Economics Bulletin

Volume 33, Issue 4

Common risk factors in commodities

Julien Chevallier
IPAG Business School (IPAG Lab)

Florian Ielpo
LOIM, CERMSEM, U. Paris 1

Ling-Ni Boon
Amundi, University Paris Dauphine

Abstract

This article aims at establishing an understanding of the common risk factors in commodity markets, as well as their interactions with equities, currencies and interest rates. Since commodity markets often exhibit cross-sectional dependency, common risk factors exist and can be identified. By using daily data from 1995 to 2012, the econometric methodology resorts to factor modeling combined with a criterion to determine the number of factors presented in Alessi et al.(2010). The operational significance of the results is to evaluate risk-adjusted performance of portfolios allocated to commodities, and to help building cross-asset strategies. Investors can then pinpoint the correlation between any two-position taken within commodity markets, and attempt to profitably exploit the common sources of risk. In turn, it should provide the researcher with an increased understanding of the risks at work in the commodity world.

Disclaimer: The views and opinions expressed in this article are those of the authors, and do not necessarily reflect the views and opinions of Lombard-Odier Investment Management.

Citation: Julien Chevallier and Florian Ielpo and Ling-Ni Boon, (2013) "Common risk factors in commodities", *Economics Bulletin*, Vol. 33 No. 4 pp. 2801-2816.
Contact: Julien Chevallier - julien.chevallier04@univ-paris8.fr, Florian Ielpo - f.ielpo@lombardodier.com, Ling-Ni Boon - lingni.boon@amundi.com.
Submitted: December 14, 2012. **Published:** November 05, 2013.

1 Introduction

Measuring common factors affecting a group of assets' returns is a natural step towards an enhanced understanding of a given investment universe. In the equity case, Fama and French (1993)'s three factors – market risk (*i.e.* excess return of the market), size (*i.e.* market capitalization) and value (*i.e.* book-to-market ratio) – helped to improve the existing knowledge on the nature of equity portfolios' performances. Litterman and Scheinkman (1991) made an equivalent contribution to the bond market-related literature, highlighting how yield curves were primarily affected by three factors: level, slope and curvature. Commodities being regarded as an alternative asset class and given their heterogeneous structure (Erb and Harvey, 2006, Kat and Oomen, 2007 a,b), little empirical evidence has been gathered so far regarding the nature of potential common factors explaining their cross-sectional dynamics. This article aims at bridging this gap, using factors analysis.

When it comes to commodities, research has been mainly focused on single assets. Pricing single commodities using stochastic discount factor-based settings such as the Capital Asset Pricing Model (CAPM) (Dusak, 1973, Bodie and Rosansky, 1980) or the Consumption CAPM (CCAPM) (Breeden, 1980) delivered poor performances. Studies applying the CCAPM on multiple commodities – such as Jagannathan (1985) and de Roon and Szymanowska (2010) – concluded that such equity-like approaches fail when being applied to commodities, as they are unable to cope with the uniqueness of commodities' features. Their findings are consistent with the fact that correlations between commodities and more traditional assets appear limited in the long run. Therefore, commodity specific factors need to be the focus of empirical experiments around such an investment universe.

A vast research effort has been thus undertaken to improve our understanding of the factors specifically affecting commodity returns. Nevertheless, many studies are still focused on a single-asset level: Stoll (1979), Hirschleifer (1988, 1989) and de Roon et al. (2000) assess the impact of systematic factors and hedging pressure on single commodity futures. Drawing on the theory of storage[1] (Working, 1949), Gorton et al. (2013) investigated the relation between commodity inventory levels and commodity futures' expected returns. Acharya et al. (2013) produced an equilibrium model of commodity markets based on a capital constraint on speculators and a hedging demand stemming from producers. Hence, the heterogeneity of commodities is usually assumed to be so strong that it makes the quest for common factors useless.

Daskalaki et al. (2012) is by far the largest study available aiming at finding factors affecting the cross-section of a group of 22 commodities[2]. Their work focuses on testing a very large number of potential models and factors, covering both equity-related factors – without much success – and a spectrum of other global factors, such as brokerage houses' leverage, monetary policy related factors or commodity hedging pressure variables. Again, they conclude by the fact that there are little pieces of evidence that commodity markets are affected by such factors – especially from a cross-sectional perspective – consistently again with the idea that commodities provide investors with an investment vehicle that offers a weak correlation to standard assets and among themselves. In a final attempt, they use Principal Component factors in two steps Fama-MacBeth regressions, finding again a low explanatory power over the time series evolution of the group of commodities that they investigate. Still, the Fama-MacBeth regression stands a good chance to fail in capturing the significance of a factor affecting only a group of variables under the scope of the investigation – and commodities are likely to be affected by sector influences. This remains a rather minor part of their research piece and our article aims at further developing this point using recent econometric tests making it possible to determine the number of factors affecting a group of variables.

In this article, we extend the work of Daskalaki et al. (2012) by using Alessi et al. (2010)'s criterion to determine the number of factors to be included in a factor model – as opposed to progressively testing PC models with an increasing number of factors. This criterion is selected for its accuracy demonstrated by the various Monte Carlo experiments available in their article. We use a large dataset of 25 commodities – including the GSCI sector indices – over the 1995-2012 period to which we add equities, foreign exchange rates and interest rates in order to gauge the specificity of commodity factors. Our analysis provides an economic identification of the common factors and analyzes their dynamics over time. Our investigations are related to previous work on factor models, as is Alessi et al. (2010), such as Bai and Ng (2002), Forni et al. (2000), Forni et al. (2009) and Stock and Watson (2005). A part of our findings is consistent with the existing literature and a part is not: first, we find that there is only one factor explaining the cross-section of commodities' returns, explaining 28% of the total variations of our dataset when the equity factors – the three of them – explain 75% of equity variations. Across all the types of datasets investigated here, the commodity dataset's common factor has thus the weakest explanatory power. A second empirical conclusion is that standard assets' factors explain very poorly commodities' returns – consistent with the previously quoted references. Finally, when considering standard and non-standard assets all together, commodities still appear as one of the 7 factors of the investment universe to which a global macro hedge fund has typically access, highlighting their importance from a global financial perspective.

The rest of the article is organized as follows. Section 2 briefly details the PCA and factor modeling approaches. Section 3 contains the empirical results. Section 4 concludes.

2 PCA and the estimation of the number of common factors

The joint behavior of commodities is analyzed through a factor model specification, making it possible for a group of variables to either rely on components that are common or specific to each of them. An r-factor *approximate factor model* is represented as such:

$$\begin{aligned} X_{it} &= \lambda_i' F_t + e_{it} \\ i = 1 \ldots N \quad &and \quad t = 1 \ldots T, \end{aligned} \tag{1}$$

where X_{it} is the return of the i^{th} asset at time t, F_t is the $r \times 1$ vector of unobservable common factors, $F_t = (F_{1t}, F_{2t}, \ldots, F_{rt})$, λ_i is the $r \times 1$ vector of factor loadings, e_{it} is the idiosyncratic component and $'$ denotes the transpose of the matrix. The following assumptions are imposed: (1) F_t and e_t are uncorrelated, so that a common factor cannot be reflected in the specificities of a given variable. (2) The matrix Ω comprised of $cov(e_{it},\, e_{jt})$ is not necessarily diagonal, allowing for serial correlation and heteroskedasticity, but the degree of correlation between the idiosyncratic components is limited, *i.e.* the largest eigenvalue of Ω, the $N \times N$ covariance matrix of the idiosyncratic component, is assumed to be bounded[3]. The common factors are assumed to be unobservable and must be thus measured from the cross-section of the variables X_{it}, and only a subset of them are expected to be statistically significant. The estimation of the number of factors is performed using the criterion provided in Alessi et al. (2010). The factors are estimated using a Principal Component Analysis. We now briefly review the key steps of the methodology.

The first r principal components are obtained by solving a minimization problem, set up to yield the sum of squared residuals:

$$V(r) = \min_{\Lambda, F^r} \frac{1}{NT} \sum_{i=1}^{N} \sum_{t=1}^{T} (X_{it} - \lambda_i^r F_t^r)^2$$

with $\Lambda = (\lambda_1, \lambda_2, \ldots \lambda_N)'$, $F = \left(F_1^k, F_2^k, \ldots, F_T^k\right)'$ with F_i^k a vector containing for date i the first k factors, subject to either $\frac{F'F}{T} = I_r$ and $\Lambda'\Lambda$ is diagonal, or $\frac{\Lambda'\Lambda}{N} = I_r$ and $F'F$ is diagonal.

Linear algebraic manipulations demonstrate that the result of this minimization problem is essentially the ordered eigenvectors corresponding to the asset return's covariance matrix.

As for Step 1, determining the number of factors to include, numerous criteria have been developed but arguably the most popular technique consists in using an information criteria. This approach is based on the idea that an $(r + 1)$-factor model can fit no worse than an r-factor model, but is less efficient. The balance between parsimony and explicability is evaluated via a loss function, defined as

$$V(r, F^r) + rg(N, T) \tag{2}$$

or

$$\log(V(r, \hat{F}^r)) + r\bar{\sigma}^2 g(N, T), \tag{3}$$

whereby $V(r, F^r) = \min_{\Lambda} \frac{1}{NT} \sum_{i=1}^{N} \sum_{t=1}^{T} (x_{it} - \lambda_i^{r'} F_t^r)^2$ is the value function, $g(N, T)$ is the penalty for over-fitting, r is a constant, and $\bar{\sigma}^2$ is a consistent estimate of $\frac{1}{NT} \sum_{i=1}^{N} \sum_{t=1}^{T} E[e_{it}]^2$, which in practice can be replaced by $V(r_{max}, \hat{F}^{r_{max}})^4$. N is the cross-section dimension while T is the time dimension. The estimated number of factors is then the r corresponding to the lowest value of the loss function among those considered.

The criterion adopted in our analysis is proposed by Alessi, Barigozzi, and Capasso (ABC), which is a refinement of the one by Bai and Ng (2002). ABC defines a refined loss function and evaluate it over a range of the constant and over random subsamples of the data. The estimated number of factors is then the number that is insensitive to neighboring values of the constant, and has no dependance on the subsamples. The purpose of the constant is to tune the penalizing power of $g(N, T)$, resulting in an estimate that is empirically more robust than when this constant c is fixed. The number of factors is estimated using the following formula, that is a modified version of Bai and Ng (2002)'s original criterion:

$$r^{\text{opt}} = \underset{0 \le r \le r_{\max}}{\text{argmin}} \; V(r, F^r) + crg(N, T). \tag{4}$$

The results presented below are those obtained by using the following penalty functions:

$$g_1(N,T) = \frac{N+T}{NT}\log(\min(\sqrt{N},\sqrt{T})^2) \tag{5}$$

$$g_2(N,T) = (N+T-k)\frac{\log(NT)}{NT}. \tag{6}$$

g_1 is frequently used in empirical work due to its stability. g_2 has been shown to have good properties when errors are cross-correlated. The choice of ABC's criteria as opposed to other methods, *e.g.* Bai and Ng (2002), Connor and Korajczyk (1993), Onatski (2009) is motivated by a Monte Carlo Study implemented on financial data comprising equities, commodities, credit spreads, interest rates, and currencies (Boon and Ielpo, 2013) and by the Monte Carlo experiments conducted in Alessi et al. (2010). The results demonstrate that ABC's criterion is superior in accuracy (overall best in the Monte Carlo study, even when cross-section and serial correlation exist in the data), and precision (less sensitive to whether linear dependencies exist in the financial data, yielding the same estimation regardless of whether the criterion is applied to the data, or the vector autoregressive residuals).

3 Empirical findings

By using this empirical approach, we tackle the issue of the factors priced in commodity markets, relying on the dataset listed in Table 1. We do so by highlighting four different stylized facts:

– We present evidence that the risk factors priced in other asset classes are only weakly priced in commodities.

– The concentration of risk factors is much weaker in commodities: the first factor of commodities explains only a weak part of this investment universe.

– By estimating the number of common factors over each dataset, we find a single common factor in commodities.

– By identifying the number of common factors in a global dataset, we find that commodities are still contributing with other asset classes to the joint evolution of financial markets.

The data comes from Bloomberg from 1995 to 2012 with a daily frequency, totalling 2,757 observations. The characteristics of the time series are given in Table 1. Note that GSCI stands for the Goldman Sachs Commodity Index (with specific sub-indices for agricultural products, metals and energy commodities). In subsequent analysis, equities and commodities are converted to returns, foreign exchange rates are converted to log-returns and interest rates to variations.

Prior to determining the number of factors using ABC's criteria, the explanatory power of the first factor in each dataset is analyzed. By doing so, we aim at analyzing the concentration of correlations in commodity markets, and comparing it to other markets. The decomposition of the amount of the variance explained by each factor is obtained by taking the ratio of the eigenvalue associated to each factor, and the sum of all eigenvalues of the entire return covariance matrix. A very concentrated market – that is a market for which correlations across assets' variations are very high – should deliver a ratio that is very

Data Composition (60 assets)	
Equities (15 assets)	Dow Jones Industrial Average, S&P 500, Nasdaq 100, S&P TSX, Mexico IPC, Brazil Bovespa, Eurostoxx, FTSE, CAC 40, DAX, IBEX 35, Swiss Market Index, Nikkei 225, Hang Seng, S&P ASX 200.
Interest rates (8 assets)	US 30Y, US 10Y, US 5Y, US 2Y, GE 30Y, GE 10Y, GE 5Y, GE 2Y.
Foreign Exchange (12 assets)	EUR, CAD, YEN, AUD, NZD, GBP, CHF, SEK, NOK, ZAR, MXN, TWD
Commodities (25 assets)	Metals: Gold, silver, platinum, aluminium, copper, nickel, zinc, lead. Fuel: WTI, Brent, Gasoil, Natural Gas, Heating Oil. Softs: Coffee, Sugar, Cocoa, Cotton. Grains: Corn, Wheat, Soybean, Rice. Financials: GSCI Agriculture, Energy, Industrial Metals, Precious Metals.

Summary statistics by asset category

Category	Mean (Annualized Return)	Median	Standard Deviation (Annualized Return)	Min	Max
Equities return	11%	1.725×10^{-4}	42%	-0.158	0.334
Interest variation in rates	108%	-2.0×10^{-4}	0.6%	-0.533	0.473
Foreign Exchange log return	25%	0	45%	-0.186	0.168
Commodities geometric return	5%	0	24%	-0.156	0.189

Table 1: Composition of the dataset and summary statistics. The dataset is made of weekly data and covers a period starting on the 19^{th} of January 1994 and ending on the 24^{th} August 2012.

close to 100%. On the contrary, when a market exhibits a low cross-asset correlation – as we suspect in the case of commodities since Gorton and Rouwenhorst (2006) – this ratio should be closer to 0%.
Such results are presented in Table 2. As expected, this concentration is quite high in the case of interest rates' variations for which the first factor obtained from the PCA explains 69% of the dataset. Equities are not far behind, as their first factor explains 49% of equities' variations. Next is the currency asset class, for which the first PCA factor explains 45% of the dataset. Finally, the first factor of commodities explains 28% of the returns on commodities. Thus, commodities obtain the weakest value in this first investigation of risk concentration. This finding is clearly in line with previous literature, underlining the diversification effects obtained when investing into commodities.
To get a general idea about which assets are priced in individual commodity returns, each commodity's returns is then regressed on each of the previously estimated first factors by using an OLS estimation of the form:

$$r_t^i = \alpha + \beta r_t^A + \epsilon_t, \tag{7}$$

where r_t^i is the close-to-close return on the i^{th} commodity, r_t^A is the return on the first factor of the asset class A, and ϵ_t has a distribution with mean 0 and variance σ_ϵ^2.

	Equities	Interest rates	FX	Commodities
Proportion of variance explained	49%	69%	45%	28%

Table 2: Explanatory Power of the First Factor

In this CAPM-like framework, the higher the absolute value of β and the R^2 associated to the regression, the higher the explanatory power of the risk factor considered. The slope coefficient[5] and R^2 are reported in Table 4. Low R^2 levels are observed when commodity assets are regressed on the first factor of equities and interest rates. Yet, the R^2 is in general higher when the regressions are run on the first factor of foreign exchange data and commodities. The former is likely, because prices of commodities are often labeled in US Dollar (i.e. the first factor among the G10 currencies). The latter suggests that despite the high heterogeneity, there exist common factors that are priced in commodity assets, which we are interested in uncovering. These findings are consistent with the idea that including commodities in a bond or an equity portfolio should yield to an increased diversification, given that those three asset classes have been pricing different risk factors over the period considered. Consistently with previous studies, commodities are pricing risk factors that are common to those priced in currency markets. These findings are consistent with the empirical conclusions obtained by Diaskalaki and Skiadopoulos (2011): the usual risk factors are useless when it comes to explaining the returns on commodities. There are a couple exceptions: copper exhibits a stronger sensitivity to equities, with a R^2 that is equal to 11.94%. Nickel, zinc and lead also exhibit an increased sensitivity to the equity factor over the period covered here. Hence, industrial metals partly price equity risk.

Now, focusing on the pure commodity dataset, we perform an identification of the first five factors driving commodities. To identify the factors, the correlation between the estimated factors and the assets whose variances the factors are trying to explain is investigated. When the assets are too numerous, then selected assets are chosen such that when the factor is regressed upon the set of assets, at least 95% R^2 is attained. By interpreting the sign and magnitude of the correlations, an attempt to label the factors is made in order to establish economic sense. We perform this identification so that to be able to compare the results obtained with ABC's test to a correlation analysis: the economic sense of our findings should be consistent with the statistical findings detailed later.

Figures 1 presents the correlations between commodity returns and the returns over each of the first five factors obtained from the PCA analysis. The correlation bar plots indicate that factor 1 is a global commodity factor, with a stronger exposure to energy commodities. Factor 2 is an industrial metals factor. Factor 3 is an oil vs. gas factor, as these two commodities have a strong and opposite correlation sign with this factor. Factor 4 is an agricultural factor. Factor 5 is another agricultural factor, opposing coffee to grain commodities. From this empirical analysis, we conclude that analyzing commodities by sector makes sense: there are common risk factors to commodity sectors such as industrial metals or agricultural products. Despite the fact that commodities from a given sector are used for various purposes, they still share common risk factors[6]. The sector feature explains part of the commodity risk

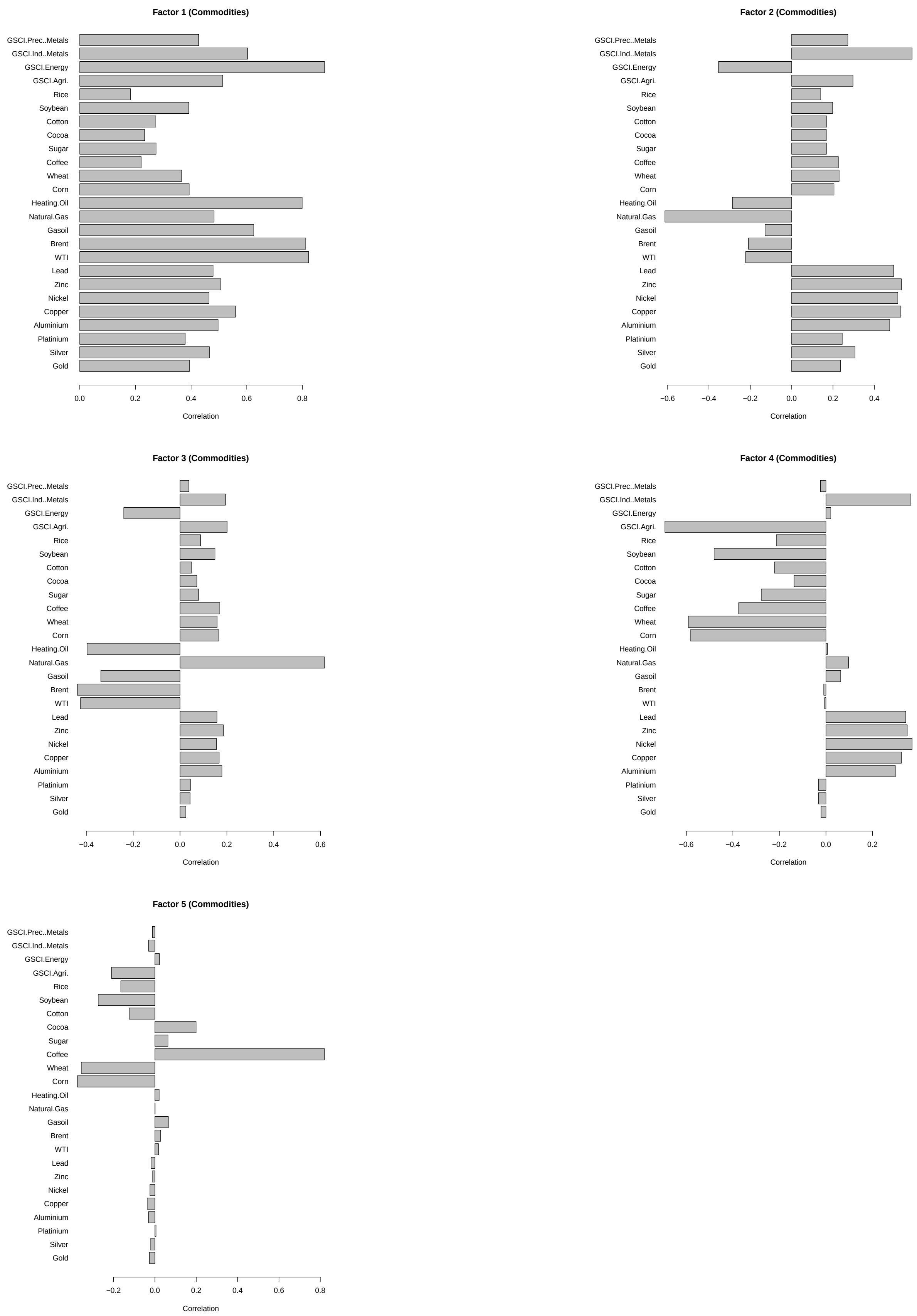

Figure 1: Factor Loadings Bar Plots for the First Five Factors of the Commodities Dataset

factor conundrum: but the factor 1, the rest of the commodity factors are not common to *all* commodities, but to subsets of them that match the usual definition of commodity sectors.

Dataset	Equities	Interest rates	FX	Commodities
Estimated Number of Factors by ABC	3	2	1	1
Proportion of Variance Explained	75%	88%	45%	28%

Table 3: Estimated Number of Factors and Proportion of Variance Explained by Dataset

Next, we estimate the number of factors that explain the cross-section of each of the asset classes considered here by using the ABC criterion. These results are presented in Table 3. Equities are estimated to have three factors, in line with Fama and French's (1993) model. These three factors explain up to 75% of equities fluctuations. Interest rates are found to be driven by two factors that explain 88% of their variations, consistently with the findings by Litterman and Scheinkman (1991)[7]. Hence, these two investment universes exhibit a strong concentration of risk around common risk factors. Only a weak part of their varitations is explained by asset-specific factors that market observers would call 'alpha' (or idiosyncratic risk). When it comes to currencies or commodities, we obtain a different picture: the estimation scheme only diagnoses a single common factor in these respective datasets, which explains less than half of the variance. Asset-specific risk factors dominate the evolution of these markets. This finding is key to investment managers. Indeed, it appears that commodities and currencies have a strong bottom-up side, and a more limited top-down side.

	Equities		Rates		Currencies		Commodities	
	β	R^2 (%)	β	R^2 (%)	β	R^2 (%)	β	R^2 (%)
Gold	0.01*	0.13	0*	0.38	-0.25*	16.68	0.09*	15.17
Silver	0.08*	3.1	0*	0.12	-0.45*	17.33	0.19*	21.9
Platinium	0.06*	3.31	0.01*	0.75	-0.26*	10.79	0.12*	14.66
Aluminium	0.09*	7.57	0.01*	2.14	-0.25*	11.62	0.14*	25
Copper	0.14*	11.94	0.02*	2.58	-0.36*	14.62	0.2*	31.91
Nickel	0.13*	5.88	0.02*	1.27	-0.34*	6.7	0.23*	22.89
Zinc	0.12*	8.24	0.02*	1.71	-0.33*	10.82	0.2*	26.92
Lead	0.13*	7.79	0.02*	1.84	-0.37*	10.82	0.21*	24.27
WTI	0.11*	3.82	0.02*	1.54	-0.33*	6.78	0.38*	64.23
Brent	0.1*	3.74	0.02*	1.22	-0.31*	6.43	0.36*	63.2
Gasoil	0.06*	1.65	0.01*	0.38	-0.27*	5.69	0.27*	38.87
Natural.Gas	0.04*	0.2	0.01	0.05	-0.16*	0.67	0.37*	24.91
Heating.Oil	0.08*	2.46	0.01*	0.72	-0.29*	5.46	0.36*	61.26
Corn	0.06*	1.96	0.01*	0.39	-0.21*	4.24	0.15*	15.6
Wheat	0.06*	1.76	0.01*	0.57	-0.22*	3.98	0.15*	13.3
Coffee	0.07*	1.48	0.01*	0.27	-0.2*	1.98	0.13*	5.37
Sugar	0.05*	1.07	0.01*	0.37	-0.2*	2.64	0.13*	7.82
Cocoa	0.05*	0.9	0.01*	0.59	-0.24*	4.77	0.1*	5.81
Cotton	0.06*	2.01	0.01*	0.92	-0.18*	2.86	0.11*	7.74
Soybean	0.06*	2.48	0.01*	0.93	-0.2*	4.97	0.13*	15.54
Rice	0.04*	1.05	0*	0.09	-0.13*	1.71	0.07*	3.52

Table 4: Slope Coefficient of Commodity Asset by First Factor of Each Dataset

Finally, in order to improve our understanding of the interactions of commodities with other asset classes, we now run the ABC criterion over the full dataset considered as a whole. Commodities are analyzed in an expanded investment universe consisting of other assets such as equities, currencies, and interest rates by compiling the datasets analyzed above into a single dataset. The purpose is to reveal commodities' influence from a global perspective. If commodities are a driving force in financial markets, then they should be linked to the common factors of such a dataset. This could be evidence of integration of commodities into global financial markets. Furthermore, the decomposition of the percentage of the variance explained by commodity-linked factors divulges the extent of such integration, and could shed light on the observation that in the long run, commodities are uncorrelated with other asset classes but are linked to them in the short run. Seven factors are estimated by using ABC's criterion, as detailed in Table 5.

These factors are identified as previously through a correlation analysis, as presented in Figure 2 to 4. Factor 1 is a factor that opposes equity to bonds, and can be considered as a 'risk appetite' factor. During bullish periods, investors are more willing to take risks, and prefer to invest into risky assets (equities) than into riskless assets (bonds). Conversely, during bearish periods, investments are diverted from equities into bonds. Factor 2 is a factor that is strongly correlated to the US Dollar[8], and negatively with commodities (labelled in USD). Factor 3 is strongly related to interest rates, whereas factor 4 is positively correlated to the Euro and negatively to the energy sector. Factor 5 is an industrial metals factor, factor 6 is an agricultural factor, and factor 7 is a Euro vs. US rates. Three of the seven factors (*i.e.* factors 4 to 6) are linked to commodities and explain about 10% of variances in the return (or about 12% of the total variances accounted for by all seven factors). This contribution is not limited to the negative relation between the US Dollar and the price of US Dollar labelled in commodities: it includes additional cross-asset factors that jointly explain the evolution of financial assets.

4 Concluding remarks

In this article, we resort to factor modeling techniques to assess the presence of common risk factors in commodities. We find evidence that corroborates the heterogeneity of commodities: the explanatory power of the first factor of the commodity dataset (28%) is lower than for other asset classes (such as equities, interest rates or currencies). When a criterion is used to determine the number of factors, only one common factor is estimated for commodities, suggesting that variances in commodity markets are largely due to idiosyncracies that cannot be accounted for by common factors.

On a brighter side, we also find by using a global dataset that commodities are reasonable driving forces of a global macro investment universe, explaining a small part (10 to 12%) of the joint mechanics of worldwide financial markets. Hence, our analysis points towards the possibility that commodities are somewhat integrated to financial markets, even though they price only weakly bond and equity risks. This is however only a static view: further research is needed to assess how commodities and traditional

Figure 2: Factor Identification - Global Macro Data (Factors 1-3)

assets actually interact from a dynamic perspective. Such a further work would need to be related to the approaches presented in Forni et al. (2009) and Forni and Gambetti (2010)[9].

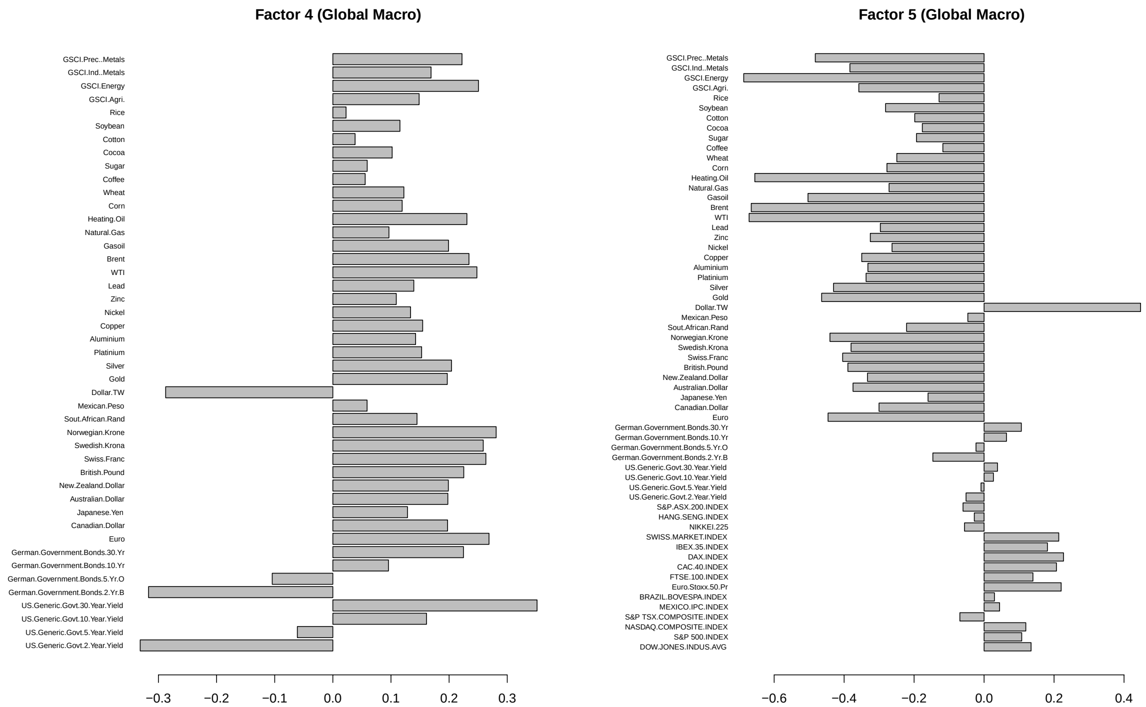

Figure 3: Factor Identification - Global Macro Data (Factors 4-5)

Factor	**1**	**2**	**3**	**4**
Label	Equities vs. Bond	US dollar	Interest Rates	Euro vs, energy
Proportion	54%	15%	7.4%	4.9%
Cumulative	54%	69%	76.4%	81.3%

Factor	**5**	**6**	**7**
Label	Indutrials Metals	Agricultural	Euro rates vs. US rates
Proportion	2.9%	2.5%	1.6%
Cumulative	84.2%	86.7%	88.3%

- “Proportion” indicates the proportion of variance explained by each principal component.
- “Cumulative” is the proportion of variance explained by principal component(s) up to that order.

Table 5: Result Summary for Global Macro Data

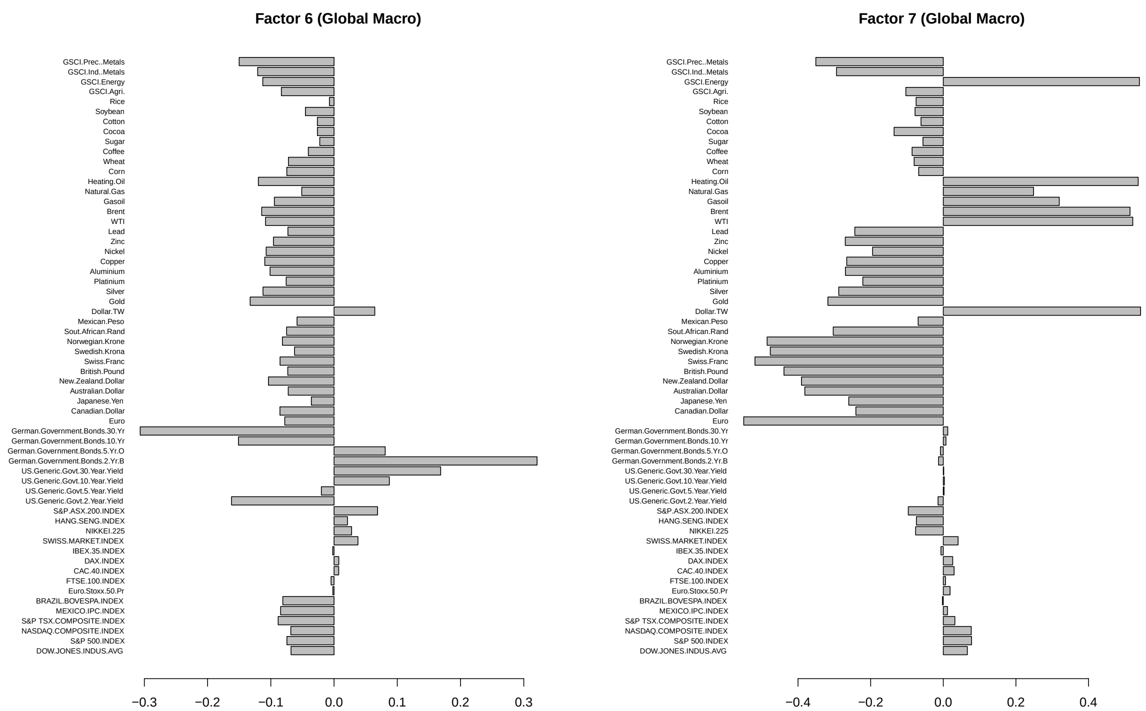

Figure 4: Factor Identification - Global Macro Data (Factors 6-7)

Notes

[1]The theory of storage states that when commodity inventory levels are high, commodity futures prices are likely to be in contango, and volatility and spot futures prices are typically low. The converse is also true.

[2]Juvenal and Petrella (2012) and Le Pen and Svi (2013) also use a factor approach to analyzing commodity shocks, that are significant price movement that cannot be explained from global fundamentals. Those shocks can be identified from the correlation between the residuals obtained from regressions of commodity returns on fundamental factors. The goal we pursue here is different as we want to identify *market* factors that are common to returns on commodities. Therefore, in this article, we ignore whether these market factors are related to fundamentals or to excess co-movement in commodities. We thank an anonymous referee for this remark.

[3]Following Bai and Ng (2002), a strict factor model is based on the assumption that the matrix Ω is a diagonal matrix, therefore not allowing for correlation between two different idiosyncratic components. For a rigorous and detail statement of the hypothesis underlying factor models, see page 196 of Bai and Ng (2002), assumptions A-D. Such hypothesis are beyond the scope of this empirical research piece, but any interested reader will find a finer level of details in the aforementioned article. Bai (2003) shows that the factor and loadings estimates are asymptotically normaly distributed.

[4]There is no general guide in selecting r_{max} in panel data analysis. In time series analysis, Schwert's rule (1989) of $r_{max} = 8int[(\frac{T}{100})^{\frac{1}{4}}]$ is occasionally used.

[5]When regressing commodity returns on the first commodity factor, the estimated slopes are numerically very similar to the corresponding PCA loadings. The difference between them comes from the potentially non-zero intercept of the regression.

[6]If commodities were driven by commodity-by-commodity specific factors, this correlation analysis would have shown that a single and different commodity asset explains each of these first five factors. This is not the case here: beyond precious metals, we find factors that identifies well to commodity sectors.

[7]They actually found three factors, but the last factor explains between 1 and 2% of interest rates variations – by using a dataset covering a different period from ours.

[8]Every currency considered here is against the US Dollar. For example, the Euro currency is in fact the Euro vs. the US Dollar cross.

[9]We thank an anonymous referee for suggesting these references.

References

Acharya, VV., Lochstoer, LA., Ramadorai, T. 2011. Limits to Arbitrage and Hedging: Evidence from Commodity Markets. *Journal of Financial Economics* 109(2):441465

Adrian, T., Etula, E., Muir, T. 2010. Financial Intermediaries and the Crosssection of Asset Returns. *Technical Report* #422, Federal Reserve Bank of New York Staff Report, USA

Alessi, L., Barigozzi, M., Capasso, M. 2010. Improved Penalization for Determining the Number of Factors in Approximate Factor Models. *Statistics and Probability Letters* 80(23-24):1806-1813

Anzuini, A., Lombardi, MJ., Pagano, P. 2010. The Impact of Monetary Policy Shocks on Commodity Prices. *Working Paper Series* #1232, European Central Bank, Germany

Bai, J. 2003. Inferential Theory for Factor Models of Large Dimensions. *Econometrica* 70:135-172.

Bai, J., Ng, S. 2002. Determining the Number of Factors in Approximate Factor Models. *Econometrica* 70(1):191 221

Balvers, RJ., Huang, D. 2009. Money and the C-CAPM. *Journal of Financial and Quantitative Analysis* 44(02):337 368

Barsky, RB., Kilian, L. 2004. Oil and the Macroeconomy since the 1970s. *Journal of Economic Perspectives* 18(4):115 134

Bessembinder, H., Chan, K. 1992. Time-Varying Risk Premia and Forecastable Returns in Futures Markets. *Journal of Financial Economics* 32(2):169 193

Bodie, Z., Rosansky, V. 1980. Diversification Returns And Asset Contributions. *Financial Analysts Journal* 48(3):26-32

Boon, L.N., Ielpo, F. 2012. Determining the Maximum Number of Uncorrelated Strategies in a Global Portfolio. Forthcoming in *the Journal of Alternative Investment*

Breeden, DT. 1980. Consumption Risk in Futures Markets. *Journal of Finance* 35(2):503-520

Connor, G., Korajczyk, RA. 1993. A Test for the Number of Factors in an Approximate Factor Model. *Journal of Finance* 48(4):1263-1291

Dasklaki, C., Skiadopoulos, G. 2011. Should Investors Include Commodities in their Portfolios After All? *Journal of Banking and Finance* 35(10):2606-2626

de Roon, F., Nijman, T., Veld, C. 2000. Hedging Pressure Effects in Futures Markets. *The Journal of Finance* 55(3): 14371456.

de Roon, F., Szymanowska, M. 2010. The Cross-Section of Commodity Futures Returns. *SSRN Working Paper* #891073, Social Science Research Network, USA

Dusak, K. 1973. Futures Trading and Investor Returns: An Investigation of Commodity Market Risk Premiums. *Journal of Political Economy* 81(6):1387-1406

Erb, CB., Harvey, C.R. 2006. The Strategic and Tactical Value of Commodity Futures. *Financial Analysts Journal* 62(2):69-97

Fama, EF., French, KR. 1993. Common Risk Factors in the Returns on Stocks and Bonds. *Journal of Financial Economics* 33(1):3-56

Forni, M., Gambetti, L. 2010. The Dynamic Effects of Monetary Policy: A Structural Factor Model Approach, *Journal of Monetary Economics* 57(2):203-216

Forni, M., Giannone, D., Lippi, M., Reichlin, L. (2009). Opening the Black Box: Structural Factor Models versus Structural VARs. *Econometric Theory*, 25:1319-1347

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

Frankel, JA. 2006. The Effect of Monetary Policy on Real Commodity Prices. *NBER Working Paper* #12713, National Bureau of Economic Research, Boston, Masachussetts, USA

Gorton, GB., Rouwenhorst, KG. 2006. Facts and Fantasies about Commodity Futures. *Financial Analysts Journal* 62(2):47-68

Gorton, GB., Hayashi, F., Rouwenhorst, KG. 2012. The Fundamentals of Commodity Futures Returns. *Review of Finance, European Finance Association*, 17(1):35-105.

Hirshleifer, D. 1988. Residual Risk, Trading Costs, and Commodity Futures Risk Premia. *Review of Financial Studies* 1(2):173 193

Hirshleifer, D. 1989. Determinants of hedging and risk premia in commodity futures markets. *Journal of Financial and Quantitative Analysis* 24(3):313 331

Hong, H., Yogo, M. 2011. What does futures market interest tell us about the macroeconomy and asset prices? *Journal of Financial Economics* 105(3): Pages 473490

Jagannathan, R. 1985. An investigation of commodity futures prices using the consumption-based intertemporal capital asset pricing model. *Journal of Finance* 40(1):175-191

Juvenal, L., Petrella, I. 2012. Speculation in the Oil Market. *Working Paper of the Federal Reserve Bank of St. Louis*

Kat, HM., Oomen, RCA. 2007a. What Every Investor Should Know About Commodities, Part I: Univariate Return Analysis. *Journal of Investment Management* 5(1):1-25

Kat, HM., Oomen, RCA. 2007b. What Every Investor Should Know About Commodities, Part II: Multivariate Return Analysis. *Journal of Investment Management* 5(3):1-25

Le Pen, Y., Sévi, B. 2013. Futures Trading and the Excess Comovement of Commodity Prices. *SSRN Working Paper* #2191659, USA.

Litterman, R., Scheinkman, J. 1991. Common Factors Affecting Bond Returns. *The Journal of Fixed Income* 1:54-61

Onatski, A. 2009. Testing hypotheses about the number of factors in large factor models. *Econometrica* 77(5):1447-1479

Schwert, GW. 1989. Why does Stock Market Volatility Change Over Time? *Journal of Finance* 44:1115-1153

Stoll, HR. 1979. Commodity futures and spot price determination and hedging in capital market equilibrium. *Journal of Financial and Quantitative Analysis* 14(4):873-894

Working, H. 1949. The theory of price of storage. *American Economic Review* 39(6):1254 1262