Futures Trading and the Excess Comovement of Commodity Prices

Yannick Le Pen, Benoît Sévi

Futures Trading and the Excess Comovement of Commodity Prices

Yannick Le Pen
Benoît Sévi

Futures trading and the excess comovement of commodity prices[*]

Yannick Le Pen[†]

Université Paris-Dauphine

Benoît Sévi[‡]

Aix-Marseille Université (Aix-Marseille School of Economics), CNRS & EHESS

First version: November 2009

This version: January 2013

[†]Université Paris-Dauphine, LEDa-CGEMP, Place du Maréchal de Lattre de Tassigny, 75775 Paris Cedex 16, France. Email : yannick.le_pen@dauphine.fr

[‡]Aix-Marseille University, Development and International Finance Department, Château La Farge – Route des Milles 13290 Les Milles Aix-en-Provence, France, Email: benoit.sevi@gmail.com

Abstract

We empirically reinvestigate the issue of excess comovement of commodity prices initially raised in Pindyck and Rotemberg (1990) and show that excess comovement, when it exists, can be related to hedging and speculative pressure in commodity futures markets. Excess comovement appears when commodity prices remain correlated even after adjusting for the impact of common factors. While Pindyck and Rotemberg and following contributions examine this issue using a relevant but arbitrary set of control variables, we use recent developments in large approximate factor models so that a richer information set can be considered and "fundamentals" are likely to be adequately modeled. We consider a set of 8 unrelated commodities along with 187 real and nominal macroeconomic variables from which 9 factors are extracted over the period 1993-2010. Our estimates provide evidence of a time-varying excess comovement which is only occasionally significant, even after controlling for heteroscedasticity. Interestingly, excess comovement is mostly significant in recent years when a large increase in the trading of commodities is observed and also in crisis periods. However, we show that this increase in trading activity alone has no explanatory power for the excess comovement. Conversely, measures of hedging and speculative pressure explain around 60% of the estimated excess comovement thereby showing the strong impact not only of the financialization process, but also the impact of behaviour of some categories of traders on the price of commodities and the fact that supply and demand variables are not the sole factors in determining equilibrium prices.

1   Introduction

The commodity markets have undergone major changes in the last fifteen years. The popularity of commodity-related financial instruments, such as commodity indices, led many observers to consider that the commodity markets are more deeply connected to the financial market. While more participants in the commodity markets may induce a better risk sharing, the financialization process is also criticised for causing a socially undesirable price volatility. Our purpose in this paper is to examine whether changes in commodity futures trading, globally, or for some specific categories of traders, may be related to an excess comovement of commodity prices, a concept that will be defined below.

We gather a large dataset of macroeconomic and financial variables from developed and emerging countries and rely on large approximate factor models to extract the most informative principal components. These factors are expected to represent the main forces driving commodity prices. They are used to filter out the returns of a set of 8 seemingly unrelated commodities and residual correlation is examined to investigate the issue of excess comovement. Our estimates show that commodity returns are mainly correlated with real aggregate variables of emerging countries, which proves the important role played by these countries in shaping commodity prices in recent years. We show that the empirical evidence of an excess comovement between commodity returns for the 1993-2010 period is only occasional whereas the findings in Pindyck and Rotemberg (1990) provide strong evidence for an excess comovement over the 1960-1985 period, without distinguishing, as we do, the time-varying behaviour of the phenomenon. The use of many variables allows us to properly filter the commodity returns thus rendering our results relatively immune to the criticism of omitted – or arbitrary selected – control variables. Interestingly, we show that measures of hedging and speculative pressure are able to explain a very significant part (around 60%) of the newly estimated time-varying excess comovement. As such, our results may be interpreted as evidence of the importance of the direction given to the market by some categories of traders and, more generally, the recent "financialization of the commodity markets", a concept that has gained interest in the academic and the political spheres in recent years, in determining commodity prices.

Commodity prices excess comovement is worth studying for several reasons. First, as noted in the seminal contribution of Pindyck and Rotemberg (1990), remaining correlation (or "comovement") may mean that "[...] commodity demands and supplies are affected by unobserved forecasts of the economic variable." (p. 1174). Second, from a theoretical angle, it can be argued that the standard model of supply and deamand is not sufficient to explain commodity returns. Third, if comovement exists and is strong enough, exporters countries may also find an interest in using commodity indexes as an additional hedging instrument beyond their initial interest in using futures and options

on the commodity they export (see Rolfo (1980), Larson *et al.* (1998) and the recent contribution by Borensztein *et al.* (2009)). Therefore, studying comovement of commodities is also particularly relevant for developing countries whose revenues sometimes heavily depend on one or two commodities (Deaton, 1999).[1]

Additional issues related to the question of comovement of commodities are in the field of finance. Consider a hedger or an investor whose aim is to invest in some commodities with a strategy based on the analysis of supply and demand fundamentals. If excess comovement exists then such a strategy may be unsuccessful. Similarly, from a portfolio management perspective, comovement would reduce diversification and make investment in commodity indexes relatively more interesting than using several futures contracts as investment vehicles. Viewed differently, it would also mean that investing in several commodity futures (see Gorton and Rouwenhorst (2006) and Gorton *et al.* (2013)) would not be as interesting as it would be without comovement.

The main novelty in our paper is that we establish an empirical relationship between the notion of excess comovement and potential explanatory variables. This issue has not been investigated in the literature so far except in Tang and Xiong (2011) but with a very different methodology that we briefly describe below. Our explanation relies on the principle developed in Hong and Yogo (2012) that traders have limited absorption capacity towards too large order flow thus providing a possible impact of trading activity variables, such as open interest, on returns. Therefore, when institutional traders invest large amounts in commodity indices, as hedgers or speculators, they also simultaneously invest in several (or many) commodity futures and these aggregated positions in futures markets may partly explain a possible excess comovement in these markets.

We use three different variables derived from traders' positions in commodity futures markets. The first variable is the total open interest in futures to first gauge the explanatory power of trading activity as a whole on the excess comovement. The next two variables are refined index of trading activity representing either the hedging or the speculative pressure in futures. While the total open interest has very limited explanatory power for the excess comovement, combining hedging and speculative pressure allows to explain about 60% of the variability of the estimated excess comovement. Our investigation of explanatory factors for an excess comovement may be viewed as a test of the price distortion due to the existence of commodity indices and coordinated investment strategies in multiple commodity markets, what Tang and Xiong (2011) coin as "Financialization".

From a methodological point of view, examining the issue of excess comovement is twofold.[2] Indeed,

---

[1]Borensztein *et al.* (2009, Table 2) provide recent estimates of the dependence of many developing countries on a very limited number of commodities for the period 2002-2007 using data from the IMF.

[2]In Deb *et al.* (1996), the first issue which should be considered when dealing with the question of excess comovement question is the concept of "unrelated commodities". In the present paper, we choose similar commodities as in Pindyck and Rotemberg (1990) and Deb *et al.* (1996). These commodities have no substitutability or complementarity and the conditions of their production or extraction are regionally distinct.

we are interested in answering the following question: do commodity prices move together beyond what fundamentals explain? Then our first concern is on how to represent these "fundamentals"? This point raises serious concerns on how to best incorporate relevant information as noted in Pindyck and Rotemberg (1990): "Indeed, a major limitation of our approach is that we can never be sure we have included all relevant macroeconomic variables and latent variables." (p. 1185) (see also Leybourne *et al.* (1994) on this issue). In this paper we gather our own set of 187 real and nominal variables from developed *and* emerging countries and resort to large approximate factor models to sum up the information contained in these data in a manageable number of variables. These factors are expected to best approximate the "fundamentals" driving commodity prices. Using factors avoids the systematic search for relevant exogenous variables when their number is large, despite pre-tests based on univariate regressions are possible in this case (see Ludvigson and Ng (2009)).

Our second methodological concern is to obtain an unbiased measure of the filtered returns correlation. As shown by Forbes and Rigobon (2002), the usual sample correlation is a biased measure of the true correlation when there is a change in volatility. As most of our commodity returns are characterised by a time-varying volatility, we use the correlation coefficient corrected for heteroscedasticity of Forbes and Rigobon (2002). We use a rolling window scheme to estimate the unbiased correlation coefficient as Kallberg and Pasquariello (2008) and then detect change in the intensity of excess comovement if any.

We think that our paper a new perspective to the analysis of commodity returns comovement. First, we use the large approximate factor model methodology to uncover the relevant factors that allow to explain commodity returns. This methodology has only been used so far in Juvenal and Petrella (2011) for modelling commodity returns but the authors do not consider variables from emerging countries allowing their sample to run from 1971. In addition, to our best knowledge, this is the first time that this methodology is used to filter out returns before analysing excess comovement. The main advantage of factors is that they allow to deal with a large number of variables while maintaining econometric tractability thereby including a richer information set of "fundamentals". Hence, we avoid limiting the information set artificially, which has been a major constraint in previous contributions. As a byproduct of our analysis, we uncover factors that best explain the commodity returns and provide an interpretation of these factors based on the idea of Ludvigson and Ng (2009) to group explanatory variables.

As a second contribution, we offer an explanation of the excess comovement in commodity returns. While previous contributions used different methodologies to assess the hypothesis of excess comovement, the issue of which variable can be related to this excess comovement has not been considered so far. Our indicators of trading activity computed using traders' positions available

from the U.S. Commodity Futures Trading Commission (hence CFTC) data performs particularly well in explaining the residual correlation – our measure of excess comovement – and thus help to highlight a possible source of this comovement. Trading activity by both speculators and hedgers appears to be highly correlated with our measure of excess comovement thereby indicating the strong role of financialization and the significant impact of demand for commodities by categories of investors.

The plan for the rest of the paper is as follows. In the next section, we review the literature related to the issue of the excess comovement of commodity prices along with recent research on the financialization of commodity market. Then, in Section 3, we present the data used for the empirical implementation. In section 4, we very briefly review the factor model methodology and compute the factors used to filter commodity returns. Next, excess comovement is evaluated in section 5 while section 6 is dedicated to the analysis of the relation between excess comovement and futures trading through a number of trading activity variables. Finally, Section 7 concludes by providing some limits and possible further extensions of our analysis.

2 Relevant literature

2.1 The concept of excess comovement

The early contribution by Cooper and Lawrence (1975) focused on the dramatic increase in commodity prices during the 1973-74 period. While all commodity prices did not increase exactly at the same time, they all reached their two-year or historical highs during this two-years period. Interestingly, the authors first raise the issue of the comovement of commodity prices as follows: "Interesting tales can be told about many of the individual commodities – the special circumstances that led to the rise in prices and to the subsequent fall. Bad weather reduced harvests of many crops here and there around the world, labor disruptions curtailed mine output, several important materials-producing countries were subject to political unrest, newly rich Arabs were buying disproportionately large amounts, and so on. But the movement in commodity prices was quite general, and while these stories are intriguing and sometimes significant, they do not fill the need for some general explanation – a common cause, or strong linkages among the commodities affected." (p. 672). After exploring the conventional demand-supply factors that could explain such a trend for commodity prices, the authors, as we do, investigate the "speculative" demand for commodities. Nevertheless, a limit to their empirical analysis is the lack of data about the speculative/hedging activity in futures markets. Indeed, while they note that: "A further indicator of the "speculative" behaviour in 1973 and 1974 was the tremendous expansion of trading in futures in a wide range of

commodities.", they also remark that: "It is possible neither empirically nor conceptually to differentiate between pure speculation and hedging by users [...]" (p. 702). Indeed available data from CFTC about the relative positions of different kinds of traders did not exist at this period.

While not the case for many concepts in economics, there is a kind of consensus on the definition of "excess comovement": comovement in excess of common effects of supply and demand determinants such as production indices, inflation, interest rates, *etc*. (Pindyck and Rotemberg, 1990). While simple in nature, this definition first raises concerns on why any comovement should be considered as excessive. Indeed, one might argue that any excessive comovement, as defined in Pindyck and Rotemberg (1990), could be related to economic agents' expectations and, as such, there would be nothing excessive in observing remaining correlation beyond fundamentals.[3] In the present work, we define "excess comovement" as in the original work of Pindyck and Rotemberg (1990), *i.e.* we exclude forward-looking aspects and focus solely on observed variables that could explain commodity returns.

As noted in the introduction, a second concern is on how to select macroeconomic and/or financial variables to represent "fundamentals". Pindyck and Rotemberg (1990) select 6 variables: the US index of industrial production, the consumer price index, the effective US $ exchange rate (the early contribution by Gilbert (1989) emphasizes the relevance of the exchange rates as an explanatory variable for commodity prices), the three-month Treasury bill interest rate, M1 and the S&P stock index. These variables are also used in Deb *et al.* (1996). To deal with the issue of omitted variables, we suggest relying on large approximate factor model which allows to enlarge significantly the set of information while preserving a sufficiently low dimension for the econometric estimation. We thus avoid the arbitrariness and computational difficulties of selecting relevant variables, in particular when the number of possible combinations is large.

Borensztein and Reinhart (1994) point out the necessity to consider well-defined supply and demand variables in order to explain commodity prices. In particular, the authors advocate the inclusion of Eastern Europe, at least for their time span (1970-1992). We consider a set of economic variables from developed and emerging countries (China, India or Brazil among others) and we assume that this will permit to filter out commodity returns more relevantly. Indeed, whilst commodity prices are the product of transactions in one particular place in the world, they are also the outcome of order flows coming from many regions worldwide and from investors arbitraging with respect to other financial places. As such, the price of crude oil, say the U.S. West Texas Intermediate (WTI), is widely accepted as a world price (see Kilian (2009) among others).

---

[3]The criticism about the "excessive" aspect of speculation on oil prices raised in Fattouh *et al.* (2012) also exists for the debate about excess comovement. Indeed, there could be nothing excessive either in speculation or in comovement if expectations from all investors were taken into account. We do not discuss this point and rather try to explain the excess comovement considered as in Pindyck and Rotemberg (1990).

2.2 Excess comovement and contagion

Comovement is a concept which may be confounded with contagion at first sight[4]. However, there is a significant difference between the two concepts. While excess comovement is defined as a remaining significant correlation once common factors are considered, contagion is defined as a significant increase in correlation following a shock in one market. At this point, two remarks are in order. First, most of the literature on contagion does not consider common factors or these factors are very simply defined. This is quite different from the excess comovement literature where "excess" means "beyond common factors" and the determination of common factors heavily condition the estimated comovement. Second, we do not need to observe an increase in correlation to validate excess comovement but rather a significant correlation most of the time or on average.

Nevertheless, one tool developed in the contagion literature proved useful for our purpose, namely the fact that sample correlation is biased upward or downward in a time-varying volatility environment. The argument is that a simultaneous increase (decrease) in the respective volatility of two variables will spuriously increase (decrease) their correlation if measured using the usual sample correlation. Forbes and Rigobon (2002) propose a bias-corrected estimator of correlation to be used when heteroscedasticity is present.[5]. As our residuals will exhibit heteroscedasticity, this unbiased estimator will be applied to evaluate properly comovement.[6] This correction has been applied recently in Kallberg and Pasquariello (2008) to examine excess comovement in sectoral indices in the US and we follow the technical implementation of these authors in our empirical analysis.

2.3 The financialization of commodities

A recent interest in commodities has emerged in economic literature, which draws some conclusions about the usefulness of commodity prices for forecasting financial variables. Part of this literature relies on CFTC data to investigate the role of speculative/hedging activity for various purposes. In the present paper, we also rely on these data to compute different measures of trading activity which are likely to explain residual correlation between commodity returns.

Hong and Yogo (2012) rely on CFTC data to investigate the informativeness of open interest for forecasting commodity returns as well as, and it is far more surprising, bond, currency and stock returns. Pollet (2004) and Driesprong *et al.* (2008) show that change in oil prices is able to predict stock market returns for both developed and emerging countries, a result also discussed in Hong and Yogo (2012) using trading data. More recently, Acharya *et al.* (2010) establish a relationship

---

[4]See Dungey *et al.* (2004) for a survey on contagion.

[5]Similar results have been provided in Boyer *et al.* (1999) or Loretan and English (2000). Recently, Campbell (2008) provided similar analysis for the Student-t distribution.

[6]Cashin *et al.* (1999) use an interesting measure of concordance which is nonparametric, but due to the absence of macroeconomic variables in their analysis, defining excess comovement is challenging.

between the default risk of energy producers and energy futures returns - in their model, the default risk can be viewed as a measure of hedging demand - and empirically show the impact of trading variables on returns. Gospodinov and Ng (2010) provide evidence that the first principal component in a large panel of commodity convenience yields has statistical predictive power for inflation in developed countries. Interestingly, these principal components, which explain commodity prices, are correlated with economic conditions in the U.S. and the fast growing economies, notably India and China (Gospodinov and Ng, 2010). The convenience yield can thus be seen as informational variables about future demand as conveyed by the futures markets.

The paper closest to ours is Tang and Xiong (2011), which investigates the financialization process of the commodity market as a potential source of the dramatic rise and fall of commodity (and in particular crude oil) prices in 2008. With a different methodology from ours, a slightly shorter period and a different set of commodities, Tang and Xiong (2011) also provide evidence of a rising comovement of commodity prices in recent years. Their "analysis focuses on connecting the large inflow of commodity index investment to the large increase of commodity price comovements in recent years by examining the difference in these comovements between indexed and off-index commodities."[7] (p. 3). This greater comovement may be due to the huge increase in commodity index related investments, from a low $ 15 billion in 2003 to a high $ 200 billion in mid-2008 (cf. Tang and Xiong (2011), p. 2). Commodity financial instruments can be used as means to reduce risk in a diversified portfolio. As such, returns may be larger in periods of lower returns in more classical financial markets such as bonds and stocks (see Gorton *et al.*, 2007). In addition, investing in commodities is shown to be an efficient hedge against inflation. However, the research strategy in Tang and Xiong (2011), who regress the S&P-GSCI on a measure of the net position change for different categories of traders, suffers from ignoring common factors that could affect the behaviour of most, if not all, commodity prices. Tang and Xiong (2011) also investigate the relationship between economic activity in emerging countries and the comovement of commodity prices using a novel time series of Chinese futures prices available since late 1990s. While commodity prices are usually thought as a global price, the authors show that the picture is more complex. Interestingly, while U.S. commodity prices exhibit a pronounced cycle, this is not the case for Chinese prices of similar commodities, thereby raising "doubt about commodity demands from China as the driver of all commodity prices in the US." (p. 15). Our regressions for commodity returns show that the demand from emerging economies does play a role in shaping the prices of U.S. non-agricultural commodity futures prices while leaving a large place for other factors.

Juvenal and Petrella (2011) partly reach similar conclusions to ours. They demonstrate the central role of demand in shaping commodity returns and the importance of speculation which is associ-

---

[7]Their research question builds on Barberis, Shleifer and Wurgler (2005) which analyzes theoretically and empirically the behaviour of newly-included stocks in a stock index. It is shown that the price comovement between the stock and the index significantly increases after the inclusion.

ated with the comovement between oil prices and other commodity prices.[8] While we use a different methodology – the authors rely on a Structural Vector Autoregressive (SVAR) model – we also find a strong role of demand in explaining commodity returns with an emphasis on demand coming from emerging countries. And once commodity returns are filtered out using common factors as fundamentals, we also deliver evidence of an association between residual correlation (excess comovement) and speculative intensity.

3   Data

Our data are a set of 8 commodity prices and another one of 187 macroecononomic variables which could influence these prices. These 8 commodities are: wheat, copper, silver, soyabeans, raw sugar, cotton, crude oil and pork bellies. They are representatives of the main classes of commodities and assumed to be unrelated as defined in Pindyck and Rotemberg (1990) [9]. All prices are cash prices except for crude oil where the current front-month contract price is taken as a proxy for the cash price. They are nominal prices in US$. Due to data limitation, in particular for macroeconomic variables from emerging countries, we consider monthly observations from 1993:03 to 2010:03. All data are extracted from DataStream.

Some papers (Palaskas and Varangis (1991), Leybourne *et al.* (1994)) consider excess comovement of nominal or real prices rather than returns and rely on a co-integration analysis. We think that returns are more appealing when dealing with risk management issues and thus consider returns excess comovement as do Pindyck and Rotemberg (1990). Returns rather than prices have also been considered recently in Ai *et al.* (2006) for main agricultural commodities (see also Malliaris and Urritia (1996)) where it is shown that storage levels can significantly explain excess comovement.

Prices and returns[10] are respectively displayed in Figures 1 and 2. The price pattern is rather similar for each commodity, except for cotton and pork bellies: a first increase in 1996 is followed by a larger one in 2008. This last increase has raised much concern on the operating of commodity market. Returns also tend to be more volatile at the end of our sample. Descriptive statistics in Table 1 show evidence of skewness and excess kurtosis and, accordingly, the Jarque-Bera test rejects the hypothesis of a Gaussian distribution for most returns. Heteroscedasticity in the data may explain this non-normality. Table 2 show sample correlations between returns and their associated p-values. There are respectively 17, 16 and 11 significant correlations at the 10%, 5% and 1

---

[8]The conclusions reached by Juvenal and Petrella (2011) are much debated by experts in the field of energy economics. In particular, Kilian rejects their results on the ground of insufficient identification hypothesis (see for instance Kilian and Murphy (2012a)). As their main final result is about the impact of financialization of commodities in the 2003-2008 and that this result is of limited interest in our setting, we take as given their first two conclusions to discuss our own results and let aside this third conclusion.

[9]These prices are unrelated in the sense that their supply or demand cross-elasticities are almost equal to zero.

[10]Return is computed as the log difference of price assuming continuous compounding.

% critical levels. These significant correlations range from 0.4438 (wheat and soya) to 0.1268 (pork bellies and soyabeans). Their average value is 0.239.[11] Interestingly, crude oil is not correlated with wheat, soya and raw sugar. The interactions between oil and agricultural commodities with the development of ethanol is not apparent in our data probably because the phenomenon is quite recent. Our aim is to analyze whether these correlations derive from a common set of variables. If some residual correlations remain significant, we will conclude in favor of an excess comovement.

To find variables explaining these commodity returns, we gather 187 real and nominal macroeconomic variables from developed and emerging countries. The composition of this data set with a short description is given in the Appendix. Our database differs from Stock and Watson (2002b) and Ludvigson and Ng (2007, 2009) datasets, which are only representative of the U.S. economy and therefore not well-suited for our purpose. Our dataset contains variables from developed (128 variables) as well as emerging countries (59 variables from China, Brazil, Korea, India, Indonesia, etc.) that have become significant players in the world economy. More specifically, these countries are often large importers of commodities and their rapid growth may have a significant impact on the world price of these commodities. Data can be separated into 103 real variables (73 for developed countries and 30 for emerging countries) and 84 nominal variables (55 for developed countries and 29 for emerging countries). Each variable is stationnarised in a proper way as described in the Appendix. Along with these data, we also use the Real Activity Index developed in Kilian (2009), which is "based on dry cargo single voyage ocean freight rates and is explicitly designed to capture shifts in the demand for industrial commodities in global business markets." (p. 1055) following a long tradition of economists who observed the correlation between economic activity and rates for ocean freight.

The inclusion of inventory data is worth a discussion. The intuition is that these kinds of variables would help in better explaining returns. For instance, Pindyck (2001) uses inventory data released weekly by the U.S. Department of Energy to model the convenience yield in the WTI crude oil market. Geman and Nguyen (2005) rely on several sources worldwide to build their own inventory series for soyabeans and use it for modelling the forward curve of this commodity. Geman and Smith (2012) purchase daily data on inventory in the LME warehouses worldwide to test the so-called Working theory of storage for base metals. Baumeister and Kilian (2012) consider a number of oil-specific inventory data for forecasting the monthly price of oil in real-time. We do not include variables on stocks in our empirical analysis. An argument mainly comes from the impossibility to gather reliable inventory data for the eight commodities of interest at a monthly frequency. Indeed, for most of them, only annual or quarterly data are available. In addition, and from a more conceptual viewpoint, the inclusion of inventory data in our empirical analysis raises the

---

[11] Pindyck and Rotemberg (1990) obtain a maximum of 0.322 and a minimum of 0.113 and an average value of 0.161 for significant correlations for the time period 04:1960-11:1985. With great caution, we could infer that correlation between commodities return has increased through time.

question of what should be used to filter returns. In examining the issue of the excess comovement, we are interested in filtering returns using fundamentals that are, at least partly, common to all commodities. In so doing, data related to the demand for commodity are relevant as they represent common fundamentals. We make the assumption that data such as inventory are specific to each commodity and therefore less able to explain commodity returns correlation.

4   Filtering commodities returns using large approximate factors model

In this Section, we first briefly review the approximate factors methodology. Recent techniques to determine the optimal number of factors are presented in appendix A. Additional developments can be found in the Bai and Ng (2008) survey on large approximate factors models. The rest of the Section is dedicated to the filtering of commodity returns using the estimated factors.

4.1   Static factors computation

We use the static factor model of Stock and Watson (2002a). The dynamic version of Forni *et al.* (2005) is not considered in our work because recent work (Boivin and Ng (2005) and D'Agostino and Giannone (2012)) show that the dynamic and the static factor models have equivalent performance especially when the dynamics of factors is unknown. In addition the dynamic factor model is best suited for forecasting purpose, which is not our aim in this paper. Finally, the static factor model is easier to estimate and is not likely to question the robustness of our results.

We dispose of a sample $\{x_{it}\}$ of $i = 1, ..., N$ cross-section units and $t = 1, ..., T$ times series observations. Each x_{it} is split into a component depending on a set of $r << N$ common factors $F_t = (f_{1t}, f_{2t}, ..., f_{rt})'$ and an idiosyncratic e_{it} part:

$$x_{it} = \lambda_i' F_t + e_{it}$$

where λ_i is the $(r \times 1)$ factor loading.

Let $X_t = (x_{1t}, ..., x_{Nt})'$, $e_t = (e_{1t}, ..., e_{Nt})'$ be the $(N \times 1)$ vectors of observations and idiosyncratic components at date t and $\Lambda = (\lambda_1, ..., \lambda_N)'$ the $(N \times r)$ matrix of factor loadings, we have the vector form notation:

$$X_t = \Lambda F_t + e_t$$

If we assume that F_t and e_t are uncorrelated and have zero mean and make the normalisation

$E(F_t F_t') = I_d$, we have:

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

where Σ and Ω respectively denote the population covariance matrices of X_t and e_t. Let $X = (X_1, X_2, ..., X_T)'$ and $e = (e_1, e_2, ..., e_T)'$ be respectively the $(T \times N)$ matrices of observations and idiosyncratic components and $F = (F_1, F_2, ..., F_T)'$ is the $(T \times r)$ matrix of factors, a representation of the model for all dates is:

$$X = F\Lambda' + e$$

As the factors F and the loading matrix Λ are not separately identifiable (see Bai and Ng (2008) for more details), constraints are imposed to obtain a unique estimate.

In classical factor analysis, F_t and e_t are assumed to be serially and cross-sectionally uncorrelated and the number of units of observation N is fixed. Stock and Watson's (2002a,b) "large dimensional approximate factor models" differs from the classical model in two ways: the idiosyncratic errors are allowed to be "weakly correlated" across i and t [12] and the sample size tends to infinity in both directions.

We assume k factors and use the principal components method to estimate the $(T \times k)$ factors matrix F^k and the corresponding $(N \times T)$ matrix Λ^k loadings. The estimates solve the following optimization problem:

$$min\, S(k) = (NT)^{-1} \sum_{i=1}^{N} \sum_{t=1}^{T} (x_{it} - \lambda_i^{k'} F_t^k)^2$$

subject to the normalization $\Lambda^{k'}\Lambda^k/N = I_k$.

This classical principal component problem is solved by setting $\hat{\Lambda}^k$ equal to the eigenvectors of the largest k eigenvalues of $X'X$. The principal components estimator of F^k is:

$$\hat{F}^k = X'\hat{\Lambda}^k/N$$

Consistency of the principal component estimator as $N, T \to \infty$ has been demonstrated by Stock and Watson (2002a) and Bai and Ng (2002). Bai (2003) shows that the factors and loadings estimates have asymptotic normal distributions. Computation of $\hat{F}^k$ requires the eigenvectors of the $(N \times N)$ matrix $X'X$. When $N > T$, a computationally simpler approach uses the $T \times T$ matrix XX'.

The next step is to determine the optimal number of factors. Methods based upon information criteria as well as the Kapetanios (2010) are described in appendix A. According to these criteria, the optimal number of factors runs from the 2 to 9. We retain the first three and nine factors which

---

[12]Although Forni *et al.* (1999) and Stock and Watson (2002a) use different sets of assumptions to characterise "weak correlations", the main idea is that cross-correlations and serial correlations have an upper bound.

explain respectively almost 20 % and 36% of the variance of the 187 macroeconomic variables.

4.2 Modelling commodity returns

Our next step consists in modelling commodity returns with the first nine estimated factors and Kilian's activity index. We accept excess comovement if commodity returns remain correlated even after verifying the contribution of selected factors and the index.

We consider several specifications.

The first specification is the linear regression of returns on the first three factors:

$$
\begin{aligned}
r_{it} &= \alpha_i + \sum_{k=1}^{3} \beta_{ik}\widehat{F}_{k,t} + u_{it} \qquad i = 1, ..., 8 \quad t = 1, ..., T \\
&= \alpha_i + \beta_i' \hat{F}_t + u_{it}
\end{aligned}
$$

where r_{it} represents the ith commodity return at date t, α_i is a constant, β_i is the vector of factor coefficients for the ith commodity and $\hat{F}_t = (\widehat{F}_{1,t}, \widehat{F}_{2,t}, \widehat{F}_{3,t})'$ the vector of then first three selected factors at date t. Results[13] are reported in Table 4. The R^2 varies from 1.24% for pork bellies to 23.47% for copper. Factors $\widehat{F}_1$ and $\widehat{F}_2$ are significant in most regressions except raw sugar and pork bellies. We obtain a higher R^2, except for raw sugar and pork bellies, than Pindyck and Rotemberg (1990). This improvement can partly be attributed to using factors computed from a large dataset. Our results for agricultural commodities returns do not substantially differ from those of Pindyck and Rotemberg (1990).[14] The ARCH-LM test provides evidence of a time-varying volatility for 5 residuals.

In a second strategy, as in Stock and Watson (2002) and Ludvigson and Ng (2009), we consider all possible combinations of the nine estimated factors and select the regression which minimises the BIC criterion for each commodity. Once each set of regressors is selected, we jointly estimate the set of 8 regressions with a SUR estimator. By doing so, we aim to find the best model from a set of common regressors for each commodity returns. This approach is intended to eliminate

---

[13]The set of equations are a SUR estimator.

[14]We then consider possible nonlinearities by assuming that factors can enter the regression in their quadratic or cubic form. We choose the specification which gives us the higher sum of $\bar{R}^2$. The set of factors is now $\bar{F}_t^{nl} = ((\widehat{F}_{1,t}, \widehat{F}_{2,t}, \widehat{F}_{3,t}, \widehat{F}_{4,t}, \widehat{F}_{2,t}^3, \widehat{F}_{4,t}^3)')$ and our set of regressions becomes :

$$
\begin{aligned}
r_{it} &= \omega_i + \sum_{k=1}^{4} \gamma_{ik}\widehat{F}_{k,t} + \omega_{i,5}\widehat{F}_{2,t}^3 + \omega_{i,6}\widehat{F}_{4,t}^3 + u_{it} \qquad i = 1, ..., 8 \quad t = 1, ..., T \\
&= \omega_i + \gamma_i' \bar{F}_t^{nl} + v_{it}
\end{aligned}
$$

Results of the specification we retained are not reported here but available upon request. We observe that the explanatory power of our factors remains rather low except for crude oil and to a lesser extent copper. Introducing factors in a nonlinear way improves slightly the explanatory power of the regressions. Therefore, we pursue with factors that enter only linearly in the equation of returns.

as much as possible residual correlation, hence strengthening our evidence of excess comovement, if any. Estimates are displayed in Table 5. Results are similar to those obtained with the first three factors, even if we can observe some improvement for crude oil, sugar and pork bellies. As previously, the ARCH-LM test rejects the null hypothesis of a constant variance for 5 residuals.

As already noticed, factors $\widehat{F}_1$ and $\widehat{F}_2$ are significant and have the same sign for almost all commodities, except cotton and pork bellies. Even if factors are not identifiable as mentioned in section 3.1, Ludvigson and Ng (2009) suggest a simple method to give them an economic interpretation. Each of the 187 original variables is regressed on a factor to measure the correlation between them. After sorting the variables along the horizontal axis (say, beginning with real variables and then with nominal variables), it is graphically possible to show the variables for which the highest R^2 are obtained. The factor can then be considered as representative of this set of variables. We separate our 187 series into developed countries/emerging countries and within each of the previous categories between real and nominal variables. A finer classification would be difficult to illustrate and is relevant, in our opinion, only when a single country is at play.[15]

Figures 3 display the R^2 for factors $\widehat{F}_1$ and $\widehat{F}_2$. $\widehat{F}_1$ can be interpreted as a real factor as it records its highest explanatory power for real variables. More precisely $\widehat{F}_1$ is mostly correlated with real variables from emerging countries. The correlation of $\widehat{F}_1$ with crude oil and copper returns can be interpreted as evidence of the growing role played by emerging countries in shaping these commodities prices. China is for instance known as an important importer of copper and heating oil. This finding corroborates the rather weak support of previous studies of the hypothesis that oil prices (e.g. Hamilton (2009), Kilian and Murphy (2012b) and the survey in Fattouh *et al.* (2012)) or agricultural prices (Hamilton and Wu, 2012) are mainly driven by speculative activity rather than by real supply and demand variables.[16] Our results are also in line with Kilian and Hicks (2012), among others, which emphasises the role of demand, including the demand from emerging countries. To confirm our conclusions, $\widehat{F}_1$ is not significant for sugar, silver or pork bellies, that is commodities for which the demand from emerging countries creates less tensions.

The interpretation of $\widehat{F}_2$ is less obvious. $\widehat{F}_2$ is highly correlated with a small number of real variables but its explanatory power for interest rates, producer and consumer price indices and monetary aggregates of developed as well as emerging countries is higher than for $\widehat{F}_1$. This leads us to interpret $\widehat{F}_2$ as representative of these latter variables. The relation between interest rates and commodity prices is discussed in a number of recent contributions (Barsky and Kilian (2001), Frankel (2006), Frankel and Rose (2010)) but its empirical support is rather weak. As argued in Frankel (2006): "[...] a negative effect of interest rates on the desire to carry commodity inventories." should ex-

---

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

[16] Kilian and Lewis (2011) further highlight the endogeneity of the real price of oil which has strong implications for the design of relevant monetary policy.

ist and while this theoretical relationship has been debated in literature, our estimates provide evidence of a relation between $\widehat{F}_2$ and commodity prices. The price indices and the monetary aggregates could also account for the effect of inflation on commodity prices.

Finally, as shown in Table 5, adding Kilian's activity index does not bring more information as it is not significant – except for copper at the 10% threshold – indicating that $\widehat{F}_1$ does a better job in modelling commodity returns. This is an interesting conclusion as the real activity index used in Kilian (2009) is known as a reference in the literature to proxy economic activity. In our view, this confirms the power of statistical factors in aggregating information from a large number of variables. We report in Table 6 the estimates from univariate regressions of the nine empirical factors on Kilian's index. As can be observed, the correlation between this index and our factors is rather weak.

5 Testing for commodity returns excess comovement

5.1 Conditional comovement

Residuals from previous regressions represent commodity returns once "fundamentals" have been adjusted for, and because we considered fundamentals through factors, we assume that they are taken into account in the most relevant way. We first evaluate residuals correlation as in Pindyck and Rotemberg (1990). Sample correlations (in the upper triangular matrix) with their p-values [17] (in the lower triangular matrix) are respectively reported in Tables 7 and 8 for residuals from the 3 factors and BIC linear filtration.

Results are quite similar for both regressions and confirm the excess comovement hypothesis. We find 7, 9 and 14 significant correlations at 1%, 5% and 10% significance levels for the three factor regressions. For the BIC selected regressions, 6, 13 and 14 correlations are respectively significant at 1 %, 5 %, and 10 %. The Breusch-Pagan LM test rejects the null hypothesis of no residual correlation in both cases. These correlations range from 0.4293 (wheat and soyabeans) to 0.1189 (silver and raw sugar) in Table 7 and are quite the same in Table 8. The level of residual correlation remains therefore quite substantial. Compared to raw returns correlations displayed in Table 2, only two correlations (silver and crude oil, cotton and crude oil) become insignificant after adjusting returns for common factors.

To summarize, even if factors have some explanatory power, at least for non-agricultural commodities, they do not succeed in filtering returns from some unidentified common determinants. For

---

[17]The p-value is computed by transforming the correlation $\hat{\rho}$ to create a t statistic having $T-2$ degrees of freedom, where T is the number of observations.

instance, we observe that even if raw sugar and pork bellies returns are not explained by the same factors as other commodities, their residuals still exhibit some correlation. In the following part of this paper, our aim is to show that residual correlation can be explained by variables not related to fundamentals, and in particular by variables related to trading activity in the commodity financial market.

5.2 Correcting for heteroscedasticity

In a second step, we proceed as in Kallberg and Pasquariello (2008) to deal with residual time-varying volatility.[18] The main idea is to correct sample correlation for the bias induced by change in volatility using the Forbes and Rigobon (2002) estimator.[19] Applied on a rolling window scheme, this estimate is able to track the true correlation and its variation through time.[20] The estimate is non-parametric and obviates the mean-reversion problem inherent in the parametric approach, such as dynamic conditional correlation model. As noted in Kallberg and Pasquariello (2008), financial literature offered a number of examples where rolling filters do perform well in comparison with more sophisticated parametric specifications.

We consider residual $\hat{u}_{i,t}$ from each commodity return equation to compute, for all pairs of non redundant returns $i \neq j$, the excess comovement measured by residual correlation:

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

Boyer *et al.* (1999), Loretan and English (2000) and Forbes and Rigobon (2002) show that this sample correlation $\hat{\rho}_{ij,t}$ is biased in case of change in volatility. $\hat{\rho}_{ij,t}$ is therefore named *conditional* correlation. Hence, in the presence of heteroscedasticity, it is not an adequate measure of excess comovement. The aforementioned authors propose a correction[21] for this bias and define an *unconditional* correlation as:

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

where the ratio $\hat{\delta}_{i,t} = \frac{var(\hat{u}_{i,t})}{var(\hat{u}_{i,t})_{LT}} - 1$ corrects the conditional correlation $\hat{\rho}_{ij,t}$ for the relative difference between the i^{ith} return short-term $var(\hat{u}_{i,t})$ and the long-term volatility $var(\hat{u}_{i,t})_{LT}$. We employ this unconditional correlation as a measure of excess comovement. As we don't 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 asset *i* (in $\hat{\rho}^*_{ij,t}$) or asset *j* (in $\hat{\rho}^*_{ji,t}$). Therefore,

---

[18]This is an issue in Pindyck and Rotemberg (1990) contribution which has been considered further in Deb *et al.* (1996) by means of the multivariate GARCH model in its BEKK form (Engle and Kroner, 1995).

[19]Dufour and Khalaf (2002) propose a test for contemporaneous correlation in SUR regression. However their test doesn't accommodate time-varying volatility.

[20]Tang and Xiong (2011) also correct correlation for time-varying volatility using Forbes and Rigobon's (2002) method but only with a limited (insignificant) impact on their estimate.

[21]This correction is valid under the assumption of no omitted variables or endogeneity.

the two unconditional correlations $\hat{\rho}^*_{ij,t}$ and $\hat{\rho}^*_{ji,t}$, which are possibly different, are available.

As suggested in King *et al.* (1994) and Kallberg and Pasquariello (2008), we compute the arithmetic mean[22] of pairwise squared adjusted correlations for each commodity i. A non-null unconditional correlation $\hat{\rho}^*_{ij,t} \neq 0$ and $\hat{\rho}^*_{ji,t} \neq 0$ whatever its sign is taken as evidence of excess comovement between commodities i and j. We use the mean of excess square correlations as a measure of excess comovement:

$$\hat{\rho}^*_{i,t} = \frac{1}{K-1} \sum_{j=1,j\neq i}^{K} (\hat{\rho}^*_{ij,t})^2$$

for all commodity returns $i = 1, ..., K$ where $K = 8$ is the number of commodities. We finally compute a total measure of excess comovement as the mean of excess square correlation for all commodities:

$$\hat{\rho}^*_t = \frac{1}{K} \sum_{i=1}^{K} \hat{\rho}^*_{i,t}$$

In this paper, we treat the covariance matrix of returns residuals as observable and construct a time series of rolling realised excess square correlation for each commodity i. We estimate $\hat{\delta}_{i,t}$ and $\hat{\rho}^*_{i,t}$ over short-term and long-term intervals of fixed length N $[t - N + 1, t]$ and gN (with $g > 1$) $[t - gN + 1, t]$. We use a rolling window of $N = 30$ observations for short-term volatility and $gN = 60$ observations for long-term volatility.

5.3 Empirical results

We compute three averages of squared correlations. The first is the average value of squared unconditional returns correlation: $\hat{\rho}^*_{ret,t} = \frac{1}{K} \sum_{i=1}^{K} \hat{\rho}^*_{ret\,i,t}$ where correlations are computed for non-adjusted returns. The second is the average value of squared conditional residual correlation: $\hat{\rho}_t = \frac{1}{K} \sum_{i=1}^{K} \hat{\rho}_{i,t}$ where $\hat{\rho}_{i,t} = \frac{1}{K-1} \sum_{j=1,j\neq i}^{K} (\hat{\rho}_{ij,t})^2$. In this case, we use residual correlations not corrected for change in volatility. The last indicator is the average squared unconditional correlation $\hat{\rho}^*_t$ as previously defined. Two main conclusions can be made from the descriptive statistics on average return and residual squared correlations displayed in Table 9. First, filtering commodity returns does not reduce correlations much as return and residual squared correlations are highly correlated and their averages almost equal. Second, squared correlations are above the significance level in almost two-thirds of the full time period which show that excess comovement is not continuous. Their averages are above the 5 % significance level.

In the same vein as the correlation plot in Tang and Xiong (2011), the chart of average squared correlations in Figure 4 gives us a finer description of excess comovement. A first result is about

---

[22]Peña and Rodríguez (2003) propose to use $1- \mid \Gamma \mid^{1/N}$ where Γ is the correlation matrix of N variables as a measure of multivariate linear dependence. However, as we have two heteroscedasticity corrected correlation, we did not apply their criterion.

the timing of excess comovement. Returns as residual correlations were above the 5 % significance level during the two periods of financial crisis in our sample: from 2000 to 2004 and from 2008 onwards. During these two periods, commodity markets were not isolated from financial markets anymore. This feature contrasts sharply with the negligible correlation between commodity returns and the S&P 500 before the 2000s as shown by Gorton and Rouwenhorst (2006). We note that the more severe the financial crisis, the higher residual correlations. As reminded by Tang and Xiong (2011), the finding of a negative correlation between commodities returns and stock returns (see Gorton and Rouwenhorst (2006) among others) has given an incentive to include commodities into portfolio assets classes to diversify risk. Hence, a negative shock on the stock market could trigger a rebalancing strategy from stock to commodity inducing a commodity returns excess comovement. This phenomenon is known as a "flight-to-quality" where low returns in classical financial markets lead to increasing the share of alternative assets in portfolios.

A second conclusion is that filtering commodity returns reduces somewhat return correlation, particularly between 2000 and 2004. This is an indication that increases in commodity returns correlation is partly due to common factors. This result can be related to Juvenal and Petrella (2011) who find that the comovement between oil prices and the prices of other commodities is related to global demand shocks. Our empirical results partially confirm their findings in that, once factors related to demand are taken into account, residual correlation is lower. In other words, demand helps in reducing the comovement between commodity prices and leads to a reduced excess comovement. This is also an important result since Pindyck and Rotemberg (1990) as it shows that time-varying excess comovement and not only unconditional comovement has to be considered.

Finally, the unconditional average squared residuals correlation is lower that the conditional one, which stresses the need for correcting the effect of change in volatility. We should therefore be cautious when looking at excess comovement as sample correlation may overestimate excess comovement and is thus biased towards rejecting the hypothesis of no excess comovement.

We now turn to our proposed explanation for the residual correlation between commodity returns.

6 Explaining excess comovement

In this section, we offer an explanation for commodity returns excess comovement. We evaluate the relation between our measure of excess comovement, *i.e.* average unconditional squared residual correlation, and trading activity variables derived from traders' positions publicly available from the CFTC. We consider three different measures that are able to gauge the dependence of excess comovement to various trading activity variables, namely the total open interest in futures markets, the hedging pressure in futures markets using the index suggested in de Roon *et al.* (2000)

and the sentiment index developed in Han (2008) that is more related to the speculative activity in futures markets.

The contribution by Tang and Xiong (2011) is also an attempt to explain the recent increase in comovement found in a number of commodity prices. These authors suggest five hypotheses to explain this comovement: (i) the financialization of commodities, (ii) the rapid growth of emerging economies, (iii) the recent world financial crisis, (iv) inflation and (v) the adoption of biofuels. To link our research question to the arguments developed in Tang and Xiong (2011), we can say that we now test for their first and third hypothesis while having considered the second and the fourth in section 4. Indeed, we showed that growth in emerging economies could be considered as a common factor leading commodity prices and as such as a factor explaining comovement. We also show that commodity returns are correlated with a factor correlated with nominal variables. Then we want to explain the estimated *excess* comovement, *i.e.* what remains once commodity returns have been adjusted for common determinants, and investigate the stability of the relation through time, in particular in periods of turbulence for financial markets.

As previously noted, commodity markets have become more connected with financial markets since the beginning of the 2000s. These tighter links could be beneficial to commodity hedgers as more investors would facilitate and reduce the cost of commodity price risk sharing. However, this positive effect is balanced by the greater sensitivity of the commodity markets to shocks originated on financial markets.

The mechanism behind the potential impact of trading activity on commodity prices is that large institutional investors may go beyond the normal absorption capacity of other market participants (speculators, hedgers) thereby influencing commodity prices when their investment capacity increases. This is the main idea behind Hong and Yogo's (2012) contribution.[23] We investigate the link between the estimated excess comovement and trading activity measures because it could naturally be argued that hedging or speculative pressure may be a significant source of simultaneous trading in comovements in futures markets for reasons beyond simple supply and demand arguments.

---

[23]"If there is excess hedging demand from producers that want to be short futures, the futures price will fall due to limited arbitrage by speculators.Conversely, if there is excess hedging demand from consumers that want to be long futures, the futures price will rise due to limited arbitrage by speculators. Because the futures price can either fall or rise in response to anticipation of higher economic activity, the futures price is a less reliable signal of future economic activity and asset prices than open interest." (Hong and Yogo (2012), p. 474) The authors provide a simple while convincing model of this assertion which is then empirically validated using similar data to ours. Note that they are yet interested in the raw open interest while we use measures of trading activity measures that make sense either for speculative or hedging pressure.

6.1 Measures of trading activity

We use three different measures related to trading activity. Our first hypothesis, that could be viewed as a pre-test, is that residual correlation could be linked to the total volume of futures contracts that could lead to delivery which is called Total Open Interest (*TOI*). In Table 10, we provide results from the linear regression of our measure of residual correlation on the growth rates of *TOI* for each commodity [24] Results show that the *TOI* growth rate only explains 4.6 % of the variability of the estimated excess comovement. These results thus indicate that the global trading activity in commodity futures markets, as measured by the *TOI*, cannot explain residual correlation. This is an important result in light of recent findings in the literature about the predictive power of open interest in the commodity futures markets for several financial variables such as stock, bond, commodity and currency returns (see Klitgaard and Weir (2004) and Hong and Yogo (2012)).

We now consider two measures of trading activity that represent either the speculative or the hedging intensity. Our first variable follows Han's (2008) investor sentiment index computed using CFTC data.[25] Following requirements of the CFTC, large traders holding positions above a specified level have 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). The COT report includes total long and short positions for both 'commercial' traders and 'noncommercial' traders as well as more detailed variables that we do not use here.[26] In other words, 'commercial' traders have to prove an interest for the physical market and are thus considered as hedgers while 'noncommercial' traders have no relation with the cash business and are simply large speculators. The sentiment proxy in Han (2008) is then the number of long noncommercial contracts minus the number of short noncommercial contracts, scaled by the total open interest in futures markets, or:

$$\text{Han Index} = \frac{\text{number of long speculative positions} - \text{number of short speculative positions}}{\text{total open interest}}$$

The *Han Index* is inspired by the literature on Investor Sentiment[27] and allows to estimate the sentiment of speculators in the futures market of interest by considering their relative long and short positions. As such, this is a directional index of speculative activity in the futures market.

---

[24] We previously checked that Total Open Interests (in logarithm) are non stationary and that our measure of residual correlation is stationary with a structural break.

[25] Han (2008) computes the net position of large speculators in S&P 500 futures contract. The author also considers another investment sentiment proxy based in Investors Intelligence's weekly survey that we do not use in the present study.

[26] Since 2006, the CFTC also releases weekly Commodity Index Traders (CIT) report on each Friday. This complements the COT report by providing more detailed categories of traders such as Index Traders who played a significant role these last years. We also do not consider the CIT as it would considerably restrain the sample period for the analysis. The CIT is used in Tang and Xiong (2011) to build their variable of trading activity which is the variation in the value of the net long position by index traders.

[27] see Baker and Wurgler (2007) for a presentation of recent contributions in this field.

We use this index for 7 out of our 8 commodities because for pork bellies data are not available for the period of interest. In an investigation of changes in exchange rates markets, Klitgaard and Weir (2004) also rely on a speculative index computed from CFTC data. Note that their measure is not scaled.

Our second measure of trading activity follows de Roon *et al.* (2000) who study the hedging pressure in futures markets. In particular, the authors show that futures risk premia depend on both own-market and cross-market hedging pressures. The variable corresponds to the difference between the number of short hedge positions and the number of long hedge positions, divided by the total number of hedge positions or:

$$\text{RNV Index} = \frac{\text{number of short hedge positions} - \text{number of long hedge positions}}{\text{total number of hedge positions}}$$

The idea behind this measure is to focus on the positions of traders who are hedgers, *i.e.* who have a cash business for the commodity. This estimate of hedging pressure is quite different to the Han Index for which the denominator is the total open interest and not the total number of speculative positions. As a consequence, we believe that these measures may be complementary in our regression analysis while matching existing literature dealing with futures market trading activity.

6.2 Empirical findings using hedging and speculative pressure

In Table 11, we report the results from the OLS estimation of the regression of the average squared unconditional correlation - our measure of commodity return excess comovement - on the *HAN* and *RNV* indices. Regressions with *Han* and *RNV* indices are respectively reported in columns *(1)* and *(2)*. Results are quite similar for both sets of indices. For each regression, about 4 indices are significant and the explanatory power is about 30%. Column *(3)* reports estimates when both indices are considered and column *(4)* estimates for significant variables. The explanatory power of the regression with both kinds of indices is quite satisfactory as it is almost equal to 60% with ten significant indices. This high explanatory power demonstrates the complementarity between the two indices and the important contribution of variables related to trading activity in explaining excess comovement in the commodity markets.

Our results validate the assumption that trading and traders' choices in the commodity financial market do indeed constitute a significant part of the explanation for the existence of an excess comovement in commodity prices. As noted in Pindyck and Rotemberg's (1990) seminal paper, "[...] traders are alternatively bullish or bearish on *all* commodities for no plausible reason"[28] (p.

---

[28]Italic original from the authors.

1173). The empirical evidence thus highlights the herding hypothesis suggested in Pindyck and Rotemberg (1990) as a possible channel for the existence of an excess comovement. Alternatively, as mentioned in the introduction, rebalancing strategy from non-commodity financial markets to commodity markets, through investment in commodity indices, would also be a possible explanation to illustrate the relation between commodity returns excess comovement and trading activity.

Regressions using either both measures of trading activity or each measure individually can explain a much larger part of the total variability than in Tang and Xiong's (2011) best regression where only 8% of the total variance is explained using control variables. It thus appears that trading activity variables have much more to say than economic activity variables at least when the variable of interest is the correlation between commodity returns.

Our findings can also be related to the conclusion in Juvenal and Petrella (2011) that the change in speculative activity (speculative shock) can be associated with a positive comovement between oil price and the prices of other commodities. Our findings are in the same vein, as we demonstrate an association between speculative activity and excess comovement for our eight commodities.

6.3   Stability of the relation

We now investigate the stability of the relation between excess comovement and our trading activity measures. As the time pattern of excess comovement exhibits notable changes and as the latter years of our sample are characterized by a severe financial crisis, we assume this relation could undergo several structural changes. We apply the standard tests in Bai and Perron (1998, 2003a,b) allowing for multiple structural breaks to the regression reported in column *(4)*. Tests for structural breaks are displayed in Table 12. The $SupF_T$ test shows evidence of three structural breaks and the estimated dates of these breaks are $\hat{T}_1 =$01:01:2000, $\hat{T}_2 =$01:05:2000 and $\hat{T}_3 =$01:02:2007 thus defining four sub-periods. These dates are reported in Figure 4. Interestingly, structural breaks coincide with important events in financial markets as $\hat{T}_1$ and $\hat{T}_3$ correspond to the beginning of the financial crisis. Average squared residual correlation is under its significance level in the first and the third periods where financial stability is observed while it is above for the second and the fourth periods of financial turmoil. Commodity return excess comovement is therefore intrinsically a feature of financial crisis at least in the last few years.

Estimates for the four different periods are reported in Table 13. They clearly show changes in the value and the significance of many coefficients between each period. The number of significant coefficients increases through time (6 for the first period, 4 for the second period, and 8 for the third and fourth periods at the 5% threshold) and their R^2 and $\bar{R}^2$ are around 60% or 40% for the first two periods but increase for the last ones. This increase is particularly important for the last period

for which we obtain a $R^2 = 91.85\%$ and a $\bar{R}^2 = 88.71\%$. It seems therefore that the role played by investors in commodity markets has significantly increased in recent years and can be linked to the rising correlation of commodity returns, specially during the last financial crisis.

7 Concluding remarks

The aim of this paper is to reconsider the question of excess comovement of commodities and to provide an explanation for this comovement, if present in the data.

Our contribution to the literature on excess comovement on commodity markets is twofold. First, we enlarge the data set of macroeconomic and financial variables compared with previous contributions thus allowing to conclude that "fundamentals" should be fully taken into account. We provide empirical evidence of a time-varying excess comovement which only is infrequently significant, at least with respect to the set of information to be considered. Second, we provide an explanation for this excess comovement based on the respective positions of hedgers and/or speculators in commodity futures markets. We obtain highly significant estimates for the two measures of trading activity that we consider thereby validating the hypothesis of the impact of market variables on the comovement of returns.

The limits of our analysis are also good topics for future research. First, we consider, as in most factor-models literature, factors as if they were being observed while they are estimated in practice. Even if this should only have a limited impact on our results, 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 (2010). Second, our analysis may be carried out using dynamic factor models (DFM) following Forni *et al.* (2005) as, for instance, in Vansteenkiste (2009). Nevertheless, the bulk of the literature has concluded with a weak improvement in using DFM and we have some doubts that for our purpose it would add much to the present analysis. In particular, d'Agostino and Giannone (2012) emphasise the limited impact of DFM for forecasting purpose with macroeconomic data.[29] Third, MIDAS regressions may be used to include more information with different frequencies. Tang and Xiong (2011) use daily and monthly regressions and MIDAS may help to reconcile both approaches, with market indices on a daily basis and macroeconomic variables either monthly or quarterly. Fourth, the issue of volatility spillover, as investigated in the penultimate section in Tang and Xiong (2011) may also be considered as a complement to our analysis on returns. Indeed, the analysis of commodity volatility comovement may have interesting implications for the purpose of financial risk management. Fifth, alternative measures of trading activity such as liquidity measures may also help to better explain the excess comovement. In this respect, the

---

[29] See also Boivin and Ng (2005) and the discussion in section 4.

recent contributions by Marshall *et al.* (2012, 2013) may be helpful in selecting appropriate liquidity measures for commodities and investigating the explanatory power of their common liquidity factor.

The "comovement in commodity prices" issue is now more than twenty years old and the debate on how commodity prices evolve according to fundamentals has never been more intense. Numerous contributions have investigated the impact of the 2008 financial crisis, the recent increase index trading, diversification incentives, etc., on commodity prices but without considering the fact that an increasing correlation between commodities, and beyond fundamentals, may simply invalidate the supply/demand model. hence by showing that the recent financialization of commodity markets plays a important role in shaping the price of several seemingly unrelated commodities, we emphasise the need for policy makers to implement stabilising mechanisms that could limit the impact of trading on food and energy commodity prices.

References

ACHARYA, V.V., LOCHSTOER, L.A., RAMADORAI, T., 2010. Limits to arbitrage and hedging: Evidence from commodity markets. Unpublished working paper, New York University.

D'AGOSTINO, A., GIANNONE, D., 2012. Comparing alternative predictors based on large-panel factor models. *Oxford Bulletin of Economics and Statistics* 74, 306-326.

AI, C., CHATRATH, A., SONG, F., 2006. On the comovement of commodity prices. *American Journal of Agricultural Economics* 88, 574-588.

ALQUIST, R., KILIAN, L., 2010. What do we learn from the price of crude oil futures? *Journal of Applied Econometrics* 25, 539-573.

ALQUIST, R., KILIAN, L., VIGFUSSON, R.J., 2012. Forecasting the price of oil. In: G. Elliott and A. Timmermann (Eds), *Handbook of Economic Forecasting 2*. Amsterdam: North-Holland, forthcoming.

ANDERSON, R.W., DANTHINE, J.-P., 1983. hedger diversity in futures markets. *Economic Journal* 93, 370-389.

BAI, J., 2003. Inferential theory for factor models of large dimensions. *Econometrica* 71, 135-172.

BAI, J., NG, S., 2002. Determining the number of factors in approximate factor models. *Econometrica* 70, 191-221.

BAI, J., NG, S., 2008. Large dimensional factor analysis. *Foundations and Trends in Econometrics* 3, 89-163.

BAI, J., PERRON, P. 1998. Estimating and testing linear models with multiple structural changes. *Econometrica* 66, 47-78.

BAI, J., PERRON, P. 2003a. Computation and analysis of multiple structural change model, *Journal of Applied Econometrics* 18, 1-22.

BAI, J., PERRON, P. 2003b. Critical values for multiple structural change tests. *Econometrics Journal* 6, 72-77.

BAKER, M., WURGLER, J., 2007. Investor sentiment in the stock market. *Journal of Economic Perspectives* 21, 129-151.

BARBERIS, N., SHLEIFER, A., WURGLER, J., 2005. Comovement. *Journal of Financial Economics* 75, 283-317.

BARSKY, R.B., KILIAN, L., 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.

BARSKY, R.B., KILIAN, L., 2004. Oil and the macroeconomy since the 1970s. *Journal of Economic Perspectives* 18, 115-34.

BAUMEISTER, C., KILIAN, L., 2012. Real-time forecasts of the real price of oil. *Journal of Business and Economic Statistics* 30, 326-336.

BOIVIN, J., NG, S., 2005. Understanding and comparing factor based forecasts. *International Journal of Central Banking* 1, 117-152.

BORENSZTEIN, E., JEANNE, O., SANDRI, D., 2009. Macro-hedging for commodity exporters. CEPR Discussion Paper, International Macroeconomics Series, no. 7513.

BORENSZTEIN, E., REINHART, C.M., 1994. The macroeconomic determinants of commodity prices. *IMF Staff Papers* 41, 236-261.

BOYER, B., GIBSON, M., LORETAN, M., 1999. Pitfalls in tests for changes in correlations. Working Paper, vol. 597R. Federal Reserve Board.

BÜYÜKŞAHIN, B., HARRIS, J.H., 2011. Do Speculators Drive Crude Oil Futures Prices? *Energy Journal* 32, 167-202.

CAMPBELL, J. Y., 2008. Estimating the equity premium. *Canadian Journal of Economics* 41, 1-21.

CASHIN, P., MCDERMOTT, C.J., SCOTT, A., 1999. The myth of comoving commodity prices. *IMF Working Paper* 169.

COOPER, R.N., LAWRENCE, R.Z., 1975. The 1972–1975 commodity boom. *Brookings Papers on Economic Activity* 3, 671-715.

DEATON, A., 1999. Commodity prices and growth in Africa. *Journal of Economic Perspectives* 13, 23-40.

DEB, P., TRIVEDI, P.K., VARANGIS, P., 1996. The excess comovement of commodity prices reconsidered. *Journal of Applied Econometrics* 11, 275-91.

DUNGEY, M., FRY, R., GONZÁLEZ-HERMOSILLO, B., MARTIN, V.L., 2005. Empirical modelling of contagion: a review of methodologies. *Quantitative Finance* 5, 9-24.

DRIESPRONG, G., JACOBSEN, B., MAAT, B., 2008. Striking oil: Another puzzle? *Journal of Financial Economics* 89, 307-327.

FATTOUH, B., KILIAN, L., MAHADEVA, L., 2012. The role of speculation in oil markets: what have we learned so far? *CEPR Discussion Paper* no. 8916.

FERRARO, D., ROGOFF, K., ROSSI, B., 2012. Can Oil Prices Forecast Exchange Rates? Unpublished manuscript. Presented at the NBER Meeting on "Economics of Commodity Markets", October 2012, Stanford.

FORBES, K., RIGOBON, R., 2002. No contagion, only interdependence: measuring stock market co-movements. *Journal of Finance* 57, 2223-2261.

FORNI, M., HALLIN, M., LIPPI, M., REICHLIN, L., 2005. The generalized dynamic factor model, one sided estimation and forecasting. *Journal of the American Statistical Association* 100, 830-840.

FRANKEL, J.A., 2006. The effects of monetary policy on real commodity prices. NBER Working Paper no. 12713, Cambridge, MA.

FRANKEL, J.A., ROSE, A.K., 2010. Determinants of Agricultural and Mineral Commodity Prices. *Harvard Kennedy School Working Paper* no. 10-038.

GEMAN, H., NGUYEN, V.-N., 2005. Soybean inventory and forward curve dynamics. *Management Science* 51, 1076-1091.

GEMAN, H., SMITH, W.O., 2012. Theory of storage, inventory and volatility in the LME base metals. *Resources Policy*, forthcoming.

GILBERT, C.L., 1989. The impact of exchange rates and developing country debt on commodity prices. *Economic Journal* 99, 773-84.

GORTON, G., HAYASHI, F., ROUWENHORST, K.G., 2013. The fundamentals of commodity futures returns. *Review of Finance*, forthcoming.

GORTON, G., ROUWENHORST, K.G., 2006. Facts and fantasies about commodity returns. *Financial Analysts Journal* 62, 47-68.

GOSPODINOV, N., NG, S., 2010. Commodity prices, convenience yields, and inflation. *Review of Economics and Statistics*, forthcoming.

HAN, B., 2008. Investor sentiment and option prices. *Review of Financial Studies* 21, 387-414.

HAMILTON, J.D., 2009. Understanding crude oil prices. *Energy Journal* 30, 179-206.

HAMILTON, J. D., WU, J.C., 2012. Effects of index-fund investing on commodity futures prices. Working paper, UCSD.

HONG, H., YOGO, M., 2012. What does futures market interest tell us about the macroeconomy and asset prices? *Journal of Financial Economics* 105, 473-490.

JUVENAL, L., PETRELLA, I., 2011. Speculation in the oil market. Working paper, Federal Reserve Bank of St. Louis.

KALLBERG, J., PASQUARIELLO, P., 2008. Time-series and cross-sectional excess comovement in stock indexes. *Journal of Empirical Finance* 15, 481-502.

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.

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.

KILIAN, L., HICKS, B., 2012. Did unexpectedly strong economic growth cause the oil price shock of 2003-2008? *Journal of Forecasting*, forthcoming.

KILIAN, L., LEWIS, L.T., 2011. Does the Fed Respond to Oil Price Shocks? *Economic Journal* 121, 1047-1072.

KILIAN, L., MURPHY, D.P., 2012a. Why agnostic sign restrictions are not enough: understanding the dynamics of oil market VAR models. *Journal of the European Economic Association* 10, 1166-1188.

KILIAN, L., MURPHY, D.P., 2012b. The role of inventories and speculative trading in the global market for crude oil. Unpublished manuscript, July.

KLITGAARD, T., WEIR, L., 2004. Exchange rate changes and net positions of speculators in the futures market. *Economic Policy Review* 10, 17-28.

LARSON, D.F., VARANGIS, P., YABUKI, N., 1998. Commodity risk management and development. World Bank Working Paper Series No. 1963, International Economics Department, Washington DC.

LEYBOURNE, S.J., LLOYD, T.A., REED, G.V., 1994. The excess comovement of commodity prices revisited. *World Economy* 22, 1747-1758.

LORETAN, M., ENGLISH, W., 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.

LUDVIGSON, S.C., NG, S., 2007. The empirical risk return relation: A factor analysis approach. *Journal of Financial Economics* 83, 171-222.

LUDVIGSON, S.C., NG, S., 2009. Macro factors in bond risk premia. *Review of Financial Studies* 22, 5027-5067.

LUDVIGSON, S.C., NG, S., 2010. A factor analysis of bond risk premia. *Handbook of Empirical Economics and Finance*, e.d. by A. Uhla and D.E.A. Giles, pp. 313-372. Chapman and Hall, Boca Raton, FL.

MALLIARIS, A.G., URRITIA, J.L., 1996. Linkages between agricultural commodity futures contracts. *Journal of Futures Markets* 16, 595-610.

MARSHALL, B.R., NGUYEN, N.H., VISALTANACHOTI, N., 2012. Commodity liquidity measurement and transaction costs. *Review of Financial Studies* 25, 599-638.

MARSHALL, B.R., NGUYEN, N.H., VISALTANACHOTI, N., 2013. Liquidity commonality in commodities. *Journal of Banking and Finance* 37, 11-20.

MILLER, G., URDINOLA, P., 2010. Cyclicality, mortality, and the value of time: the case of coffee price fluctuations and child survival in Colombia. *Journal of Political Economy* 118, 113-155.

PALASKAS, T.B., VARANGIS, P.N., 1991. Is there excess co-movement of primary commodity prices? A co-integration test. World Bank Working Paper Series No. 758, International Economics Department, Washington DC.

PINDYCK, R.S., 2001. The dynamics of commodity spot and futures markets: A primer. *Energy Journal* 22, 1-29.

PINDYCK, R.S., ROTEMBERG, J.J., 1990. The excess co-movement of commodity prices. *Economic Journal* 100, 1173-89.

POLLET, J.M., 2004. Predicting asset returns with expected oil price changes. Unpublished working paper, University of Illinois.

RAUSSER, G.C., WALRAVEN, N.A., 1990. Linkages among commodity futures markets and dynamic welfare analysis. *Review of Economics and Statistics* 72, 631-639.

ROLFO, J., 1980. Optimal hedging under price and quantity uncertainty: the case of a cocoa producer. *Journal of Political Economy* 88, 100-116.

DE ROON, F.A., NIJMAN, T.E., VELD, C., 2000. Hedging pressure effects in futures markets. *Journal of Finance* 55, 1437-1456.

STOCK, J.H., WATSON, M.W., 2002a. Forecasting using principal components from a large number of predictors. *Journal of the American Statistical Association* 97, 1167-1179.

STOCK, J.H., WATSON, M.W., 2002b. Macroeconomic forecasting using diffusion indexes. *Journal of Business and Economic Statistics* 20, 147-62.

TANG, K., XIONG, W., 2010. Index investment and financialization of commodities. NBER Working Paper no. 16385, Cambridge, MA.

VANSTEENKISTE, I., 2009. How important are common factors in driving non-fuel commodity prices? A dynamic factor analysis. European Central bank Working Paper no. 1072, July, Frankfurt.

Tables

Table 1
Descriptive statistics for the 8 commodities monthly returns over the period 1993:03-2010:03.

	Wheat	Copper	Silver	Soyabeans	Raw Sugar	Cotton	Crude Oil	Pork Bellies
Mean	0.0014	0.0059	0.0073	0.0025	0.0052	0.0020	0.0069	0.0049
Maximum	0.2056	0.2243	0.2665	0.2069	0.2231	0.2281	0.3320	0.5225
Minimum	-0.2308	-0.4099	-0.2885	-0.4574	-0.2382	-0.2540	-0.4367	-0.5346
Std. Dev.	0.0796	0.0783	0.0776	0.0860	0.0865	0.0850	0.1092	0.1732
Skewness	-0.2086	-0.6844	-0.3053	-1.0893	-0.0825	-0.1109	-0.4060	-0.0690
Kurtosis	3.0794	6.8325	4.5219	7.1735	3.3749	3.4968	4.1207	4.3201
Jarque-Bera	1.5400	141.46***	22.96***	189.31***	1.43	2.52	16.36***	15.04***
Nb of Obs	205	205	205	205	205	205	205	205

Notes: (i) Monthly returns are computed as price log differences. (ii) Commodity prices are cash prices except crude oil where the current month contract price is taken as a proxy for the cash price. (iii) ***, ** and * respectively denotes rejection of the null hypothesis of a Gaussian distribution at 1%, 5 % and 10 % levels.

Table 2
Correlation between the 8 commodities monthly returns over the period 1993:03-2010:03.

	Wheat	Copper	Silver	Soyabeans	Raw Sugar	Cotton	Crude Oil	Pork Bellies
Wheat	1	0.2832 ***	0.1777 **	0.4438 ***	0.0149	0.2510 ***	0.1031	0.1039
Copper	0.0000	1	0.3211 ***	0.2030 ***	0.2042 ***	0.2644 ***	0.3706 ***	0.0487
Silver	0.0108	0.0000	1	0.1116	0.1382 **	0.1019	0.1644 **	0.1463 **
Soyabeans	0.0000	0.0035	0.1110	1	0.0863	0.3978 ***	0.0354	0.1268 *
Raw Sugar	0.8315	0.0033	0.0482	0.2183	1	0.1465 **	0.0653	0.0184
Cotton	0.0003	0.0001	0.1461	0.0000	0.0360	1	0.1857 ***	0.0137
Crude Oil	0.1412	0.0000	0.0185	0.6142	0.3526	0.0077	1	-0.0014
Pork Bellies	0.1380	0.4879	0.0363	0.0701	0.7938	0.8450	0.9836	1

Note: The lower triangular matrix reports correlation while the upper one reports the p-values. ***, ** and * respectively denotes significance at 1%, 5 % and 10 % levels.

Table 3
Summary statistics for estimated static factors $\widehat{F}_{t,i}$ for $i = 1, ..., 9$.

factor i	ρ_1	ρ_2	ρ_3	R_i^2
1	0.1614	0.1256	0.3176	0.0951
2	0.1357	0.0805	0.3110	0.1581
3	-0.0748	0.0145	-0.0294	0.1988
4	-0.0285	-0.0694	0.1866	0.2310
5	-0.1439	-0.0966	0.0950	0.2600
6	0.2546	0.0328	-0.0091	0.2874
7	0.1012	0.3234	0.3844	0.3124
8	0.3405	0.4066	0.1768	0.3357
9	-0.0065	-0.0413	-0.1447	0.3576

Note: For $i = 1,..,9$, $\hat{F}_{it}$ is estimated by the method of principal components using a panel of data with 187 indicators of economic activity from 1993:03 to 2010:03 (205 time series observations). The data are transformed (taking logs and differenced where appropriate) and standardized prior to estimation. ρ_i denotes the ith autocorrelation. The 95% confidence bounds are ± 0.1397. The relative importance of the common component, R_i^2, is calculated as the fraction of total variance in the data explained by factors 1 to i.

Table 4
Modelling the 8 commodities returns: the three factors regressions - time period 1993:03-2010:03.

	Wheat	Copper	Silver	Soyabeans	Raw Sugar	Cotton	Crude Oil	Pork Bellies
Intercept	0.0011	0.0055	0.0073	0.0025	0.0052	0.0020	0.0069	0.0049
	(0.19)	(1.14)	(1.36)	(0.42)	(0.86)	(0.34)	(1.03)	(0.40)
$\widehat{F}_1$	-0.0396	-0.0928	-0.0279	-0.0407	-0.0286	-0.0566	-0.1241	0.0193
	(-2.21)	(-5.93)	(-1.59)	(-2.12)	(-1.45)	(-3.00)	(-5.67)	(0.49)
$\widehat{F}_2$	-0.0194	-0.0782	-0.0427	-0.0136	-0.0250	-0.0420	-0.1447	-0.0724
	(-0.88)	(-4.06)	(-1.98)	(-0.57)	(-1.03)	(-1.81)	(-5.38)	(-1.49)
$\widehat{F}_3$	0.0106	0.0749	0.0115	0.0665	0.0121	0.0280	-0.0011	0.0115
	(0.38)	(3.13)	(0.43)	(2.26)	(0.40)	(0.97)	(-0.03)	(0.19)
R^2	0.0282	0.2347	0.0322	0.0472	0.0165	0.0619	0.2335	0.0124
$\overline{R}^2$	0.0136	0.2233	0.0178	0.0330	0.0018	0.0479	0.2220	-0.0024
ARCH-LM (2)	14.62*	0.3362	4.8693***	13.85*	3.5515	14.11*	4.7081***	12.20*

Notes: (i) This table reports OLS estimates of the regression of the 8 commodities monthly returns with the variables reported in left column. Commodities are reported in upper row. A constant is always included in the regression and $\widehat{F}_i$ denotes the ith factor. (ii) t-statistics are reported in parenthesis under the estimates. ***, **, and * respectively denotes rejection of the null hypothesis of no significance at the 1%, 5% and 10% levels. (iii) For the ARCH_LM, ***, **, and * respectively denotes rejection of the null hypothesis of no ARCH effect at the 1%, 5% and 10% levels.

Table 5
Modelling the 8 commodities returns: the BIC minimizing regressions - time period 1993:03-2010:03.

	Wheat	Copper	Silver	Soyabeans	Raw Sugar	Cotton	Crude Oil	Pork Bellies
Intercept	0.0014	0.0059	0.0073	0.0025	0.0052	0.0020	0.0069	0.0049
	(0.25)	(1.24)	(1.37)	(0.42)	(0.89)	(0.34)	(1.15)	(0.41)
$\widehat{F}_1$	-0.0252	-0.0843***				-0.0383**	-0.1245***	
	(-1.58)	(-5.76)				(-2.20)	(-6.42)	
$\widehat{F}_2$		-0.0682***	-0.0330	-0.0683**			-0.1420***	
		(-3.76)	(-1.58)	(-2.32)			(-5.97)	
$\widehat{F}_3$		0.0669***						
		(3.11)						
$\widehat{F}_4$							0.1081	
							(3.32)	
$\widehat{F}_5$					0.1019***			
					(3.18)			
$\widehat{F}_6$								
$\widehat{F}_7$							0.1521***	
							(4.12)	
$\widehat{F}_8$							0.1672***	
							(4.36)	
$\widehat{F}_9$								0.1809**
								(2.28)
R^2	0.0224	0.2396	0.0180	0.040	0.0499	0.0377	0.3759	0.0211
$\overline{R}^2$	0.0176	0.2282	0.0132	0.0353	0.0452	0.0330	0.3602	0.0163
ARCH-LM (2)	16.03*	0.23	4.49**	0.04	1.2092e-004	12.88*	14.73*	12.15*

Notes: This table reports OLS estimates of the regression of the 8 commodities monthly returns with the variables reported in left column. Commodities are reported in upper row. A constant is always included in the regression and $\widehat{F}_i$ denotes the ith factor. (ii) t-statistics are reported in parenthesis under the estimates. ***, **, and * respectively denotes rejection of the null hypothesis of no significance at the 1%, 5% and 10% levels. (iii) For the ARCH_LM, ***, **, and * respectively denotes rejection of the null hypothesis of no ARCH effect at the 1%, 5% and 10% levels.

Table 6
Regression of the Kilian real activity index on each of the 9 factors

	$\widehat{F}_1$	$\widehat{F}_2$	$\widehat{F}_3$	$\widehat{F}_4$	$\widehat{F}_5$	$\widehat{F}_6$	$\widehat{F}_7$	$\widehat{F}_8$	$\widehat{F}_9$
Coefficient	-17.54***	-21.05***	-4.52	26.11***	-19.97*	11.53	61.95***	-41.35***	10.26
t-stat	(-3.11)	(-3.03)	(-0.51)	(2.67)	(-1.93)	(1.07)	(5.96)	(-3.65)	(0.85)
R^2	0.0455	0.0433	0.0013	0.0341	0.0180	0.0056	0.1491	0.0616	0.0036
$\bar{R}^2$	0.0408	0.0385	-0.0036	0.0293	0.0131	0.0007	0.1449	0.0570	-0.0013

Notes: *Coefficient* reports the estimated coefficient of each factor and *t-stat* its Student statistic. ***, **, and * respectively denotes rejection of the null hypothesis of no significance at the 1%, 5% and 10% levels.

Table 7
Correlation between residuals from the 3 factors linear model.

	Wheat	Copper	Silver	Soyabeans	Raw Sugar	Cotton	Crude Oil	Pork Bellies
Wheat	1.0000	0.2361 ***	0.1558 **	0.4293 ***	0.0047	0.2109 ***	0.0309	0.1048
Copper	0.0007	1	0.2791 ***	0.1251 *	0.1831 ***	0.1568 **	0.2045 ***	0.0376
Silver	0.0257	0.0001	1	0.0889	0.1189 *	0.0630	0.0926	0.1383 **
Soyabeans	0.0000	0.0739	0.2051	1	0.0663	0.3729 ***	-0.0335	0.1294 *
Raw Sugar	0.9462	0.0086	0.0895	0.3448	1	0.1195 *	0.0063	0.0141
Cotton	0.0024	0.0247	0.3695	0.0000	0.0880	1	0.0857	0.0072
Crude Oil	0.6606	0.0033	0.1868	0.6337	0.9291	0.2216	1	-0.0278
Pork Bellies	0.1347	0.5926	0.0479	0.0644	0.8412	0.9187	0.6921	1
Breusch Pagan LM test					688.98			
p-value					1			

Note: The upper triangular matrix reports correlation while the lower reports the p-values. ***, ** and * respectively denotes significance at 1%, 5 % and 10 %.

Table 8
Correlation between residuals from the BIC minimizing regressions.

	Wheat	Copper	Silver	soyabeans	Raw Sugar	Cotton	Crude Oil	Pork Bellies
Wheat	1	0.2395 ***	0.1636 **	0.4248 ***	0.0022	0.2287 ***	0.0209	0.1101
Copper	0.0005	1	0.2841 ***	0.1352 **	0.1492 **	0.1766 **	0.2027 ***	0.0442
Silver	0.0191	0.0000	1.0000	0.0946	0.1252 *	0.0754	0.0839	0.1411 **
soyabeans	0.0000	0.0532	0.1774	1.0000	0.0537	0.3716 ***	-0.0606	0.1375 **
Raw Sugar	0.9753	0.0328	0.0736	0.4446	1.0000	0.1458 **	-0.0167	0.0228
Cotton	0.0010	0.0113	0.2829	0.0000	0.0370	1	0.0828	0.0336
Crude Oil	0.7661	0.0036	0.2319	0.3880	0.8116	0.2378	1	-0.0362
Pork Bellies	0.1159	0.5289	0.0436	0.0493	0.7456	0.6321	0.6065	1
Breusch Pagan LM test					713.78			
p-value					1			

Note: The upper triangular matrix reports correlation while the lower reports the p-values. ***, ** and * respectively denotes significance at 1%, 5 % and 10 %.

Table 9
Descriptive statistics on returns and residual squared correlations - time period 1998:02 to 2010:03 (146 observations)

	$\hat{\rho}^*_{ret,t}$	$\hat{\rho}^*_t$
μ	0.1844**	0.1792**
σ	0.0342	0.0271
$F\rho^{*2}$	59.59%	58.22%
C_ρ		0.9743

Notes: (i) This table reports summary statistics on squared unconditional return correlation $\hat{\rho}^*_{ret,t}$ and squared unconditional residual correlation $\hat{\rho}^*_t$. (ii) $F\rho^{*2}$ is the mean percentage of squared unconditional correlation significant at the 5 % level using the t-square ratio test $\hat{t}^2_{ijt} = (\hat{\rho}^*_{ijt})^2[1 - \hat{\rho}^*_{ijt}]^{-1}(N - 2) \sim F(1, N - 2)$. (iii) ***, ** and * respectively denotes significance at 1%, 5 % and 10 %. (iv) C_ρ is the correlation between $\hat{\rho}^*_{ret,t}$ and $\hat{\rho}^*_t$.

Table 10

Regression of average excess residual correlation on Total Open Interest (in log first difference) - time period 1998:02 to 2010:03 (146 observations).

	$\hat{\rho}_t^{*2}$
Intercept	0.1798***
	(79.74)
Wheat	0.0029
	(0.16)
Copper	0.0328
	(1.24)
Silver	-0.0228
	(-1.10)
Soyabeans	-0.0064
	(-0.36)
Sugar	-0.0273
	(-1.21)
Cotton	-0.0430*
	(-1.73)
Crude oil	-0.0295
	(-0.69)
R^2	0.0461
$\overline{R}^2$	-0.0023

Notes: (i) t-statistics are reported in parenthesis. Standard errors are computed with the White (1980) covariance matrix. (ii) ***, ** and * respectively denotes significance at the 1 %, the 5 % and the 10 % significance levels.

Table 11
Regression of average excess residual correlation on Han and RNV indices - time period 1998:02 to 2010:03 (146 observations).

	$\hat{\rho}_t^{*2}$			
	(1)	(2)	(3)	(4)
Intercept	0.1699***	0.1942***	0.2335***	0.2303***
	(45.49)	(27.50)	(23.80)	(25.13)
Han Wheat	-0.0354**		-0.0554	
	(-2.46)		(-1.48)	
Han Copper	-0.0499***		-0.1181**	-0.1679***
	(-3.43)		(-2.03)	(-3.52)
Han Silver	0.0183		0.2344***	0.2124***
	(1.42)		(6.58)	(6.43)
Han Soyabeans	0.0203		-0.0503	
	(1.30)		(-1.31)	
Han Sugar	0.0566***		0.3820***	0.3421***
	(4.40)		(7.88)	(7.32)
Han Cotton	0.0407***		0.2907***	0.1851***
	(4.62)		(4.30)	(4.03)
Han Crude oil	0.0191		-0.1451	
	(0.61)		(-1.51)	
RNV Wheat		0.0561***	-0.0519	
		(4.43)	(-1.28)	
RNV Copper		0.0227*	-0.1186**	-0.1713***
		(1.88)	(-2.14)	(-3.60)
RNV Silver		0.0342***	0.2113***	0.2078***
		(2.57)	(5.98)	(6.12)
RNV Soyabeans		-0.0135	-0.0731***	-0.0447***
		(-1.36)	(-2.61)	(-5.52)
RNV Sugar		-0.0240**	0.3382***	0.3024***
		(-2.28)	(7.67)	(7.10)
RNV Cotton		-0.0329***	0.3033***	0.1973***
		(-4.27)	(4.26)	(3.96)
RNV Crude oil		-0.0746*	-0.2499**	-0.0682**
		(-1.94)	(-2.34)	(-2.01)
R^2	0.2939	0.2798	0.6253	0.5981
$\overline{R}^2$	0.2581	0.2433	0.5852	0.5683

Notes: (i) t-statistics are reported in parenthesis. Standard errors are computed with the White (1980) covariance matrix. (ii) ***, ** and * respectively denotes significance at the 1 %, the 5 % and the 10 % significance levels.

Table 12
Tests for structural breaks

$SupF_T(1)$	$SupF_T(2)$	$SupF_T(2)$	$SupF_T(2\mid1)$	$SupF_T(3\mid2)$
354.20***	331.07***	350.39***	83.34***	118.56***
UD max	*WD max* (5%)	*WD max* (10%)	*WD max* (1%)	
354.20***	354.20***	354.20***	354.20***	

Notes: (i) $SupF_T(l)$ reports the test statistic of 0 versus l structural breaks. $SupF_T(l+1\mid l)$ reports the test statistics of $l+1$ versus l structural breaks. *UD max* and *WD max* are the *"double maximum tests"* of no structural break versus an upper bound $M = 5$ structural breaks (*see Bai and Perron (1998, 2003a) for more details*). (ii) ***, ** and * respectively denotes significance at the 1 %, the 5 % and the 10 % significance levels. Critical values are taken from Bai and Perron (2003b).

Table 13
Regression of average excess residual correlation on Han and RNV indices - Parameters estimates
with three structural breaks.

	Dates of a break			
	$\hat{T}_1$	$\hat{T}_2$	$\hat{T}_3$	
	01:2000	05:2004	02:2007	
Confidence Intervals	(11:2000-05:2000)	(05:2004-07:2004)	(12:2006-03:2007)	
	Parameters estimates			
	$\hat{\rho}_t^{*2}$			
Period	(1)	(2)	(3)	(4)
Intercept	0.1589***	0.2045***	0.1780***	0.2598***
	(18.74)	(29.78)	(67.51)	(10.66)
Han Copper	-0.1216***	0.0321	0.1304**	-0.1681
	(-2.62)	(0.89)	(2.23)	(-1.01)
Han Silver	0.0377	0.0309	0.0418***	0.3920***
	(1.00)	(1.28)	(3.24)	(2.66)
Han Sugar	-0.0111	0.0847**	0.0881**	1.0261***
	(-0.22)	(2.24)	(2.46)	(5.67)
Han Cotton	0.0799**	0.1614***	0.0159	0.5401***
	(2.30)	(3.56)	(0.68)	(3.07)
RNV Copper	-0.1218***	0.0294	0.1070**	-0.2297
	(-3.13)	(0.86)	(2.10)	(-1.37)
RNV Silver	0.0196	0.0462*	0.0424***	0.3987***
	(0.69)	(1.95)	(3.91)	(3.31)
RNV Soyabeans	-0.0308***	0.0031	0.0076***	-0.1669***
	(-2.96)	(0.54)	(2.85)	(-4.09)
RNV Sugar	-0.0084	0.0780**	0.0866**	0.9320***
	(-0.19)	(2.23)	(2.45)	(4.43)
RNV Cotton	0.0751**	0.1607***	0.0185	0.7109***
	(2.08)	(3.28)	(0.85)	(3.89)
RNV Crude Oil	0.0657**	-0.0295	-0.0688***	-0.2833**
	(2.03)	(-1.38)	(-3.08)	(-2.32)
Nobs	36	40	33	37
R^2	0.6157	0.5878	0.6678	0.9185
$\overline{R}^2$	0.4620	0.4187	0.5168	0.8871

Notes: (i) t-statistics are reported in parenthesis. Standard errors are computed to Andrews (1991) covariance matrix allowing for heterogeneity and autocorrelation in the residuals. (*see Bai and Perron (1998, 2003a) for more details*)(ii) ***, ** and * respectively denotes significance at the 1 %, the 5 % and the 10 % significance levels.

Figures

Figure 1
 Commodity monthly prices over the period 1993:02-2010:03 (206 observations).
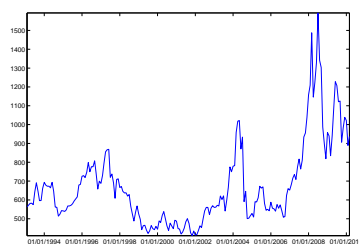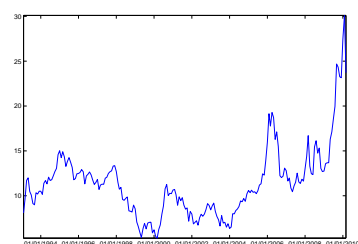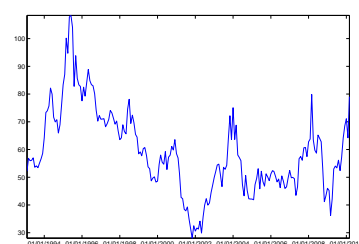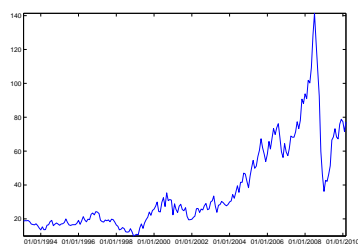

(a) *Wheat*

(b) *Copper*
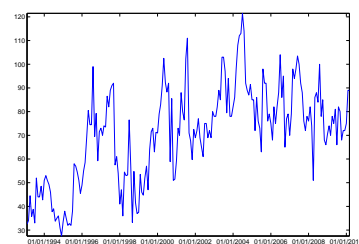

(c) *Silver*

(d) *Soyabeans*
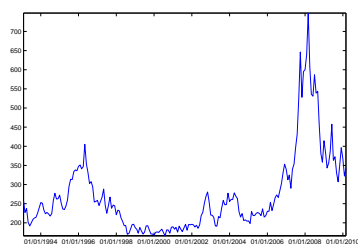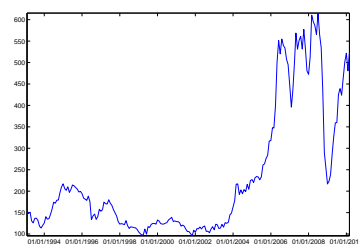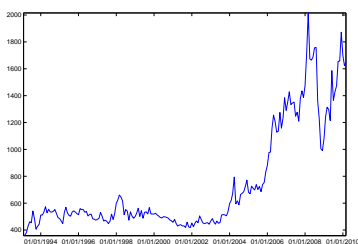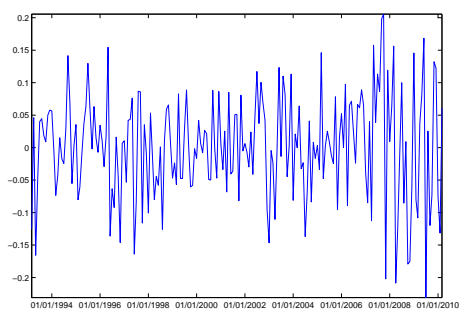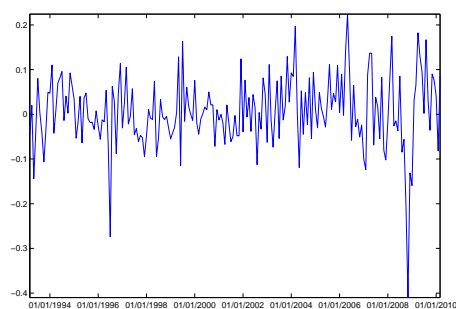

(e) *Sugar*

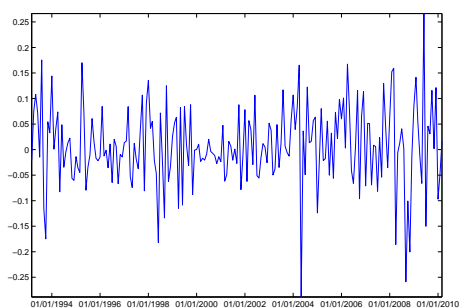(f) *Cotton*

(g) *Oil*

(h) *Pork Bellies*

Figure 2
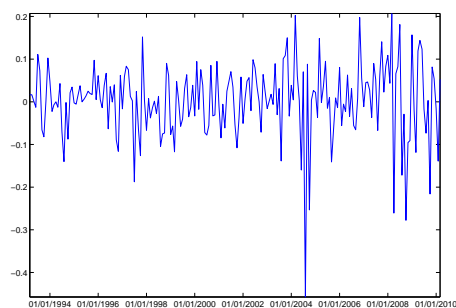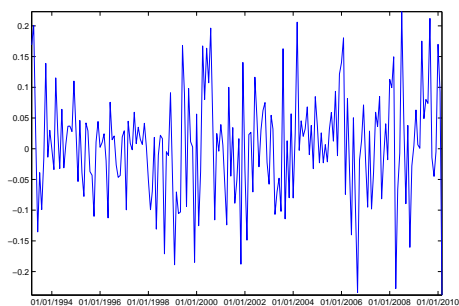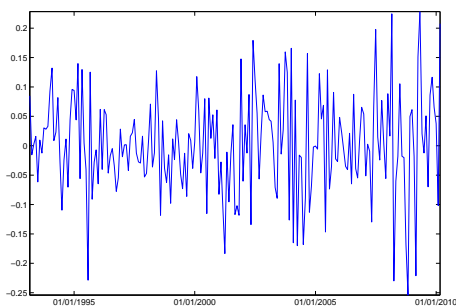Commodity monthly log returns over the period 1993:03-2010:03 (205 observations).
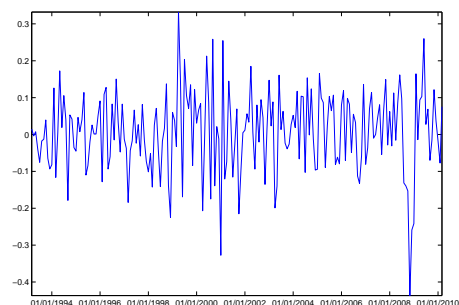

(a) *Wheat*

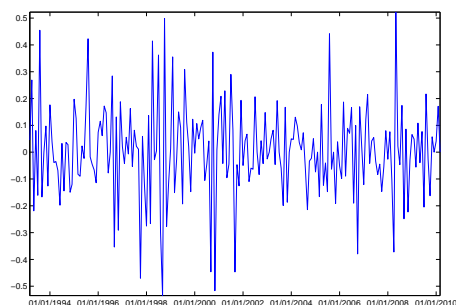(b) *Copper*

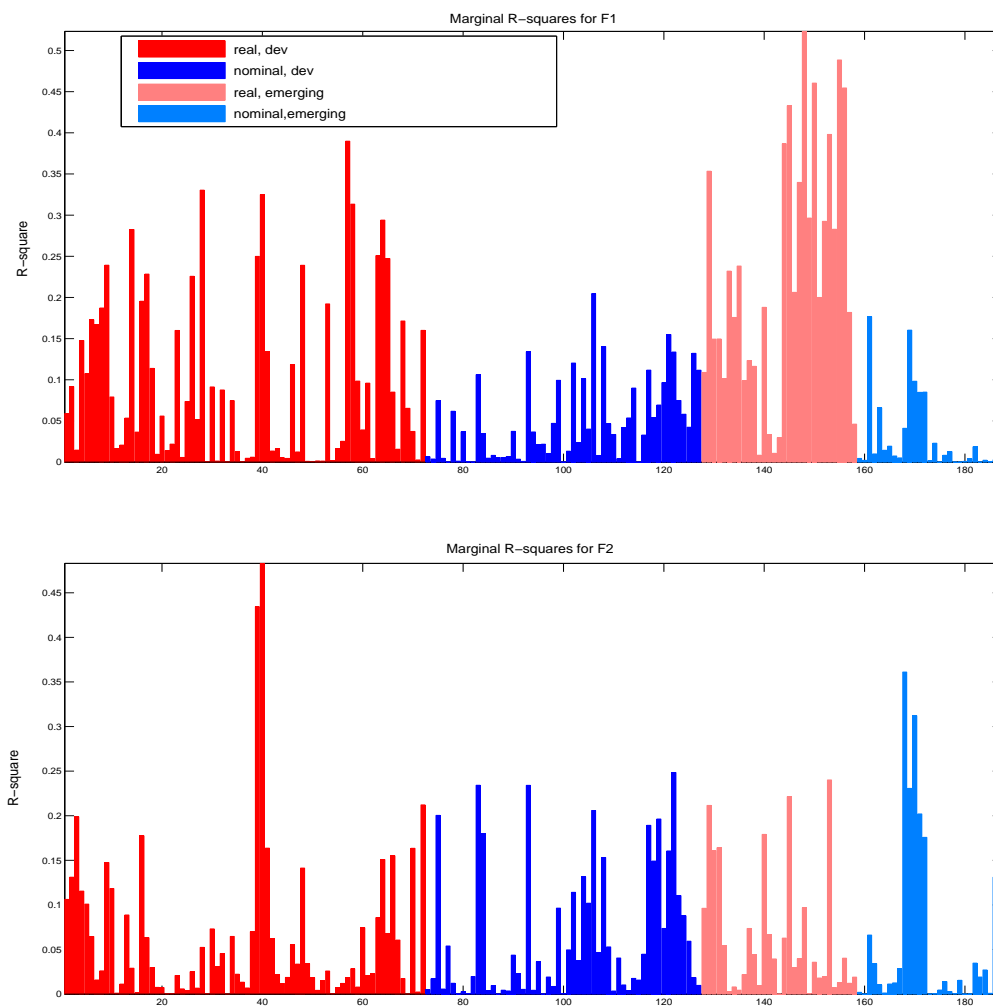(c) *Silver*

(d) *Soyabeans*

(e) *Sugar*

(f) *Cotton*

(g) *Oil*

(h) *Pork bellies*

Figure 3
Marginal R^2 of macroeconomic and financial variables regressed on the first two estimated factors.

Note: Each panel shows the R^2 from regressing the series number given on the x-axis onto each individual factor $\widehat{F}_i$. 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).
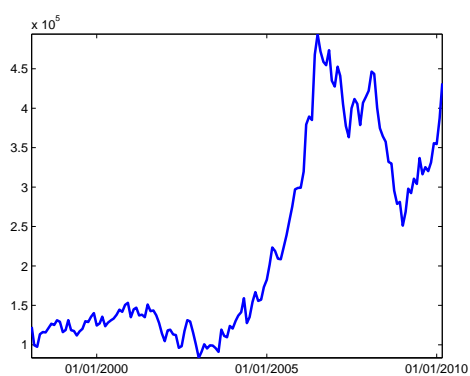
Figure 4

Mean excess squared correlation for raw returns and SUR residuals - time period 1998:02 to 2010:03 (146 observations)
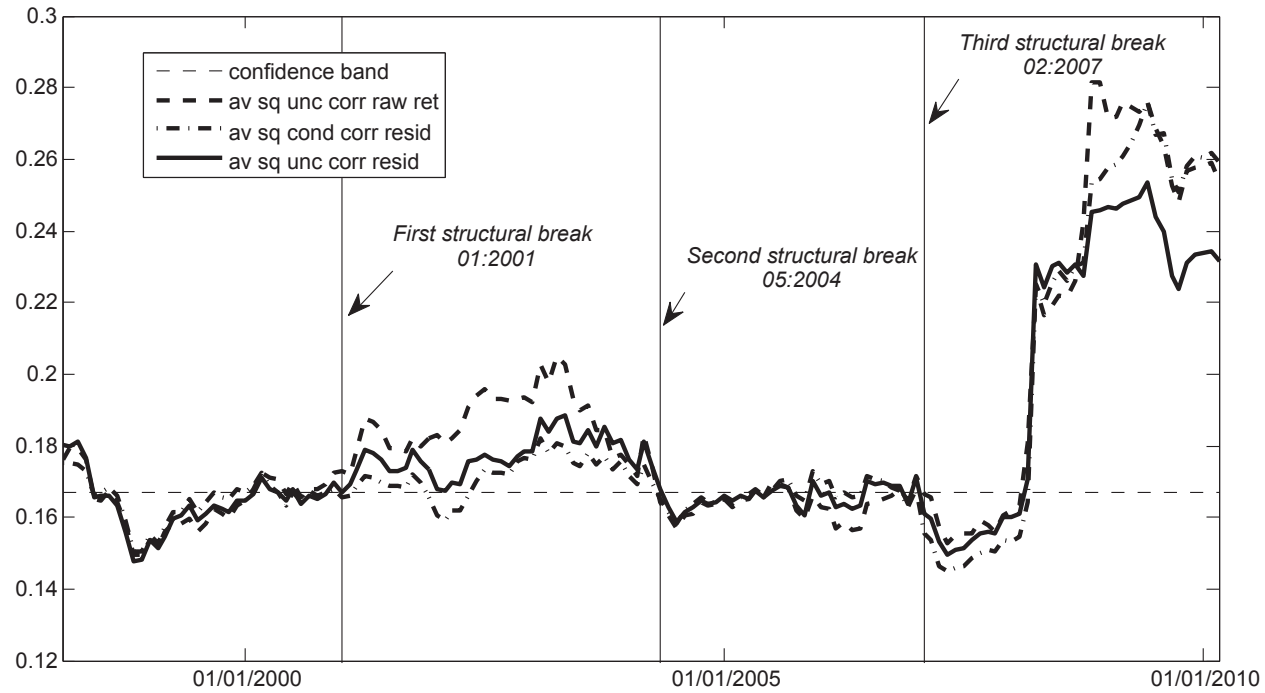
Notes: (i) "av sq unc corr raw ret" is the average squared unconditional correlation of raw returns: $\rho^*_{ret,t}$. (ii) "av sq cond corr resid " is the average square conditional residual. (iii) "av sq unc corr resid " is ρ^*_t. (iv) The confidence band the minimal value above which square correlation is significant at 5 % level. It is computed from the *t*-square ratio test $\hat{t}^2_{ijt} = (\hat{\rho}^*_{ijt})^2[1 - \hat{\rho}^*_{ijt}]^{-1}(N-2) \sim F(1, N-2)$.

Figure 5
Total open interest. 1998:02-2010:03 (146 observations).
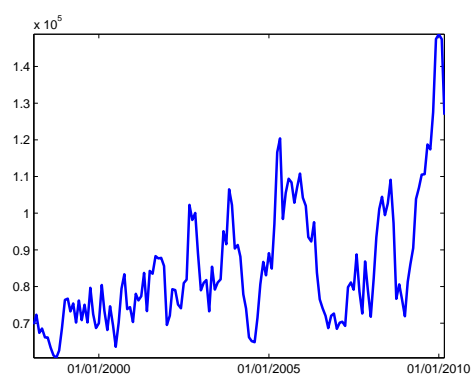
(a) *Wheat*

(b) *Copper*

(c) *Silver*

(d) *Soyabeans*

(e) *Sugar*

(f) *Cotton*

(g) *Oil*

Figure 6
De Roon hedging index. 1998:02-2010:03 (146 observations).

(a) *Wheat*

(b) *Copper*

(c) *Silver*

(d) *Soyabeans*

(e) *Sugar*

(f) *Cotton*

(g) *Oil*

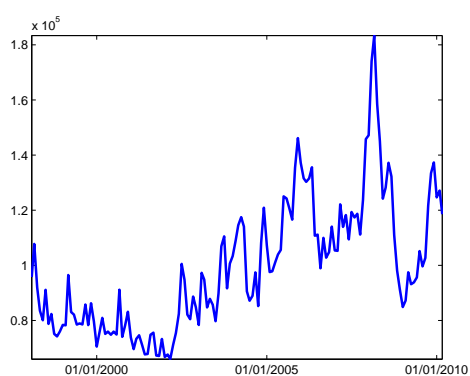**Figure 7**
Han index. 1998:02-2010:03 (146 observations).
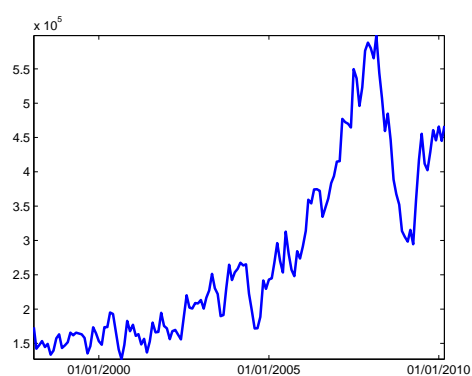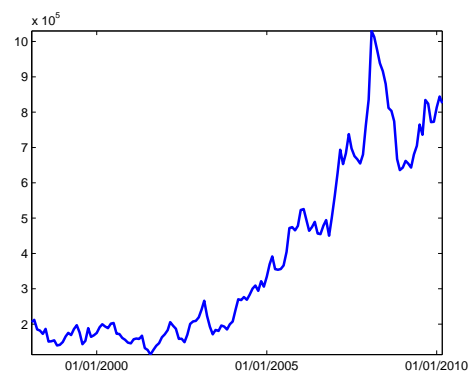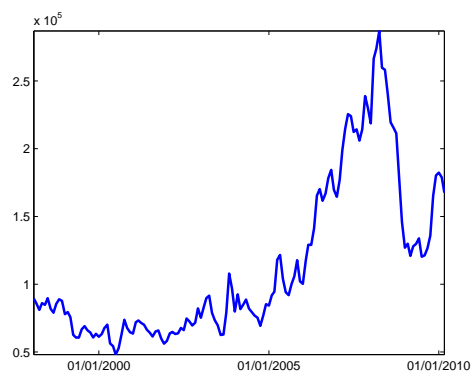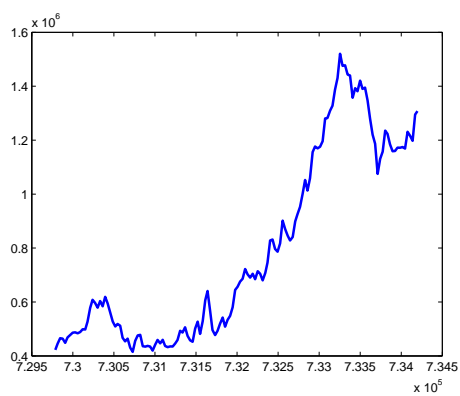

(a) *Wheat*

(b) *Copper*

(c) *Silver*

(d) *Soyabeans*
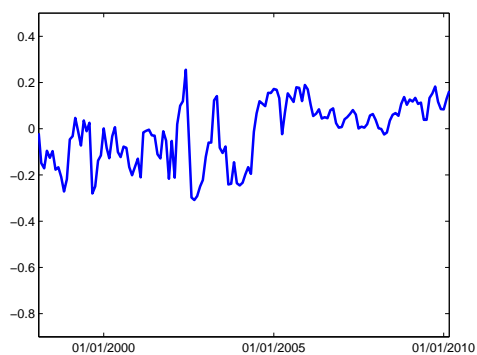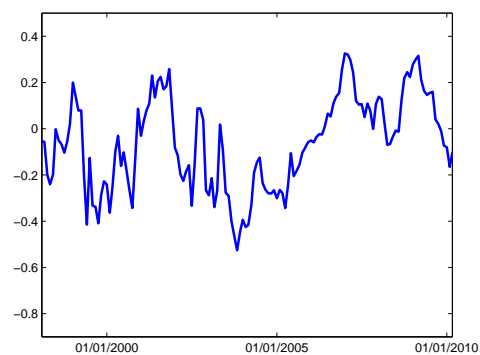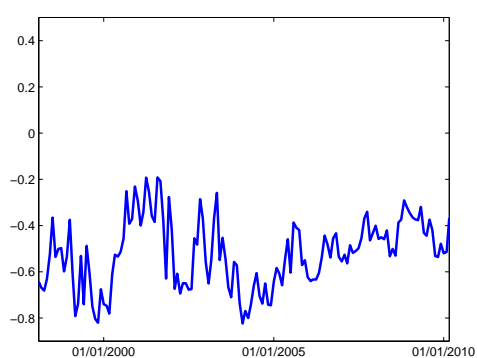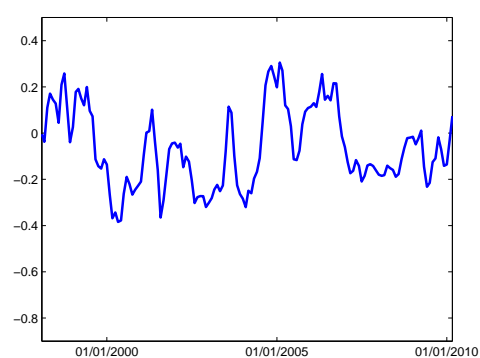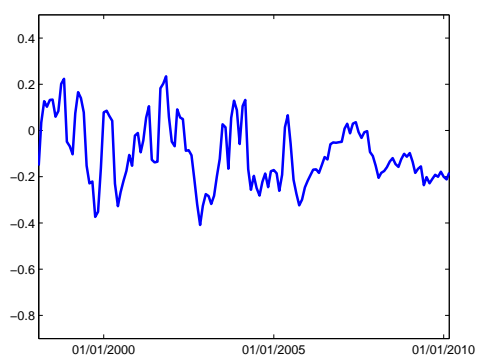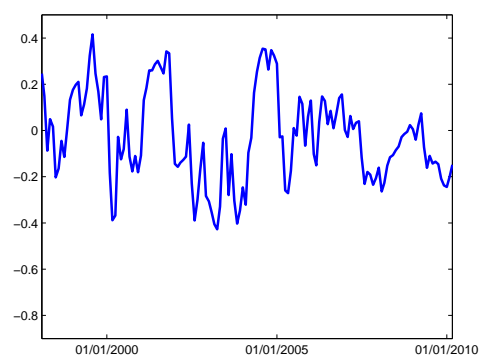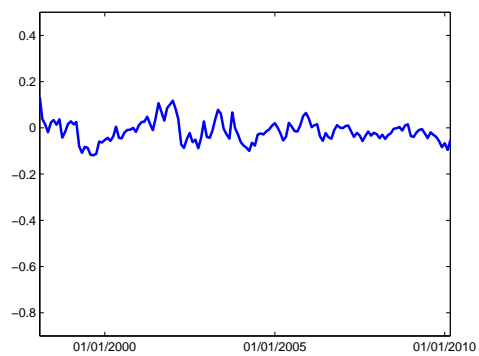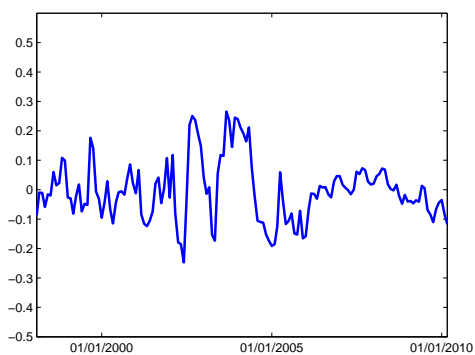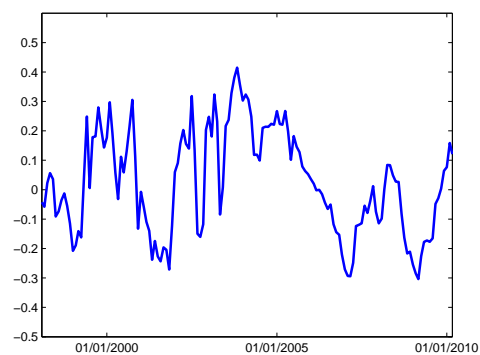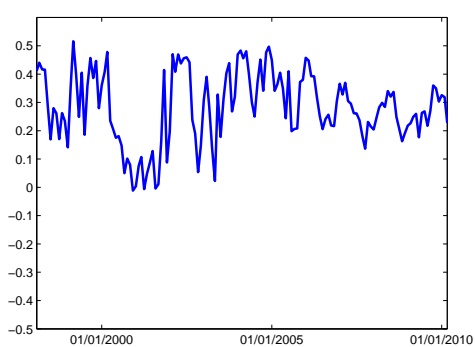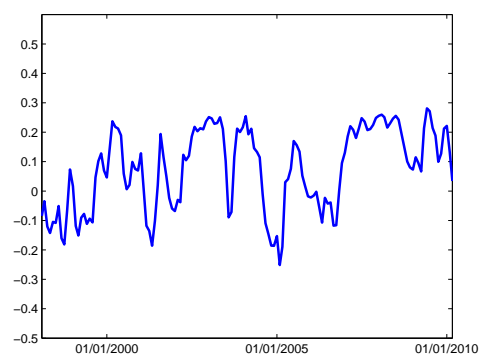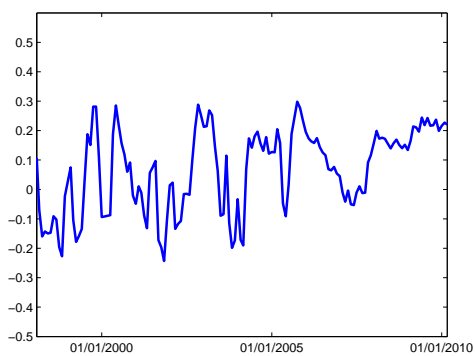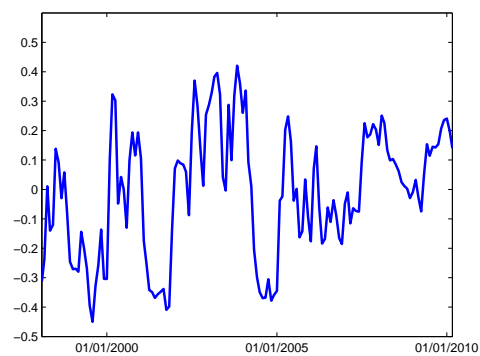
(e) *Sugar*

(f) *Cotton*
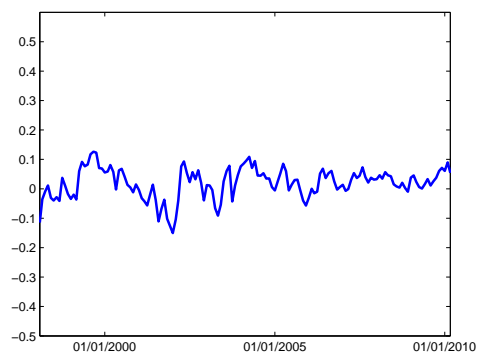
(g) *Oil*

Appendix A: Estimating the number of factors

Bai and Ng (2002) proposed two kinds of information criteria to select the number of common factors. If we note $\hat{S}(k) = (NT)^{-1} \sum_{i=1}^{N} \sum_{t=1}^{T} (x_{it} - \hat{\lambda}_i^{k'} \hat{F}_t^k)^2$ the sum of squared residuals (divided by NT) when k factors are estimated, these information criteria have the following expressions:

$$PCP_i(k) = S(k) + k\bar{\sigma}^2 g_i(N, T)$$
$$IC_i(k) = ln(S(k)) + k g_i(N, T)$$

where $g(N, T)$ is a penalty function[30] and $\bar{\sigma}^2$ is equal to $S(k_{max})$ for a pre-specified value k_{max}. The estimated number of factors $\hat{k}$ minimises the aforementioned information criteria.

We also apply the sequential test by Kapetanios (2010) to determine the number of factors. This test is based on the property that if the true number of factors is k_0, then, under some regularity conditions, the first k_0 eigenvalues of the population covariance matrix Σ will increase at rate N while the others will remain bounded. If we denote by $\hat{\lambda}_k, k = 1, ..., N$ the N eigenvalues of the sample covariance matrix $X'X$, the difference $\hat{\lambda}_k - \hat{\lambda}_{k^{max}+1}$ will tend to infinity for $k = 1, ..., k_0$ but remain bounded for $k = k_0 + 1, ..., k_{max}$ where k_{max} is some finite number such that $k_0 < k_{max}$. 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 ($H_{1,k} : k_0 > k$) is therefore tested with the test statistics $\hat{\lambda}_k - \hat{\lambda}_{k^{max}+1}$. If there is no factor structure, $\hat{\lambda}_k - \hat{\lambda}_{k^{max}+1}$ properly normalized by a sequence of constant $\tau_{N,T}$ should converge to a law limit. In the presence of factors, it should tend to infinity. The law limit as the rate of convergence $\tau_{N,T} \to \infty$ have to be estimated by resampling technique. The test is sequential. In a first step, we test ($H_{0,k} : k_0 = k = 0$) against ($H_{1,k} : k_0 > 0$). If we reject the null hypothesis, then we consider the null ($H_{0,k} : k_0 = k + 1 = 1$). We stop once we cannot reject the null hypothesis. Kapetanios (2010) called this algorithm the MED (maximal eigenvalue distribution) algorithm.

The estimated[31] numbers of factors are displayed in Table 14. There is clearly no agreement on the optimal number of factors. This result is similar to previous empirical studies, which show that there is a great instability in determining the correct number of factors[32]. According to the information criteria by Bai and Ng (2002), the optimal number of factors runs from the 2 to 9. The sequential test by Kapetanios (2010) returns a number of factors equal to 2. Additional information on the autocorrelation and the explanatory power of the estimated factors $\widehat{F}_t$ are displayed in Table 3. The first 3 factors only explain 20% of the variance of the 187 time series, while we reach 36% with 9 factors. Hence, we choose to consider the set of the first 9 factors as potential set of regressors. Factors' autocorrelations (up to 3 lags) provided in Table 3 show that most of them are persistent.

---

[30] Penalty functions suggested by Bai and Ng (2002) are: $g_1(N, T) = \frac{N+T}{NT} ln(\frac{NT}{N+T})$, $g_2(N, T) = \frac{N+T}{NT} ln(C_{NT}^2)$, $g_3(N, T) = \frac{ln(C_{NT}^2)}{C_{NT}^2}$, $g_4(N, T) = (N + T - k)\frac{ln(NT)}{NT}$

[31] We use the Matlab routine provided by Ng for the estimation and information criteria. The Kapetanios test routine is ours.

[32] See for instance the empirical applications in Kapetanios (2010) which show that there may be some variations in the estimations of the number of factors according to the selection criterion.

Table 14
Static factors selection results

Method	No of static 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

Note: 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 $g_i(N, T)$.

Appendix B: list of the 187 variables considered in the computation of the common factors

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

Series Number	Short name	Mnemonic	Tran	Description
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

Series Number	Short name	Mnemonic	Tran	Description
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

Series Number	Short name	Mnemonic	Tran	Description
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

Series Number	Short name	Mnemonic	Tran	Description
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

Series Number	Short name	Mnemonic	Tran	Description
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

Series Number	Short name	Mnemonic	Tran	Description
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

Emerging countries

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

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