This item was submitted to [Loughborough's Research Repository](#) by the author.
Items in Figshare are protected by copyright, with all rights reserved, unless otherwise indicated.

What drives commodity returns? Market, sector or idiosyncratic factors?

PLEASE CITE THE PUBLISHED VERSION

<https://doi.org/10.1111/obes.12334>

PUBLISHER

Wiley

VERSION

AM (Accepted Manuscript)

PUBLISHER STATEMENT

This is the peer reviewed version of the following article: MA, J., VIVIAN, A.J. and WO HAR, M.E., 2019. What drives commodity returns? Market, sector or idiosyncratic factors? Oxford Bulletin of Economics and Statistics, 82 (2), pp.311-330, which has been published in final form at <https://doi.org/10.1111/obes.12334>. This article may be used for non-commercial purposes in accordance with Wiley Terms and Conditions for Use of Self-Archived Versions.

LICENCE

CC BY-NC-ND 4.0

REPOSITORY RECORD

Ma, J., Andrew Vivian, and Mark E. Wohar. 2019. "What Drives Commodity Returns? Market, Sector or Idiosyncratic Factors?". Loughborough University. <https://hdl.handle.net/2134/38159>.

What Drives Commodity Returns? Market, Sector or Idiosyncratic Factors?

Abstract

This paper examines the relationship between 43 commodity returns using a dynamic factor model with time varying stochastic volatility. The dynamic factor model decomposes each commodity return into a common (market), sector-specific and commodity-specific component. It enables the variance attributed to each component to be estimated at each point in time. We find the return variation explained by the common factor has increased substantially for the recent period and is statistically significant for the vast majority of commodities since 2004 (at each point in time). This phenomenon is strongest for non-perishable products. We link the amount of variation explained by the common factor to economic variables.

(JEL C12, C32, E20, G12, G15)

Keywords: dynamic factor model; commodity markets; common variation; financialization

What Drives Commodity Returns? Market, Sector or Idiosyncratic Factors?

1. Introduction

Economic shocks have an impact that often go far beyond the immediate asset, product or country from which they emanate. For example, the impact of economic shocks will be greater if commodity markets are integrated and share a common component. The extent to which commodity returns are globally determined has received relatively little attention and little rigorous analysis has been conducted. This topic is of great importance though, given that strong links between markets could result in global shocks having a dramatic impact upon commodities as witnessed during the recent financial crisis of 2008-2010. Market-wide transmission of shocks also have policy implications for commodity producers and extractors who could find that diversification benefits are not as large as anticipated. While a large literature examines commodity price co-movements (building on Pindyck and Rotemberg, 1990), a much smaller body of literature examines commodity returns and whether these are correlated with each other.¹

In stark contrast to the literature on commodity price co-movement, studies using commodity returns generally suggest that the market is highly segmented and any common component is very small (see for example Erb and Harvey, 2006). However, there are numerous channels which could lead to common movement in commodity returns. For example, Hong and Yogo (2012) provide a model and evidence that futures open interest predicts commodity returns and other asset returns. Further, since the turn of the millennium there have been dramatic

¹ There is a large literature examining international co-movements of commodity prices that builds on the seminal paper of Pindyck and Rotemberg (1990). This area of inquiry tends to focus upon the economic determinants of co-movement and tries to resolve the puzzle that commodity price co-movement is much greater than can be explained by economic fundamentals. Saadi (2011) provides a review of commodity price co-movement in international markets. Despite a vast literature, the empirical evidence has not reached a consensus on the important variables in determining commodity co-movements. Not surprisingly, each candidate variable (such as aggregate demand, inflation, interest rates and exchange rates) influences the movement of commodity prices during at least one period, some more in one period than other periods.

developments within commodity markets. In particular, there has been an influx of speculative investors who view commodities purely as financial assets; there has been rapid growth in the money under management in commodity investment funds. Such investment inflows can induce return co-movement across a wide range of commodities (Basak and Pavlova, 2016). Hence individual commodity returns can be at least partly determined by agents who respond to factors other than those facing the commodity producer or commodity consumer (Gorton and Rouwenhorst, 2006; Tang and Xiong, 2012). There is also some evidence of commonality in liquidity across commodities which would induce return co-movements via both the demand channel (co-ordinated buying and selling by institutional investors) and the supply channel (liquidity withdrawal during large market declines) [see Marshall et al., 2013]. Hence, it is plausible that the extent of co-movement has changed substantially recently and that it also exhibits variation over time.

The main objective of this paper is to model the dynamic behaviour of commodity returns in the context of a sample of 43 commodities drawn from eight sectors over the period January 1984-June 2017.² Specifically, our dynamic factor model decomposes commodity returns into three components: a common (or market) factor, a sectoral factors and a commodity-specific factor. Note that the common factor and sectoral factor are orthogonal (in the population). Importantly, the common factor is the common empirical factor across these 43 commodity returns; thus it is not based on value-weighted commodity indices, such as S&P GSCI, and hence is not driven by the largest component (WTI and Brent Crude) or segment (Energy);³ this is in contrast to much literature where the commodity market factor is based on a value-weighted index

² Please note this model requires the series to be stationary and hence we cannot apply this approach to commodity prices.
³ The S&P GSCI index is driven by Energy sector commodities which typically comprise at least 60% of the overall index with Crude Oil (WTI and Brent Crude) typically accounting for at least 40% of the overall index.

(see e.g. Tang and Xiong, 2012).

Our innovative development of a dynamic factor model which extends those implemented in other contexts (e.g. Stock and Watson, 2002; Neely and Rapach, 2011) enables variance contributions of each component of commodity return to be estimated at each point in time and statistical significance assessed. These time varying factor contributions are estimated in the dynamic factor model by allowing for time-varying stochastic volatility. Therefore, a key contribution of the paper is to estimate at each point in time how much of the variance of the overall return is attributable to each factor and establish confidence intervals. For example, if the common factor becomes more important for the overall return variance then this indicates that commodities become a less effective tool for diversification; this insight would be of interest both to investment managers seeking to diversify portfolios of financial assets and also to commodity producers considering diversifying their production; please note that an increase in the common factor's variance contribution will be linked to a decrease in the variance contribution of the sector factor or the individual factor or both. The flexibility of our model specification does not require us to impose any abrupt structural breakpoint, which is often estimated with a large amount of uncertainty (see for example Eo and Morley (2015)). Rather, our flexible approach allows data to speak for the potential time-varying co-movement among commodity returns. Since the common and sector factors are orthogonal in the population, our confidence intervals are relatively narrow; hence our analysis is unlikely to be subject to the weak identification issue reported in prior variance decomposition applications to exchange rates (Balke et al., 2013) or valuation ratios (Ma and Wohar, 2014).

Our sample includes a wide cross-section of 43 commodities, which covers sectors from Industrial Metals to Cereals and from Energy to Raw Materials. This enables us to examine the

role of market-wide and sectoral factors across a wide spectrum of different commodities. A preview of our results reveals that i) the common factor of commodity returns can explain a statistically significant amount of the return variance for the vast majority of commodities (29 out of 43) at each point in time since 2004. Further, we find that the importance of the market component is generally substantially larger for non-agricultural products. For durable commodities (for example metal and energy commodities) the market-wide component of returns is generally more important than the commodity-specific component over recent years. In contrast, those that are in fixed or inelastic supply remain highly driven by the commodity-specific component. ii) changes in the variance contribution of the common factor are related to global demand and iii) the common factor itself is correlated with the US (trade-weighted) exchange rate and with commodity market indicators such as hedging pressure (net commercial long position / open interest). The result for aggregate hedging pressure supports the hedging pressure hypothesis (Keynes, 1930)⁴ at the commodity market level and indicates that commercial traders pay an insurance premium and thus an overall (net) long position signals a negative commodity market return. The result for the US exchange rate is consistent with Gourinchas and Rey (2007) that the dollar as a safe haven currency appreciates during periods of weak global demand and hence is associated with negative commodity returns. These findings complement prior literature, such as Kilian (2009), that have documented the key role of the demand shocks for the historical commodity prices, especially the recent rising commodity prices due to the growth of emerging economies.⁵

⁴ This theory is also often referred to as the theory of normal backwardation.
⁵ A tangential strand of literature examines different ideas within commodity futures markets. Yang (2013) identifies a “slope” factor in the cross section of commodity futures returns. They find that high-basis commodity futures have higher loadings on this factor than low-basis commodity futures. Szymanowska, Roon, Nijman and van den Goorbergh. (2014) identify two types of risk premia in commodity futures returns (i.e. spot premia related to the risk in the underlying commodity, and term premia related to changes in the basis). Daskalaki, Kostakis and Skiadopoulos (2014) examines whether common

The outline of the paper is as follows. Section 2 provides a short review of previous literature. In sections 3 the econometric techniques used are presented. A description of the data is presented in section 4. The empirical results from the factor model and the cross-commodity analysis is presented in section 5. Section 6 concludes.

2. Review of the empirical literature

Since the beginning of 2000, there were dramatic increases in commodity prices that continued until mid-2008. None of the proposed theories brought forth in the literature thus far have been able to fully explain this phenomenon. It is thought that one major factor that contributes to this increase in commodity prices has been global economic growth, especially for such commodities as energy and metals (see e.g., Alquist and Coibion (2014)).

One theme that is receiving strong recent attention is "financialization" of commodity markets (see for example Irwin and Sanders, 2011; Tang and Xiong, 2012). That is, the idea that financial investors now consider commodities as a distinct asset class. The financialization of commodity markets can be beneficial. For example, in the crude oil market, if investors are willing to take long forward exposures then they can increase their potential for risk sharing and hedging. It is argued that price stability and the process of price discovery may also be improved with these developments. However, financialization may also bring about the possibility for noise trading and momentum strategies (Miffre and Rallis, 2007) to increasingly affect prices especially if there are a large proportion of active speculative investors in the market. It is still debated as to whether such trading played a role in the acceleration of commodity prices between 2000 and 2008.

factors in the cross-section of individual commodity futures returns are priced and their results suggest that commodity markets exhibit considerable heterogeneity.

Since the early 2000s, commodities have emerged as an asset class alongside stocks and bonds. Studies, using data starting prior to 2000s, find that return correlations between commodity and stock returns have been increasing (Greer, 2000; Gorton and Rouwenhorst, 2006) as have the return correlations between crude oil and other commodities (Tang and Xiong, 2012; Silvennoinen and Thorp, 2013). However, generally return correlations among commodities in different sectors are small (Erb and Harvey, 2006). A number of studies have reported decreasing or non-increasing trends in return correlations between commodities and stocks before the recent financial crisis (Chong and Miffre, 2010; Buyuksahin, Haigh, and Robe, 2010). However, recent work finds that greater participation by speculators, especially hedge funds, is associated with a higher correlation between commodity and equity indices (Buyuksahin and Robe, 2014, Bruno et al., 2017), that global economic activity impacts co-movements between agricultural markets and equities (Bruno et al., 2017) and that commodity co-movements increased temporarily during the financial crisis period (Charlot et al., 2016).

The above mentioned characteristics of commodity returns suggest an opportunity for diversification. Institutional investors and hedge funds have started allocating funds in commodities intensively through trading commodity indices such as Standard & Poor's Goldman Sachs Commodity Index (S&P GSCI) and the Dow-Jones UBS Commodity Index (DJUBS), and commodity derivatives. As a result, time-varying correlations in commodity markets became increasingly of interest for empirical work (Silvennoinen and Thorp, 2013; Olson et al., 2014). Commodity returns also exhibit time-varying volatility (Vivian and Wohar, 2012; Silvennoinen and Thorp, 2013). Vivian and Wohar (2012) provide evidence of structural breaks in unconditional volatility of commodity returns. Structural breaks are largely idiosyncratic and occur throughout the 1985-2011 sample period; structural breaks are not concentrated in the financial

crisis period or when financialisation occurred. Silvennoinen and Thorp (2013) model conditional volatility of commodity returns in a GARCH framework. They find factors such as Bond spread, T-Bill and Open Interest are determinants of conditional volatility for some commodities.

Poncela, Senra and Sierra (2014) extend the empirical evidence on commodity comovement in two dimensions. First, they attempt to determine the extent of comovements in 44 monthly nonfuel commodity *return* series over the period 1992-2012, while previous literature focuses upon commodity *prices*. They provide evidence that the links among commodity returns, or co-movement, have increased since the end of 2003, as suggested by those who have discussed the financialization hypothesis (Masters, 2008; Tang and Xiong, 2012; among others). Second, they highlight the role of uncertainty (VIX) and macroeconomic fundamentals as potential drivers of non-energy comovement in the short term. However, given that they extract the commodity common factor using the principal components methodology which implicitly assumes a constant factor volatility, our approach of permitting a time-varying stochastic volatility is much more general and would provide a more accurate estimate of the commodity returns factors given prevalent evidence of structural instability of the commodity returns volatility (see for example Vivian and Wohar (2012)). Moreover, in their study, they break the sample in 2003 in order to establish the evidence of increasing co-movement among commodity returns, and the breakpoint was chosen without any formal structural break test. In fact, even when the standard structural break test is employed often the breakpoint is estimated with a large amount of uncertainty (see for example, Eo and Morley), and thus imposing the breakpoint in the subsequent estimation subjects the results to potential misspecifications. One of the advantages of our dynamic factor model with time-varying stochastic volatility is to avoid these above issues while at the same allowing for time-varying co-movement among the commodity returns.

Factor models have been used for the last 40 years in econometrics. Static factor models assume that all factors are given contemporaneously (See Stock and Watson (2002a,b)). The approximate factor model of Chamberlain (1983) and Chamberlain and Rothschild (1983) allows for idiosyncratic components which are cross-sectionally weakly correlated (approximate factor model), but are static. The dynamic factor models were first employed by Sargent and Sims (1977) and Geweke (1977), and the original version has mutually orthogonal idiosyncratic components. An intermediate model allowing for some dynamics is the restricted factor model where q dynamic factor are expressed with r static factors, $q < r$ (Forni *et al.*, 2005; Bai and Ng, 2007). A dynamic version of the approximate factor model, called the generalized dynamic factor model, has weekly cross-correlated idiosyncratic components (see Forni and Lippi, 2001; Forni *et al.*, 2000). Byrne, Fazio, and Fiess (2013), Chen, Jackson, Kim, and Resiandini (2013), West and Wong (2014), have applied factor models to a panel of commodity prices while Karstanje, van der Wel and van Dijk (2017) use such models to examine commodity price levels and price curves. In contrast very little work so far has applied such models to commodity returns with Poncela, Senra and Sierra (2014) being a notable exception.

3. A Dynamic Factor Model with Time-Varying Parameter and Stochastic Volatility

We extend the typical constant parameter Dynamic Factor Model (DFM) to a DFM with time-varying stochastic volatility (denoted by DFM-SV), following closely the work by Del Negro and Otrok (2008) and Bhatt, Kishor, and Ma (2017) ⁶. The DFM-SV decomposes the

⁶ We did estimate the most general model as in Del Negro and Otrok (2008) with time varying loadings as well as stochastic volatility, but we find very little time variations in factor loadings using our commodity data. We do find a great deal of variations in volatility, as will be discussed in details below. As a result, we maintain the assumption of constant loadings but allow for time varying stochastic volatility in the version of the model presented below. Similarly, Mumtaz and Surico (2012) also assume constant loadings while allowing for time varying stochastic volatility in their study of international inflation dynamics. Notably, they also allow for time varying autoregressive parameters, which, however, is not explored in our study for the same reason as for the loadings.

variations of each commodity return into three components: common market movement, sector-specific movement, and each individual commodity's idiosyncratic movement. Specifically the model is set up as below:

$$y_{i,t} = \lambda_i \cdot g_t + \gamma_i \cdot s_{j,t} + e_{i,t} \quad (1)$$

Here $y_{i,t}$ is the monthly return of commodity i ($i = 1, 2, \dots, n$, and n is the total number of commodities); g_t is the common market factor that affects all commodities markets and the economic forces behind this factor may be global economic activity that tends to affect all commodities markets. The loading parameter for the common factor g_t is λ_i for commodity i . The sector-specific factor $s_{j,t}$ captures the movement specific to sector j ($j = 1, 2, \dots, m$ and m is the number of sectors) and its loading parameter is γ_i . And $e_{i,t}$ is the idiosyncratic commodity-specific factor.

Stack the loading parameters for both the common market factor and the sector factor using vector $\beta_i = [\lambda_i \ \gamma_i]$ for each commodity return. As typical in the DFM, for identification purpose we assume that the common market factor, sector factors, and the commodity factors are orthogonal to each other. The time variations in these loading parameters permit timing varying contributions of various factors to commodities returns. As a result, the variance decomposition of each commodity return is:

$$Var(y_{i,t}) = \lambda_i^2 \cdot Var(g_t) + \gamma_i^2 \cdot Var(s_{j,t}) + Var(e_{i,t}) \quad (2)$$

In order to identify these factors we further assume certain time series dynamics for these factors. The market factor follows a stationary AR(p) process with a stochastic volatility:

$$g_t = \phi_1^g g_{t-1} + \phi_2^g g_{t-2} + \dots + \phi_p^g g_{t-p} + \sqrt{\exp(h_t^g)} \cdot \varepsilon_t^g \quad (3)$$

Where $\varepsilon_t^g \sim i.i.d.N(0, \sigma_g^2)$. The stochastic volatility follows a simple random walk process:

$$h_t^g = h_{t-1}^g + \sigma_g^h \cdot v_t^g, v_t^g \sim i.i.d.N(0,1) \quad (4)$$

Similarly, each sector factor also follows a stationary AR(l) process:

$$s_{j,t} = \phi_{j,1}^s s_{j,t-1} + \phi_{j,2}^s s_{j,t-2} + \dots + \phi_{j,l}^s s_{j,t-l} + \sqrt{\exp(h_{j,t}^s)} \cdot \varepsilon_{j,t}^s \quad (5)$$

Where $\varepsilon_{j,t}^s \sim i.i.d.N(0, \sigma_{j,s}^2)$. Its stochastic volatility follows a random walk process:

$$h_{j,t}^s = h_{j,t-1}^s + \sigma_{j,s}^h \cdot v_{j,t}^s, v_{j,t}^s \sim i.i.d.N(0,1) \quad (6)$$

Finally, each idiosyncratic commodity factor also follows a stationary AR(q) process:

$$e_{i,t} = \phi_{i,1} e_{i,t-1} + \phi_{i,2} e_{i,t-2} + \dots + \phi_{i,q} e_{i,t-q} + \sqrt{\exp(h_{i,t})} \cdot \varepsilon_{i,t} \quad (7)$$

Where $\varepsilon_{i,t} \sim i.i.d.N(0, \sigma_i^2)$. The stochastic volatility follows a random walk process:

$$h_{i,t} = h_{i,t-1} + \sigma_i^h \cdot v_{i,t}, v_{i,t} \sim i.i.d.N(0,1) \quad (8)$$

The volatility shocks are assumed to be orthogonal to each other. Note that the loading parameters and their corresponding factors shock variances are not separately identifiable. To normalize we first restrict the shocks to market and sector factors so that

$\sigma_g^2 = \sigma_{1,s}^2 = \dots = \sigma_{m,s}^2 = 1$. Also because in this model the factors volatilities are time varying, we

follow the approach in Del Negro and Otrok (2008) and impose the restriction that the time-varying volatility h 's in eq. (4), (6) and (8) all start from zero as their initial values, i.e.,

$h_0^g = h_{j,0}^s = h_{i,0}$ for $j = 1, 2, \dots, m$ and $i = 1, 2, \dots, n$. Similarly, to normalize the signs of loading

parameters and factors we restrict the factor loading parameter of the common market factor for

the first commodity to be positive, and the factor loading parameter of the sector factor to be

positive for the first commodity in each of the eight sectors. Finally, because the means of factors

are not separately identifiable, we demeaned all returns data before estimating the model.⁷

4. Data

Our empirical analysis is based upon a broad cross-section of 43 commodities over the January 1984 to June 2017 sample period. There are three main reasons for selecting this data. Firstly, an important consideration is to ensure that there is a broad cross-section of commodities included in the sample. Our sample includes commodities from a range of different sectors from Energy to Cereals, from Industrial Raw Materials to Precious Metals and from Other Food and Beverages to Industrial Metals. Secondly, our data sources are the widely used IMF and Thomson Datastream datasets. Thirdly, we wanted to include a reasonably sized comparator period to examine against the post-2000 period during which the Financialisation of commodities occurred; this will help to identify more clearly if and when a change in commodity market dynamics occurred.

In the empirical analysis we use commodity return defined as the natural log difference of commodity prices which are denoted by lower case. i.e. $y_t - y_{t-1} = \ln(Y_t / Y_{t-1})$. Note all raw data is de-meaned prior to estimation since the means of factors are not separately identifiable; consequently by construction all estimated factors have an expected mean of zero. The expected correlation between factors is also zero.

Descriptive statistics are provided in table 1. The mean of all factors is close to its expected value of zero. However, there are substantial differences in the standard deviations of the factors. This reflects the fact that the variances of the factor estimates are dynamic and depend upon the

⁷ The estimation algorithm for the model is discussed in detail in the online appendix (as suggested by an anonymous referee).

persistence of the factors themselves.⁸ There are also substantial differences in the percentiles of the factor distributions for the same reason. For example the Energy sector lower quartile is -4.336 and upper quartile is 4.631; in contrast there is much less variation for the Industrial Raw Materials sector where the lower quartile is -0.144 and upper quartile is 0.114. These results suggest that there is substantial differences in the time-series properties of the sectoral factors.

[INSERT TABLE 1:]

5. Econometric results

5.1 Common and sectoral factor estimates – stochastic volatility model

First, we present the estimates of the common and sectoral factors. These are shown in Figure 1. Recall that by construction the common factors and sectoral factors are orthogonal (in the population). The common factor identifies the one common component amongst all commodities, while the sectoral factors captures variation within a sector that is not attributable to the common factor. The common factor displays substantial time-variation and passes through its mean (zero) numerous times. In fact, as expected, the common factor of returns is not very persistent. It also appears there are a number of substantial shocks to the common factor which seem to be especially prevalent since the Financialisation of commodities and during the Financial Crisis period. Note that the common factor we estimate is not based on value-weighting and hence is not driven by the largest commodities (e.g. energy). This is in contrast to some empirical literature which uses a value-weighted commodity index (such as the S&P GSCI commodity indices). Hence when we compare the common factor to the S&P GSCI commodity there are some broadly similar movements but also periods of deviation such that the correlation between them is

⁸ Hence despite the market and sectoral factors being subject to shocks of unit variance, the variance of the factors themselves can differ substantially.

around 0.5. In contrast the common factor moves more closely the first principal component of the 43 commodity returns ($\rho = c.0.9$) and the equally weighted return across the 43 commodities ($\rho = c.0.85$); nevertheless the common factor is smoother than these other indices.

[INSERT FIGURE 1:]

The sectoral factor captures the comovement within a sector once the common factor is accounted for; i.e. the sectoral factor is obtained once the common factor is subtracted out. Most of the sectoral factors are also not very persistent and cross their mean of zero multiple times. In contrast to the common factor, the sectoral factors do not tend to have a general increase in the magnitude (or number) of shocks after Financialisation occurred or during the Financial Crisis period. In fact, for some sectors, such as Other Food and Beverages, and Industrial Raw Materials, the largest shocks clearly occur prior to 2000.

5.2 Variance decompositions on average by sector – stochastic volatility model

Table 2 presents estimates from the stochastic volatility model for the variance decomposition grouped by sector for the full sample. The contributions of each component overall are about 10% for the common factor, about 19% the sector factor and about 68-70% for the individual factor.⁹ However, there are a few differences across sectors.¹⁰ Firstly, the contribution of the common factor is much smaller for Livestock and Other Food and Beverage, which keeps the overall common factor contribution down to about 10%. Secondly, the importance of the individual factor is lowest on average at less than 50% for Energy and Precious Metals, the two sectors which are arguably the most liquid markets, the greatest interest from the financial press

⁹ We did estimate a constant loading and constant volatility model, which for the full sample provided similar results but which by definition cannot shed light on the time-variation in the importance of the factors.

¹⁰ There are also differences within sectors since the lowest and highest value for each sector are reported in brackets in Table 2.

and the largest content of macroeconomic news. Thirdly, it is worth noting that the importance of the sectoral factor differs dramatically from less than 2.5% in Other Food and Beverages, Livestock and Industrial Raw Materials to over 50% in Energy and Precious Metals. Finally, within a sector there is substantial variation in the importance of each factor; this is examined in more detail and depth in the next section.

[INSERT TABLE 2:]

5.3 Cross-sectional variation in Variance decompositions – stochastic volatility model

We next examine the proportion of the commodity-level return variance that can be explained by each factor with a focus on the common factor. These are displayed in Figure 2 for the common factor, while the results for the sector factor and the commodity-specific factor are available upon request. Table 3 presents variance decomposition results for the common factor over time. We consider results for the full-sample and two sub-samples (1984-1998 and 2004-2017). The final period is of special interest since the impact of speculative investors really gained impetus since 2004;¹¹ the sub-samples are of approximately equal length but with a break between them since there may have been gradual change in the dynamics prior to 2004. We discuss the results for the common factor in detail given this is one of the key contributions of the study.

5.3.1 Common Factor

Figure 2 interestingly shows that for the large majority of commodities (over 30 out of 43) we cannot reject the hypothesis that the common factor explains none of the variance in returns prior to about 1995. However, there is a substantial increase in point estimates for the proportion

¹¹ Commodity-index related instruments are estimated to have increased to \$200 billion by mid-2008 from \$15 billion in 2003 (CFTC, 2008; Tang and Xiong, 2012).

of variance explained by the common factor after 2000 (see Figure 2). For the large majority of commodities (29 out of 43) the contribution of the common factor is statistically significant from 0 for the last 10 years. However, there is also variation in the influence of the common factor even for commodities based in the same sector.

From Table 3 we note that for the industrial metals the common factor typically contributes less than 10% of total variation prior to 2000. However, since 2004 this has been about 30% for Aluminium, Copper, Lead, Nickel and Zinc, however, for Tin and especially Uranium there has been much less of an increase. For all industrial metals except Uranium the common component is statistically significant from 0 since 2005. For Energy commodities, similar to industrial metals, there has also been an increase in the influence of the common factor over time. The common factor could explain less than 5% of the return variance on average prior to 2003 but this has substantially increased to 10%-20% for 2004-2017. Most commodities also display significance of the common factor from 0 since 2005, given that the 95% confidence intervals exclude 0.

For Cereal commodities, the increase in the importance of the common factor is very pronounced in all cases except for rice; in terms of statistical significance the confidence intervals do not include 0 for any cereal except rice after about 2005. Since 2004 the common factor can on average explain more than 25% for Soybeans, Barley and Maize, which may be linked to the use of cereals as biofuels in recent years.

For Other Food and Beverage commodities, overall the common factor can explain on average less than 5% of return variation prior to 2003. However, since 2004 the common factor can explain about 10% of return variation for the majority of commodities. For several commodities there is a spike in the contribution of the common factor to around 40% around the Financial Crisis period, however, this appears to be short-lived and imprecisely estimated, since

the confidence bounds appear to increase. However, the common factor is significantly different from 0 since 2005 for only some of the commodities. The results for Livestock are rather similar to Other Food and Beverages. Figure 2 Panel B, indicates that for Livestock the point estimate is generally less than 5%, there is not typically a substantial increase in the mean point estimate of the common factor in recent years, nor is this generally statistically significant from 0 for any substantial period of time.

For precious metals the importance of the common factor approximately doubles from the 1984-1998 period to the 2004-2017 period. For example for Gold (Platinum) it increases from about 3% (8.5%) to over 7% (19%) post 2004. The figure indicates a substantial increase in the importance of the common factor around the time of the Financial Crisis, however, this tends to dissipate somewhat towards the end of our sample. For all precious metals the contribution of the common component is different from 0 since the financial crisis.

For Vegetable oils, there is a dramatic increase in the importance of the common factor for the 2004-2017 period (about 32% on average) relative to the pre-1998 period (about 7% on average). The increase in the importance of the common factor is more than triple for all each commodity. The increased usage of vegetable oils as biofuels likely is a determinant of this substantial increase in the explanatory power of the common factor for variance in the post 2004-period, which are consistently statistically different from 0 since 2004. With respect to industrial raw materials, there is a general increase in the amount of variance explained by the common factor during the financial crisis period. The common factor is statistically different from 0 since around 2005 onwards apart from for Logs – Soft (and perhaps Timber – Hard).

Overall there seems to be an upward trend in variance contribution of the common factor for many durable commodities and consequently the common factor has become both

economically and statistically important for many durable commodities since around 2004. These commodities such as industrial metals, oil and vegetable oils can be easily stored for long periods with substantial storage capacity available. Such commodities are much more demand-driven. In contrast a trend is less apparent for Livestock or many Food and Beverage commodities. This may reflect that these commodities are typically more difficult to store and therefore prices are much more supply inelastic and supply-driven.

[INSERT FIGURE 2 AND TABLE 3:]

5.3.2 Sectoral Factor and Individual Factor

Next, we summarise results for the sectoral factor.¹² In general there is not a strong trend in the sectoral effect over time including during the Financial crisis or since Financialisation of commodities occurred; thus there is very little support that Financialisation increased co-movements at the sector level. However, the sector factor appears to be important throughout the sample for many commodities (especially those in the Energy and Precious Metals sector). For most commodities in the Industrial Metals, Energy, Cereals and Precious Metals groups the sector factor contributes positively and is statistically different from 0 (and typically also from 1) for almost all of the sample period. For the other sectors the sectoral component is typically very weak indeed and rarely statistically different from 0.

In terms of the individual factor, the conventional wisdom is that commodity returns are largely idiosyncratic (see e.g., Erb and Harvey, 2006) and much prior literature models them as if they were entirely idiosyncratic (see e.g., Bidarkota and Crucini, 2000). We can report that there is substantial variation in the point estimates, however, there are also some noticeable trends.¹³ In

¹² The full results are available upon request and are in the same style as those for the common factor.

¹³ The full results are available upon request and are in the same style as those for the common factor.

particular there is generally a downward trend in the individual contribution for industrial metals precious metals and for some vegetable oils since at least 2000. For the remaining commodities there is generally little clearly apparent trend in the point estimates for the individual factor, although on average there is some, albeit more modest, decline for Other Food and Beverages, Livestock and Industrial Raw Materials.¹⁴

An important finding is that for a large majority commodities the individual factor explains statistically less than all of the variance (less than 100%) of returns, which complements the findings that the common factor has increased in importance since about 2004. This is especially true for commodities which are non-perishable such as Precious Metals and Vegetable Oils. These results indicate that our analysis is not plagued by the weak identification issue documented by Balke, Ma and Wohar (2013, 2015), which can impact variance decompositions of various types. As illustrated through their work, the weak identification often plagues the type of decomposition specification that involves nonlinearity in parameters or correlated shocks or both. The complex nature of these types of model specifications is such that often data do not provide sufficient information to allow us to well identify some of the key parameters. The resulting weak identification leads to wide confidence intervals of the decomposition. The likely reasons here for us to avoid this issue is that the return is a linear function of parameters and factors, and that by construction each factor is uncorrelated with the other factors, whereas in other applications the components are correlated with each other and often nonlinearity in parameters are also involved. As a result, these linearity and orthogonality specifications help us better identify the variance contributions of factors components, and also facilitates a more intuitive interpretations of these

¹⁴ For some commodities this could reflect that it takes considerable time for supply to adjust to market factors and thus returns in the short term will depend heavily on current crop yields, which are largely idiosyncratic (commodity-specific); this is especially true for coffee and cocoa trees where a crop is not produced until three years after planting (Josephs, 2014).

orthogonal components than the correlated ones otherwise. The important implication of this is that the sector and common factors of the return are important or have become important for a large number of commodities and when conducting analysis on these commodities it will be vital to model or take account of the sectoral and market-wide element of their returns.

5.4 Relationship of the Common Factor to other Variables

We now proceed to examine the connections of the common factor of the return to economic and financial market indicators given the focus of our study on this issue and the previous results indicating the importance of the common factor has grown considerably over recent years.¹⁵ Table 4 reports results investigating the relationship of the common factor of commodity returns from the stochastic volatility model with other variables. We model $g_t = f(Z_t)$, where g is the common factor and Z is a vector of candidate predictors variables, which include all the fundamental variables reported to be determinants of variation in commodity prices by prior literature. These include the US exchange rate return (ΔS_t), global demand (GD_t), inflation (π_t), real T-Bill rate ($TBill_t$) and the real long-term bond rate (LTB_t). We also include two measures of uncertainty; first a measure of economic uncertainty ($\sigma_{EC,t}$) from Bloom (2009) and second a measure of financial market uncertainty, the change in the VIX index (ΔVIX_t). We also examine the lag of the common factor (g_{t-1}) since it could exhibit some persistence¹⁶. Given the importance of China as a user of commodities, we examine some Chinese economic variables namely industrial production growth ($\Delta IP_{CH,t}$), inflation ($\pi_{CH,t}$), change in real effective exchange rate ($S_{CH,t}$) and change in retail sales ($\Delta RS_{CH,t}$).

¹⁵ This paper focuses on a detailed examination of the common return factor component given the increasing importance of this factor. A detailed examination of the sector and individual return components is a possible avenue for future research.

¹⁶ Although two lags are allowed for the common factor in MCMC estimation the second lag comes out being small and insignificant and thus is not included in regressions here.

Finally, we consider commodity market measures. The first measure is hedging pressure (HP_t) which is computed as the average of net long commercial positions to total open interest. The remaining measures are more related to speculation and financialization. The average change in open interest (ΔOI_t) captures the increase in activity in commodity markets for which financialization and speculative traders comprise a large proportion of the increase in recent years. We also include a measure of excessive speculation (ES_t) following Tadesse (2014),¹⁷ which is measured as: (net long commercial positions – net long non-commercial positions)/ open interest. The last measure covers the change in non-reported long positions scaled by open interest (ΔNRL_t), these are trading positions held by smaller participants (including some hedge funds), which tend to be more speculative in nature.

For multivariate regressions we provide estimates of the contribution of each variable to the variance of the dependent variable. We use the concept of Shapley (1953) derived from cooperative game theory to determine the exact contribution of each explanatory variable to the regression R-square; this is expressed as a percentage.¹⁸ We also use a general to specific modelling approach in the multivariate regressions to avoid concerns of model overfitting. As such in the first step we estimate a multivariate regression with all candidate variables. We then include the variables which had an absolute t-value of 1.00 or higher in the second multivariate regression, which we report.¹⁹

Table 4 Panel 1.1 provides evidence from bivariate regressions that many of the fundamentals are statistically significantly related to the common factor of commodity returns.²⁰

¹⁷ This is similar to the Working T measure.
¹⁸ The Shapley-Owen R-squared is estimated by the REGO command in STATA, which draws on code from <http://www.uni-leipzig.de/~rego/>. See Huettner and Sunder (2012) for further details and motivation of the procedure.
¹⁹ Alternative estimation results are available upon request.
²⁰ Prior to estimation, the variables of these regressions were standardized by dividing by the standard deviation of the variable.

The signs on the relationships are also generally intuitively acceptable. For example there is a positive relationship between the common factor and global demand. This is consistent with increased economic activity pulling up commodity prices. The most striking results from Panel 1 is the strong relationship between the change in the trade-weighted exchange rate and the common factor. An increase in the exchange rate here is an appreciation of the dollar and is negatively related to commodity returns. The change in the exchange rate strongly related to the common factor and by itself has an R^2 of almost 30%.

Table 4 Panel 1.2 reports results from multivariate regression, which can explain 63% of the variation in the common factor. The main variables related to the common factor in this setting are the change in the US exchange rate, hedging pressure, change in open interest and the lag of the common factor. All these variables are significant at the 1% level. The Shapley decomposition of R-squared indicates that the change in the US exchange rate is the most important variable connected to the common factor with the lagged common factor second. Hedging pressure contributes about 16% to R-squared and change in open interest about 8%. The Chinese economic variables are not the most important although Chinese inflation is statistically significant at the 5% level and although the Chinese real exchange rate contributes to the overall R-squared it has a low t-value and is far less important than the US exchange rate. Overall, there is limited support for Chinese macro-financial variables being strongly linked to commodity movements.

[INSERT TABLE 4 AROUND HERE:]

Table 4 Panel 2 reports results for the 2004:1 – 2017:6 sample period. In the bivariate setting most explanatory variables are statistically significant, although the change in the US exchange rate has the largest R-squared. In the multivariate setting it is hedging pressure and especially the change in the exchange rate which are the most important variables; both these

variables always have a Shapley decomposition of more than 19%. Further the coefficient on the lagged common factor is less important during this period with lower coefficient estimates and t-statistics. The adjusted R-squared is 0.7 which indicates a very good fit from the model.

Overall, the results suggest that apart from its own first lag, hedging pressure and the change in the exchange rate have the strongest connections to the common factor with the latter being of crucial importance over 2004-2017. On the surface it is not surprising that the exchange rate has an important common effect on the return to commodities since all commodities are denominated in a common currency, US \$'s, it is perhaps surprising however that prior work primarily documents a link between the exchange rate and gold (Pukthuanthong and Roll, 2011) or energy (Olson et al., 2014) rather than with the commodity index. The strong co-movement between the common commodity factor and the US dollar exchange rates reported in this paper may present a strong evidence in support of a fundamental story of the exchange rate and the view that global demand plays an important role in commodities markets. Specifically, Gourinchas and Rey (2007) note that US dollar tends to become stronger (weaker) when there is adverse (good) news about the global economy relative to US and investment funds come into (leave) the safe haven. Engel and West (2005) demonstrate that exchange rate predict fundamentals and are therefore forward-looking; Chen, Rogoff and Rossi (2010) provide further evidence that exchange rates are forward looking, in particular for future commodity prices. Since our exchange rate data is quoted as the value of dollar in terms of foreign currencies, the negative correlations in our regressions imply that the stronger (weaker) dollar is related to weaker (stronger) commodities markets that can be attributed to a weaker (stronger) global demand. In general, the variation in the exchange rate reflects contemporaneous economic developments and the expectations of future economic activities. Hence at least part of the exchange rate fluctuation reflects change of the

fundamentals for the rest of the world relative to the US.

5.5 Relationship of the Aggregate Factor Variation Contributions to other Variables

Our analysis enables us to track the amount of variance attributed to each component over time. In this section we examine the what is connected with the change in the average variance contribution from the common factor. We investigate the change in contributions since the level of contributions are highly persistent. We use the same economic and financial market explanatory variables which we investigated as being connected to the common factor.

Table 5 reports results for the change in the contribution of the common factor. The important variables differ from those of the common factor itself. For the full-sample (Panel 1) bivariate results indicate that the significant variables are Global Demand, Real T-Bill, Real Long-term Bond, the Chinese real exchange rate and the lagged dependent variable. However, in a multivariate setting only Global Demand, Inflation and the lagged dependent variable are significant, with the lagged dependent variable providing the vast majority of explanatory power.

The results for the latest sub-sample 2004:1-2017:12 (Panel 2) are broadly consistent with those from the full-sample. The main change is that the global demand effect has become stronger, both in terms of its coefficient and in terms of the Shapley R^2 and that while US inflation was important in the bivariate setting it was not in the multivariate one. The goodness of fit also increases in the latest sub-sample to 75%.²¹ Finally, the lagged dependent variable provides less of the explanatory power during the later sub-period compared to the full-sample. The Chinese economic variables show some marginal statistical significance but in terms of economical

²¹ In results that are available on request we find that for earlier sub-periods (1984:1-1993:12 and 1994:1-2003:12) only the lagged dependent variable is robustly related to the change in variance contributed by the common factor. These are omitted in the interests of brevity.

importance only contribute a small amount of the total R^2 . The positive sign for global demand is especially interesting since it suggests that there is more common movement in expansions rather than recessions while conventional wisdom suggests that correlations increase during bad times (for commodities see Charlot et al., 2016); however our result complements Buyuksahin and Robe (2014) who find less co-movement of commodities with equities during periods of financial stress.²²

[INSERT TABLE 5 AROUND HERE:]

5.6 Robustness Tests for Regression Results

For robustness purposes we also investigated i) the impact of outliers on our main results and ii) if the change in the average contribution of the individual factor is a reflection of the results for the common factor. Results are available upon request. In terms of the tests of outliers, we winsorized our dependent variables at the 2.5% and 97.5% level and re-estimated the main regressions again. These results are qualitatively very similar to the results reported in Tables 4 and 5. In terms of the average contribution of the individual factor, these are used to verify the results from the change in contribution of the common factor, since if there are differences then this would imply that the sector factor has a role to play. Overall the key variables for explanatory power are consistent with Table 5, namely the lagged dependent variable is the most important with global demand and inflation being of secondary importance.

6. Conclusion

A large body of literature treats commodities as being largely idiosyncratic and especially overlooks the potential importance of common movements in commodities. Prior work has

²² Buyuksahin and Robe ascribe this to increased frictions and constraints on market participants which discourages arbitrage and thus inhibits the speed of adjustment of market prices.

1
2
3 focused on commodity prices. In contrast, the dynamics of realized commodity returns have
4
5 received relatively little attention until recently (notable exceptions include Poncela, Senra and
6
7 Sierra 2014). In this paper the focus is on commodity returns and examining the extent to which
8
9 these are driven by common factors and whether this component is statistically different from zero.
10
11

12 Firstly, this paper provides a dynamic factor model which allows for stochastic volatility;
13
14 this model generates a common factor and sectoral factors that i) capture common comovement
15
16 and ii) are orthogonal. We apply the model to 43 commodities over the 1984M1–2017M6 sample
17
18 period. The common factor of commodity returns is relatively precisely estimated (and has
19
20 relatively narrow confidence intervals). The stochastic volatility of the common factor of
21
22 commodity returns varies over time and is especially high from 2005 onwards. This method
23
24 enables us to estimate at each point in time how much of the return variation is due to each factor.
25
26
27

28 Secondly, this paper examines the drivers of movements in commodities. The common
29
30 factor now has a statistically significant role to play in explaining movements in returns (along
31
32 with the individual factor). Hence this analysis does not appear to be subject to the weak
33
34 identification issue outlined in Balke, Ma and Wohar (2013, 2015) which has plagued much prior
35
36 work examining variance decompositions. Crucially we find the importance of the common factor
37
38 for driving returns increases sharply after 2004. We find the common factor contributes positively
39
40 and statistically significantly to the variance of commodity returns for 29 out of 43 commodities
41
42 especially over the post-2004 sample period. Consequently, we add to a growing literature that
43
44 emphasises that the way in which commodity markets operate has evolved since the
45
46 financialization of commodities (Silvennoinen and Thorp, 2013). While we find some evidence of
47
48 a link between the common factor of commodity returns and financial market measures (change
49
50 in open interest and hedging pressure), we do not find a strong relationship between the aggregate
51
52
53
54
55
56
57
58
59
60

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

change in return variance attributable to the common factor and the financial market measures.
Hence the rise in importance of the common factor as a driver of commodity returns cannot be fully ascribed to financial market developments.

One of our main contributions is to demonstrate that the common factor of commodity returns drives an increasing proportion of the variance in many commodities after 2004. Hence commodities have become a less effective tool for diversification over recent years; this insight applies both to investment managers seeking to diversify portfolios of financial asset and also to commodity producers considering diversifying their production.

For Peer Review

References

- Alquist, R. and Coibion, O. (2014), "Commodity-Price Comovement and Global Economic Activity," NBER working paper No. 20003.
- Bai, J., and Ng, S. (2007), "Determining the Number of Primitive Shocks in Factor Models," *Journal of Business and Economic Statistics*, 25: 52–60.
- Balke, N., Ma, J., Wohar, M. (2013), "The contribution of economic fundamentals to movements in exchange rates," *Journal of International Economics*, 90: 1-16
- Balke, N., Ma, J., Wohar, M. (2015), "A Bayesian Analysis of Stock Price Decomposition", *Macroeconomic Dynamics*, 19(4): 728-752.
- Basak, S. and Pavlova, A. (2016),. "A model of financialization of commodities". *The Journal of Finance*, 71(4), pp.1511-1556.
- Bhatt, Vipul, Kishor, N Kundan, and Jun Ma, (2017) "The impact of EMU on bond yield convergence: Evidence from a time-varying dynamic factor model" *Journal of Economic Dynamics and Control*, (82): 206-222.
- Bidarkota, P. and M. J. Crucini (2000), "Commodity Prices and the Terms of Trade," Review of International Economics, 8(4), 647-666.
- Bloom, N. (2009), "The impact of uncertainty shocks", *Econometrica*, 77(3): 623-685.
- Bruno, V. G., Büyükaşahin, B., & Robe, M. A. (2017). The financialization of food?. *American Journal of Agricultural Economics*, 99(1), 243-264.
- Buyuksahin, B., Haigh, M. S., and Robe, M. A. (2010), "Commodities and equities: Ever a 'market of one'?" *Journal of Alternative Investments*, 12: 76-95.
- Büyükaşahin, B., and Robe, M. A. (2014). Speculators, commodities and cross-market linkages. *Journal of International Money and Finance*, 42, 38-70.
- Byrne, J.P., Fazio, G, and Fiess, N. (2013), "Primary commodity prices: Co-movements, common factors and fundamentals," *Journal of Development Economics*, 101: 16-26.
- CFTC. (2008), "Staff Report on Commodity Swap Dealers & Index Traders with Commission Recommendations." Commodity Futures Trading Commission (September).
- Chamberlain, G. (1983), "Funds, Factors, and Diversification in Arbitrage Pricing Models," *Econometrica*, 51: 1305–1323.
- Chamberlain, G., Rothschild, M. (1983), "Arbitrage, Factor Structure, and Mean-Variance Analysis on Large Asset Markets," *Econometrica*, 51: 1281–1304.
- Charlot, P., Darné, O., & Moussa, Z. (2016). Commodity returns co-movements: Fundamentals or "style" effect?. *Journal of International Money and Finance*, 68, 130-160.
- Chen, S., Jackson, J.D., Kim, H., and Resiandini, P. (2013), "What drives commodity prices?" working paper, AUWP 2013-03.
- Chen, Yu-Chin, Rogoff Kenneth S., and Barbara Rossi (2010) "Can Exchange Rates Forecast Commodity Prices? The Quarterly Journal of Economics 125 (3): 1145-1194.
- Chong, J., Miffre, J. (2010), "Conditional correlation and Volatility in Commodity Futures and Traditional Asset Markets," *Journal of Alternative Investments*, 12: 61-75.
- Del Negro, M., Otrok, C. (2008), "Dynamic Factor Models with Time-Varying Parameters: Measuring Changes in International Business Cycles," *Federal Reserve Bank of New York Staff Report*, 326.
- Engel, C. and West, K.D. (2005), "Exchange rates and fundamentals", *Journal of Political Economy* 113(3): 485-517.

- 1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
- Eo, Y. and J. Morley (2015), "Likelihood-ratio-based confidence sets for the timing of structural breaks," *Quantitative Economics* 6: 463-497.
 Erb, C. B., Harvey, C. R. (2006), "The strategic and tactical value of commodity futures", *Financial Analysts Journal*, 62: 69-97.
 Forni, M., Hallin, M., Lippi, M., and Reichlin, L. (2000), "The generalized dynamic-factor model: identification and estimation," *The Review of Economics and Statistics*, 82: 540-554.
 Forni, M., Hallin, M., Lippi, M., and Reichlin, L. (2005), "The Generalized Dynamic Factor Model: One-Sided Estimation and Forecasting," *Journal of the American Statistical Association*, 100: 830-840.
 Forni, M., Lippi, M. (2001), "The Generalized Dynamic Factor Model: Representation Theory," *Econometric Theory*, 17: 1113-1141.
 Geweke, J. (1977), "The dynamic factor analysis of economic time-series models," in D.J. Aigner and A.S. Goldberger (Eds.), *Latent Variables in Socio- Economic Models*, (North-Holland, Amsterdam), Chapter 16: 365-383.
 Gorton, G. B., Rouwenhorst, G. K., (2006), "Facts and fantasies about commodity futures," *Financial Analysts Journal*, 62: 47-68.
 Gourinchas, P., H. Rey (2007), "From World Banker to World Venture Capitalist: U.S. External Adjustment and the Exorbitant Privilege," in NBER book *G7 Current Account Imbalances: Sustainability and Adjustment*, edited by Richard H. Clarida, University of Chicago Press, 11-66.
 Greer, R. (2000), "The nature of commodity index returns," *Journal of Alternative Investments*, 3(1): 45-53.
 Hong, H., Yogo, M., 2012. "What does futures market interest tell us about the macroeconomy and asset prices?", *Journal of Financial Economics*, 105(3), pp.473-490.
 Huettner, F. and Sunder, M. (2012), "Axiomatic arguments for decomposing goodness of fit according to Shapley and Owen values", *Electronic Journal of Statistics*, 6: 1239-1250.
 Irwin, S. H. and Sanders, D. R. (2011), "Index Funds, Financialization, and Commodity Futures Markets", *Applied Economic Perspectives and Policy*, 33(1): 1-31.
 Josephs, L. (2014), "Commodity Outliers Attract Investors" *Wall Street Journal*, 25 September 2014. <http://online.wsj.com/articles/commodity-outliers-attract-investors-1411686300>
 Karstanje, D., M. van der Wel, and D. van Dijk (2017), "Common factors in commodity futures curves," working paper, available at SSRN: <http://ssrn.com/abstract=2558014>.
 Keynes, J. (1930), *A Treatise on Money*, Vol. 2, Macmillan, London.
 Kilian, L. (2009), "Not All Oil Price Shocks Are Alike: Disentangling Demand Supply Shocks in the Crude Oil Market," *American Economic Review*, 99:3, 1053-1069.
 Ma, J. and Wohar, M.E., (2014), Determining what drives stock returns: Proper inference is crucial: Evidence from the UK, *International Review of Economics & Finance*, 33, 371-390.
 Marshall, B.R., Nguyen, N.H. and Visaltanachoti, N., (2013), Liquidity commonality in commodities, *Journal of Banking & Finance*, 37(1), 11-20.
 Masters, M. (2008), Testimony before the committee on homeland security and governmental affairs, *United States Senate*, May 20.
 Miffre, J., Rallis, G. (2007), Momentum strategies in commodity futures markets, *Journal of Banking & Finance*, 31(6): 1863-1886.
 Mumtaz, H., P. Surico (2012), "Evolving International Inflation Dynamics: World and Country-

- Specific Factors,” *Journal of European Economic Association*, 10, 4, 716-734.
- Neely, C., Rapach, D. (2011), “International Comovements in Inflation Rates and Country Characteristics,” *Journal of International Money and Finance*, 30: 1471-1490.
- Olson, E., Vivian, A., and Wohar, M.E. (2014) “The Relationship Between Energy and Equity Markets: Evidence From Volatility Impulse Response Functions,” *Energy Economics*, 41: 95-109.
- Pindyck, R.S., Rotemberg, J.J. (1990), “The excess co-movement of commodity prices,” *The Economic Journal*, 100: 1173-1187.
- Poncela, P., Senra, E, and Sierra, L.P. (2014), “Common dynamics of nonenergy commodity prices and their relation to uncertainty,” *Applied Economics*, 46(30): 3724-3735.
- Pukthuanthong, K., & Roll, R. (2011). “Gold and the Dollar (and the Euro, Pound, and Yen),” *Journal of Banking & Finance*, 35(8), 2070-2083.
- Saadi, H. (2011), “Price comovements in international markets and their impacts on price dynamics,” In *Methods to Analyze Agricultural Commodity Price*, ed. Isabelle Piot-Lepetit and Robert M’Barek. Springer, New York.
- Sargent, T.J., Sims, C.A. (1977), “Business cycle modeling without pretending to have too much a priori economic theory,” in *New Methods in Business Research*, Christopher A. Sims (Ed.), Federal Reserve Bank of Minneapolis, Minneapolis.
- Shapley, L. (1953), “A value for n-person games,” in Kuhn A.H. and Tucker A.W., eds., *Contributions to the Theory of Games*, Vol. 2 (Princeton University Press).
- Silvennoinen, A., Thorp. S. (2013), “Financialization, crisis and commodity correlation dynamics,” *Journal of International Financial Markets, Institutions and Money*, 24: 42-65.
- Stock, J.H., Watson, M.W. (2002a), “Macroeconomic Forecasting Using Diffusion Indexes,” *Journal of Business & Economic Statistics*, 20: 147–162.
- Stock, J.H., Watson, M.W. (2002b), “Forecasting Using Principal Components from a Large Number of Predictors,” *Journal of the American Statistical Association*, 97: 1167–1179.
- Szymanowska, M., Roon, F., Nijman, T., & Goorbergh, R. (2014), An anatomy of commodity futures risk premia, *The Journal of Finance*, 69(1), 453-482.
- Tadesse, G., Algieri, B., Kalkuhl, M. and von Braun, J. (2014). Drivers and triggers of international food price spikes and volatility. *Food Policy*, 47, 117-128.
- Tang, K., Xiong, W. (2012), “Index investment and the financialization of commodities,” *Financial Analysts Journal*, 68: 54-74.
- Vivian, A., Wohar, M. (2012) “Commodity Breaks in Volatility,” *Journal of Financial Markets, Institutions and Money*, 22(2): 395-422.
- West, K.D., Wong, (2014), “A factor model for co-movements of commodity prices,” *Journal of International Money and Finance*, 42(C): 289-309.
- Yang, F., (2013), Investment shocks and the commodity basis spread. *Journal of Financial Economics*, 110(1), 164-184.

TABLE 1. Descriptive Statistics for the Stochastic Volatility Dynamic Factor Model

Monthly Data: 1984M1 – 2017M6

	Mean	SD		Mean	SD
MARKET FACTOR	0.004	1.165	INDIVIDUAL FACTORS (CONT.)		
SECTORAL FACTORS			OTHER FOOD & BEVERAGES		
1. IND. METALS	-0.073	2.174	BANANAS	0.005	16.345
2. ENERGY	0.005	2.830	COCOA	-0.010	5.288
3. CEREALS	0.020	1.418	COFFEE	-0.004	7.567
4. OTHER FOOD & BEVS.	-0.011	0.346	TEA	-0.006	7.386
5. LIVESTOCK	-0.002	0.796	FISHMEAL	-0.008	4.712
6. PREC. METALS	-0.015	1.092	ORANGES	0.000	12.288
7. VEG. OILS	-0.033	0.643	SALMON	-0.008	5.777
8. IND. MATERIALS	-0.002	0.127	SUGAR	0.005	16.345
INDIVIDUAL FACTORS			LIVESTOCK		
IND. METALS			BEEF	0.002	3.785
ALUMINUM	0.065	4.137	HIDES	-0.001	6.235
COPPER	0.074	3.950	LAMB	0.001	3.413
LEAD	0.099	4.954	PORK	0.000	10.669
NICKEL	0.092	6.497	POULTRY	0.000	2.148
TIN	0.039	4.615	PRECIOUS METALS		
ZINC	0.091	4.092	GOLD	0.045	1.790
URANIUM	0.001	6.889	SILVER	0.071	3.903
ENERGY			PLATINUM	0.043	4.547
FUEL OIL	-0.013	4.778	VEG. OILS		
PROPANE	-0.011	10.491	OLIVE OIL	0.101	6.906
HEATING OIL	-0.011	1.408	RAPESEED OIL	0.186	4.898
GASOLINE	-0.013	8.637	PALM OIL	0.156	2.384
OIL - WTI	-0.011	6.290	SOYBEAN OIL	0.174	6.061
CEREALS			SUNFLOWER OIL	-0.011	4.414
BARLEY	-0.038	5.370	IND. MATERIALS		
MAIZE	-0.059	2.560	COTTON	-0.002	5.181
RICE	-0.002	6.014	LOGS – SOFT	0.000	6.908
SOYBEANS	-0.045	3.123	RUBBER	-0.009	5.650
WHEAT	-0.037	4.657	TIMBER – HARD	0.001	4.767
			WOOL – COARSE	-0.003	4.325

Notes: This table provides descriptive statistics for all the factors. Please note mean is the sample average of the variable. SD is the standard deviation.

**TABLE 2. Variance decomposition:
Time-Varying Volatility Model – Averages across sectoral groups 1984:1-2017:6**

<i>GROUPING</i>	<i>Common factor</i>	<i>Sectoral Factor</i>	<i>Individual factor</i>
<i>ALL</i>	9.42%	19.98%	68.85%
<i>INDUSTRIAL METALS</i>	13.86% [0.56%; 23.96%]	21.47% [0.29%; 32.52%]	62.24% [47.04%; 98.66%]
<i>ENERGY</i>	9.42% [6.30%; 12.43%]	52.46% [28.90%; 84.50%]	35.68% [5.39%; 56.98%]
<i>CEREALS</i>	12.14% [0.46%; 13.87%]	26.17% [0.24%; 57.29%]	60.01% [25.78%; 98.99%]
<i>OTHER FOOD & BEVERAGES</i>	4.07% [0.17%; 8.26%]	1.58% [0.55%; 3.97%]	93.35% [89.91%; 98.76%]
<i>LIVESTOCK</i>	1.50% [0.10%; 5.50%]	2.20% [0.66%; 4.91%]	95.50% [87.74%; 98.83%]
<i>PRECIOUS METALS</i>	8.85% [4.88%; 12.95%]	54.74% [40.65%; 66.18%]	34.40% [27.90%; 43.77%]
<i>VEG. OILS</i>	17.22% [4.92%; 29.67%]	24.77% [0.52%; 47.58%]	55.21% [30.25%; 94.13%]
<i>INDUSTRIAL RAW MATERIALS</i>	9.53% [0.21%; 22.58%]	0.82% [0.35%; 1.53%]	88.58% [75.46%; 99.17%]

Notes: This table reports results from commodity-level variance decompositions into components due to the Common factor, the Sectoral factor and the Individual factor. These are then averaged across commodities in each sector and for all commodities. In brackets are the lowest and highest value amongst the individual commodities in a sector.

TABLE 3. Variance decompositions of Common Factor – Time Varying Volatility Model – Averages across sub-periods.

	OVERALL	1984-1998	2004-2017		OVERALL	1984-1998	2004-2017
<u>IND. METALS</u>				<u>OTHER FOOD & BEVERAGES (CONT.)</u>			
ALUMINIUM	15.29%	6.07%	25.27%	SALMON	8.19%	7.83%	9.39%
COPPER	23.96%	13.86%	35.22%	SUGAR - CSCE	5.03%	1.89%	9.55%
LEAD	11.45%	5.53%	18.76%	AV. OTHER FOOD & BEVERAGES	4.07%	2.31%	6.78%
NICKEL	14.59%	11.46%	19.99%	<u>LIVESTOCK</u>			
TIN	15.09%	12.37%	18.93%	BEEF	1.09%	0.56%	1.91%
ZINC	16.03%	7.82%	25.78%	HIDES	0.43%	0.06%	0.95%
URANIUM	0.56%	0.53%	0.53%	LAMB	5.50%	1.73%	11.00%
AV. IND. METALS	13.85%	8.24%	20.64%	PORK	0.39%	0.11%	0.78%
<u>ENERGY</u>				POULTRY	0.10%	0.01%	0.24%
FUEL OIL	9.26%	4.61%	16.87%	AV. LIVESTOCK	1.50%	0.49%	2.98%
PROPANE	6.30%	3.72%	10.73%	<u>PRECIOUS METALS</u>			
HEATING OIL	9.04%	4.24%	16.65%	GOLD	4.88%	3.08%	7.41%
GASOLINE	10.07%	4.47%	18.72%	SILVER	8.72%	6.01%	12.66%
OIL – WTI	12.43%	7.66%	20.36%	PLATINUM	12.95%	8.56%	19.73%
AV. ENERGY	9.42%	4.94%	16.67%	AV. PRECIOUS METALS	8.85%	5.88%	13.27%
<u>CEREALS</u>				<u>VEG. OILS</u>			
BARLEY	11.38%	3.97%	20.36%	OLIVE OIL	18.56%	5.12%	38.27%
MAIZE	13.87%	8.81%	20.80%	RAPESEED OIL	21.47%	8.60%	39.98%
RICE	0.46%	0.13%	0.90%	PALM OIL	29.67%	13.89%	52.54%
SOYBEANS	26.74%	18.15%	39.72%	SOYBEAN OIL	11.50%	4.21%	22.46%
WHEAT	8.27%	6.41%	11.66%	SUNFLOWER OIL	4.92%	2.69%	8.02%
AV. CEREALS	12.14%	7.49%	18.69%	AV. VEG. OILS	17.22%	6.90%	32.25%
<u>OTHER FOOD & BEVERAGES</u>				<u>IND. MATERIALS</u>			
BANANAS	0.17%	0.01%	0.40%	COTTON	7.58%	4.96%	11.87%
COCOA	4.92%	2.34%	8.98%	LOGS – SOFT	0.21%	0.08%	0.37%
COFFEE	8.26%	3.73%	15.37%	RUBBER	22.58%	23.53%	25.09%
TEA	0.92%	0.29%	1.77%	TIMBER - HARD	4.77%	0.68%	9.89%
FISHMEAL	4.17%	2.05%	6.95%	WOOL - COARSE	12.49%	7.62%	20.15%
ORANAGES	0.90%	0.32%	1.79%	AV. IND. MATERIALS	9.53%	7.37%	13.48%
				AV. ALL	9.52%	5.34%	15.32%

TABLE 4. The Connections of the Common Factor to Economic and Financial Indicators

	Panel 1.1: Bivariate			Panel 1.2: Multivariate			
	1984:1-2017:6	Common Factor			Common Factor		
		β	t	r-sq	β	T	r-sq Shapley
$\Delta\sigma_{ECt}$		-0.066	-1.13	0.00			
ΔVIX_t		-0.152	-1.67*	0.02	-0.074	-1.75*	1.89
GD_t		0.167	2.87***	0.03			
π_t		-0.121	-1.83*	0.01	-0.202	-2.51**	1.74
$TBILL_t$		0.004	0.06	0.00	0.384	3.12***	2.44
LTB_t		0.024	0.39	0.00	-0.259	-1.97**	1.59
ΔS_t		-0.540	-5.60***	0.29	-0.532	-4.06***	31.65
$\Delta IP_{CH,t}$		0.080	1.66*	0.01			
$\pi_{CH,t}$		0.075	1.75*	0.00	0.126	2.00**	1.65
ΔSCH_t		-0.153	-2.07**	0.02	0.140	1.01	8.95
$\Delta RS_{CH,t}$		-0.084	-1.55	0.01	-0.062	-1.35	0.45
ΔOI_t		0.300	4.49***	0.09	0.169	4.22***	7.81
HP_t		-0.458	-8.94***	0.20	-0.280	-6.58***	16.12
ES_t		0.056	1.23	0.00			
ΔNRL_t		-0.167	-0.42	0.00			
g_{t-1}		0.532	5.22***	0.28	0.297	5.62***	0.63 25.70

	Panel 2.1: Bivariate			Panel 2.2: Multivariate			
	2004:1-2017:6	Common Factor			Common Factor		
		β	t	r-sq	β	T	r-sq Shapley
$\Delta\sigma_{ECt}$		-0.099	-0.91	0.01			
ΔVIX_t		-0.241	-1.66*	0.03	-0.069	-1.24	2.24
GD_t		0.166	2.24***	0.02	0.117	1.27	1.13
π_t		-0.121	-0.84	0.01	-0.381	-1.95*	1.38
$TBILL_t$		0.520	2.99***	0.07	0.506	3.21***	4.88
LTB_t		0.496	2.23***	0.05	-0.667	-1.72*	2.34
ΔS_t		-0.996	-6.22***	0.49	-0.645	-4.54***	38.76
$\Delta IP_{CH,t}$		0.383	1.88*	0.03	0.059	0.85	1.75
$\pi_{CH,t}$		0.298	1.05	0.01			
ΔSCH_t		-1.378	-3.72***	0.00			
$\Delta RS_{CH,t}$		-0.241	-1.69*	0.02	-0.144	-1.66*	1.36
ΔOI_t		0.616	4.55***	0.19	0.167	2.49**	9.19
HP_t		-0.898	-6.10***	0.27	-0.494	-5.63***	19.66
ES_t		0.016	0.16	0.00			
ΔNRL_t		-0.051	-0.54	0.00			
g_{t-1}		0.540	5.09***	0.30	0.183	3.43***	0.70 17.31

Notes: Table 4 Panel 1 reports results from simple bivariate time-series regressions and Panel 2 reports results from multivariate regressions. The column β reports the slope coefficient, t is the t-statistic value for the slope coefficient. R-sq is the coefficient of determination. * is significant at 10% level, ** is significant at 5% level and *** is significant at the 1% level. $\Delta\sigma_{ECt}$ is the change in US economic uncertainty, ΔVIX_t is the change in the VIX Index. GD_t is Global Demand, π_t is US inflation. $TBILL_t$ is real US treasury bill rate. LTB_t is the real US Long-term Bond. ΔS_t is the change in the US exchange rate. $\Delta IP_{CH,t}$ is the growth in Chinese Industrial Production, $\pi_{CH,t}$ is the change in Chinese Consumer Prices (a measure of Chinese inflation), ΔSCH_t is the change in the Chinese Real Exchange Rate, $\Delta RS_{CH,t}$ is the change in Chinese Retail Sales. Commodity market measures are average changes across 17 commodities with CFTC data. ΔOI_t is the average change in open interest, HP_t is the average of hedging pressure (net commercial long position / open interest), ES_t is a measure of excessive speculation and ΔNRL_t is a measure of the change in non-reported long positions. g_{t-1} is the one period lag of the Common Factor.

TABLE 5. The Connections of the Change in Contribution of the Common Factor to Economic and Financial Indicators

1984:1-2017:6	Panel 1.1: Bivariate			Panel 1.2: Multivariate			
	Change Cont. CF			Change Cont. CF			
	β	t	r -sq	β	t	r -sq	Shapley
$\Delta\sigma_{ECt}$	0.022	0.41	0.00	-0.020	-0.65		0.20
ΔVIX_t	0.058	1.20	0.00	.			
GD_t	0.288	5.77***	0.08	0.054	1.84*		6.23
π_t	0.048	0.94	0.00	0.071	2.19**		6.08
$TBILL_t$	-0.206	-3.73***	0.04	.			
LTB_t	-0.306	-5.57***	0.09	.			
ΔS_t	-0.013	-0.25	0.00	.			
$\Delta IP_{CH,t}$	-0.039	-0.83	0.00	-0.012	-0.45		0.15
$\pi_{CH,t}$	0.049	0.84	0.00	.			
ΔSCH_t	0.193	2.70***	0.04	.			
$\Delta RSCH_t$	0.020	0.42	0.00	.			
ΔOI_t	-0.185	-0.35	0.00	0.042	1.30		0.14
HP_t	0.018	0.37	0.00	0.032	1.24		1.04
ES_t	0.056	1.20	0.00	-0.021	-0.78		0.10
ΔNRL_t	0.062	1.27	0.00	.			
$\Delta CONT_{CF,t-1}$	0.850	30.84***	0.74	0.779	20.28***	0.70	86.06

2004:1-2017:6	Panel 2.1: Bivariate			Panel 2.2: Multivariate			
	Change Cont. CF			Change Cont. CF			
	β	t	r -sq	β	t	r -sq	Shapley
$\Delta\sigma_{ECt}$	0.051	0.71	0.00	-0.042	-1.07		0.22
ΔVIX_t	0.120	1.80*	0.02	.			
GD_t	0.326	5.77***	0.18	0.081	2.13**		12.89
π_t	0.510	7.50***	0.27	.			
$TBILL_t$	0.208	1.85*	0.02	.			
LTB_t	-0.299	-2.59***	0.03	.			
ΔS_t	-0.011	-0.13	0.00	.			
$\Delta IP_{CH,t}$	-0.115	-0.92	0.00	-0.118	-1.88*		0.93
$\pi_{CH,t}$	0.269	2.26***	0.02	0.151	1.86*		1.41
ΔSCH_t	0.120	0.70	0.00	-0.152	-1.68*		0.36
$\Delta RSCH_t$	0.059	0.62	0.00	.			
ΔOI_t	-0.040	-0.43	0.00	.			
HP_t	0.122	1.23	0.01	0.064	1.25		1.60
ES_t	0.016	0.18	0.00	.			
ΔNRL_t	0.167	2.08***	0.02	.			
$\Delta CONT_{CF,t-1}$	0.848	18.40***	0.73	0.805	15.32***	0.75	82.58

Notes: Table 5 Panel 1 reports results from simple bivariate time-series regressions and Panel 2 reports results from multivariate regressions. The multivariate regressions are the second estimation which retains all variables with a t-value $\geq |1|$ from an initial estimation which included all explanatory variables. $\Delta CONT_{CF}$ designates the change in the Contribution of the Common Factor and is the dependent variable; the first lag of this is included as an explanatory variable. The column β reports the slope coefficient, t is the t-statistic value for the slope coefficient. R-sq is the coefficient of determination. * is significant at 10% level, ** is significant at 5% level and *** is significant at the 1% level. See Table 4 for variable descriptions.

Figure 1. Plots of Common and Sector factors

Panel 1: Common Factor and Sectoral Factors

Panel 2: Common Factor and Alternative Commodity Indices

Notes: In panel 1, the central line is the point estimate with the above and below giving the 95% confidence intervals.

Figure 2: Panel A: Plots Variance Contribution of the Common Factor

Notes: Figure 2 plots the variance contribution of the commodity common factor to individual asset returns over time. The centre solid black line is the point estimate with the dot and dash lines above and below giving the 95% confidence intervals.

Figure 2: Panel B: Plots Variance Contribution of the Common Factor (cont.)

Notes: Figure 2 plots the variance contribution of the commodity common factor to individual asset returns over time. The centre solid black line is the point estimate with the dot and dash lines above and below giving the 95% confidence intervals.

Figure 2: Panel C: Plots Variance Contribution of the Common Factor (cont.)

Notes: Figure 2 plots the variance contribution of the commodity common factor to individual asset returns over time. The centre solid black line is the point estimate with the dot and dash lines above and below giving the 95% confidence intervals.

Data Appendix

Global / US Global Economic Variables

$\Delta\sigma_{ECT}$ is the change in US economic uncertainty from the work of Bloom and downloaded from

<http://www.policyuncertainty.com/>

ΔVIX_t is the change in the VIX Index. VIX Index taken from <http://www.cboe.com/micro/vix/historical.aspx> from January 1990 to June 2017.

GD_t is Global Demand / MARKET_ECON_ACTIVITY is from Kilian (2009) and downloaded from his website: <http://www-personal.umich.edu/~lkilian/>

ΔS_t is the change in the (Broad) Trade-weighted US exchange Rate from the St Louis Fed series TWEXBMTH (available over ... only)

π_t is US inflation from IMF series I64...F.

TBILL_t is US treasury bill from IMF series I60C and is turned into a real series by adjusting for inflation.

LTB is the US Long-term Bond from IMF series I61 and is turned into a real series by adjusting for inflation.

Chinese Economic Variables

$\Delta IP_{CH,t}$ is the growth in Chinese Industrial Production from Datastream series CHIPTOT.H.

$\pi_{CH,t}$ is the change in Chinese Consumer Prices (a measure of Chinese inflation) from Datastream series CHCONPRCF.

$\Delta S_{CH,t}$ is the change in the Chinese Real Exchange Rate from Datastream series CHOCC011P.

$\Delta RS_{CH,t}$ is the change in Chinese Retail Sales from Datastream series CHRETUOTA.

Commodity Market Variables

Our variables from Commodity Futures Trading Commission (CFTC) are based on averages of the respective variables across the following 17 commodities with contract numbers in parentheses: WHEAT-SRW (001601 and 001602), SOYBEANS (005601 and 005602), SOYBEAN OIL (007601), NATURAL GAS (023651), COTTON NO. 2 (033661), ROUGH RICE (039601 and 039781), LEAN HOGS (054642), LIVE CATTLE (057642), CRUDE OIL - LIGHT SWEET (067651), COCOA (073732), PLATINUM (076651), SUGAR NO. 11 (080732), COFFEE C (083731), SILVER (084691), COPPER-GRADE #1 (085692), GOLD (088691) AND GASOLINE (111659 and 111652).

ΔOI_t is the average change in open interest (OI) based on data from the CFTC:
$$\Delta OI_t = \frac{1}{N} \sum_{n=1}^{N=17} \left(\frac{OI_t}{OI_{t-1}} - 1 \right)$$

HP_t is the average of hedging pressure. It is calculated as net commercial long position (NCL) / open interest. The net

commercial long position equals commercial long positions minus commercial short positions.
$$HP_t = \frac{1}{N} \sum_{n=1}^{N=17} \frac{NCL_t}{OI_t}$$

ES_t is a measure of excessive speculation. It is the average of the difference in net non-commercial long positions (NNCL) and

net commercial long positions scaled by open interest.
$$ES_t = \frac{1}{N} \sum_{n=1}^{N=17} \frac{(NNCL_t - NCL_t)}{OI_t}$$

ΔNRL_t measures the change in non-reported long positions (NRL) scaled by open interest.
$$\Delta NRL_t = \frac{1}{N} \sum_{n=1}^{N=17} \left(\frac{NRL_t}{OI_t} - \frac{NRL_{t-1}}{OI_{t-1}} \right)$$

g_{t-1} is the one period lag of the Common Factor.

ONLINE APPENDIX TO ACCOMPANY:

What Drives Commodity Returns? Market, Sector or Idiosyncratic Factors?

Contents:

Page 2 – Bayesian MCMC Estimation Algorithm estimation details to augment Section 3 of main paper (Text Appendix A.3 – A.4).

Page 8 - Selection of the mixing distribution as referred to in the Estimation Algorithm (Table A.1)

Page 9 – Full Descriptive Statistics (Table A.2)

Page 11 - Robustness of Regression Results to Outliers (Table A.6)

A.3. Bayesian MCMC Estimation Algorithm

The above DFM-SV is estimated using the Monte Carlo Markov Chain (MCMC) Bayesian estimation method by breaking the whole model into several blocks and making random draws from the conditional densities. Since most part of the model is linear and Gaussian the standard Gibbs-Sampling algorithm as outlined in Kim and Nelson (1999) readily applies. However, the time-varying stochastic volatility introduces nonlinear/non-Gaussian feature to the model and the standard Kalman Filter is no longer applicable. Therefore we employ the procedure developed by Kim, Shephard, and Chib (1998) that has been widely applied in the literature to make draws of the stochastic volatility (see, e.g., Stock and Watson (2007), and Primiceri (2005)). Their approach is based on a mixture of normal densities to approximate the underlying non-Gaussian density in making draws of the stochastic volatility. As a result, such an estimation algorithm allows us to conveniently make random draws of the parameters and state variables from conditional densities. After the algorithm converges numerical integrations are easily taken to yield the marginal distributions of model parameters as well as the states.

The MCMC Gibbs-Sampling algorithm is outlined below:

- (1) Make a random draw of the market and sectoral factors from the conditional distribution

$$f\left(\{g_t\}_{t=1}^T, \left\{\{s_{j,t}\}_{t=1}^T\right\}_{j=1}^m \mid \{\beta_i\}_{i=1}^n, \boldsymbol{\varphi}_g, \{\boldsymbol{\varphi}_{j,s}\}_{j=1}^m, \{\boldsymbol{\varphi}_{i,e}\}_{i=1}^n, \{\sigma_i^2\}_{i=1}^n, \{h_t^g\}_{t=1}^T, \left\{\{h_{j,t}^s\}_{t=1}^T\right\}_{j=1}^m, \left\{\{h_{i,t}\}_{t=1}^T\right\}_{i=1}^n\right), \text{ where } \boldsymbol{\beta}_i \text{ is}$$

the vector containing loading parameters as defined in eq. (2), $\boldsymbol{\varphi}_g = (\phi_1^g, \phi_2^g, \dots, \phi_p^g)$,

$\boldsymbol{\varphi}_{j,s} = (\phi_{j,1}^s, \phi_{j,2}^s, \dots, \phi_{j,l}^s)$ for $j = 1, 2, \dots, m$, $\boldsymbol{\varphi}_{i,e} = (\phi_{i,1}, \phi_{i,2}, \dots, \phi_{i,q})$ for $i = 1, 2, \dots, n$.

- (2) Make a random draw of the AR parameters and variance parameters for the commodity factor from the conditional distribution $f(\boldsymbol{\varphi}_{i,e}, \sigma_i^2 \mid \{g_t\}_{t=1}^T, \{s_{j,t}\}_{t=1}^T, \boldsymbol{\beta}_i, \{h_{i,t}\}_{t=1}^T)$. Note that due to the orthogonality property, these draws can be made one by one without loss of generality for each $i = 1, 2, \dots, n$ and

its corresponding $j = 1, 2, \dots, m$. To ensure a stationary process any draw of parameters that implies nonstationarity is discarded.

- (3) Make a random draw of the loading parameters from the conditional distribution

$f(\beta_i | \{g_t\}_{t=1}^T, \{s_{j,t}\}_{t=1}^T, \Phi_{i,e}, \sigma_i^2, \{h_{i,t}\}_{t=1}^T)$. Again due to the orthogonality property, these draws can be made one by one for each series for $i = 1, 2, \dots, n$.

- (4) Make a random draw of the AR parameters of the market and sector factors from the conditional distributions $f(\phi_g | \{g_t\}_{t=1}^T, \{h_t^g\}_{t=1}^T)$ and $f(\phi_{j,s} | \{s_{j,t}\}_{t=1}^T, \{h_{j,t}^s\}_{t=1}^T)$ for $j = 1, 2, \dots, m$. To ensure a stationary process any draw of parameters that implies nonstationarity is discarded.

- (5) Make a random draw of the stochastic volatility for the market factor from the conditional distribution $f(\{h_t^g\}_{t=1}^T, \sigma_g^h | \{g_t\}_{t=1}^T, \phi_g)$, for the sector factors from the conditional distribution

$f(\{h_{j,t}^s\}_{t=1}^T, \sigma_{j,s}^h | \{s_{j,t}\}_{t=1}^T, \phi_{j,s})$, for commodity factors from the conditional distribution

$f(\{h_{i,t}\}_{t=1}^T, \sigma_i^h | \{g_t\}_{t=1}^T, \{s_{j,t}\}_{t=1}^T, \beta_i, \Phi_{i,e})$

- (6) Repeat step (1) through (5) for $(D+S)$ number of times. Here D is the initial burn-in draws needed to be discarded and the results are based on the saved S number of draws once the convergence is achieved.

We follow Neely and Rapach (2011) and use 2 lags for all factors, i.e., $p=l=q=2$. Given initial parameters the above procedure can be started and iterated for $(D+S)$ number of times, and we set D to 2000 and S to 8000. The summary statistics are based on every 8th sampled draws after the initial burn in, and the results are robust to different positions of these sampled draws. We also increased the number of draws by ten times and the results remain robust. Below we explain further details of each of the above steps along with the priors employed in the estimation.

A.4.1. Draws of the market and sector factors

To reduce dimension of the state vector so as to facilitate the estimation, we follow Kim and Nelson (1999) and plug eq. (1) into (7) and this results in the following state-space representation of the above factor model:

Measurement Equation:

$$y_{i,t}^* = \mathbf{H}_t \cdot \mathbf{Z}_t + \sqrt{\exp(h_{i,t})} \cdot \varepsilon_{i,t} \quad (9)$$

Transition Equation:

$$\mathbf{Z}_t = \mathbf{F} \cdot \mathbf{Z}_{t-1} + \boldsymbol{\zeta}_t \quad (10)$$

Where, $y_{i,t}^* = (1 - \phi_{i,1}L - \phi_{i,2}L^2)y_{i,t}$, $\mathbf{H}_t = (\lambda_i, -\lambda_i\phi_{i,1}, -\lambda_i\phi_{i,2}, \mathbf{0}', \gamma_i, -\gamma_i\phi_{i,1}, -\gamma_i\phi_{i,2}, \mathbf{0}')$ and $\mathbf{0}$ is the zero vector with an appropriate dimension, and the state vector

$\mathbf{Z}_t = (g_t, g_{t-1}, g_{t-2}, s_{1,t}, s_{1,t-1}, s_{1,t-2}, \dots, s_{j,t}, s_{j,t-1}, s_{j,t-2}, \dots, s_{m,t}, s_{m,t-1}, s_{m,t-2})$ and here the i -th commodity belongs to its corresponding j -th sector. The matrix in the transition equation is

$$\mathbf{F} = \begin{pmatrix} \phi_1^g & \phi_2^g & 0 & & & & & & & & \\ 1 & 0 & 0 & & & & & & & & \\ 0 & 1 & 0 & & & & & & & & \\ & & & \phi_{1,1}^r & \phi_{1,2}^r & 0 & & & & & \\ & & & 1 & 0 & 0 & & & & & \\ & & & 0 & 1 & 0 & & & & & \\ & & & & & & \square & & & & \\ & & & & & & & \phi_{j,1}^r & \phi_{j,2}^r & 0 & \\ & & & & & & & 1 & 0 & 0 & \\ & & & & & & & 0 & 1 & 0 & \\ & & & & & & & & & & \square \\ & & & & & & & & & & & \phi_{m,1}^r & \phi_{m,2}^r & 0 \\ & & & & & & & & & & & 1 & 0 & 0 \\ & & & & & & & & & & & 0 & 1 & 0 \end{pmatrix}, \text{ where the blank areas are all}$$

zeros, and the shock vector is

$\zeta_t = \left(\sqrt{\exp(h_t^g)} \cdot \varepsilon_t^g, 0, 0, \sqrt{\exp(h_{1,t}^s)} \cdot \varepsilon_{1,t}^s, 0, 0, \sqrt{\exp(h_{j,t}^s)} \cdot \varepsilon_{j,t}^s, 0, 0, \sqrt{\exp(h_{m,t}^s)} \cdot \varepsilon_{m,t}^s, 0, 0 \right)$ with its variance matrix Q_t .

Given the above state-space representation and conditioning upon relevant parameters values, we implement the Kalman Filter and Carter and Kohn (1994) algorithm “filter forward, sample backwards” to draw the latent factors Z_t . See Kim and Nelson (1999) for details of the algorithm.

A.4.2. Draws of the model parameters in factors dynamics

Given a draw of the market and sector factors obtained from the last step and conditioning upon the loading parameters we compute the commodity factors $e_{i,t}$ from (1). Given factors the model parameters in factors dynamics such as the *AR* parameters and variance parameters in eq. (3), (5), and (7) can simply be obtained through linear regressions. The only complication is that in all these cases since the volatilities are time varying and hence the regression errors have a heteroskedasticity structure we need to first re-scale the variables in these equations to make the errors homoscedastic. In particular, divide both sides of eq. (3) by $\sqrt{\exp(h_t^g)}$, eq. (5) by $\sqrt{\exp(h_{j,t}^s)}$, and eq. (7) by $\sqrt{\exp(h_{i,t}^s)}$. After the transformation we make draws of the *AR* parameters from their posterior distributions based on (3), (5), and (7). For all these cases we use the conjugate priors and set the priors as the following:

$(\phi_{i,1}, \phi_{i,2}) \sim N(\mathbf{0}, \mathbf{I}_2)$, where $\mathbf{I}_2$ is the identity matrix and this is the same prior for all sets of the *AR*(2) parameters of all factors; $\sigma_i^2 \sim IG(0,0)$, where *IG* denotes the Inverted-Gamma distribution, and this setting-up ensures a diffuse prior for the variance parameters.

A.4.3. Draws of the loading parameters

To remove the serial auto-correlations in the error variables in eq. (1) we again subtract lagged

variables from both sides of eq. (1) and end up with the following linear regression for each commodity i :

$$y_{i,t}^* = \mathbf{K}_t \cdot \boldsymbol{\beta}_i + \sqrt{\exp(h_{i,t})} \cdot \varepsilon_{i,t} \quad (11)$$

Where, $y_{i,t}^* = (1 - \phi_{i,1}L - \phi_{i,2}L^2)y_{i,t}$, $\mathbf{K}_t = (g_t - \phi_{i1} \cdot g_{t-1} - \phi_{i2} \cdot g_{t-2}, s_{j,t} - \phi_{i1} \cdot s_{j,t-1} - \phi_{i2} \cdot s_{j,t-2})$, $\boldsymbol{\beta}_i = [\lambda_i \quad \gamma_i]$.

Given factors, time-varying volatility, and other parameters, equation (11) is nothing but a simple linear regression with known heteroskedasticity structure. Divide this equation by $\sqrt{\exp(h_{i,t})}$ to convert it to a linear regression with homoscedasticity and then the loading parameter $\boldsymbol{\beta}_i$ are easily drawn from its posterior distribution. We again employ the conjugate prior distribution with diffuse prior information:

$$\boldsymbol{\beta}_i \sim N(\mathbf{0}, \infty \cdot \mathbf{I}_2).$$

A.4.4. Draws of time-varying stochastic volatility

Since conditional on factors and corresponding parameters the time-varying volatilities for factors are independent from each other, we will first discuss the sampling algorithm for the market factor, and the same procedure applies to the sector and commodity factors volatilities.

First given a draw of the market factor and corresponding parameters, it is straightforward to compute the error variable in eq. (3):

$$g_t^* = g_t - \phi_1^g g_{t-1} - \phi_2^g g_{t-2} = \sqrt{\exp(h_t^g)} \cdot \varepsilon_t^g \quad (12)$$

Squaring and taking a natural logarithm of both sides of the above equation gives:

$$g_t^{**} = h_t^g + \varsigma_t^g \quad (13)$$

Where, $g_t^{**} = \ln(g_t^*)^2$, $\varsigma_t^g = \ln((\varepsilon_t^g)^2)$, and the time-varying volatility h_t^g follows a simple random walk process as stated in eq. (4), which is repeated here for reference:

$$h_t^g = h_{t-1}^g + \sigma_g^h \cdot v_t^g, v_t^g \sim i.i.d.N(0,1) \quad (4)$$

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

First note that the shocks ζ_t^g and v_t^g are independent. However, the shock ζ_t^g in the measurement equation is not normally distributed, and its distribution is in fact $\ln \chi^2(1)$. Kim, Shephard, and Chib (1998) provides an approach based on a mixture of normal densities to approximate this non-normal distribution when utilizing the Kalman Filter to draw the stochastic volatility. Specifically, they suggest using seven normal densities with means $m_k - 1.2704$ and variances τ_k^2 , for $k = 1, \dots, 7$, and the component probabilities are θ_k . They carefully choose these values to closely replicate the exact density of $\ln \chi^2(1)$. Table A.1 is replicated from Kim, Shephard, and Chib (1998) and reports these values.

Conditional on knowing g_t^{**} and the component probabilities the above state space model is approximately linear and Gaussian, and as a result we can apply the standard Carter and Kohn (1994) algorithm to sample the volatility. Given a sample of volatility, the component probabilities can then be updated using the Bayesian rule. The general procedure here follows Primiceri (2005), Koop and Korobilis (2010), and a correction procedure suggested by Del Negro and Primiceri (2013).

TABLE A.1: Selection of the mixing distribution to be $\ln \chi^2(1)$

ω	$\Pr(\omega = k)$	m_k	τ_k^2
1	0.00730	-10.12999	5.79596
2	0.10556	-3.97281	2.61369
3	0.00002	-8.56686	5.17950
4	0.04395	2.77786	0.16735
5	0.34001	0.61942	0.64009
6	0.24566	1.79518	0.34023
7	0.25750	-1.08819	1.26261

TABLE A.2 Descriptive Statistics for the Stochastic Volatility Dynamic Factor Model

Monthly Data: 1984M1 – 2017M6

	<i>Mean</i>	<i>SD</i>	<i>Min</i>	<i>PCT10</i>	<i>PCT25</i>	<i>PCT50</i>	<i>PCT75</i>	<i>PCT90</i>	<i>Max</i>
MARKET FACTOR	0.004	1.165	-10.937	-1.041	-0.445	-0.031	0.550	1.151	4.332
SECTORAL FACTORS									
1. INDUSTRIAL METALS	-0.073	2.174	-7.916	-2.379	-1.233	0.035	1.128	2.482	7.168
2. ENERGY	0.005	2.830	-11.276	-3.289	-1.485	-0.018	1.596	3.248	10.678
3. CEREALS	0.020	1.418	-5.416	-1.589	-0.684	0.051	0.846	1.496	5.796
4. OTHER FOOD & BEVERAGES	-0.011	0.346	-1.274	-0.503	-0.238	0.003	0.234	0.396	1.266
5. LIVESTOCK	-0.002	0.796	-2.735	-0.933	-0.462	0.012	0.430	0.930	2.946
6. PRECIOUS METALS	-0.015	1.092	-4.027	-1.310	-0.686	-0.037	0.633	1.424	3.744
7. VEGETABLE OILS	-0.033	0.643	-2.425	-0.776	-0.311	-0.031	0.252	0.669	2.876
8. INDUSTRIAL RAW MATERIALS	-0.002	0.127	-0.369	-0.163	-0.090	-0.007	0.083	0.162	0.401
INDIVIDUAL FACTORS									
INDUSTRIAL METALS									
ALUMINUM	0.065	4.137	-28.565	-3.909	-1.803	-0.146	2.054	4.595	15.141
COPPER	0.074	3.950	-15.906	-3.812	-1.594	-0.056	1.784	3.693	20.838
LEAD	0.099	4.954	-22.714	-4.800	-2.248	-0.245	2.925	5.417	20.691
NICKEL	0.092	6.497	-24.474	-6.674	-3.048	-0.249	2.865	7.290	54.501
TIN	0.039	4.615	-24.658	-4.792	-2.171	-0.151	2.205	5.165	17.290
ZINC	0.091	4.092	-22.045	-4.294	-2.287	0.012	2.131	4.977	12.051
URANIUM	0.001	6.889	-31.336	-6.241	-2.331	-0.111	2.303	6.871	37.379
ENERGY									
FUEL OIL	-0.013	4.778	-39.074	-1.108	-0.274	0.004	0.365	1.478	38.674
PROPANE	-0.011	10.491	-56.130	-9.856	-4.429	-0.214	4.717	10.158	99.871
HEATING OIL	-0.011	1.408	-10.792	-1.571	-0.738	0.010	0.694	1.522	6.200
GASOLINE	-0.013	8.637	-28.830	-10.244	-4.605	0.336	4.483	10.088	35.812
OIL - WTI	-0.011	6.290	-28.387	-7.703	-3.341	0.186	3.437	7.209	23.232
CEREALS									
BARLEY	-0.038	5.370	-17.911	-6.043	-2.835	0.017	2.646	5.922	27.273
MAIZE	-0.059	2.560	-17.649	-2.673	-1.046	-0.010	0.917	2.303	9.747
RICE	-0.002	6.014	-28.338	-6.703	-2.702	-0.137	2.287	6.043	41.106
SOYBEANS	-0.045	3.123	-16.664	-3.960	-1.821	0.022	1.808	3.375	9.823
WHEAT	-0.037	4.657	-18.875	-5.614	-2.720	-0.312	2.702	5.868	18.398

Notes: This table provides descriptive statistics for all the factors. Please note mean is the sample average of the variable. SD is the standard deviation, MIN is the minimum value, then we report following percentiles (PCT) of the distribution: 10th percentile (PCT10), lower quartile (PCT25), median (PCT50), upper quartile (PCT75) and 90th percentile (PCT90). Finally the maximum is reported.

TABLE A.6. Robustness of Regression Results to Outliers

	Panel A.1: Multivariate				Panel A.2: Multivariate			
	Common Factor			Shap	Common Factor			Shap
	1984:1-2017:6				2004:1-2017:6			
	β	t	r-sq		β	t	r-sq	
$\Delta\sigma_{ECt}$	.							
ΔVIX_t	.							
GD_t	0.038	0.91		3.13				
π_t	-0.055	-1.30		0.98				
$TBILL_t$	0.144	2.24**		1.08	0.268	3.03***		7.27
LTB_t	-0.048	-0.59		1.02				
ΔS_t	-0.272	-7.68***		31.31	-0.332	-4.79***		31.62
$\Delta IP_{CH,t}$	.				0.072	1.23		2.09
$\pi_{CH,t}$	0.027	0.91		0.61				
$\Delta S_{CH,t}$	.							
$\Delta RS_{CH,t}$	.				-0.072	-1.2		1.37
ΔOI_t	0.090	2.98***		6.21	0.119	2.12**		10.25
HP_t	-0.273	-8.12***		30.74	-0.511	-6.73***		33.55
ES_t	.							
ΔNRL_t	.							
g_{t-1}	0.234	3.92***	0.54	23.70	0.076	1.09	0.61	13.86
Panel B.1: Multivariate					Panel B.2: Multivariate			
Change in Cont. of CF					Change in Cont. of CF			
1984:1-2017:6					2004:1-2017:6			
	β	t	r-sq	Shap	β	t	r-sq	Shap
$\Delta\sigma_{ECt}$	.				.			
ΔVIX_t	.				.			
GD_t	0.050	1.76*		5.73	0.080	2.12**		11.89
π_t	0.071	2.25**		5.70	.			
$TBILL_t$	.				-0.021	-0.45		1.63
LTB_t	.				.			
ΔS_t	.				.			
$\Delta IP_{CH,t}$	-0.013	-0.5		0.14	-0.126	-2.01**		1.00
$\pi_{CH,t}$	.				0.132	1.60		1.57
$\Delta S_{CH,t}$	.				-0.146	-1.64		0.35
$\Delta RS_{CH,t}$	.				.			
ΔOI_t	0.039	1.27		0.13	.			
HP_t	0.032	1.26		1.10	0.071	1.39		1.67
ES_t	.				.			
ΔNRL_t	.				.			
$\Delta CONT_{CF,t-1}$	0.774	20.50***	0.69	87.20	0.797	15.14***	0.75	81.89

Table 6 reports results with the dependent variable winsorized at the 0.025 and 0.975 level. Panel A reports results for the common factor as the dependent variable, and B is for the change in contribution of the common factor (Change in Cont. CF). Panel .1 reports results from the multivariate regression for the full sample and Panel .2 reports results from multivariate regressions for the 2004:1-2017:6 sub-period. These multivariate regressions are the second estimation which retains all variables with a t-value $\geq |1|$ from the initial estimation which included all explanatory variables. The column β reports the slope coefficient, t is the t-statistic value for the slope coefficient. R-sq is the coefficient of determination. * is significant at 10% level, ** is significant at 5% level and *** is significant at the 1% level. See Table 4 for variable descriptions.