1} \end{pmatrix}$$
$$+ \begin{pmatrix} \beta_{11} & \beta_{12} & \beta_{13} & \beta_{14} \\ \beta_{21} & \beta_{22} & \beta_{23} & \beta_{24} \\ \beta_{31} & \beta_{32} & \beta_{33} & \beta_{34} \\ \beta_{41} & \beta_{42} & \beta_{43} & \beta_{44} \end{pmatrix}\begin{pmatrix} cr_{t-2} \\ cp_{t-2} \\ wh_{t-2} \\ be_{t-2} \end{pmatrix} + \underset{(4x4)}{\mathbf{G}}\begin{pmatrix} \varepsilon_{1t} \\ \varepsilon_{2t} \\ \varepsilon_{3t} \\ \varepsilon_{4t} \end{pmatrix}$$

where cr_t, cp_t, wh_t and be_t indicate the logarithm of crude oil, copper, wheat and beef price, respectively. $c_{.}$, $\delta_{..}$, $\gamma_{..}$, $\alpha_{..}$ and $\beta_{..}$ are unknown parameters to b estimated.
For the presence of autoregressive unobserved factors, this kind of models is also called Dynamic-Factor (DF) model and, whenever the observation equation takes itself an autoregressive form (VAR), it has been more recently designated as Factor Augmeneted VAR (FAVAR) model (Bernanke et al, 2005).

Due to the linear specification, model (6a,b) can be estimated with Maximum Likelihood Estimation (MLE) using the diffused Kalman Filter to obtain the prediction error form of the log-likelihood function (Harvey, 1989).[4] This is feasible an i.i.d. normal distribution is assumed for both ν_{kt} and ε_{it}. While this assumption is here maintained for ν_{kt}, the presence of volatility interdependence would suggest a more articulated specification of $\mathbf{G}$ ($\mathbf{G} \neq \mathbf{I}$). In particular, this matrix could be specified in order to have a MGARCH structure. This FAVAR-MGARCH specification might capture most of the complex common movements across commodity prices, its estimation via the Kalman Filter MLE is unfeasible due to the non-linearity implied by the MGARCH structure and for the highly computational complexity involved (Carnero and Eratalay, 2014; Maekawa and Setiawan, 2014). Nonetheless, a two-step estimation procedure is here followed: a consistent estimation of $\mathbf{G}$ ($\hat{\mathbf{G}}$) is firstly obtained with a conventional MGARCH model and, then, $\hat{\mathbf{G}}$ is entered within model (6a,b) to perform the Kalman Filter MLE estimation.

4. The empirical analysis

Data here used are the monthly commodity prices (in US$) reported in the IMF Primary Commodity Price Database. The series cover the period from January 1980 to April 2016. Therefore, for any price series 436 observations are available. The empirical investigation initially focuses on the univariate analysis aiming to identify the stochastic properties of the 4 price series and their common features in this respect. Then, the multivariate analysis is introduced by firstly estimating conventional models expressing price interdependence in levels and volatility (VAR/VECM and MGARCH models). Then the estimation of the DF/FAVAR model (6a,b) is reported and discussed.

4.1. Univariate analysis

The upper part of Table 1 reports the statistical tests performed to identify the time series properties of the commodity prices. Unit roots, conditional heteroskedasticity (ARCH effects) and persistence (long memory or fractional integration) are tested in sequence. It emerges that all prices show the same stochastic features. First of all, once the proper specification has been selected (number of lags, and of the presence of drift and deterministic trend), all price series show a unit root.[5]
Price series also presents similarities in terms of long memory or fractional cointegration. In such cases price series are neither I(0) nor I(1) but rather I(d) processes, with $0<d<1$ (Wei and Leuthold, 1998). Fractional integration implies that price series, though not behaving as random walks, still keep the memory of a shock for a long period.[6] Test results indicate that stationarity can be excluded in all cases while the I(1) hypothesis actually show quite low p-values with two

[4] The Kalman filter is a recursive procedure for computing estimates of unobserved state variables based on observations that depend on these state variables. Given a prior distribution on the initial value of the state variables and a model describing the likelihood of the observations as a function of the true values, the Kalman filter generates updated posterior distributions for these state variables in accordance with the Bayes' rule.

[5] The ADF test is reported in Table 1. However, also KPPS and PP tests are performed. The former assumes stationarity as the null hypothesis while the latter is expected to be more robust under heteroskedasticity. Results of these tests fully correspond with what obtained with ADF tests and are available upon request.

[6] Fractional integration is here tested following the approach originally proposed by Geweke and Porter-Hudak (1983) and then modified by Phillips (1999a,b). This test is based on a particular representation of the stochastic process generating the price series called ARFIMA(p,d,q) (Autoregressive Fractionally Integrated Moving Average) model, where p and q express, as usual, the orders of auto-regressive and the moving-average parts, respectively, and d the order of (fractional) integration. The procedure proposed by Phillips (1999a,b), and here adopted, tests the value of parameter d thus distinguishing stationary, unit-root and fractionally integrated processes. This procedure produces two test statistics, one for the null d=0 and one for d=1. If d=0 is accepted the series is stationary; if d=1 is accepted the series has an unit root. If both are rejected (namely, $0<d<1$), then fractional integration (long memory) is accepted.

commodities (copper and wheat) for which it is lower than 10%. All series thus seem to move according to a process that is very close to be either an I(1) process or a mean-reverting series where the effects of one-time shock takes a very long time to vanish.

At the same time, for no price series the presence of conditional heteroskedasticity can be excluded thus suggesting the generalized presence of volatility clusters, that is, periods characterized by higher price variances.

The fact that all 4 price series show very similar stochastic processes, however, does not necessarily mean they follow some common movement. To provide an initial evidence supporting this hypothesis, Table 1 also reports the pair-wise correlation coefficients of estimated residuals of the 4 ADF unit-root test equations. For all pairs of commodity prices we observe a significant and positive correlation coefficient. This suggests that whenever the individual autoregressive stochastic processes are taken into account, there remains a residual part of price dynamics that shows some commonality across series.

4.2. Multivariate analysis

The lower part of Table 1 reports results concerning the multivariate analysis of price dynamics. These results highlight some relevant features of the alleged common movement among the 4 commodity prices. First of all, the trace cointegration test would indicate that these commodity prices are cointegrated but, in fact, they may move along two different cointegration vectors. Consequently, cointegration by itself does not guarantee an univocal economic interpretation of price linkages. Even if the consider only the first extracted cointegration relationship, the consequent VECM[7] estimation confirms that though a long-run relationship among prices can be identified its interpretation is not straightforward and, somehow, counterintuitive. First of all, the sign of the cointegration vector coefficients would suggest that copper moves in the opposite direction compared to other commodities while, in fact, they are expected to move in the same direction.

Secondly, all adjustment parameters show the correct sign and are statistical significant. Therefore, no price behaves as an exogenous driver of all the others. All prices respond to the others and this is not easily interpretable especially when crude oil, the main candidate to be the driving price, is considered. The Granger causality tests make the picture emerging from this VECM estimation even more puzzling. Crude oil and wheat prices are caused by all other prices, while cooper price is independent from beef price (as could be expected) and beef price only depends on wheat price.

If we admit that price common movement not only concerns their levels but also the respective volatilities, the complete picture we obtain is a further confirmation of the multifaceted and hardly interpretable interdependence among prices. DCC-MGARCH[8] estimates reported in Table 1 suggest that price volatility varies over time for all series, as indicated by the adjustment parameters, but this movement is common only for crude oil and copper, while this volatility transmission is not confirmed in the case of the two agricultural commodities.

The conventional multivariate analysis supports the hypothesis of some common movement among these commodity prices. It is less evident, however, whether this commonality is the consequence of some clear and economic meaningful interdependence, thus of shocks transmission. No univocal and unambiguous causal direction can be identified. As discussed, alternative to the price interdependence hypothesis, the common latent factor hypothesis does not imply any direct linkage among commodity prices. In such circumstance, price interdependence is the spurious outcome of the fact that prices actually respond to the same exogenous signals.

The common latent factor hypothesis is reflected in the DF/FAVAR model (6a,b) whose estimation is reported in Table 2. As discussed in section 3, the **G** matrix is specified according to the

[7] Vector Error Correction Model.

[8] Dynamic Conditional Correlation Multivariate Generalized AutoRegressive Conditional Heteroskedasticity.

abovementioned DCC-MGARCH estimation in order to take volatility spillover into account. Estimates confirm that all prices are significantly driven by the long-term latent factor (z_1) and the sign confirms the expectation as all prices move in concordance with the variations of the factor. For the short-term latent factor (z_2), on the contrary, only one price (beef) shows a significant relationship and this seems fully consistent with the fact that the autoregressive structure of the VAR should already take short-term responses into account and z_2 is expected to capture only the unexplained part of this variation.
The most interesting outcome of this approach, however, is the estimation of the latent factors themselves. They are displayed in Figure 2.[9] Estimated patterns are fully consistent with latent variables' construction in equation (6a). z_2 behaves as a zero-mean stationary autoregressive process. Nonetheless, the last decade signals larger deviations from the mean value as a consequence of the increased price volatility. z_1 is consistent with a random walk with drift and it is able to capture what can be intended as the varying long-term equilibrium price with the respective jumps (or structural breaks). It emerges that such long-term price remains quite stable until late eighties then it jumped to higher values in early nineties to drop again to the initial values in late nineties. The most relevant change, however, occurred in the early part of years 2000 when this long-term equilibrium price reached a much higher value that is then maintained, though with a significant variability, for the rest of the observed period.

5. Concluding remarks

This paper focuses on the common movement of commodity prices. Understanding where this common movement comes from is empirically challenging. Most literature concentrates on the possible interdependence among prices and, above all, on price transmission in both levels and volatility as a consequence of some real or financial causal linkage. Nonetheless, finding support to this interdependence hypothesis is complex particularly when heterogeneous and apparently unrelated commodities are considered, as in the present case.
Therefore, the main objective of the present work is to provide empirical support to an alternative hypothesis, that of common latent factors. Results suggest that a latent factor expressing the non-stationary long-term equilibrium price can be identified and its estimation indicates that, beside the increased price volatility, the last decade experienced a significant and permanent rise of the long-term price levels.
The implications of this result are remarkable but evidently requires further research efforts. From an economic perspective, the key implication concerns the possible theoretical motivation underlying this latent factor. As pointed out by Gilbert (1995), it is expected to capture all long-term market imbalances and, more importantly, those processes and factors that generate these imbalances in all these commodity markets. Therefore, supply and demand structural changes should be involved in order to elaborate a model whose empirical counterpart is the estimated FAVAR specification.
From a policy perspective, this result opens the very relevant issue of whether and to what extent this rising long-term equilibrium commodity prices can be interpreted as the macroscopic evidence of a major change in the long-term perspective about these markets and the underlying natural resources endowment. The main idea would be that, mostly depending on the intense demand growth coming from emerging economies, the days of abundant resources are over and we are entering a new era of shortage and raising prices (Grantham, 2011). In terms of policy action, the main implication of this result is the urgent need of measures favouring a generalised higher supply growth without inducing any additional and unsustainable pressure on the global environmental resources. In this respect, especially in food production, that of *sustainable intensification* has

[9] Figure reports estimates starting from 1981:1 as the first 12 observations (months) are used to initialize and stabilize the Kalman Filter estimation of the latent factors.

become a much debated strategy aiming at improving natural resources' yields while reducing the negative impact of the respective production activities on the environment (Godfray, 2105). Whether this strategy is actually feasible for all sectors here involved (from agriculture, to mining and energy sector) can be seriously questioned. The policy implication of any argument against the sustainable intensification strategy, however, is perhaps more problematic as it would imply even more urgent actions to reduce demand growth rate and/or to accelerate all processes of replacement, substitution and recycling within natural resourced-based production processes.
Needless to say, therefore, finding evidence supporting the idea that we are entering a new era of scarcity is of great social and political relevance. For this same reason, however, any result in this respect has to be interpreted with caution and should require further confirmations and sounder theoretical justifications.

References

Baffes, J. and Haniotis, T. (2016) What explains agricultural price movements? *Journal of Agricultural Economics* 67(3): 706–721.

Bai, J. 2013. Panel Data Model and Factor Analysis. In D. Acemoglu, M. Arellano and D. Eddie (eds.), *Advances in Economics and Econometrics. Volume III: Econometrics*, New York: Cambridge University Press, 437-484.

Bernake, B.S., Boivin, J. and Eliasz, P. (2005). Measuring the effects of monetary policy: a factor-augmented vector autoregressive (FAVAR) approach. *The Quarterly Journal of Economics* 120(1): 387-422.

Bobenrieth, E. S. A., Bobenrieth, R. A. and Wright, B. D. (2014). Bubble Troubles? Rational Storage, Mean Reversion, and Runs in Commodity Prices. In: J-P. Chavas, D. Hummels, B.D. Wright (eds.), The Economics of Food Price Volatility. University of Chicago Press, Chicago.

Brooks, C. and Prokopczuk, M. (2013). The Dynamics of Commodity Prices. *Quantitative Finance* 13(4): 527-542.

Bruno, V.G., Büyükşahin, B., Robe, M. A. (2017). The Financialization of Food? *American Journal of Agricultural Economics*, 99(1): 243-264.

Carnero, M.A. and Eratalay, M.H. (2014). Estimating VAR-MGARCH models in multiple steps. *Studies in Nonlinear Dynamics & Econometrics* 18(3): 339-365.

Chen, S-L., Jackson, J.D., Kim, H. and Resiandini, P. (2014). What Drives Commodity Prices. *American Journal of Agricultural Economics* 96(5): 1455-1468.

Esposti, R. (2016). I prezzi delle commodities si muovono davvero insieme' E, soprattutto, perché? *ARE – AgriRegioniEuropa* 12(47): 22-27 http://agriregionieuropa.univpm.it/

Esposti, R. and Listorti, G. 2013. Agricultural price transmission across space and commodities during price bubbles. *Agricultural Economics* 44: 125–139.

Geweke, J. and Porter-Hudak, S. (1983). The Estimation and Application of Long Memory Time Series Models, *Journal of Time Series Analysis* 4(4): 221-238.

Gilbert, C.L. (1995). Modelling market fundamentals: a model of aluminum market. *Journal of Applied Econometrics* 10: 385-410.

Godfray, H.C.J. (2015). The debate over sustainable intensification. *Food Security* 7: 199-2008.

Grantham, J. (2011). Time to Wake Up: Days of Abundant Resources and Falling Prices Are Over Forever. GMO Quarterly Letter, April, Boston.

Kim, T-H., Pfaffenzeller, S., Rayner, T. and Newbold, P. (2003). Testing for linear trend with application to relative primary commodity prices. *Journal of Time Series Analysis* 24(5): 539-551.

Le Mouël, C., Forslund, A. (2017). How can we feed the world in 2050? A review of the responses from global scenario studies. *European Review of Agricultural Economics* (forthcoming).

Listorti, G. and Esposti, R. (2012). Horizontal price transmission in agricultural markets: fundamental concepts and open empirical issues. *Bio-based and Applied Economics* 1(1): 81-108.

Maekawa, K. and Setiawan, K. (2014). Estimation of Vector Error Correction Model with GARCH Errors: Monte Carlo Simulation and Application. Paper Presented at the International Conference on Economic Modelling – ECOMOD 2014, July 16-18, Bali.

Martin, V., Hurn, S. and Harris, D. (2013). *Latent Factor Models*. Econometric Modelling with Time Series. Cambridge University Press, New York.

Ohashi, K. and Okimoto, T. (2016). Increasing Trends in the Excess Comovement of Commodity Prices. CAMA Working Paper 9/2016, Crawford School of Public Policy, Australian National University, Acton.

Phillips, P.C.B. (1999a). Discrete Fourier Transforms of Fractional Processes. Cowles Foundation for Research in Economics, Discussion Paper No. 1243, Yale University.

Phillips, P.C.B. (1999b). Unit Root Log Periodogram Regression. Cowles Foundation for Research in Economics, Discussion Paper No. 1244, Yale University.

Piot-Lepetit I. and M'Barek, R. (eds.) (2011). *Methods to Analyse Agricultural Commodity Price Volatility*. Berlin: Springer Verlag.

Routledge, B.R., Seppi, D.J. and Spatt, C.S. (2000). Equlibrium Forward Cruves for Commodities. *The Journal of Finance* 55(3): 1297-1338.

Schleicher, C. (2003). Structural Time Series Models with Common Trends and Common Cycles. Computing in Economics and Finance n. 108, Society for Computational Economics.

Schwartz, E.S. (1997). The Stochastic Behavior of Commodity Prices: Implications for Valuation and Hedging. *The Journal of Finance* 52(3): 923-973.

Schwartz, E.S. and Smith, J.E. (2000). Short-Term Variations and Long-Term Dynamics in Commodity Prices. *Management Science* 46(7): 893-911.

Sørensen, C. (2002). Modelling Seasonality in Agricultural Commodity Futures. *The Journal of Future Markets* 22(5): 393-426.

Valera, H.G.A. and Lee, J. 2016. Do rice prices follow a random walk? Evidence from Markov switching unit root tests for Asian markets. *Agricultural Economics* 47: 683-695.

Wei A. and Leuthold, R.M. (1998). Long Agricultural Future Prices: ARCH, Long Memory or Chaos Processes? OFOR Paper n. 98-03, Department of Agricultural Economics, University of Illinois at Urbana-Champaign.

Table 1 – Univariate and multivariate analysis on the 4 commodity prices

Univariate Analysis

	ADF[a]	ARCH[b]	Fractional integration[c]	
			t (H_0: d=0)	z (H_0: d=1)
Crude oil	-2.258	79.034*	0.000	0.152
Cooper	-2.982	83.997*	0.000	0.068
Wheat	-2.971	43.018*	0.000	0.089
Beef	-2.390	36.042*	0.000	0.126

Correlation coefficient of estimated residuals of ADF unit-root test equations

	Crude oil	Cooper	Wheat	Beef
Crude oil	1.000			
Cooper	0.328*	1.000		
Wheat	0.187*	0.218*	1.000	
Beef	0.145*	0.103*	0.174*	1.000

Multivariate Analysis

Cointegration[d]						MGARCH[e]	
Rank	Trace statistic		Vector	Adjustment	Short-run Granger causality test	Correlation	Adjustment
0	63.32	Crude oil	1	-0.027*	Copper: 14.74* Wheat: 29.71* Beef: 13.40*	Copper: 0.319* Wheat: 0.014 Beef: 0.076	$\lambda_1 = 0.011$
1	47.66	Cooper	0.549	-0.018*	Crude oil: 12.25* Wheat: 23.71* Beef: 4.48	Wheat: 0.106 Beef: 0.040	$\lambda_2 = 0.956$*
2	19.11*	Wheat	-4.379*	-0.023*	Crude oil: 11.49* Copper: 20.97* Beef: 12.22*	Wheat: -0.038	
		Beef	-0.653*	-0.006*	Crude oil: 3.91 Beef: 10.62 Wheat: 22.92*		

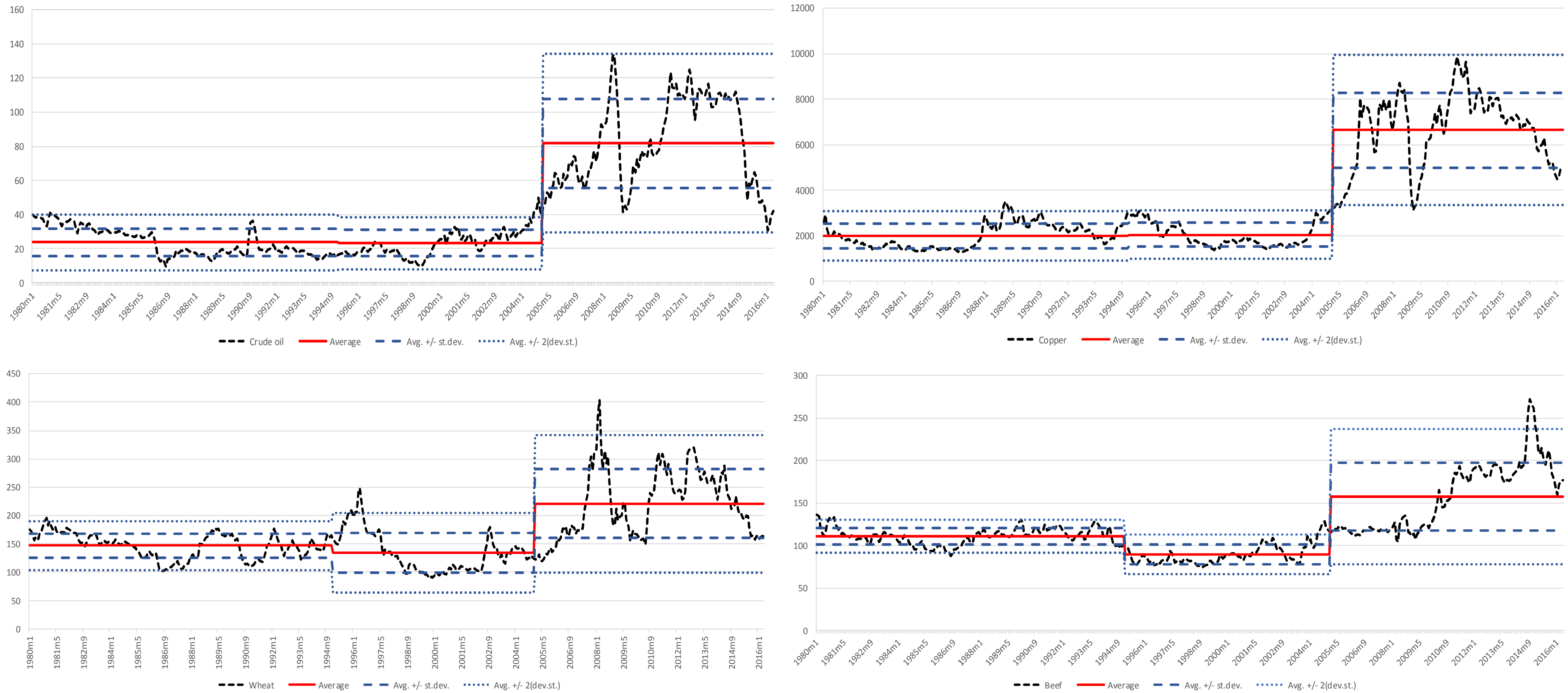
[a] *Augmented Dickey Fuller (ADF) unit-root test with 12 lags and deterministic trend*
[b] *Lagrange Multiplier (LM) test performed on the residuals of the ADF unit-root test equations and 12 lags*
[c] *Test of fractional integration according to Phillips (1999a,b); p-values are reported*
[d] *The VECM model is estimated on the first cointegration vector extracted with 6 lags and a deterministic trend within the cointegration vector.*
[e] *DCC-MGARCH (1,1) model*
**Statistically significant at 5% level*

Table 2 – FAVAR model (6a,b) estimates (estimated standard error in parenthesis)

Equation Coefficient	Estimate		Equation Coefficient	Estimate	
Factor 1 (z_1)			*Wheat* (wh)		
c_1	1.983*	(0.014)	c_{wh}	160.67*	(63.06)
Factor 2 (z_2)			¥$_{1wh}$	0.245*	(0.086)
¢$_1$	0.759*	(0.256)	¥$_{2wh}$	-0.480	(0.307)
¢$_2$	0.224	(0.414)	~$_{31}$	0.629*	(0.248)
Crude oil (cr)			$_{31}$	-0.709*	(0.234)
c_{cr}	135.85*	(50.49)	~$_{32}$	0.008*	(0.004)
¥$_{1cr}$	0.641*	(0.041)	$_{32}$	0.003	(0.003)
¥$_{2cr}$	0.199	(0.264)	~$_{33}$	0.441*	(0.112)
~$_{11}$	0.742*	(0.170)	$_{33}$	-0.239	(0.122)
$_{11}$	-0.373*	(0.163)	~$_{34}$	0.588*	(0.199)
~$_{12}$	0.003*	(0.001)	$_{34}$	-0.025	(0.176)
$_{12}$	0.002	(0.001)	*Beef* (be)		
~$_{13}$	0.017	(0.023)	c_{be}	57.75	(63.10).
$_{13}$	-0.059	(0.032)	¥$_{1be}$	0.256*	(0.121)
~$_{14}$	0.238*	(0.050)	¥$_{2be}$	-0.633*	(0.351)
$_{14}$	-0.108	(0.059)	~$_{41}$	0.212	(0.138)
Copper (cp)			$_{41}$	-0.318*	(0.103)
c_{cp}	359.81*	(126.1)	~$_{42}$	0.002	(0.003)
¥$_{1cp}$	0.684*	(0.103)	$_{42}$	0.001	(0.002)
¥$_{2cp}$	0.993	(0.949)	~$_{43}$	0.110*	(0.051)
~$_{21}$	0.128*	(0.027)	$_{43}$	-0.084	(0.091)
$_{21}$	-0.068	(0.088)	~$_{44}$	0.089	(0.048)
~$_{22}$	0.659*	(0.172)	$_{44}$	0.168	(0.148)
$_{22}$	0.243	(0.189)			
~$_{23}$	0.228	(0.129)			
$_{23}$	-0.201	(0.127)			
~$_{24}$	0.364	(0.313)			
$_{24}$	0.309	(0.290)			

**Statistically significant at 5% level*

Figure 1 – Monthly international prices of the 4 commodities (in US$) from January 1980 (1980:1) to April 2016 (2016:4): sub-period statistics (1980m1-1994m12; 1995m1-2004m12; 2005m1-2016m4)



Source: IMF

Figure 2 – Estimated Long (z_1) and Short-term (z_2) dynamic latent factors (1981m1=100)