Fundamental and Financial Influences on the Co-movement of Oil and Gas Prices

Derek Bunn, Julien Chevallier,** Yannick Le Pen,*** and Benoît Sévi*****

ABSTRACT

As speculative flows into commodity futures are expected to link commodity prices more strongly to equity indices, we investigate whether this process also creates increased correlations amongst the commodities themselves. Considering U.S. oil and gas futures, we investigate whether common factors, derived from a large international data set of real and nominal macroeconomic variables by means of the large approximate factor models methodology, are able to explain both returns and whether, beyond these fundamental common factors, the residuals remain correlated. We further investigate a possible explanation for this residual correlation by using some proxies for trading intensity derived from CFTC publicly available data, showing most notably that the proxy for speculation in the oil market increases the oil-gas correlation. We thus identify the central role of financial activities in shaping the link between oil and gas returns.

Keywords: Oil Futures, Gas Futures, Common Factors, Speculation, Excess Comovement

http://dx.doi.org/10.5547/01956574.38.2.dbun

1. INTRODUCTION

This paper investigates whether traders' behavior in futures markets may play a role in shaping the link between oil and natural gas prices beyond fundamentals. Specifically, we first estimate how international macroeconomic variables combined into 'factors' can explain the percentage change in oil and gas prices.[1] After taking account of these macroeconomic factors, we then test for the impact of speculation and hedging behavior on the relationship between oil and gas returns. This provides new insights on the link between the prices of these two energy commodities.

We combine a large data set of global macroeconomic variables into a small number of factors that can explain a substantial fraction of the variation in the data. We then show that just three of these factors can explain a significant proportion of oil and natural gas returns in the U.S.,

1. In contrast to much of the literature on the relationship between oil and natural gas prices, we choose to examine returns rather than prices. The former are more likely to deliver relevant analyses for the risk management activities that are of central interest to participants in energy markets.

* Corresponding author. London Business School, Sussex Place, Regent's Park, London NW1 4SA, United Kingdom. E-mail: dbunn@london.edu.

** IPAG Business School (IPAG Lab), 184 boulevard Saint-Germain, 75006 Paris, France and Université Paris 8 (LED), 2 rue de la Liberté, 93526 Saint-Denis cedex, France.

*** Université Paris-Dauphine, PSL Research University, LEDa, 75016 Paris, France.

**** Université de Nantes, LEMNA, BP 52231 Chemin de la Censive du Tertre, 44322 Nantes Cedex, France.

which we take as an indication of common macroeconomic fundamentals driving oil and gas price dynamics. In a second step, we show that the correlation between the unexplained parts of the returns—residuals after filtration by the factors—is significantly related to our proxy for speculative activity in the oil market. This is consistent with the "financialization" conjecture for related energy commodities and the widely-accepted hypothesis of the oil market as a leading commodity market.

There is a substantial literature on the links between oil and gas prices. In general, the conventional view was one of a strong linkage, as developed in Serletis and Herbert (1999), mainly in the U.S., because of the history of product substitution between gas and fuel oil (e.g. for power generation, industrial boilers). Furthermore, especially within Continental Europe and South-East Asia, the development of gas pipelines, and more recently LNG facilities, by the upstream oil producers have generally been underwritten with long-term gas contracts index-linked to crude oil prices.

Oil and gas markets exhibit, however, highly different market characteristics. Oil markets are part of broader international markets (Kilian (2009)), while natural gas markets remain essentially regional (Li et al. (2014)). Surplus production of natural gas can often arise, since it may be a by-product of oil and thereby considered secondary. Gas supply is more inelastic than oil in the short-term, partly because of production and delivery logistics (Villar and Joutz (2006)); likewise gas demand is less elastic because of its substantial component of residential heating (Ewing et al. (2002)) compared to the high transportation component in oil. Finally, more recent data suggests that linkages may have weakened with the advent of shale gas and, looking beyond the U.S., with the continuing deregulation of energy markets worldwide (Ramberg and Parsons (2012)). From existing research, however, the strength of the relationship between oil and gas remains an open empirical question.

This debate has been revived with an increased focus upon financial practices (see Smith (2009) or Brunetti et al. (2011)). Whilst the evolution and volatility of commodity prices have always presented hedging and risk management concerns to producers and consumers, the so-called "financialization" of commodities through the active involvement of investors and speculators adds a new ingredient to the complexity of their price formations (Rouwenhorst and Tang (2012), Fattouh et al. (2013)).[2] This theme of increased investment and speculative activity in commodities became especially topical in relation to oil price behavior in the first decade of this century, with a view emerging that the financial effects may be substantial in linking commodity price indices to speculative volumes and to equity indices, but only alongside the changes in global economic fundamentals.[3]

For individual commodities, however, whilst questions still remain on the relative effects of fundamental drivers and financial market activity, there is, in addition, a more subtle aspect relating to the changes in the relationships amongst the commodities themselves. If two or more commodities are part of the same asset class, traded perhaps as part of a commodity index, then it is plausible to expect that extra financial activity will further increase their correlation beyond that already attributable to their product fundamentals. This expectation now appears to follow as a conjecture from various strands of theoretical and empirical research. For instance, as a result of

2. A special issue of the *Journal of International Money and Finance* is devoted to the topic. The editorial introduction (Arezki et al. (2014)) surveys the recent work on the "financialization" of commodity markets. See also Manera et al. (2013), Kim (2015), Lehecka (2014), Adams and Glück (2015), Etienne et al. (2015) or Liu et al. (2016) that are relevant references on this question.

3. See, among others, Hamilton (2009), Kilian (2009), Büyüksahin and Harris (2011), Tang and Xiong (2012), Alquist and Gervais (2013), Singleton (2014), Juvenal and Petrella (2015) or Manera et al. (2016).

information frictions and adaptive learning from prices, more financial trading may increase the link between commodities and equity market indices, inducing a pro-cyclical tendency (e.g. Singleton (2014)), which would plausibly also manifest a greater co-movement amongst the commodities involved. Furthermore, capital frictions have been shown to influence risk premia in commodity futures (e.g. Acharya et al. (2013)) and the consequent limits to arbitrage may again influence the correlations amongst commodities in the same asset class. More directly, it has been shown in general that as the objective function of investors becomes compromised by a need to outperform benchmark indices, index-focused trading increases the correlations between asset prices comprising the indices (e.g. Basak and Pavlova (2016)). Overall, the question therefore emerges of whether traders' behavior in futures markets may play a role in shaping the link between oil and gas prices.

In this paper, we take account of fundamentals and then test for the impact of financial aspects on the relationship between oil and gas returns. This provides new clarity on the factors influencing the link between these two energy commodities. In a first step, we use a large number of macroeconomic variables to explain oil and gas returns, thereby modeling the latent state of the macroeconomy through fundamental variables that traditionally affect the oil and gas markets. The commodity returns are therefore determined by supply and demand, as well as by a broader array of macroeconomic factors. Such an approach reconciles both impacts of real and nominal variables on commodity prices. In addition, our panel of countries is diversified sufficiently to include variables from emerging countries, which have become major consumers of commodities as part of their growth.

To uncover common factors, we rely on the large approximate factor model methodology following Stock and Watson (2002a,b). Large approximate factor models have been used in a number of economic and financial applications.[4] In particular, Gospodinov and and Ng (2013) rely on this technique for the purpose of extracting common factors from 23 commodity convenience yields and show the predictive power of these factors for forecasting inflation. We extract the factors from a large data set of global macroeconomic variables and methodically search for those that are best able to explain oil and gas returns in the U.S. futures markets. We then show that a few factors can explain a significant proportion of both returns, which we take as an indication of common fundamentals for oil and gas dynamics. Interestingly, we find that the factor with the highest explanatory power for oil is mostly connected with real macroeconomic variables from emerging countries.[5]

In the second step, we show that the correlation between the unexplained parts of the returns—residuals after filtration by the factors—is significantly related to our proxy for speculative activity in the oil market and, to a lesser extent, to the proxy for hedging intensity in the gas market. The first result is consistent with the financialization conjecture for related energy commodities. The second result highlights the specificity of the hedging demand in the gas market.

As in recent analyses (see Büyüksahin and Robe (2014) or Manera et al. (2016) and references therein), we use the commercial vs. non-commercial positions, as contained in the Commitments of Traders (CoT) weekly report of the U.S. Commodity Futures Trading Commission (CFTC). From this data, our empirical results provide evidence that speculative activity coincides

4. Excellent surveys of this literature are Stock and Watson (2006) and Bai and Ng (2008).

5. This finding about oil partly appears in a related paper by two of the authors (Le Pen and Sévi (2011)) where the approximate factor model methodology is used along with the same data set of 187 real and nominal variables. Note, however, that the sole purpose of their paper was to uncover the determinants of U.S. oil futures returns and that no regression was performed to relate these determinants to factors associated with other commodities and no link was made with financial variables such as positions in futures markets, as is done in the present article.

with an increase in the correlation between oil and gas returns. Such results about the importance of speculative activity are in line with the theoretical analysis developed in Basak and Pavlova (2016).

The remainder of the paper is structured as follows. The next Section briefly reviews the existing empirical studies dealing with the oil-gas relationship. In Section 3, we present the data and in Section 4 we briefly review the approximate factor modeling methodology. Section 5 reports our empirical findings about the determinants of oil and gas returns using estimated factors and Section 6 contains the analysis of residual correlation. In Section 7, we analyze the explanatory power of trading activity proxies to model the co-movement between oil and gas returns. Finally, Section 8 provides concluding remarks.

2. LITERATURE

This section briefly reviews the cointegration studies that have previously comprised the main methodology for analysing the adjustments between oil and gas prices. In such a changing and multifaceted context of fundamental influences, empirical analysis has unsurprisingly revealed mixed results concerning the existence of a long-term relationship between gas and oil prices. Thus, the early cointegration analysis by Serletis and Herbert (1999) identified shared trends among the U.S. Henry Hub natural gas price and the fuel oil price using daily price data over the short 1996–1997 period. Villar and Joutz (2006) found qualitatively similar results making use of monthly data during 1989–2005 for the Henry Hub natural gas price and the West Texas Intermediate (WTI) crude oil price. In both papers, the authors provided empirical evidence of a stable relationship between oil and gas prices, despite some transient periods when they have appeared to 'decouple'. However, later, by using error-correction models and common cycle tests, Serletis and Rangel-Ruiz (2004) claimed that Henry Hub and WTI did not have common price cycles, and suggested that the apparent progressively decoupling of U.S. energy prices was a result of deregulation.

Other related studies rely on error-correction models. Among them, Bachmeier and Griffin (2006) found evidence of market integration among the primary energy fuels in the U.S. over the 1990–2004 period. In this context, their analysis found that oil and natural gas prices were cointegrated. In the same vein, Brown and Yücel (2008) demonstrated that movements in crude oil prices have a prominent role in shaping natural gas prices in the U.S., once other price drivers such as weather, seasonality, storage, and production disruptions have been properly taken into account. Yet, Hartley et al. (2008) found evidence that the link between natural gas and crude oil prices in the U.S. was indirect, acting through competition at the margin between natural gas and residual fuel oil (being the price of the main competitive oil product). More precisely, the residual fuel oil price caused movements in the natural gas price, while the converse was not true. More recently, Hartley and Medlock (2014) emphasized the role of exchange rates beyond standard variables in shaping the relationship between oil and gas prices.[6]

Despite this large body of work on market integration between oil and gas, the cointegration approach appears too restrictive for our purposes. Evidence of cointegration reveals the exis-

6. For completeness, we observe several cointegration studies that investigate the relationship between oil and gas prices in the U.K., where a fully liberalized, actively traded, gas market has existed since the early 1990s. Panagiotidis and Rutledge (2007) found a linear relationship between U.K. gas prices and Brent oil prices during 1996–2003, whilst Asche et al. (2006) looked more generally at the relationship between gas and other energy prices. They found that over the 1995–1998 period there was a single energy market in the U.K., with crude oil price being the leader.

Figure 1: Crude Oil (left plot) and Natural Gas (right plot) Prices

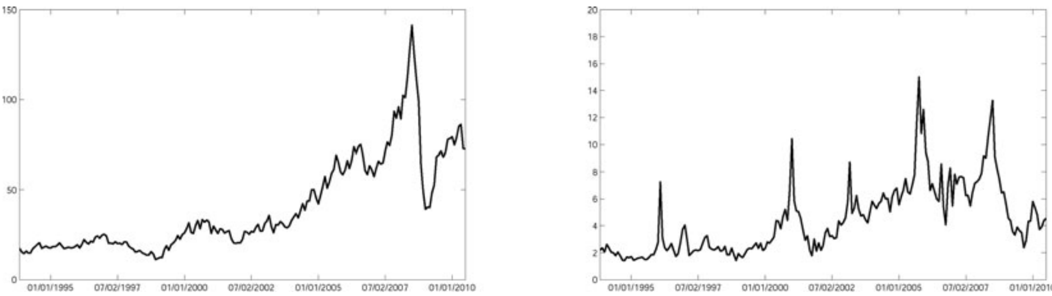
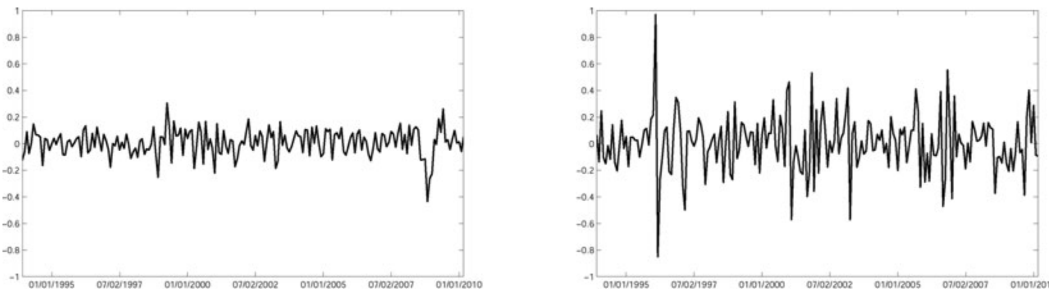


Figure 2: Crude Oil (left plot) and Natural Gas (right plot) Returns



tence of common stochastic trends but cannot tell us the precise nature of these trends. In seeking to go beyond tests for cointegration and the estimation of dynamic error-correction models, we are therefore looking to identify what may be the common determinants of co-movements for these two commodities, amongst an extensive set of global macroeconomic variables in order to provide effective pre-filtering before estimating proxies for financial activities.

3. DATA

We look at the main U.S. oil and natural gas prices. The natural gas futures are the Henry Hub prices in $/MMBTU, whilst the crude oil futures are the WTI prices in $/BBL. We use the quoted price for the nearby-contract for each commodity. We have 196 monthly observations from November 1993 to February 2010. Returns are computed as log-difference of prices. Prices and returns are displayed in Figure 1 and Figure 2, respectively.

The descriptive statistics for the returns are reported in Table 1. Oil returns have the higher average and the lower standard error. Both returns show evidence of excess kurtosis. Crude oil returns have negative skewness, distinct from the case of natural gas returns for which the distribution is found to be approximately symmetric. Accordingly, the Jarque-Bera test rejects normality in each case. The correlation between oil and gas returns is positive (0.2095) and significant at 1%.

Whilst cointegration tests established that oil and gas prices are linked in the long run (Brown and Yücel (2008), Hartley and Medlock (2014)), we identify factors extracted from 187 global macroeconomic variables to find the common macro drivers of these two series. Our dataset differs in its composition from the widely known datasets of Stock and Watson (2005) and Ludvigson and Ng (2009) which consist of U.S. national time series only.[7] Since our purpose is to

7. The original dataset in Stock and Watson (2005) covers the period 1959:01 to 2003:12. It is slightly extended in Ludvigson and Ng (2009) to cover the period 1964:01 to 2007:12.

Table 1: Oil and Natural Gas Monthly Returns Descriptive Statistics

	Crude Oil	Natural Gas
Mean	0.0077	0.0040
Max.	0.3045	0.9694
Min.	-0.4340	-0.8496
Std. Dev.	0.0991	0.2176
Skewness	-0.5770	0.0555
Kurtosis	4.6766	5.5875
Jarque-Bera Stat.	33.83*	54.77*
$\hat{\rho}$	0.2095***	
p-value	0.0032	
Nb. of Obs.	196	

Notes:(i) Monthly returns are the log difference of prices. (ii) ***,**, and * respectively denote a rejection of the appropriate null hypothesis at the 1%, 5% and 10% levels.

explain crude oil and natural gas returns, we include data from the main developed economies (128 variables) but also from emerging countries (59 variables). Therefore our dataset is representative of the world economy and, in particular, high-level demand from emerging countries will be accounted for in the information conveyed by the estimated factors. These variables can be classified into 103 real variables (73 for developed countries, 30 for emerging countries) and 84 nominal variables (55 for developed countries and 29 for emerging countries).

Thus, the inclusion of variables in our analysis is guided by two principles: (i) to gather, as far as possible, a balanced panel between developed and emerging countries, and (ii) to limit the adverse effect of the noise created by uninformative variables on factor estimations (Boivin and Ng (2006)). All data are extracted from Thomson Financial DataStream. The list of the 187 time series is given in the Appendix, where a coding system indicates how the data are transformed to ensure stationarity. Also, data are standardized prior to model estimation.

4. THE LARGE APPROXIMATE FACTOR METHODOLOGY

We consider N macroeconomic variables x_{it} and make the assumption that each of them depend on a small set of common factors F_t and an idiosyncratic component:

$$x_{it} = \lambda_i' F_t + e_{it},\ i = \{1,\dots,N\},$$

where $F_t = (F_{1t},\dots,F_{rt})'$ is the vector of $r << N$ common factors, $\lambda_i = (\lambda_{1i},\dots,\lambda_{ri})'$ are the factor loadings and e_{it} is the idiosyncratic error. We dispose of T observations for all variables. With $X_t = (x_{1t},\dots,x_{Nt})'$, $e_t = (e_{1t},\dots,e_{Nt})'$ and $\Lambda = (\lambda_1,\dots,\lambda_N)'$, we can write:

$$X_t = \Lambda F_t + e_t$$

Under the assumption that F_t and e_t are zero mean and uncorrelated, and with the normalization $E(F_t F_t') = I_d$, we obtain the usual decomposition of the covariance matrix Σ of X_t:

$$\Sigma = \Lambda\Lambda' + \Omega$$

where $\Omega = E(e_t e_t')$.

In classical factor analysis, N is fixed, and F_t and e_t are serially and cross-sectionally uncorrelated. Stock and Watson's (2002a,b) 'large dimensional approximate factor model' alleviates these assumptions as N and T tend to infinity and the idiosyncratic errors e_{it} can be 'weakly correlated' across i and t.[8]

Factors and loadings can be estimated by a principal components analysis.[9] If we assume k factors, the estimates $\hat{F}_k$ and loadings matrix ^{k} minimize the sum of squared residuals :

$$minS(k) = (NT)^{-1} \sum_{i=1}^{N} \sum_{t=1}^{T} (x_{it} - \lambda_i^{k\prime} F_{kt})^2$$

subject to $\Lambda^{k\prime} \Lambda^k / N = I_k$ and $F_k' F_k$ being diagonal.

Estimates of F_{kt} for all dates are given by $\hat{F}_k = X^k / N$ where X is the $T \times N$ matrix with t^{th} row X_t' and ^{k} equal to $\sqrt{N}$ times the eigenvectors of the largest k eigenvalues of $X'X$. Computation of $\hat{F}_k$ then simply requires the eigenvectors of the $N \times N$ matrix $X'X$. When $N > T$, a simpler approach uses the $T \times T$ matrix XX' . Stock and Watson (2002a) and Bai and Ng (2002) demonstrate the consistency of the principal components estimator, as N and $T \to \infty$ and Bai (2003) gives its asymptotic distribution.

In the following step, we first apply the information criteria of Bai and Ng (2002) to select the number of factors. These are extensions to factor models of usual information criteria and have the general expression:

$$\begin{aligned} PCP_i(k) &= \hat{S}(k) + k\bar{\sigma}^2 g_i(N,T) \\ IC_i(k) &= ln(\hat{S}(k)) + kg_i(N,T) \end{aligned}$$

where $\hat{S}(k)$ is the sum of squared residuals (divided by NT) for k factors, $\bar{\sigma}^2$ is equal to $\hat{S}(k_{max})$ for a pre-specified value k_{max}, and $g_i(N,T)$ is a penalty function. We allow $k_{max} = 20$ factors and apply the four penalty functions $g_i(N,T), i = 1, \ldots, 4$ of Bai and Ng (2002). The optimal number of factors minimizes these information criteria.

We also apply Kapetanios (2010) sequential test as a complement. If the true number of factors is k_0, the k_0 eigenvalues (in decreasing order) of the covariance matrix Σ increase at rate N while the others are bounded under some regularity conditions. Thus, the difference between the eigenvalues $\hat{\lambda}_k - \hat{\lambda}_{k_{max}+1}$ will tend to infinity for $k = 1, \ldots, k_0$ but be bounded for $k = k_0 + 1, \ldots, k_{max}$ where k_{max} is some finite number such that $k_0 < k_{max}$. A test statistics based on $\hat{\lambda}_k - \hat{\lambda}_{k_{max}+1}$ is used to discriminate the null hypothesis that the true number of factors k_0 is equal to k ($H_{0,k}: k_0 = k$) against the alternative hypothesis ($H_{1,k}: k_0 > k$). If there is no factor structure, $\hat{\lambda}_k - \hat{\lambda}_{k_{max}+1}$, properly normalized, converge to a finite limit while it should go to infinity if there is a factor structure. We begin with ($H_{0,k}: k_0 = k = 0$) against ($H_{1,k}: k_0 > 0$) and and increase k_0 until we cannot reject the null hypothesis. Kapetanios (2010) refers to this algorithm as MED (Maximal Eigenvalue Distribution). Finally, we include, as an additional selection criteria, the lowest number of factors which explain at least 50% of the variance of the dataset of macroeconomic variables.

8. Although Forni and al. (1999) and Stock and Watson (2002) use different sets of assumptions to characterize 'weak correlations', the main idea is that cross-correlations and serial correlations have an upper bound.

9. We adopt the static approach following D'Agostino and Giannone (2012) who show that there is no clear advantage of using dynamic factor models.

Table 2: Static Factors Selection

Method	Nb of factors
MED	2
IC_1	3
IC_2	2
IC_3	20
IC_4	20
PCP_1	9
PCP_2	7
PCP_3	20
PCP_4	20
More than 50%	17

Notes: MED denotes the number of factors given by the Maximum Eigenvalue Distribution algorithm. IC_i and PCP_i denote, respectively, the number of factors given by the information criteria IC and PCP estimated with the penalty function $g_i(N,T)$. "More than 50% " is the lowest number of factors which explain more than 50 % of the variance of the initial dataset.

Results are displayed in Table 2 and we observe that there is no clear agreement on the on optimal number of factors. This observation is consistent with previous empirical studies, which also show substantial variations in the estimated number of factors. In our case, according to the information criteria of Bai and Ng (2002), the optimal number of factors runs from 2 to 20, whilst the sequential test of Kapetanios (2010) selects 2 factors.

Additional diagnostics on the autocorrelation and the explanatory power of the estimated factors $\hat{F}_{it},\ i = 1,\ldots,20$, are reported in Table 3. We notice that the first 2 factors only explain 16% of the variance of the 187 time series, whilst we reach 36% with 9 factors. The number of factors required to explain at least 50 % of the variance of the macroeconomic database is 17. Hence, we choose to consider two sets of factors to model the oil and and gas returns: the set of the first 9 factors, given by the PCP_1 information criterion. For robustness, we also consider the set of the first 17 factors. The factors autocorrelations (up to 3 lags) provided in Table 3 show that most factors are persistent, as commonly observed in such analyses.

5. FACTOR ANALYSIS OF OIL AND GAS RETURNS

We therefore commence modeling the oil and gas returns with the first 9 factors. In a preliminary step, we evaluate the explanatory power of each of these factors for oil and gas returns by means of a simple linear regression of each return on a single factor. Following this, as $\hat{F}_{3t}$ and $\hat{F}_{9t}$ had very low explanatory power—R^2 are around 1%—, we eliminate them from our set of regressors. We then select the multivariate BIC criterion for minimizing a subset from all possible combinations of the remaining 7 factors (as in Stock and Watson (2002) and Ludvigson and Ng (2009)) and obtain $\{\hat{F}_{1t},\hat{F}_{2t},\hat{F}_{7t}\}$ as our best combination. As a robustness check, we enlarge our modeling to all possible combinations of the first 17 factors, and still reassuringly obtain the same set $\{\hat{F}_{1t},\hat{F}_{2t},\hat{F}_{7t}\}$. We thereby conclude that the set $\{\hat{F}_{1t},\hat{F}_{2t},\hat{F}_{7t}\}$ represents the main common driver for oil and gas returns.

Table 3: Summary Statistics for $\hat{F}_{it}$ for $i = 1,\ldots,20$.

Factor i	ρ_1	ρ_2	ρ_3	R_i^2
1	0.2005	0.1521	0.3196	0.0977
2	0.0999	0.0335	0.3120	0.1622
3	-0.0078	0.0188	-0.0419	0.2034
4	-0.0738	-0.0914	0.1531	0.2358
5	-0.2214	-0.0805	0.1178	0.2657
6	0.1833	0.0369	0.0361	0.2930
7	0.0729	0.2731	0.2678	0.3188
8	0.4038	0.5004	0.3322	0.3422
9	0.0018	-0.0248	-0.0339	0.3640
10	-0.0990	-0.0539	0.1740	0.3841
11	-0.1949	-0.1373	0.0582	0.4041
12	0.0399	0.1753	0.1779	0.4231
13	-0.2680	0.0920	0.0276	0.4418
14	-0.0289	0.1719	0.1209	0.4589
15	-0.2331	-0.0417	0.1317	0.4748
16	-0.2550	0.1589	-0.0136	0.4905
17	-0.2853	0.0669	0.2717	0.5057
18	-0.2953	0.0264	0.0673	0.5204
19	-0.4048	0.1815	-0.0780	0.5346
20	-0.3225	0.0624	0.1267	0.5482

Notes: Factors $\hat{F}_{it}$ are estimated by the method of principal components from a panel of 187 economic variables transformed (taking logs and first difference where appropriate) and standardized prior to estimation. ρ_i denotes the *ith* autocorrelation. $Ri2$ is the fraction of total variance in the data jointly explained by factor $\hat{F}_{1t}$ to $\hat{F}_{it}$.

We further enlarge our specification to take into account some specific features of oil and gas markets. To test for the impact of weather on returns, we follow Brown and Yücel (2008) and Hartley and Medlock (2014) and compute deviations from cooling degree days (CDD) and heating degree days (HDD) in each month. These variables are denoted $\text{HDDdev}_t = \text{HDD}_t - \overline{\text{HDD}}_t$ and $\text{CDDdev}_t = \text{CDD}_t - \overline{\text{CDD}}_t$ where $\overline{\text{HDD}}_t$ and $\overline{\text{CDD}}_t$ are the monthly average over the time period 1963–1992. HDDdev_t and CDDdev_t both have a positive and highly significant impact on gas returns while CDDdev_t has a positive but only marginally significant effect on oil returns. We also find a very significant and negative impact of the monthly dummies for July and August on gas returns.[10] Our results are consistent with Hartley and Medlock (2014) who found a positive impact on oil and gas returns for both HDDdev_t and CDDdev_t and monthly seasonality.[11] We estimate the following SUR model:

10. We report in Table 7 in the appendix estimations of linear regressions of oil and gas returns on CDDdev, HDDdev and monthly dummies. We find that June, July and August are significant for gas return but not for oil return. Slight discrepancies with respect to Hartley and Medlock's(2014) results are likely to be due to the seasonal component that is partly captured by the factors.

11. Following theoretical developments in Pindyck (1994) and recent empirical analysis in Mu (2007), we also check for the effect of variations in inventory levels but this variable is not significant. The same is true for the OPEC spare capacity in the case of oil. All our results are available from the authors and, for the purpose of facilitating replication, data and Matlab programs are provided on the Yannick Le Pen's homepage at https://sites.google.com/site/yannicklepen/.

Table 4: Fitting Oil and Natural Gas Returns

	r_oil (1.a)	r_gas (1.b)	r_oil (2.a)	r_gas (2.b)
Intercept	0.0077	0.0040	0.0032	0.0372**
	(1.34)	(0.27)	(0.47)	(2.01)
$\widehat{F}_{1t}$	-0.1221***	-0.0361	-0.1209***	-0.0360
	(-6.68)	(-0.75)	(-6.67)	(-0.79)
$\widehat{F}_{2t}$	-0.1481***	-0.1582***	-0.1473***	-0.1758***
	(-6.59)	(-2.68)	(-6.59)	(-3.12)
$\widehat{F}_{7t}$	-0.1469***	-0.2575***	-0.1397***	-0.2734***
	(-4.13)	(-2.76)	(-3.94)	(-3.00)
CCDdev			0.0005*	0.0024***
			(1.82)	(2.93)
HDDdev			0.00001	0.0009***
			(0,09)	(3.22)
June				-0.0738
				(-1.35)
July				-0.1328**
				(-2.46)
August				-0.1929***
				(-3.34)
R^2	0.3494	0.0730	0.3603	0.1692
$\overline{R}^2$	0.3393	0.0585	0.3435	0.1336
Arch-LM (2)	3.17	27.69**	1.90	24.88***
Residual correl.	0.0961		0.0616	
p-value	0.1802		0.3910	

Notes: (i) Columns (1.a) and (1.b) report the OLS estimates of the regression of crude oil and natural gas returns on the contemporaneous variables named in the left column. (ii) Columns (2.a), (2.b) report the SUR estimates for the oil and gas returns. (iii) *t*-statistics are reported in parenthesis under the estimates. A constant whose estimate is reported in the second row is always included in the regression. (iv) *CDDdev* and *HDDdev* are CCD and HDD minus their average for each months for the time period 1963–1993. (v) Arch-LM (2) stands for Engle's ARCH Lagrange Multiplier test with two lags.(vi) For each test ***, **, and * respectively denote rejection of the null hypothesis at the 1%, 5% and 10% levels.

$$\begin{cases} r_{oil,t} = \alpha_1 + \beta_{1,1}\hat{F}_{1t} + \beta_{1,2}\hat{F}_{2t} + \beta_{1,7}\hat{F}_{7t} + \theta_{1,1}CDDdev_t + \theta_{1,2}HDDdev_t + u_{1,t} \\ r_{gas,t} = \alpha_2 + \beta_{2,1}\hat{F}_{1t} + \beta_{2,2}\hat{F}_{2t} + \beta_{2,7}\hat{F}_{7t} + \theta_{2,1}CDDdev_t + \theta_{2,2}HDDdev_t + \omega_2 D_t + u_{2,t} \end{cases}$$

where D_t contains the monthly dummies.

Estimates are reported in Table 4. $\hat{F}_{2t}$ and $\hat{F}_{7t}$ are significant for both returns while $\hat{F}_{1t}$ is only significant for oil. We are able to explain a substantial amount of oil return (35%) but much less of gas return (7%) when using factors only. Adding weather variables and monthly dummies markedly increases the R^2 for gas returns—the increase is about 10%—but not for oil returns. This denotes the local feature of gas in comparison to oil and the associated importance of seasonality for gas return compared to oil. Engle's ARCH LM tests shows that there remains significant serial correlation in squared gas residuals. Residual correlation for the whole sample is equal to 0.0616 and is not significant. Filtering oil and gas returns has therefore eliminated the correlation between them. We will show in the next section, however, that residual correlation varies substantially through time and could be biased by volatility.

Regarding the economic interpretations of the common factors, $\hat{F}_{1t}$, $\hat{F}_{2t}$ and $\hat{F}_{7t}$, we follow the approach of Ludvigson and Ng (2009). Recall that these factors are not "identified" from an econometric viewpoint, as we have to impose enough restrictions to estimate them (the number of decompositions is, by nature, infinite). Thus, we first classify our macroeconomic variables into four groups according to the combination of real/nominal variables and developed/emerging countries. Then, we regress each of the 187 macroeconomic variables on a single factor and display in Figure 3 the R^2 along the horizontal axis with the macroeconomic variables sorted.[12]

From Figure 3, we observe that $\hat{F}_{1t}$ can easily be interpreted as a real factor, since it records its highest explanatory power for real variables. More precisely, $\hat{F}_{1t}$ is mostly associated with real variables from emerging countries. As previously noted in Table 3, $\hat{F}_{1t}$ is significant for oil returns but not for gas.[13] The correlation of $\hat{F}_{1t}$ with oil returns can be interpreted as an evidence of the growing share of emerging countries in world oil demand during the time period considered. This finding is consistent with previous studies, such as Hamilton (2009), Kilian (2009) or Aastveit et al. (2015), who show that demand from emerging countries has been responsible for the increase in oil price in the 2000s. More importantly, because we include in our database a number of Asian variables, it seems that their explanatory power is quite large and supports the view of dynamic effects following their demand shocks (see the discussion in Aastveit et al. (2015) on this specific issue). $\hat{F}_{7t}$ can also be interpreted as a real factor as it reaches its highest R^2 for a small set of real variables from developed countries (most notably western housing starts and car registrations). Interestingly, from a demand perspective, these latter variables are associated with heating and transport.

$\hat{F}_{2t}$ has higher explanatory power with some nominal variables, essentially money supply from emerging countries, compared with the other two factors. We thus interpret $\hat{F}_2$ to represent these latter variables even if it is also correlated with a very small subset of real variables from developed countries. The relation between nominal variables and commodity prices is discussed in a number of recent contributions (Barsky and Kilian (2002), Frankel (2006)) although the empirical evidence is fairly mixed. Our results bring some support to the hypothesis of a relationship between commodity prices and monetary variables.

To summarize, there are substantial indications from factors $\hat{F}_{2t}$ and $\hat{F}_{7t}$ that gas has a linkage with oil due to common economic and other global drivers, but, from factor $\hat{F}_{1t}$, oil also has its own distinct global economic driver linked to the growth of the emerging economies. This finding is in line with our earlier remark that the oil market is a world market when the gas market is distinctly regional.

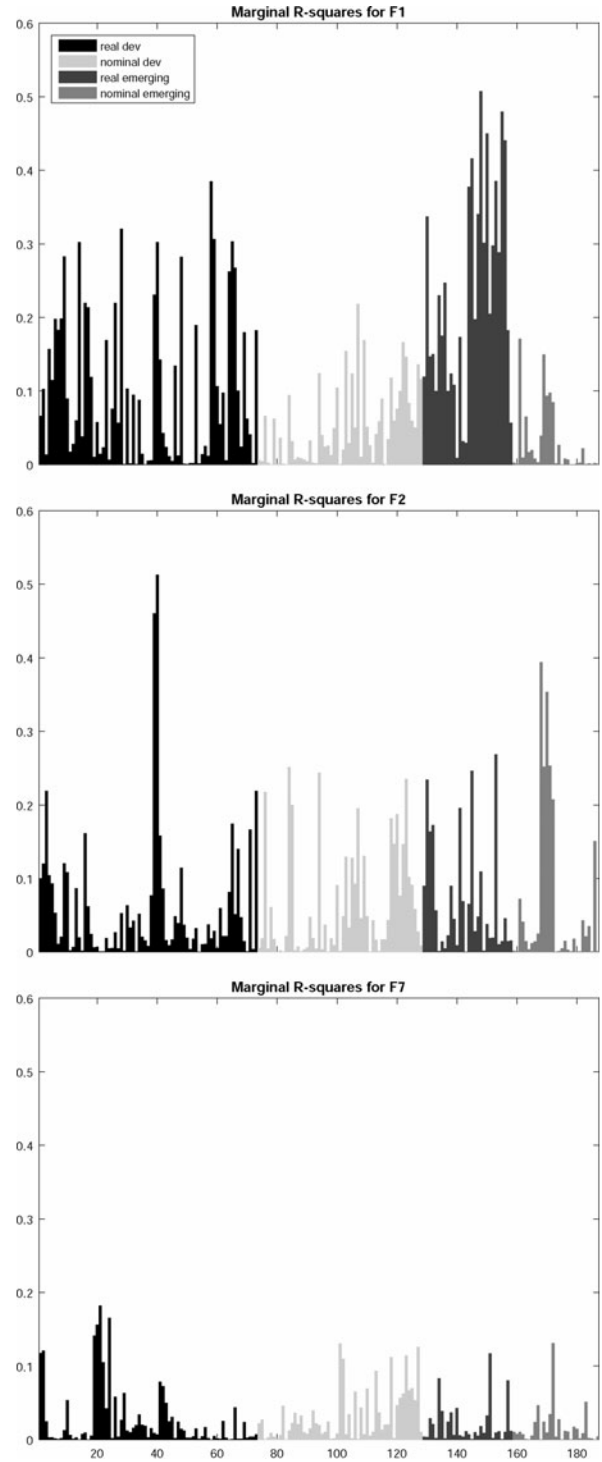
6. A TIME-VARYING AND UNBIASED RESIDUAL CORRELATION MEASURE

In this section, we proceed, as in Kallberg and Pasquariello (2008), to compute an unbiased time-varying measure of residual correlation. The objective is to estimate the sample correlation, corrected for the effects of heteroscedasticity by using Forbes and Rigobon (2002)'s methodology. When applied on a rolling basis, this estimation technique is able to track the true correlation. Note that the resulting estimate is nonparametric. As observed by Kallberg and Pasquariello (2008),

12. Ludvigson and Ng (2009) rely on a finer classification, but they only use U.S. variables. We do not think that this methodology is applicable when several economies are considered if we want to preserve some interpretability.

13. Recall that factors are not identified, unless we impose some constraints to estimate them. Therefore, the sign of the coefficient of $\hat{F}_{1t}$ in the crude oil return equation has no meaning *per se*.

Figure 3: Relationship Between $\hat{F}_{1t}$, $\hat{F}_{2t}$ and $\hat{F}_{7t}$ and the 187 macroeconomic variables



Notes: Each panel shows the R^2 from regressing the series number given on the x-axis onto each individual factor $\hat{F}_i$. The series are detailed in the Appendix, and sorted as they appear in the Figure (real variables for developed countries, nominal variables for developed countries, real variables for emerging countries, nominal variables for emerging countries).

financial research contains various empirical applications where such rolling filters are found to perform quite well in comparison with parametric specifications.

The usual residual correlation is defined as:

$$\hat{\rho}_{ijt} = \frac{\mathrm{cov}(\hat{u}_{i,t},\hat{u}_{j,t})}{[\mathrm{var}(\hat{u}_{i,t})\mathrm{var}(\hat{u}_{j,t})]^{1/2}}$$

Boyer et al. (1999), Loretan and English (2000) and Forbes and Rigobon (2002) show that with time-varying volatility, the usual sample correlation may be biased. In this case, the sample correlation is defined as *conditional correlation*. These authors propose a correction for this bias, and define an *unconditional* correlation measure under the assumption of no omitted variables or endogeneity. The unconditional correlation has the following expression:

$$\hat{\rho}^{*}_{ij,t} = \frac{\hat{\rho}_{ij,t}}{[1+\hat{\delta}_{i,t}(1-(\hat{\rho}^{2}_{ij,t})]^{1/2}}$$

where the ratio $\hat{\delta}_{i,t} = \frac{\mathrm{var}(\hat{u}_{i,t})}{\mathrm{var}(\hat{u}_{i,t})_{LT}} - 1$ corrects the conditional correlation $\hat{\rho}_{ij,t}$ by the relative difference between short-term volatility $\mathrm{var}(\hat{u}_{i,t})$ and the long-term volatility $\mathrm{var}(\hat{u}_{i,t})_{LT}$ of the $iith$ return. As we do not make any *ex ante* assumption on the direction of propagation of shocks from one commodity to another, we alternatively assume that the source of these shocks is commodity i (in $\hat{\rho}^{*}_{ij,t}$) or commodity j (in $\hat{\rho}^{*}_{ji,t}$). Therefore, $\hat{\rho}^{*}_{ij,t}$ and $\hat{\rho}^{*}_{ji,t}$ may be different. Thus, if the source of a shock is oil (index 1):

$$\hat{\rho}^{*}_{12,t} = \frac{\hat{\rho}_{12,t}}{[1+\hat{\delta}_{1,t}(1-(\hat{\rho}^{2}_{12,t}))]^{1/2}}$$

and, if the source of a shock is gas (index 2):

$$\hat{\rho}^{**}_{12,t} = \frac{\hat{\rho}_{12,t}}{[1+\hat{\delta}_{2,t}(1-(\hat{\rho}^{2}_{12,t}))]^{1/2}}$$

Ultimately, we use the average of these two unconditional correlations:

$$\hat{\rho}^{*}_{t} = \frac{1}{2}(\hat{\rho}^{*}_{12,t} + \hat{\rho}^{**}_{12,t})$$

The estimate $\hat{\rho}_{12,t}$ is derived from a short-time rolling window $[t-L+1,t]$ of fixed length L. Long-term volatilities are estimated over rolling windows of length gL (with $g>1$) $[t-gL+1,t]$, which gives us the two correcting parameters $\hat{\delta}_{1,t}$ and $\hat{\delta}_{2,t}$. We set $L=30$ and $g=2$.[14]

We apply this methodology to estimate the time-varying unconditional correlations of oil and gas returns ρ^{*Ret}_{t} and residuals ρ^{*Res}_{t} for the time period from October 1998 to February 2010. Descriptive statistics on these correlations are reported in Table 5. Even though the correlation

14. We consider several alternative values for L and g and find estimates for correlation that are very similar.

Table 5: Descriptive Statistics on Return and Residual Correlations

	$\hat{\rho}_t^{*Ret}$	$\hat{\rho}_t^{*Res}$
Mean	0.3436*	0.1627
Std. Dev.	0.2432	0.2476
$F\rho^*$	55.47	24.82
C_ρ	0.9511	

Notes: $\hat{\rho}_t^{*Ret}$ is the average oil and gas returns unconditional correlation and $\hat{\rho}_t^{*Res}$ is the average residuals unconditional correlation. (i) $F\rho^*$ is the percentage of correlations significant at the 5% level using the t-ratio test $\hat{t} = \hat{\rho} / \sqrt{\frac{(1-\hat{\rho}^2)}{L-2}} \sim t(N-2)$. (iii) * denotes a significant correlation at the 5% level. (iv) C_ρ is the correlation between $\hat{\rho}_t^{*Res}$ and $\hat{\rho}_t^{*Ret}$.

between ρ_t^{*Ret} and ρ_t^{*Res} is high, filtering oil and gas returns with our estimated factors has noticeably reduced the level of correlation between both returns. The average of $\hat{\rho}_t^{*Ret}$ is above the 5% critical value for significance test, but not the average of ρ_t^{*Res}. Furthermore, while return correlations are significant 55% of the time, residual correlations are significant 25% of the time only.

Residual correlations remain significant for some time periods, evident in Figure 4 where ρ_t^{*Ret} and ρ_t^{*Res} are plotted. We can broadly distinguish two periods during which oil and gas residual correlations are significant, namely from October to December 1998, and more significantly from October 2003 to August 2005. Thus, during these two periods, the correlation of oil and gas returns cannot be satisfyingly explained by the macroeconomic variables, as represented by our set of common factors.

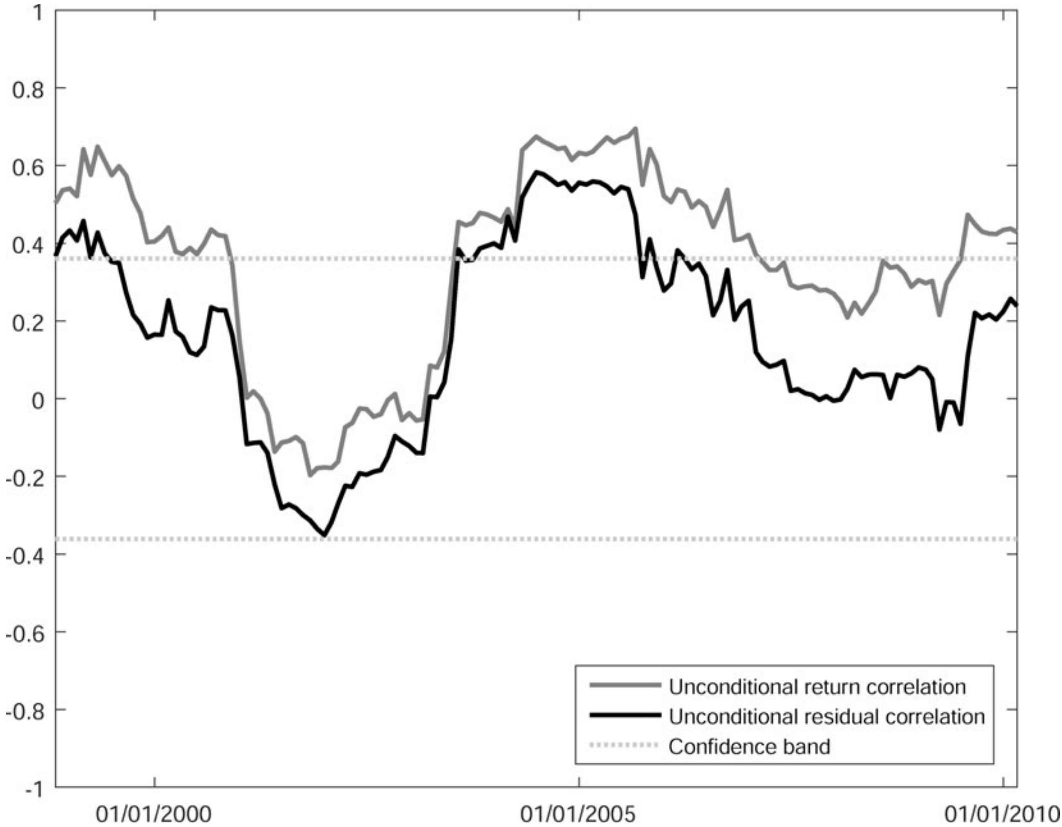
7. FINANCIAL IMPACTS ON THE US OIL-GAS RESIDUAL CORRELATION

As a consequence, the apparent residual correlation of the oil and gas returns poses an open question and, in this context, we explore the applicability of the highly popularized "financialization of commodities" theme. More precisely, we investigate how the degree of trading activity in oil and gas futures can influence co-movements between both commodities. Whilst a vast literature has studied the influence of speculation on the rise in oil prices over the 2000–2008 period (see, among others, Smith (2009), Büyüksahin and Harris (2011), Manera et al. (2013) or Singleton (2014)), there is no consensus view, as emphasized by Fattouh et al. (2013) and Alquist and Gervais (2013).[15] Moreover, except in Tang and Xiong (2012), the impact of financial activities on commodity price co-movements has not been investigated. We perform such an analysis for the oil-gas pair and provide new clarity on the importance of financial aspects.[16]

15. The empirical results in Alquist and Gervais (2013) provide evidence that changes in the positions of financial firms in oil futures markets do not have predictive power for oil returns. See also the empirical analysis in Kim (2015).

16. The empirical approach in Tang and Xiong (2012) radically differs from our approach. Tang and Xiong (2012) compute the correlation between oil returns and on- and off-index non-energy commodities to conclude about the impact of index trading in commodity futures markets. The authors ignore the impact of fundamentals on commodity returns while these fundamentals could be responsible for changes in correlations.

Figure 4: Time Varying Correlation between Oil and Natural Gas Returns and Residuals

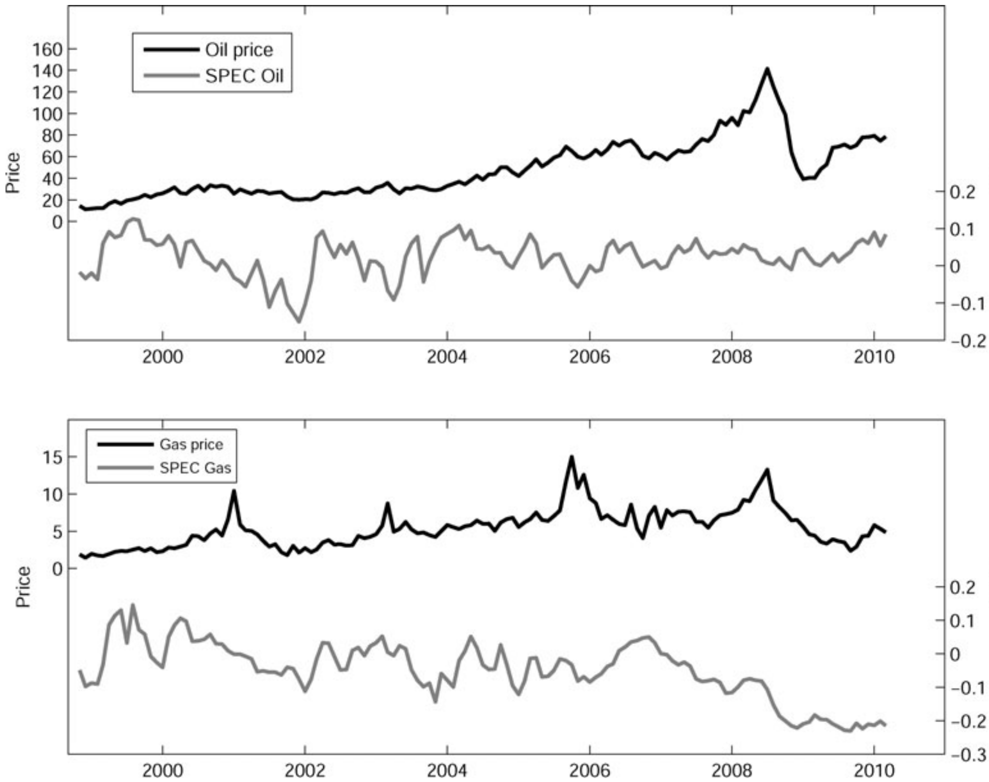


Notes: (i) 'Unconditional return correlation' is $\hat{\rho}_t^{*Ret}$. 'Unconditional residual correlation' is $\hat{\rho}_t^{*Res}$. (ii) The confidence band represents the minimal value above which correlation is significant at the 5% level. It is computed from the t ratio test $\hat{t} = \hat{\rho} / \sqrt{\frac{(1-\hat{\rho}^2)}{N-2}} \sim t(N-2)$.

Our analysis relies on CFTC data to proxy the intensity of speculative activity as undertaken in many studies dealing with the impact of speculation (see the discussion in Büyüksahin and Robe (2014)).[17] The CFTC requires large traders holding positions above a specified level to report their positions on a daily basis. Then, the CFTC aggregates the reported data, and releases the breakdown of each Tuesday's open interest in its Commitments of Traders Report (CoT). This report contains the number of long positions and the number of short positions for both 'commercial' traders and 'non-commercial' traders. The CFTC defines a noncommercial trader as any trader who does not use oil or gas futures contracts to hedge. Traders looking to hedge are typically the producers and consumers of the physical commodity. The noncommercial traders, known as speculators, seek profit by taking positions in the futures market in the hope of gaining from changes in price of the commodity, and do not intend to engage in physical delivery. The boundary between

17. Manera et al. (2016), whose purpose is to relate oil price volatility to the speculative activity in U.S. futures markets, also discuss at length some issues on CFTC data and rely on three different proxies for speculation to increase the robustness of their results. We follow the same idea in considering two different measures of trading activity.

Figure 5: Energy Price and SPEC Variable for Crude Oil (top plot) and Natural Gas (bottom plot)



these two types of traders, however, is somewhat blurred as recently discussed in Büyüksahin and Harris (2011). We, nevertheless, show the significance of our explanatory variables based on these categories.

Our main variable as a proxy for the degree of speculative activity follows Han (2008). We label *SPEC oil* and *SPEC gas* the number of long non-commercial contracts minus the number of short non-commercial contracts scaled by the total open interest for oil and gas futures, respectively.[18] Han (2008) suggests that this measure is a relevant global investor sentiment index when used along with data for the S&P 500 futures. As the index computes the net position of large speculators, it is, indeed, well-suited for proxying speculative intensity. In some respects, Han's (2008) index is reminiscent of the so-called Working's T index measuring excessive speculative positions that has been widely used in research.[19] We plot both *SPEC* indices along with commodity price in Figure 5. From this Figure no visible relationship is apparent between speculative intensity and price behavior, being consistent with the view of a limited impact of speculation on oil and gas prices (Fattouh et al. (2013), Kim (2015)). However, for the reasons developed in the intro-

18. Monthly data are computed from CFTC weekly data using linear interpolation. Our measure considers all maturities.

19. Recent contributions include Manera et al. (2016) or Bruno et al. (2016). We reproduce our empirical analysis using the Working's T index of speculation but our results, available upon request, show that the index is not significant to explain the oil-gas residual correlation.

Table 6: Residual Correlation and Trading Activity

	$\hat{\rho}_t^{*Res}$	
Intercept	0.1218*	0.1412**
	(1.78)	(2.04)
SPEC Oil	2.0097***	
	(2.99)	
SPEC Gas	0.0790	
	(0.27)	
HEDG Oil		-1.4891
		(-1.63)
HEDG Gas		-0.5665**
		(-2.35)
R^2	0.1706	0.1356
$\overline{R}^2$	0.1582	0.1227

Notes: (i) $\hat{\rho}_t^{*Res}$ is the unconditional average residual correlation.(ii) *SPEC Oil* and *SPEC Gas* are the speculative trading activity proxies for oil and gas. *HEDG Oil* and *HEDG Gas* are the proxies for hedging pressure in futures markets.(iii) ***, ** and * respectively denotes significance at the 1%, 5% and 10% levels.

duction we expect the *SPEC* variables to be positively correlated with the co-movements between oil and gas returns.

Results from the regression of the unconditional average residual correlation (i.e. the $\hat{\rho}_t^{*Res}$) on the two contemporaneous exogenous variables *SPEC oil* and *SPEC gas* are reported in the first column of Table 6. As expected, the estimated coefficients are both positive. However, only the proxy for speculation in the oil futures market significantly explains the residual correlation. The explanatory power for the regression is around 17% showing that financial aspects explain a material part of the residual correlation between the filtered oil and gas return series. At this stage, two comments are in order. First, the high significance of the estimated parameter for oil leaves no doubt about the strong link between co-movements and speculative activity in energy futures markets. Second, our estimates reinforce the strong role of the oil market as a leading commodity market. Indeed, increasing speculation in the gas market does not have an impact of the same magnitude than in the case of oil on the correlation of residuals. Overall, our results support the conjecture developed in the introduction, namely that investment from speculators in several commodity markets at the same time makes this category of traders key economic agents for the purpose of linking commodity returns in futures markets.

As an alternative measure of trading activity in futures market, we use the hedging pressure index developed in de Roon et al. (2000). Interestingly, despite this measure is being computed as the difference between the number of short hedge positions and the number of long hedge positions, divided by the total number of hedge positions, thus making use of hedging positions only, it is evidently strongly and negatively correlated with the Han's (2008) *SPEC* index for each commodity.[20] Indeed, the correlation between hedging pressure and the SPEC index is -0.96 for oil

20. Again, linear interpolation is used to form monthly data and contemporaneous variables are used for the regression.

and -0.94 for gas. Empirical results, however, are qualitatively different from those with the *SPEC* index. The estimates for the hedging pressure proxies *HEDG oil* and *HEDG gas* are reported in the last column of Table 6. Consistent with the large negative correlation between *SPEC* and *HEDG* variables, we see that the estimated values are negative for oil and gas. Nevertheless, only the index proxying the hedging pressure in the gas futures market is significant. Also, the explanatory power is slightly reduced with an R^2 around 13.5%. The intuitive interpretation for these empirical findings lies in the intrinsic nature of hedgers who operate in the futures market for commodity-specific risk management purpose, whereas speculators are more likely to invest in a commodity index to diversify their overall financial portfolio. Therefore, variations in hedging activity are prone to decorrelate commodity returns, as our results somewhat indicate. This is verified in the case of natural gas but not in the case of oil, which is likely to be due to the size of the oil market with respect to any other energy commodity (see Asche et al. (2006)).

Overall, our empirical analysis provides strong evidence that the strength of the speculative activity is an important determinant of co-movements between oil and gas returns even once fundamentals have been taken into account. This is an original finding. It adds the subtle aspect of comovement within the commodities to the widely debated issue of the potential influence of speculators on the commodity prices themselves.

8. CONCLUSION

A consequence of the research which has suggested that increased financial engagement in commodity futures will link commodity returns more closely to equity indices (Tang and Xiong (2012), Büyüksahin and Harris (2011), Singleton (2014), Adams and Glück (2015)) and that index-focused investment by itself may increase the correlations amongst the assets within the index (Basak and Pavlova (2016)), it is expected that financial flows into commodities may also manifest increased correlations between actively traded commodities. We tested this hypothesis on U.S. oil and gas futures and found significant support. Our main finding is that speculation, with its focus on index trading, increases the correlation between oil and gas. This is a highly plausible effect that is consistent with the "financialization" considerations.

The methodological challenge in obtaining these results is substantial. Since commodities are global products, they generally have a complex set of fundamental drivers, and this is certainly the case for oil and gas. Oil itself requires careful structural modeling (Hamilton (2009), Kilian (2009)) and the theme of oil-gas linkage has been a lengthy and on-going debate amongst energy economists (Brown and Yücel (2008), Hartley et al. (2008), Ramberg and Parsons (2012), Hartley and Medlock (2014)). We therefore undertook a comprehensive fundamental macroeconomic filtration of oil and gas returns before seeking to associate financial activity with the residual correlation. From a large dataset of macroeconomic and financial variables, we found that three factors can explain a significant proportion of both returns, indicating common economic fundamentals for oil and gas dynamics. This is the first study explaining oil and gas returns with factors extracted from a large dataset in the Stock and Watson (2002a,b) tradition, but ours also includes more international variables from emerging economies. Indeed, we find that the factor with the highest explanatory power for oil is mostly connected with real macroeconomic variables from emerging countries, and this was the one key factor that was not shared in common with gas. Given that gas markets tend to be more local with lower gas penetration in developing countries, this is a very plausible result.

Whilst the large dataset factor filtration was effective, it is an area for further methodological refinement, as it is crucial for the subsequent residual estimations. Thus, we considered, as in most of the factor-models literature, the factors as if they were observed, whilst they are actually estimated. Despite this, the assumption should only have a limited impact on our results. However it could be relevant to investigate the small sample case using some simulation techniques as in Ludvigson and Ng (2007, 2009 and 2010) and Gospodinov and Ng (2013). The evolving nature of these fundamentals is more challenging, as dynamic representations may become necessary. Overall, however, the analysis undertaken here appears to give robust and consistent results to the subtle question of estimating the financial effects on commodity inter-correlations in the context of complex global fundamentals.

APPENDIX

A1. Testing for Seasonality and Reaction to Abnormal Temperatures

Additional results using monthly dummies and CDDdev and HDDdev defined as monthly CCD and HDD minus their 1963–1993 monthly average, respectively, calculated following Hartley and Medlock (2014). Data on both variables are collected from the National Climatic Data Center (NCDC). We refer the interested reader to Hartley and Medlock (2014) for an exposition of why monthly dummies should be considered along with HDD and CDD variables. The estimation sample is November 1993–February 2010 (196 monthly observations).

Table 7: Regression of Oil and Gas Return on CCDdev, HHDdev and Monthly Dummies

	r_oil	r_gas	r_oil	r_gas
Intercept	-0.0003	0.0056	-0.0089	0.0776
	(-0.03)	(0.33)	(-0.32)	(1.33)
CDDdev	0.0008***	0.0013*	0.0009*	0.0027***
	(2.69)	(1.89)	(2.55)	(3.80))
HDDdev	0.0000	0.0006**	0.0000	0.0009**
	(-0.37)	(2.00)	(-0.00)	(2.75)
January			0.0294	-0.0002
			(1.00)	(-0.00)
February			0.0235	-0.1225
			(0.72)	(-1.43)
March			0.0420	-0.0686
			(0.99)	(-0.89)
April			0.0385	-0.0616
			(1.14)	(-1.03)
May			0.0322	-0.0762
			(0.82)	(-1.15)
June			0.0201	-0.1412**
			(0.60)	(-2.21)
July			0.0041	-0.1579**
			(0.12)	(-2.14)
August			-0.0119	-0.2259***
			(-0.30)	(-2.56)
September			-0.0002	-0.0424
			(-0.00)	(-0.62)
October			-0.0413	0.0065
			(-1.00)	(0.07)
November			-0.0322	-0.0378
			(-0.89)	(-0.49)
R^2	0.0240	0.0314	0.0919	0.1127
$\bar{R}$	0.0139	0.0214	0.0270	0.0493

Note: (i) CDDdev and HDDdev are CCD and HDD minus their average for each months for the time period 1963–1993. (ii) ***,**, and * respectively denote a rejection of the appropriate null hypothesis at the 1%, 5% and 10% levels.

A2. List and Treatment of Macroeconomic and Financial Variables

This appendix provides a list of the 187 variables that are used in the empirical analysis to compute factors. For each variable, we provide a short name, a mnemonic code as well as a description of the variable.

In the Trans column, we report the transformation used to ensure the stationarity of each variable. ln denotes the logarithm, Δln and $\Delta^2 ln$ denote the first and second difference of the logarithm, lv denotes the level of the series, and Δlv denotes the first difference of the series.

Developed countries				
Series Number	Short name	Mnemonic	Trans	Description
Industrial production				
1	IP: US	USIPTOT_G	$\Delta \ln$	US INDUSTRIAL PRODUCTION - TOTAL INDEX VOLA (2002=100)
2	IP: US	USIPMFGSG	$\Delta \ln$	US INDUSTRIAL PRODUCTION - MANUFACTURING (SIC) VOLA (1997=100)
3	IP: Canada	CNIPTOT.C	$\Delta \ln$	CN GDP - INDUSTRIAL PRODUCTION CONN
4	IP: France	FRIPMAN.G	$\Delta \ln$	FR INDUSTRIAL PRODUCTION - MANUFACTURING VOLA
5	IP: France	FRIPTOT_G	$\Delta \ln$	FR INDUSTRIAL PRODUCTION EXCLUDING CONSTRUCTION VOLA INDEX (2005=100)
6	IP: Germany	BDIPTOT_G	$\Delta \ln$	BD INDUSTRIAL PRODUCTION INCLUDING CONSTRUCTION VOLA (2005=100)
7	IP: UK	UKIPTOT.G	$\Delta \ln$	UK INDEX OF PRODUCTION - ALL PRODUCTION INDUSTRIES VOLA (2003=100)
8	IP: UK	UKIPMAN.G	$\Delta \ln$	UK INDUSTRIAL PRODUCTION INDEX - MANUFACTURING VOLA (2003=100)
9	IP: Japan	JPIPTOT_G	$\Delta \ln$	JP INDUSTRIAL PRODUCTION - MINING & MANUFACTURING VOLA (2005=100)
Orders and capacity utilization				
10	Capacity utilization: US	USCUMANUG	Δlv	US CAPACITY UTILIZATION - MANUFACTURING VOLA
11	Manufct. new ord.: US	USNOCOGMC	$\Delta^2 \ln$	US MANUFACTURERS NEW ORDERS - CONSUMER GOODS AND MATERIALS CONN (base 1982)
12	Manufct. new ord.: US	USBNKRTEQ	$\Delta \ln$	US MANUFACTURERS NEW ORDERS,NONDEFENSE CAPITAL GOODS SADJ (base 1982)
13	New orders: Canada	CNNEWORDB	$\Delta \ln$	CN NEW ORDERS: ALL MANUFACTURING INDUSTRIES (SA) CURA
14	Manufct. ord.: Germany	BDNEWORDE	$\Delta \ln$	BD MANUFACTURING ORDERS SADJ (2000=100)
15	Manufct. ord.: Japan	JPNEWORDB	$\Delta \ln$	JP MACHINERY ORDERS: DOM.DEMAND-PRIVATE DEMAND (EXCL. SHIP) CURA
16	Operating ratio: Japan	JPCAPUTLQ	Δlv	JP OPERATING RATIO - MANUFACTURING SADJ (2005=100)
17	Business failures: Japan	JPBNKRPTP	$\Delta \ln$	JP BUSINESS FAILURES VOLN
Housing start				
18	Housing permits: US	USHOUSETOT	$\ln$	US HOUSING AUTHORIZED VOLN
19	Housing permits: Canada	CNHOUSE.O	$\ln$	CN HOUSING STARTS: ALL AREAS (SA, AR) VOLA
20	Housing permits: Germany	BDHOUSINP	$\ln$	BD HOUSING PERMITS ISSUED FOR BLDG.CNSTR.: BLDG.S-RESL, NEW VOLN
21	Housing permits: Australia	AUHOUSE_A	$\ln$	AU BUILDING APPROVALS: NEW HOUSES CURN
22	Housing permits: Japan	JPHOUSSTF	$\ln$	JP NEW HOUSING CONSTRUCTION STARTED VOLN
Car sales				
23	Car registration: US	USCAR_P	$\ln$	US NEW PASSENGER CARS - TOTAL REGISTRATIONS VOLN
24	Car registration: Canada	CNCARSLSE	$\ln$	CN PASSENGER CAR SALES:TOTAL SADJ
25	Car registration: France	FRCARREGP	$\ln$	FR NEW CAR REGISTRATIONS VOLN
26	Car registration: Germany	BDRVNCARP	$\ln$	BD NEW REGISTRATIONS - CARS VOLN
27	Car registration: UK	UKCARTOTF	$\ln$	UK CAR REGISTRATIONS VOLN
28	Car registration : Japan	JPCARREGF	$\ln$	JP MOTOR VEHICLE NEW REGISTRATIONS: PASSENGER CARS EXCL.BELOW 66
Consumption				
29	Consumer sentiment: US	USUMCONEH	$\Delta \ln$	US UNIV OF MICHIGAN CONSUMER SENTIMENT - EXPECTATIONS VOLN (base 1966=100)
30	Pers. cons. exp.: US	USPERCONB	$\Delta \ln$	US PERSONAL CONSUMPTION EXPENDITURES (AR) CURA
31	Pers. saving: US	USPERSAVE	Δlv	US PERSONAL SAVING AS % OF DISPOSABLE PERSONAL INCOME SADJ
32	Retail sale: Canada	CNRETTOTB	$\Delta \ln$	CN RETAIL SALES: TOTAL (ADJUSTED) CURA
33	Household confidence: France	FRCNFCONQ	Δlv	FR SURVEY - HOUSEHOLD CONFIDENCE INDICATOR SADJ
34	Household confidence: Germany	BDCNFCONQ	Δlv	BD CONSUMER CONFIDENCE INDICATOR - GERMANY SADJ
35	Retail sales: UK	UKRETTOTB	$\Delta \ln$	UK RETAIL SALES (MONTHLY ESTIMATE, DS CALCULATED) CURA
36	Household confidence: UK	UKCNFCONQ	Δlv	UK CONSUMER CONFIDENCE INDICATOR - UK SADJ
37	Retail sales: Australia	AURETTOTT	$\Delta \ln$	AU RETAIL SALES (TREND) VOLA
38	Household confidence: Australia	AUCNFCONR	Δlv	AU MELBOURNE/WESTPAC CONSUMER SENTIMENT INDEX NADJ
39	Household expenditure: Japan	JPHLEXPWA	$\Delta \ln$	JP WORKERS HOUSEHOLD LIVING EXPENDITURE (INCL. AFF) CURN
40	Retail sales: Japan	JPRETTOTA	$\Delta \ln$	JP RETAIL SALES CURN

Series Number	Short name	Mnemonic	Trans	Description
Wages and labor				
41	Av. hourly real earnings: US	USWRIM_D	Δ ln	US AVG HOURLY REAL EARNINGS - MANUFACTURING CONA (base 82-84)
42	Av. overtime hours: US	USOOL024Q	Δ ln	US OVERTIME HOURS - MANUFACTURING, WEEKLY VOLA
43	Av. wkly hours : US	USHKIM_O	Δ ln	US AVG WKLY HOURS - MANUFACTURING VOLA
44	Purchasing manager index: US	USPMCUE	Δ ln	US CHICAGO PURCHASING MANAGER DIFFUSION INDEX - EMPLOYMENT NADJ
45	Av. hourly real earnings: Canada	CNWAGES.A	Δ ln	CN AVG.HOURLY EARN- INDUSTRIAL AGGREGATE EXCL. UNCLASSIFIED CURN
46	Labor productivity: Germany	BDPRODVTQ	Δ ln	BD PRODUCTIVITY: OUTPUT PER MAN-HOUR WORKED IN INDUSTRY SADJ (2005=100)
47	wages: Germany	BDWAGES.F	Δ ln	BD WAGE & SALARY,OVERALL ECONOMY-ON A MTHLY BASIS(PAN BD M0191)
48	Labor productivity: Japan	JPPRODVTE	Δ ln	JP LABOR PRODUCTIVITY INDEX -ALL INDUSTRIES SADJ
49	wages index: Japan	JPWAGES_E	Δ ln	JP WAGE INDEX: CASH EARNINGS - ALL INDUSTRIES SADJ
Unemployment				
50	U rate: US	USUNEM15Q	Δ^2 ln	US UNEMPLOYMENT RATE - 15 WEEKS & OVER SADJ
51	U rate: US	USUNTOTQ_pc	Δ^2 ln	US UNEMPLOYMENT RATE SADJ
52	Employment: Canada	CNEMPTOTO	Δ^2 ln	CN EMPLOYMENT - CANADA (15 YRS & OVER, SA) VOLA
53	U all: Germany	BDUNPTOTP	Δ ln	BD UNEMPLOYMENT LEVEL (PAN BDFROM SEPT 1990) VOLN
54	U rate: UK	UKUNTOTQ_pc	Δ^2 ln	UK UNEMPLOYMENT RATE SADJ
55	Emp: Australia	AUEMPTOTO	Δ ln	AU EMPLOYED: PERSONS VOLA
56	U all: Australia	AUUNPTOTO	Δ ln	AU UNEMPLOYMENT LEVEL VOLA
57	U rate: Japan	JPUNTOTQ_pc	Δlv	JP UNEMPLOYMENT RATE SADJ
International trade				
58	Exports: US	USI70_A	Δ ln	US EXPORTS CURN
59	Exports: EU	EKEXPGDSA	Δ ln	EK EXPORTS TO EXTRA-EA17 CURN
60	Exports: France	FREXPGDSB	Δ ln	FR EXPORTS FOB CURA
61	Exports: Germany	BDEXPBOPB	Δ ln	BD EXPORTS FOB CURA
62	Exports: UK	UKI70_A	Δ ln	UK EXPORTS CURN
63	Exports: Australia	AUEXPG&SB	Δ ln	AU EXPORTS OF GOODS & SERVICES (BOP BASIS) CURA
64	Exports: Japan	JPEXPGDSB	Δ ln	JP EXPORTS OF GOODS - CUSTOMSBASIS CURA
65	Imports: US	USIMPGDSB	Δ ln	US IMPORTS F.A.S. CURA
66	Imports: EU	EUOXT_09B	Δ ln	EU IMPORTS CURA
67	Imports: France	FRIMPGDSB	Δ ln	FR IMPORTS FOB CURA
68	Imports: Germany	BDIMPGDSB	Δ ln	BD IMPORTS CIF (PAN BD M0790) CURA
69	Imports: UK	UKIMPBOPB	Δ ln	UK IMPORTS - BALANCE OF PAYMENTS BASIS CURA
70	Imports: Australia	AUIMPG&SB	Δ ln	AU IMPORTS OF GOODS & SERVICES (BOP BASIS) CURA
71	Imports: Japan	JPOXT009B	Δ ln	JP IMPORTS CURA
72	Terms of trade: UK	UKTOTPRCF	Δ ln	UK TERMS OF TRADE - EXPORT/IMPORT PRICES (BOP BASIS) NADJ
73	Terms of trade: Japan	JPTOTPRCF	Δ ln	JP TERMS OF TRADE INDEX NADJ
Money and credit				
74	Money supply: US	USM0_B	Δ^2 ln	US MONETARY BASE CURA
75	Money supply: US	USM2_B	Δ^2 ln	US MONEY SUPPLY M2 CURA
76	Money supply: France	FRM2_A	Δ ln	FR MONEY SUPPLY - M2 (NATIONAL CONTRIBUTION TO M2) CURN
77	Money supply: France	FRM3_A	Δ ln	FR MONEY SUPPLY - M3 (NATIONAL CONTRIBUTION TO M3) CURN
78	Money supply: Germany	BDM1_A	Δ ln	BD MONEY SUPPLY-GERMAN CONTRIBUTION TO EURO M1(PAN BD M0790)
79	Money supply: Germany	BDM3_B	Δ ln	BD MONEY SUPPLY-M3 (CONTRIBUTION TO EURO BASIS FROM M0195) CURA
80	Money supply: UK	UKM1_B	Δ ln	UK MONEY SUPPLY M1 (ESTIMATE OF EMU AGGREGATE FOR THE UK) CURA
81	Money supply: UK	UKM3_B	Δ ln	UK UK MONEY SUPPLY M3(ESTIMATE OF EMU AGGREGATE FORTHE UK) CURA
82	Money supply: Australia	AUM1_B	Δ ln	AU MONEY SUPPLY - M1 CURA
83	Money supply: Australia	AUM3_B	Δ^2 ln	AU MONEY SUPPLY - M3 (SEE AUM3...OB) CURA
84	Money supply: Japan	JPM1_A	Δ ln	JP MONEY SUPPLY: M1 (METHO-BREAK, APR. 2003) CURN
85	Money supply: Japan	JPM2_A	Δ ln	JP MONEY SUPPLY: M2 (METHO-BREAK, APR. 2003) CURN

Series Number	Short name	Mnemonic	Tran	Description
Money and credit - continuation				
86	Credit: US	USCOMILND	$\Delta^2 \ln$	US COMMERCIAL & INDUSTRIAL LOANS OUTSTANDING (BCI 101) CONA (base 2005)
87	Credit: US	USCILNNCB	Δlv	US COMMERCIAL & INDL LOANS, NET CHANGE (AR) (BCI 112) CURA
88	Credit: US	USCRDNRVB	$\Delta^2 \ln$	US NONREVOLVING CONSUMER CREDIT OUTSTANDING CURA
89	Credit: US	USCSCRE_Q	$\Delta^2 \ln$	US CONSUMER INSTALLMENT CREDIT TO PERSONAL INCOME (RATIO) SADJ
90	Credit: France	FRBANKLPA	$\Delta^2 \ln$	FR MFI LOANS TO RESIDENT PRIVATE SECTOR CURN
91	Credit: Germany	BDBANKLPA	$\Delta^2 \ln$	BD LENDING TO ENTERPRISES & INDIVIDUALS CURN
92	Credit: UK	UKCRDCONB	$\Delta^2 \ln$	UK TOTAL CONSUMER CREDIT: AMOUNT OUTSTANDING CURA
93	Credit: Australia	AUCRDCONB	$\Delta^2 \ln$	AU FINANCIAL INTERMEDIARIES: NARROW CREDIT - PRIVATE SECTOR CURA
94	Credit: Japan	JPBANKLPA	$\Delta^2 \ln$	JP AGGREGATE BANK LENDING (EXCL. SHINKIN BANKS) CURN
Stock index				
95	Stock index: US	USSHRPRCF	$\Delta \ln$	US DOW JONES INDUSTRIALS SHARE PRICE INDEX (EP) NADJ
96	Stock index: France	FRSHRPRCF	$\Delta \ln$	FR SHARE PRICE INDEX - SBF 250 NADJ
97	Stock index: Germany	BDSHRPRCF	$\Delta \ln$	BD DAX SHARE PRICE INDEX, EP NADJ
98	Stock index: UK	UKOSP001F	$\Delta \ln$	UK FTSE 100 SHARE PRICE INDEXNADJ (2005=100)
99	Stock index: Japan	JPSHRPRCF	$\Delta \ln$	JP TOKYO STOCK EXCHANGE - TOPIX (EP) NADJ (1968=100)
Interest rate				
100	Interest rate: US	USFEDFUN	Δlv	US FEDERAL FUNDS RATE (AVG.)
101	Interest rate: US	USCRBBAA	Δlv	US CORPORATE BOND YIELD - MOODY'S BAA, SEASONED ISSUES
102	Interest rate: US	USGBOND	Δlv	US TREASURY YIELD ADJUSTED TO CONSTANT MATURITY - 20 YEAR
103	Interest rate: France	FRPRATE	Δlv	FR AVERAGE COST OF FUNDS FOR BANKS / EURO REPO RATE
104	Interest rate: France	FRGBOND	Δlv	FR GOVERNMENT GUARANTEED BOND YIELD (EP) NADJ
105	Interest rate: Germany	BDPRATE	Δlv	BD DISCOUNT RATE / SHORT TERM EURO REPO RATE
106	Interest rate: Germany	BDGBOND	Δlv	BD LONG TERM GOVERNMENT BOND YIELD - 9-10 YEARS
107	Interest rate: UK	UKPRATE	Δlv	UK BANK OF ENGLAND BASE RATE (EP)
108	Interest rate: UK	UKGBOND	Δlv	UK GROSS REDEMPTION YIELD ON 20 YEAR GILTS (PERIOD AVERAGE) NADJ
109	Interest rate: Australia	AUPRATE	Δlv	AU RBA CASH RATE TARGET
110	Interest rate: Australia	AUBOND	Δlv	AU COMMONWEALTH GOVERNMENT BOND YIELD 10 YEAR (EP)
111	Interest rate: Japan	JPPRATE	Δlv	JP OVERNIGHT CALL MONEY RATE, UNCOLLATERALISED (EP)
112	Interest rate: Japan	JPGBOND	Δlv	JP INTEREST-BEARING GOVERNMENT BONDS - 10-YEAR (EP)
Exchange rate				
113	Exchange rate: DM to US $	BBDEMSP	$\Delta \ln$	GERMAN MARK TO US $ (BBI) - EXCHANGE RATE
114	Exchange rate: SK to US $	SDXRUSD	$\Delta \ln$	SD SWEDISH KRONOR TO US $ (BBI, EP)
115	Exchange rate: £ to $	UKDOLLR	$\Delta \ln$	UK £ TO US $ (WMR) - EXCHANGE RATE
116	Exchange rate: Yen to $	JPXRUSD	$\Delta \ln$	JP JAPANESE YEN TO US $
117	Exchange rate: Aus.$ to US $	AUXRUSD	$\Delta \ln$	AU AUSTRALIAN $ TO US $ (MTH.AVG.)
Producer price index				
118	PPI: US	USPROPRCE	$\Delta \ln$	US PPI - FINISHED GOODS SADJ
119	PPI: Canada	CNPROPRCF	$\Delta \ln$	CN INDUSTRIAL PRICE INDEX: ALL COMMODITIES NADJ
120	PPI: Germany	BDPROPRCF	$\Delta \ln$	BD PPI: INDL. PRODUCTS, TOTAL, SOLD ON THE DOMESTIC MARKET NADJ (2005=100)
121	PPI: UK	UKPROPRCF	$\Delta \ln$	UK PPI - OUTPUT OF MANUFACTURED PRODUCTS (HOME SALES) NADJ
122	PPI: Japan	JPPROPRCF	$\Delta \ln$	JP CORPORATE GOODS PRICE INDEX: DOMESTIC - ALL COMMODITIES NADJ
Consumer price index				
123	CPI: US	USCONPRCE	$\Delta \ln$	US CPI - ALL URBAN: ALL ITEMS SADJ
124	CPI: Canada	CNCONPRCF	$\Delta \ln$	CN CPI NADJ
125	CPI: France	FRCONPRCE	$\Delta \ln$	FR CPI SADJ
126	CPI: Germany	BDCONPRCE	$\Delta \ln$	BD CPI SADJ
127	CPI: UK	UKD7BT_F	$\Delta \ln$	UK CPI INDEX 00 : ALL ITEMS- ESTIMATED PRE-97 2005=100 NADJ
128	CPI: Japan	JPCONPRCF	$\Delta \ln$	JP CPI: NATIONAL MEASURE NADJ

Series Number	Short name	Mnemonic	Trans	Description
Emerging countries				
Industrial production				
129	IP: Brasil	BRIPTOT_G	Δ ln	BR INDUSTRIAL PRODUCTION VOLA index 2002=base
130	IP: China (cement)	CHVALCEMH	Δ ln	CH OUTPUT OF INDUSTRIAL PRODUCTS - CEMENT VOLN
131	IP: India	INIPTOT_H	Δ ln	IN INDUSTRIAL PRODN. (EXCLUDING CONSTRUCTION & GAS UTILITY) VOLN index
132	IP: India	INIPMAN_H	Δ ln	IN INDUSTRIAL PRODUCTION: MANUFACTURING VOLN index
133	IP: Korea	KOIPTOT.G	Δ ln	KO INDUSTRIAL PRODUCTION VOLA (2005=100)
134	IP: Mexico	MXIPTOT_H	Δ ln	MX INDUSTRIAL PRODUCTION INDEX VOLN
135	IP: Mexico	MXIPMAN_H	Δ ln	MX INDUSTRIAL PRODUCTION INDEX: MANUFACTURING VOLN
136	IP: Philippines	PHIPMAN_F	Δ ln	PH MANUFACTURING PRODUCTION NADJ 2000 prices
137	IP: South Africa	SAIPMAN.G	Δ ln	SA INDUSTRIAL PRODUCTION (MANUFACTURING SECTOR) VOLA
Orders and capacity utilization				
138	Operating ratio: Brazil	BRCAPUTLR	Δlv	BR CAPACITY UTILIZATION - MANUFACTURING NADJ
139	Mach. ord.: Korea	KONEWORDA	Δ ln	KO MACHINERY ORDERS RECEIVEDCURN
140	Manufct. prod capa.: Korea	KOCAPUTLF	Δlv	KO MANUFACTURING PRODUCTION CAPACITY NADJ (2005=100)
Consumption				
141	Retail sales: Korea	KORETTOTF	Δ ln	KO RETAIL SALES NADJ (2005=100)
Wages and labor				
142	Labor cost: Brazil	BRLCOST.F	Δ ln	BR UNIT LABOR COST NADJ
Unemployment				
143	U rate: Korea	KOUNTOTQ_pc	Δlv	KO UNEMPLOYMENT RATE SADJ
International trade				
144	Exports: Brazil	BREXPBOPA	Δ ln	BR EXPORTS (BOP BASIS) CURN
145	Exports: China	CHEXPGDSA	Δ ln	CH EXPORTS CURN
146	Exports: India	INI70_A	Δ ln	IN EXPORTS CURN
147	Exports: Indonesia	IDEXPGDSA	Δ ln	ID EXPORTS FOB CURN
148	Exports: Korea	KOEXPGDSA	Δ ln	KO EXPORTS FOB (CUSTOMS CLEARANCE BASIS) CURN
149	Exports: Philippines	PHEXPGDSA	Δ ln	PH EXPORTS CURN
150	Exports: Singapore	SPEXPGDSA	Δ ln	SP EXPORTS CURN
151	Exports: Tawan	TWEXPGDSA	Δ ln	TW EXPORTS CURN
152	Imports: Brazil	BRIMPBOPA	Δ ln	BR IMPORTS (BOP BASIS) CURN
153	Imports: China	CHIMPGDSA	Δ ln	CH IMPORTS CURN
154	Imports: Indonesia	IDIMPGDSA	Δ ln	ID IMPORTS CIF CURN
155	Imports: Korea	KOIMPGDSA	Δ ln	KO IMPORTS CIF (CUSTOMS CLEARANCE BASIS) CURN
156	Imports: Singapore	SPIMPGDSA	Δ ln	SP IMPORTS CURN
157	Imports: Taïwan	TWIMPGDSA	Δ ln	TW IMPORTS CURN
158	Terms of trade: Brazil	BRTOTPRCF	Δ ln	BR TERMS OF TRADE NADJ (2006=100)

Series Number	Short name	Mnemonic	Tran	Description
Money and credit				
159	Money supply: Brazil	BRM1_A	Δ ln	BR MONEY SUPPLY - M1 (EP) CURN
160	Money supply: Brazil	BRM3_A	Δ ln	BR MONEY SUPPLY - M3 (EP) CURN
161	Money supply: China	CHM0_A	Δ ln	CH MONEY SUPPLY - CURRENCY IN CIRCULATION CURN
162	Money supply: China	CHM1_A	Δ ln	CH MONEY SUPPLY - M1 CURN
163	Money supply: India	INM1_A	Δ ln	IN MONEY SUPPLY: M1 (EP) CURN
164	Money supply: India	INM3_A	Δ ln	IN MONEY SUPPLY: M3 (EP) CURN
165	Money supply: Indonesia	IDM1_A	Δ ln	ID MONEY SUPPLY: M1 CURN
166	Money supply: Indonesia	IDM2_A	Δ^2 ln	ID MONEY SUPPLY- M2 CURN
167	Money supply: Korea	KOM2_B	Δ^2 ln	KO MONEY SUPPLY - M2 (EP) CURA
168	Money supply: Mexico	MXM1_A	Δ ln	MX MONEY SUPPLY: M1 (EP) CURN base=end of period
169	Money supply: Mexico	MXM3_A	Δ^2 ln	MX MONEY SUPPLY: M3 (EP) CURN
170	Money supply: Philippines	PHM1_A	Δ ln	PH MONEY SUPPLY - M1 (METHO BREAK AT 12/03) CURN
171	Money supply: Philippines	PHM3_A	Δ^2 ln	PH MONEY SUPPLY - M3 (METHO BREAK AT 12/03) CURN
172	Money supply: Russia	RSM2_A	Δ^2 ln	RS MONEY SUPPLY- M2 CURN
Stock index				
173	Stock index: Brazil	BRSHRPRCF	Δ^2 ln	BR BOVESPA SHARE PRICE INDEX (EP) NADJ
174	Stock index: Hong-Kong	HKSHRPRCF	Δ ln	HK HANG SENG SHARE PRICE INDEX (EP) NADJ (31 july 1964 =100)
Exchange rate				
175	Exchange rate: Br.R. to US $	BRXRUSD	Δ^2 ln	BR BRAZILIAN REAIS TO US DOLLAR (AVG)
176	Exchange rate: Ch.Y. to US $	CHXRUSD	Δ^2 ln	CH CHINESE YUAN TO US DOLLAR (AVERAGE AMOUNT)
177	Exchange rate: In.R. to US $	INXRUSD	Δ^2 ln	IN INDIAN RUPEES PER US DOLLAR (RBI)
178	Exchange rate: Id.R. to US $	IDXRUSD	Δ^2 ln	ID INDONESIAN RUPIAHS TO US DOLLAR
179	Exchange rate: Mx.P. to US $	MXXRUSD	Δ^2 ln	MX MEXICAN PESOS TO US $-CENTRAL BANK SETTLEMENT RATE (AVG)
180	Exchange rate: RS.R. to US $	RSXRUSD	Δ^2 ln	RS RUSSIAN ROUBLES TO US $ NADJ
Consumer price index				
181	CPI: Brazil	BRCPIGENF	Δ^2 ln	BR CPI - GENERAL NADJ
182	CPI: China	CHCONPRCF	Δ ln	CH CPI NADJ
183	CPI: India	INCONPRCF	Δ ln	IN CPI: INDUSTRIAL LABOURERS(DS CALCULATED) NADJ (2001=100)
184	CPI: Korea	KOCONPRCF	Δ ln	KO CPI NADJ (2005=100)
185	CPI: Mexico	MXCONPRCF	Δ^2 ln	MX CPI NADJ (JUN 2002=100)
186	CPI: Philippines	PHCONPRCF	Δ ln	PH CPI NADJ
187	CPI: Russia	RSCONPRCF	Δ^2 ln	RS CPI NADJ

ACKNOWLEDGEMENTS

This article has greatly benefited from the guidance of the editor and the valuable comments and suggestions of two anonymous reviewers of this journal. We also are grateful to participants at the 34th IAEE International Conference in Stockholm and the 51st EWGCFM Annual Meeting in London for their numerous remarks on earlier versions of this paper.

REFERENCES

Aastveit, K.A., H.C. Bjørnland and L.A. Thorsrud (2015. "What drives oil prices? Emerging versus developed economies. *Journal of Applied Econometrics* 30: 1013–1028. http://dx.doi.org/10.1002/jae.2406

Acharya, V.V., L.A. Lochstoer and T. Ramadorai (2013). "Limits to arbitrage and hedging: Evidence from commodity markets. *Journal of Financial Economics* 109: 441–465. http://dx.doi.org/10.1016/j.jfineco.2013.03.003

Adams, Z. and T. Glück (2015). "Financialization in commodity markets: A passing trend or the new normal?" *Journal of Banking & Finance* 60: 93–111. http://dx.doi.org/10.1016/j.jbankfin.2015.07.008

d'Agostino, A. and D. Giannone (2012). "Comparing alternative predictors based on large-panel factor models." *Oxford Bulletin of Economics and Statistics* 74: 306–326. http://dx.doi.org/10.1111/j.1468-0084.2011.00642.x

Alquist, R. and O. Gervais (2013). "The role of financial speculation in driving the price of crude oil." *The Energy Journal* 34: 35–52. http://dx.doi.org/10.5547/01956574.34.3.3

Arezki, R., P. Loungani, R. van der Ploeg and A.J. Venables (2014). "Understanding international commodity price fluctuations." *Journal of International Money and Finance* 42: 1–8. http://dx.doi.org/10.1016/j.jimonfin.2013.08.002

Asche, F., P. Osmundsen and M. Sandsmark (2006). "The UK market for natural gas, oil and electricity: Are the prices decoupled?" *The Energy Journal* 27: 27–40. http://dx.doi.org/10.5547/ISSN0195-6574-EJ-Vol27-No2-2

Bachmeier, L.J. and J.M. Griffin (2006). "Testing for market integration, crude oil, coal and natural gas." *The Energy Journal* 27: 55–71. http://dx.doi.org/10.5547/ISSN0195-6574-EJ-Vol27-No2-4

Bai, J. (2003). "Inferential theory for factor models of large dimensions." *Econometrica* 71: 135–172. http://dx.doi.org/10.1111/1468-0262.00392

Bai, J. and S. Ng (2002). "Determining the number of factors in approximate factor models." *Econometrica* 70: 191–221. http://dx.doi.org/10.1111/1468-0262.00273

Bai, J. and S. Ng (2008). "Large dimensional factor analysis." *Foundations and Trends in Econometrics* 3: 89–163. http://dx.doi.org/10.1561/0800000002

Barsky, R.B. and L. Kilian (2002). "Do we really know that oil caused the great stagflation? A monetary alternative." In: Ben S. Bernanke and Kenneth Rogoff (Eds.) *NBER Macroeconomics Annual 2001*, vol. 16: 137–183. Cambridge, MA: MIT Press.

Basak, S. and A. Pavlova (2016). "A model of financialization of commodities," *Journal of Finance*, forthcoming.

Basak, S. and A. Pavlova (2013). "Asset prices and institutional investors," *American Economic Review* 103: 1728–1758. http://dx.doi.org/10.1257/aer.103.5.1728

Boivin, J. and S. Ng (2006) "Are more data always better for factor analysis?" *Journal of Econometrics* 132: 169–194. http://dx.doi.org/10.1016/j.jeconom.2005.01.027

Boyer, B., M. Gibson and M. Loretan (1999). "Pitfalls in tests for changes in correlations." Federal Reserve Board Working Paper, vol. 597R available at http://www.federalreserve.gov/pubs/ifdp/1997/597/ifdp597.pdf

Brown, S.P.A. and M.K. Yücel (2008). "What drives natural gas prices?" *The Energy Journal* 29: 45–60. http://dx.doi.org/10.5547/ISSN0195-6574-EJ-Vol29-No2-3

Bruno, V.G., B. Büyüksahin and M.A. Robe (2014). "The financialization of food?" SSRN Working Paper available at http://papers.ssrn.com/sol3/papers.cfm?abstract_id = 2323064

Büyüksahin, B. and J.H. Harris (2011). "Do Speculators Drive Crude Oil Futures Prices?" *The Energy Journal* 32: 167–202. http://dx.doi.org/10.5547/ISSN0195-6574-EJ-Vol32-No2-7

Büyüksahin, B. and M.A. Robe (2014). "Speculators, commodities and cross-market linkages." *Journal of International Money and Finance* 42: 38–70. http://dx.doi.org/10.1016/j.jimonfin.2013.08.004

Brunetti, C., B. Büyüksahin and J.H. Harris (2011). "Speculators, prices and market volatility. SSRN Working Paper available at http://papers.ssrn.com/sol3/papers.cfm?abstract_id = 1736737. http://dx.doi.org/10.2139/ssrn.1736737.

Etienne, X.L., S.H. Irwin and P. Garcia (2015). New evidence that index traders did not drive large bubbles in grain futures markets. Working Paper, University of Illinois at Urbana-Champaign. http://www.farmdoc.illinois.edu/irwin/research/2015_Index_Trading_final.pdf

Ewing, B.T., F. Malik and O. Ozfidan (2002). "Volatility transmission in the oil and natural gas markets." *Energy Economics* 24: 525-538. http://dx.doi.org/10.1016/S0140-9883(02)00060-9

Fattouh, B., L. Kilian and L. Mahadeva (2013). "The role of speculation in oil markets: What have we learned so far?" *The Energy Journal* 34: 7–33. http://dx.doi.org/10.5547/01956574.34.3.2

Forbes, K. and R. Rigobon (2002). "No contagion, only interdependence: measuring stock market co-movements." *Journal of Finance* 57: 2223–2261. http://dx.doi.org/10.1111/0022-1082.00494

Forni, M., M. Hallin, M. Lippi and L. Reichlin (2005). "The generalized dynamic factor model, one sided estimation and forecasting." *Journal of the American Statistical Association* 100: 830–840. http://dx.doi.org/10.1198/016214504000002050

Frankel, J.A. (2006). "The effects of monetary policy on real commodity prices." NBER Working Paper no. 12713, Cambridge, MA. http://www.nber.org/papers/w12713.pdf

Gospodinov, N. and S. Ng (2013). "Commodity prices, convenience yields, and inflation." *Review of Economics and Statistics* 95: 206–219. http://dx.doi.org/10.1162/REST_a_00242

Hamilton, J.D. (2009). "Causes and consequences of the oil shock of 2007-08," *Brookings Papers on Economic Activity* 1 (Spring): 215–261. http://www.brookings.edu/ /media/Projects/BPEA/Spring2009/2009a_bpea_hamilton.PDF

Han, B. (2008). "Investor sentiment and option prices." *Review of Financial Studies* 21: 387–414. http://dx.doi.org/10.1093/rfs/hhm071

Hartley, P.R. and K.B. Medlock (2014). "The relationship between crude oil and natural gas prices: The role of the exchange rate" *The Energy Journal* 35: 25–44. http://dx.doi.org/10.5547/01956574.35.2.2

Hartley, P.R., K.B. Medlock and J.E. Rosthal (2008). "The relationship of natural gas to oil prices." *The Energy Journal* 29: 47–65. http://dx.doi.org/10.5547/ISSN0195-6574-EJ-Vol29-No3-3

Juvenal, L. and I. Petrella. (2015) "Speculation in the oil market." *Journal of Applied Econometrics* 30: 621–649. http://dx.doi.org/10.1002/jae.2388

Kallberg, J. and P. Pasquariello (2008). "Time-series and cross-sectional excess comovement in stock indexes." *Journal of Empirical Finance* 15: 481–502. http://dx.doi.org/10.1016/j.jempfin.2007.06.002

Kapetanios, G. (2010). "A testing procedure for determining the number of factors in approximate factor models with large datasets." *Journal of Business and Economic Statistics* 28: 397–409. http://dx.doi.org/10.1198/jbes.2009.07239

Kilian, L. (2009). "Not all oil price shocks are alike: Disentangling demand and supply shocks in the crude oil market." *American Economic Review* 99: 1053–1069. http://dx.doi.org/10.1257/aer.99.3.1053

Kim, A. (2015). "Does futures speculation destabilize commodity markets?" *Journal of Futures Markets* 35: 696–714. http://dx.doi.org/10.1002/fut.21716

Lehecka, G.V. (2014). "Have food and financial markets integrated?" *Applied Economics* 46: 2087–2095. http://dx.doi.org/10.1080/00036846.2014.894634

Le Pen, Y. and B. Sévi (2011) "Macro factors in oil futures returns." *International Economics* 126–127: 13–38. http://dx.doi.org/10.1016/S2110-7017(13)60034-7

Li, R., R. Joyeux and R.D. Ripple (2014). "International natural gas market integration." *The Energy Journal* 35: 159–179. http://dx.doi.org/10.5547/01956574.35.4.7

Liu, L., Y. Wang, C. Wu and W. Wu (2016). "Disentangling the determinants of real oil prices." *Energy Economics*, forthcoming. http://dx.doi.org/10.1016/j.eneco.2016.04.003

Loretan, M. and W. English (2000). "Evaluation "correlation breakdowns" during periods of market volatility." In: Bank of International Settlements (BIS), International Financial Markets and the Implications for Monetary and Financial Stability. BIS, Switzerland. https://www.bis.org/publ/confer08k.pdf

Ludvigson, S.C. and S. Ng (2007). "The empirical risk return relation: A factor analysis approach." *Journal of Financial Economics* 83: 171–222. http://dx.doi.org/10.1016/j.jfineco.2005.12.002

Ludvigson, S.C. and S. Ng (2009). "Macro factors in bond risk premia." *Review of Financial Studies* 22: 5027–5067. http://dx.doi.org/10.1093/rfs/hhp081

Ludvigson, S.C. and S. Ng (2010). "A factor analysis of bond risk premia." In: *Handbook of Empirical Economics and Finance*, A. Uhla and D.E.A. Giles (Eds), pp. 313–372. Chapman and Hall, Boca Raton, FL.

Mu, X. (2007). "Weather, storage, and natural gas price dynamics: Fundamentals and volatility." *Energy Economics* 29: 46–63. http://dx.doi.org/10.1016/j.eneco.2006.04.003

Manera, M., M. Nicolini and I. Vignati (2013). "Financial speculation in energy and agriculture futures markets: A multivariate GARCH approach." *The Energy Journal* 34: 55–82. http://dx.doi.org/10.5547/01956574.34.3.4

Manera, M., M. Nicolini and I. Vignati (2016). "Modelling futures price volatility in energy markets: Is there a role for financial speculation?" *Energy Economics* 53: 220–229. http://dx.doi.org/10.1016/j.eneco.2014.07.001

Panagiotidis, T. and E. Rutledge (2007). "Oil and gas markets in the UK: Evidence from a cointegrating approach." *Energy Economics* 29: 329–347. http://dx.doi.org/10.1016/j.eneco.2006.10.013

Pindyck, R.S. (1994). "Inventories and the short-run dynamics of commodity prices." *Rand Journal of Economics* 25: 141–159. http://www.jstor.org/stable/2555858. http://dx.doi.org/10.2307/2555858.

Ramberg, D.J. and J.E. Parsons (2012). "The weak tie between natural gas and oil prices." *The Energy Journal* 33: 13–35. http://dx.doi.org/10.5547/01956574.33.2.2

de Roon, F.A., T.E. Nijman and C. Veld (2000). "Hedging pressure effects in futures markets." *Journal of Finance* 55: 1437–1456. http://dx.doi.org/10.1111/0022-1082.00253

Rouwenhorst, K.G. and K. Tang (2012). "Commodity investing." Yale ICF Working Paper no. 06-12. http://depot.som.yale.edu/icf/papers/fileuploads/2704/original/Yale_ICF_Working_Paper_No06-12_Geert_final.pdf

Serletis, A. and J. Herbert (1999). "The message in North American energy prices." *Energy Economics* 21: 471–483. http://dx.doi.org/10.1016/S0140-9883(99)00015-8

Serletis, A. and R. Rangel-Ruiz (2004). "Testing for common features in North American energy markets." *Energy Economics* 26: 401–414. http://dx.doi.org/10.1016/j.eneco.2004.04.007

Singleton, K.J. (2014). "Investor flows and the 2008 boom/bust in oil prices." *Management Science* 60: 300–318. http://dx.doi.org/10.1287/mnsc.2013.1756

Smith, J.L. (2009). "World oil: Market or mayhem?" *Journal of Economic Perspectives* 23: 145–164. http://dx.doi.org/10.1257/jep.23.3.145

Stock, J.H. and M.W. Watson (2002a). "Forecasting using principal components from a large number of predictors." *Journal of the American Statistical Association* 97: 1167–1179. http://dx.doi.org/10.1198/016214502388618960

Stock, J.H. and M.W. Watson (2002b). "Macroeconomic forecasting using diffusion indexes." *Journal of Business and Economic Statistics* 20: 147–62. http://dx.doi.org/10.1198/073500102317351921

Stock, J.H. and M.W. Watson (2005). "Implications of dynamic factor models for VAR analysis." NBER Working Paper no. 11467, Cambridge, MA. http://www.nber.org/papers/w11467.pdf

Stock, J.H. and M.W. Watson (2006). "Forecasting with many predictors." In: G. Elliott, C.W.J. Granger, A. Timmermann (Eds.), *Handbook of Economic Forecasting*, vol. 1, Elsevier. http://dx.doi.org/10.1016/S1574-0706(05)01010-4

Tang, K. and W. Xiong (2012). "Index investment and financialization of commodities," *Financial Analysts Journal* 68, 54–74. https://www.cfainstitute.org/learning/products/publications/faj/Pages/faj.v68.n6.5.aspx. http://dx.doi.org/10.2469/faj.v68.n6.5.

Villar, J.A. and F.L. Joutz (2006). "The relationship between crude oil and natural gas prices," *Report of the Energy Information Administration*, Office of Oil and Gas, Washington D.C. http://www.uprm.edu/aceer/pdfs/CrudeOil_NaturalGas.pdf