FUNDAÇÃO GETULIO VARGAS
ESCOLA DE ECONOMIA DE SÃO PAULO

TOMMASO MOMOLI

FINANCIALIZATION OF THE COMMODITY FUTURE MARKETS:

A SVAR MODEL APPROACH

SÃO PAULO
2016

FUNDAÇÃO GETULIO VARGAS
ESCOLA DE ECONOMIA DE SÃO PAULO

TOMMASO MOMOLI

FINANCIALIZATION OF THE COMMODITY FUTURE MARKETS:

A SVAR MODEL APPROACH

Dissertação apresentada à Escola de Economia de São Paulo da Fundação Getulio Vargas, como requisito para obtenção do título de Mestre Profissional em Economia.

Campo do Conhecimento:
International Master in Finance

Orientador Prof. Dr. Pedro Luiz Valls Pereira e Martjin Boons

SÃO PAULO
2016

Momoli, Tommaso.

Financialization of the commodity future markets : a SVAR model approach / Tommaso Momoli. - 2016. 45 f.

Orientadores: Pedro Luiz Valls Pereira e Martjin Boons

Dissertação (MPFE) - Escola de Economia de São Paulo.

1. Mercado futuro de mercadorias. 2. Finanças. 3. Bolsa de mercadorias - Índices. 4. Modelos econométricos. I. Pereira, Pedro L. Valls (Pedro Luiz Valls). II. Boons, Martjin. III. Dissertação (MPFE) - Escola de Economia de São Paulo. IV. Título,

CDU 336.76

TOMMASO MOMOLI

FINANCIALIZATION OF THE COMMODITY FUTUTRE MARKETS:

A SVAR MODEL APPROACH

Dissertação apresentada à Escola de Economia de São Paulo da Fundação Getulio Vargas, como requisito para obtenção do título de Mestre Profissional em Economia.

Campo do Conhecimento:
International Master in Finance

Data de Aprovação:
25 / 01 / 2017.

Banca Examinadora:

Prof. Dr. PEDRO LUIZ VALLS PEREIRA

Prof. Dr. MARTJIN BOONS

Prof. Dr. LUIS CAMPOS E CUNHA

To my family
Antonia, Alberto
and Andrea.
To Tatiana.

RESUMO

Trata-se de um estudo sobre o impacto dos investimentos em índices no mercado futuro de commodities. Os modelos aplicados, enfocam a Análise Causal e a Função de Resposta ao Impulso através de uma ortogonalização do Vetor de Auto Regressão (SVAR), permitindo extrair a correlação lead / lag entre o Índice e o Primeiro Retorno próximo para diferentes Setores Futuros e, A choques em diferentes equações. O estudo é dividido em três períodos diferentes, para refletir antes e depois da Financialização e, em seguida, após a introdução no mercado da nova geração de índices de commodities. Os resultados mostram um comportamento diferente dos parâmetros ao longo do tempo com uma ênfase particular para os Commodities mais negociados para liderar os outros.

PALAVRAS CHAVE: Índices de commodities, Futuros, Causalidade de Granger, IRF ortogonalizado

ABSTRACT

This is a study regarding the impact of the index investments in the Commodity Future Market. The models applied, focus on the Causal Analysis and the Impulse Response Function through an orthogonalisation of the Vector of Auto Regression (SVAR), this allow to extract lead/lag correlation between the Index and First nearby Return for different Futures Sectors and in addition response to shocks in different equation. The study is divided in three different period, to reflect before and after the Financialization and then after the introduction in the market of the new generation of commodity Indexes. The results show a different behaviors of the parameters throughout time with a particular emphasis for the most traded Commodities to lead the others.

KEYWORDS: Commodity Indexes, Futures, Granger Causality, Orthogonalised IRF

INDEX

INTRODUCTION

≪The synchronized rise and fall in prices of oil and a broad set of non-energy commodities in 2006-2008 has stimulated increasing public attention to commodities markets≫1

Which is the role of indexes in the commodity future markets? Did buying pressure from new index investment created a massive bubble in commodity futures prices? One group of studies finds evidence that commodity index investment directly or indirectly had an impact on commodity futures prices.

Most investors have no desire to take delivery of hogs, corn, oil or any other commodity, they simply want to profit from price changes. Purchasing futures contracts is one way to achieve this objective, this process is called financialization.

As a result of the financialization process, the price of an individual commodity is no longer simply determined by its supply and demand or from the intent of hedging. Instead, commodity prices are also determined by a whole set of financial factors, such as the aggregate risk appetite for financial assets, and investment behavior of diversified commodity index investors.

On one hand, the presence of these investors can lead to a more efficient sharing of commodity price risk; on the other hand, their portfolio rebalancing can spill over price volatility from outside to commodities markets and also across different commodities.

Commodity index products have a variety of forms including managed funds, ETFs, ETNs, and OTC return swaps2.

My purpose is to study the impact of the index investment in the price of commodity in three different period. Before and after 1999 when the Dow Jones-UBS Commodity Index was launched, a standard benchmark for commodities investing, and after 2008 which is roughly the time when significant index investment started to flow into commodities markets after the crisis.

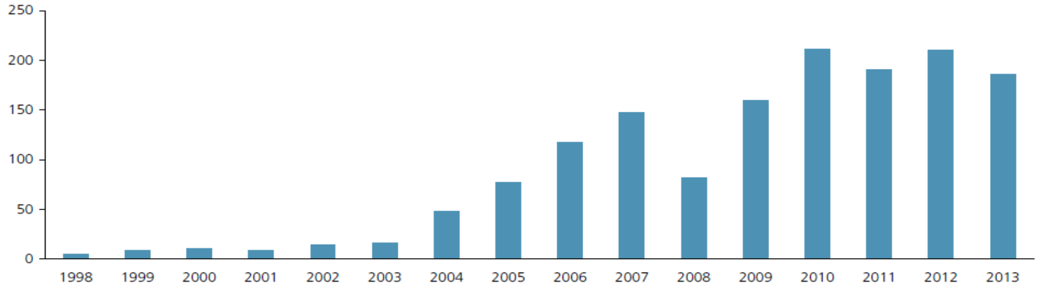
Commodities institutional investments rose from $18 billion in 2003 to more than

---

1 Tang, Ke, and Wei Xiong. "Index investment and the financialization of commodities." (2012).

2 Stoll, Hans R., and Robert E. Whaley. "Commodity index investing and commodity futures prices." (2010)

$200 billion in 2010. Notional Amount tracking commodity Indices could be seen in Table(1)



3*Source UBS Bloomberg (USD Billions)

To identify the effects of growing commodity index investment, I will try to analyze how does the parameters changed over time. Initially it will be applied to sectorial commodity future returns a VAR Model, a useful device to analyse causation links among variables. Subsequently a Structural VAR which allows me to take in consideration the contemporaneous variable and eventually estimating and interpreting impulse response functions and instantaneous correlations among the relevant variables which will be applied between Indexes and the most relevant sectors.

3 UBS Bloomberg SPGSCI Constant Maturity Termsheet (2014)

LITERATURE REVIEW

This Literature review has no claim to present a complete review of all studies made on the relationship between commodity markets and Index investment considering the stakes that has been involved.

Many of the recent studies focused on the recent controversy about commodity markets and split the researcher into two sides, those who believe index funds were responsible for a bubble in commodity futures prices and those who do not believe index funds were behind the run-up in commodity futures prices.

Testing for the impact of index investment on commodity prices is challenging in addition the financial industry developed new products for investing in commodities through long-only index funds, over-the-counter (OTC) swap agreements, exchange traded funds, and other structured products and so collect accurate data became more challenging. These data made available by the CFTC (commodities future trading commission) could be some time more adapt than others, IID (Index Investment Data) is the preferred for index total position (Irwin and Sanders 2013). This easily lead to additional investigation and debate.

The side that believe that commodity index investment was a major driver of the 2007-2008 spike in commodity futures prices is composed by a number of hedge fund managers, commodity end-users, policy-makers, and some economists. One of the first is the Hedge fund manager Michael W. Masters who argues that unprecedented buying pressure from index investors created a massive bubble in commodity futures prices and this bubble was transmitted to spot prices through arbitrage linkages between futures and spot prices exceeding so fundamental values(Masters 2008). Others argue that commodity markets were not fully integrated with financial markets prior to the development of commodity index investments and, ≪The increasing presence of index investors in commodities markets precipitated a fundamental process of financialization amongst the commodities markets, through which commodity prices now become more correlated with the prices of financial assets and each other≫ (Tang and Xiong 2010) they conclude that index investment has

increasingly impacted commodity futures prices (although in a non-bubble manner). The researchers who have debated more are Irwin and Sanders claiming limited traces of this linkage, moving important criticisms on the data set and the methodologies of these studies. Irwin and Sanders tests, fail to reject the null hypothesis that commodity index positions have no impact on futures prices (Irwin and Sanders 2011). Another study analyzed the positions of commodity index funds questioning if it has to be considered speculation, hedging, or something altogether, coming to the conclusion that index investments is not speculative due to a motivation of portfolio diversification benefit and since it’s long only, passive and fully collateralized (Stoll and Whaley 2010).

Only commodity index rolls have a little future price impact, a relevant consideration in the next chapters.

More recent study also conclude that there is little evidence that passive index investment caused a massive bubble in commodity futures prices (Irwin and Sanders 2012) even with high frequency data (Irwin and Sanders 2014).

Therefore, it is unclear if there was a bubble in commodity futures price from 2007-2008, and even less clear whether one was caused by index funds.

CHAPTER I

From VAR to SVAR models Structural analysis

Economic, financial and business variables are not only autocorrelated, but often cross-related because of various time delays. There is therefore the need to study models that take into account the inter-temporal relationships between variables. In the analysis of multivariate historical series, the autoregression vector model (VAR) is extensively used. Historically the VAR approach was proposed by Sims in 1980, as an alternative to simultaneous equation models. From a conceptual point of view VARs processes are the multivariate generalization of autoregressive process (AR). A VAR process is therefore a system in which each variable is regressed on a set of deterministic variables (trend are omitted from the representation that follows, for simplicity), on n-lag delays of itself and each of the other variables included in the system. The focus is, initially, on the reduced form which is represented as a closed system, in which all variables are explicitly modeled.The simplest case that can be entertained is a bivariate VAR, where there are only two variables, y_{1t} and y_{2t}, each of whose current values depend on different combinations of the previous values of both variables, and error terms

$$\begin{bmatrix} y_{1t} \\ y_{2t} \end{bmatrix} = \begin{bmatrix} \alpha_1 \\ \alpha_2 \end{bmatrix} + \begin{bmatrix} \varphi_{11} & \varphi_{12} \\ \varphi_{21} & \varphi_{22} \end{bmatrix} \begin{bmatrix} y_{1t-1} \\ y_{1t-1} \end{bmatrix} + \begin{bmatrix} u_{1t} \\ u_{2t} \end{bmatrix} \tag{1.1}$$

where u_{it} is a white noise disturbance term with $E(u_{it}) = 0$, (i =1,2). As should already be evident, an important feature of the VAR model is its flexibility and the ease of generalisation.

Typically VAR representations focus on a limited number of variables and it is important to pay attention to the dynamics of the variables included to capture the effects of omitted variables. Immediately you notice that the application of all the analysis in the field of methodologies VAR requires as a necessary condition the stationary autoregressive representation. We rewrite in matrix form the generic order VAR *p*:

$$y_t = A(L)y_{t-1} + \varepsilon_t \tag{1.2}$$

the stationarity condition is verified if and only if $|A(L) - \lambda I| = 0$ meaning that λ values are less than 1 in absolute term.

In my analysis we provide a function augmented Dickey-Fuller to test time-series

for the stationarity property, and another routine co-integrating augmented

Dickey-Fuller to carry out the tests if the series are co-integrated. The general model VAR (p) contains many parameters that can be difficult to interpret, due to complex interactions and feedback between the variables in the model. Therefore, the dynamic properties of a VAR (p) are often synthesized through various types of structural analysis. My main focuses are on : (1) the causal analysis; (2) the impulse response functions.

1.1 Analysis of causality according to Granger

The VAR models are widely used for the analysis of causality. In general, in the empirical analysis of economic data, the cause-effect relationships are very complex to establish. If two X and Y variables are highly correlated, we can say that they have a clear tendency to move together, but, in the absence of further information, we can not say more about the direction of causality. It means that one cannot say with certainty which are the underlying causal links; It could be a variable that causes the performance of the other or vice versa, or that there is a third variable Z, not observable, that is the cause of both.
Given the system estimate a test to check if the variable y1t causes y2t variable can be carried out by checking the joint significance of the y1t lag structure in the equation that explains y2t. The test is conducted considering a likelihood ratio or, alternatively, a simple statistical F. This type of verification has enjoyed a certain reputation in the past, but it seems difficult to give an interpretation to the structural results.
In general it can be stated that the interpretation problems arise from the fact that the (1.1) is a reduced form representation which, by its nature, lends itself poorly to provide support to structural considerations. In a VAR area, the structural considerations are based usually on an analysis of the functions of the impulse response. Before moving on the analysis of these aspects, it is necessary to better analyze the Granger causality.

The Granger causality analysis has the aim to evaluate the predictive power of a variable to the other system variables. According to Granger, if a variable or group of variables, y1, assists in improving forecasts of another variable or group of variables, y2, then y1 causes y2.

Consider for example the stationary VAR(p) the :

$$\begin{bmatrix} y_{1t} \\ y_{2t} \end{bmatrix} = \begin{bmatrix} \varphi_{11,1} \varphi_{12,1} \\ \varphi_{21,1} \varphi_{22,1} \end{bmatrix} \begin{bmatrix} y_{1,t-1} \\ y_{2,t-1} \end{bmatrix} + \cdots + \begin{bmatrix} \varphi_{11,p} & \varphi_{12,p} \\ \varphi_{21,p} & \varphi_{22,p} \end{bmatrix} \begin{bmatrix} y_{1,t-p} \\ y_{2,t-p} \end{bmatrix} + \begin{bmatrix} u_{1t} \\ u_{2t} \end{bmatrix} \quad (1.3)$$

With $u_t \sim \mathrm{N}\,(0\,,\Sigma)$ We state that y_{1t}do not Granger cause y_{2t} if $\varphi_{21,i} = 0$ for

i=1,2….,p. And similarly if $\varphi_{12,i} = 0$ for i=1,2…p. y_{2t} do not Granger Cause y_{1t}. This concept are applied to the short run causality, in case of co-integration between variables long run causality is always present.

Granger-causality really means only a correlation between the current value of one variable and the past values of others; it does not mean that movements of one variable cause movements of another.

VARs are unrestricted reduced form models, useful as a starting step in order to guide the specification of a dynamic structural model. In this light, they are useful devices to analyse causation links among variables and to guide the researcher in deciding which series, among the observed variables, are truly exogenous.

Some caution in interpreting the results of non-causality tests is necessary. First of all, results are usually very sensitive to the information set being used in the application i.e. the set of series being included in the VAR: there is always the risk of finding "spurious" causation links deriving from omitted variables.

Causality Implies simply chronological order of movements in the series, could be stated that movements in one variable appears to lead a lag correlated one.[4]

[4] Amisano, Gianni, and Carlo Giannini. *Topics in structural VAR econometrics*. Springer Science & Business Media, 2012.

1.2 Impulse Response Function

To illustrate the concept of impulse response function we can rewrite the system in a compact form as follows, where L is the *lag operator*:

$$y_t = A(L)y_t + \mathcal{E}_t \tag{1.4}$$

$$A(L) = A_1L + A_2L^2 + .. + A_pL^p \tag{1.5}$$

Assuming *I-A(L)* is invertible, with *B(L) = (I-A(L))*$^{-1}$, we can get the moving average representation of the VAR (VMA)

$$y_t = \mathcal{E}_t + B_1\mathcal{E}_{t-1} + B_2\mathcal{E}_{t-2} + \ldots + B_s\mathcal{E}_{t-s} \tag{1.6}$$

We can then interpret as follow matrix B_s:

$$B_s = \frac{\partial y_{t+s}}{\partial \mathcal{E}_t} \tag{1.7}$$

In other words the ij element of Bs identifies the consequences of an increase of one unit in the innovations on the j-th variable VAR on the value of the i-th variable VAR to time t+s, maintaining zero all other innovations in all possible dates between t and t+s. This partial derivative only makes sense if it can be assumed that shocks on different variables are not correlated. Otherwise, if the variables are correlated, you would have a variance-covariance matrix of non-diagonal errors and therefore biased results.
The impulse response function describes the effects of a temporary shock (duration is a period) to the VAR variable j on variable i, but allows contemporary correlation between the error terms, hence is not a valid assumption if the error terms are correlated.

1.3 Structural VAR

It has been seen as Sims (1980) used the VAR(p) model

$$y_t = \varphi_1 y_{t-1} + \varphi_2 y_{t-2} + \ldots + \varphi_p y_{t-p} + \varepsilon_t \qquad \tilde{\varepsilon}_t \sim N(0, \Sigma_{\mathcal{E}}) \tag{1.8}$$

in order to analyse the effects of a y_{it} variable of the vector y_t of a shock incorporated in any of the $\mathcal{E}_t$ vector disorders. As indicated, since these disorders are generally related to each other at zero lag due to the dispersion of Σ matrix, it is not possible, without making additional assumptions, distinguish the effects of a y_{it} variable of the vector y_t from a specific shock disorder from others which, although they originate from the same shock, they spread about other disorders by $\Sigma_{\mathcal{E}}$ and they also impact on the y_t.

To analyze the VAR approach the problem of identification consider the structural form of the system (1.1) can be considered as a reduced form:

$$y_t = \sum_{i=0}^{p} A_i\, y_{t-1} + Cu_t \qquad (u_t)\sim \mathrm{N}(0, \mathrm{I}) \qquad (1.9)$$

Here shocks are orthogonal to each other, then you can calculate and interpret correctly the impulse response functions. The ut are interpreted as "primitive" shocks, deprived of common causes, and not related to each other. However, it does not impose the restriction that the individual shock enter into one and only one equation: the matrix B could not be diagonal. This interpretation specifies a structural model with traditional features, in fact, the stochastic components of the various equations can be related to each other. This correlation is generated by the fact that different equations have one or more shocks in common. The ut are the answer to these innovations that are of interest to the researcher, and it is only with respect to these innovations, perpendicular to each other, it makes sense to analyze the impulse response functions.

1.3.1 Identification by triangulation or Cholesky decomposition

As we have already stressed, a VAR model has to be considered as a reduced form model where no explanations of the instantaneous relationships among variables are provided. These instantaneous relationships are naturally hidden in the correlation structure of the Σ matrix, and left completely uninterpreted.

This because even if I estimate my VAR under hypothesis of Homoskedasticity and uncorrelated errors, my errors are serial uncorrelated, not contemporary uncorrelated.

The solution proposed by Sims (1980) to the identification problem is to consider B = I and A_0 lower triangular with unit diagonal elements, being able to have so exact identification of the VAR.

So calling A_0 the lower Cholesky decomposition of Σ, it let us achieve a new error terms that are orthogonals and contemporary uncorrelated, its Variance-Covariance matrix will be diagonal,

$A_0\mathcal{E}_t = \mathcal{U}_t$ the new terms is called fundamental economic shock.

This bring to contemporary uncorrelated errors which allows to estimate the equation

Computationally:

to simplify the system considering a VAR autoregressive of order 1:

$$\begin{bmatrix} y_{1t} \\ y_{2t} \end{bmatrix} = \begin{bmatrix} \alpha_1 \\ \alpha_2 \end{bmatrix} + \begin{bmatrix} \varphi_{11} & \varphi_{12} \\ \varphi_{21} & \varphi_{22} \end{bmatrix} \begin{bmatrix} y_{1t-1} \\ y_{2t-1} \end{bmatrix} + \begin{bmatrix} \varepsilon_{1t} \\ \varepsilon_{2t} \end{bmatrix} \tag{1.10}$$

With

$$\begin{bmatrix} \mathcal{E}_{1t} \\ \mathcal{E}_{2t} \end{bmatrix} \sim N \left[\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \sigma_{11} & \sigma_{12} \\ \sigma_{21} & \sigma_{22} \end{pmatrix} \right] \tag{1.11}$$

Being σ_{12} and σ_{21} different from zero, the residuals so different from 0, structural shocks cannot be taken in consideration, with respect to which to calculate the impulse response functions.

The matrix A_0 since is a lower Cholesky decomposition, is a triangular matrix.

$A_0 = \begin{bmatrix} \delta_{11} & 0 \\ \delta_{21} & \delta_{22} \end{bmatrix}$ and we are going to multiply as implied before our system by this matrix.

With u_{1t} obtained assuming that residues from the first equation of the VAR coincide with the structural innovations, while u_{2t} is obtained as the residue of an OLS regression of ε_{2t} on u_{1t} and is therefore, by construction, orthogonal to u_{2t}.

To derive the impulse response functions considering the following autoregressive representation:

$$A_0 \begin{bmatrix} y_{1t} \\ y_{2t} \end{bmatrix} = A_0 \begin{bmatrix} \alpha_1 \\ \alpha_2 \end{bmatrix} + A_0 \begin{bmatrix} \varphi_{11} & \varphi_{12} \\ \varphi_{21} & \varphi_{22} \end{bmatrix} \begin{bmatrix} y_{1t-1} \\ y_{2t-1} \end{bmatrix} + A_0 \begin{bmatrix} \varepsilon_{1t} \\ \varepsilon_{2t} \end{bmatrix} \qquad (1.12)$$

$$\begin{bmatrix} u_{1t} \\ u_{2t} \end{bmatrix} \sim N \left[\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \right] \qquad (1.13)$$

As mentioned before the uncorrelated errors allows to estimate equation by equation and the model represents a regression that take in consideration the co-movements between the two series and try to explain if the change in the first variable cause movements on the second.

We have to point out some limitations of this structuralization, due to The fact that triangular structural model imposes the recursive causal ordering (y_1 , y_2,· · · y_p) . The ordering means that the contemporaneous values of the variables to the left affect the contemporaneous values of the variables to the right but not vice-versa. For example, the ordering (y_1, y_2 y_3) imposes the restrictions: y_{1t} affects y_{2t} and y_{3t} but y_{2t} and y_{3t}do not affect y_{1t}; y_{2t} affects y_{3t} but y_{3t} does not affect y_{2t} For a VAR(p) with n variables there are n! possible recursive causal orderings. Which ordering to use in practice depends on the context and what theory can be used to justify a particular ordering.

This hypothesis has strong implications from a statistical point of view. We make the impulse response functions dependent on the ordering of the variables in the VAR. A SVAR model can be used to identify shocks imponing restriction the coefficient (Amisano and Giannini 2012).

CHAPTER II

A future contract is a legally binding agreement for the delivery of a commodity in the future at an agreed upon price. The contracts are standardized by the CME, and regulated by the Commodity Futures Trading Commission, as to quantity, quality, time and place of delivery. Only the price is variable

The value traded is given by the equation: [future price*contract size]=value traded.

The excess returns of an investment on a non rolling day could be express like this:

$$\mathbf{R_{i,t}} = \mathbf{ln}\left(\boldsymbol{F_{i,t,T_1}}\right) - \mathbf{ln}\left(\boldsymbol{F_{i,t-1,T_1}}\right) \tag{2.1}$$

An important feature when we speak about commodity futures pricing are the fundamentals of backwardation and contango used to define the position and the shape of the futures curve.

When the futures price of a commodity is expected to appreciate coming closer to maturity we speak about Backwardation or “inverted market” (longer-term futures prices are lower than short-term futures prices)

$$\boldsymbol{S_0} > \boldsymbol{F_1} > \boldsymbol{F_2} > \cdots \tag{2.2}$$

In Contango or “carrying charge market” instead we expect a drop in the future price, considered by many investors to be ‘normal’ (longer-term futures prices are higher than short-term futures price)

$$\boldsymbol{S_0} < \boldsymbol{F_1} < \boldsymbol{F_2} < \cdots \tag{2.3}$$

As clearly explained by (Miffre, 2013) this characteristics of the forward curve are influenced by hedging pressure and theory of the storage.

The hedging pressure hypothesis cause backwardation when commodity producers are more predisposed to hedge than commodity consumers and processors,

That can be called convenience yield, a premium paid by the market for physical asset in an environment of uncertain future, Contango arises in the opposite case, when consumers and processors of a commodity outnumber producers.

The theory of storage influences the curve slope looking at the incentive or

disincentive to hold inventories and thus spot commodity.

When inventories are high, commodity futures markets are contangoed, representing a cost to the investor and giving an incentive to the inventory holders to buy the spot commodity at a cheaper price. When inventories are low, commodity futures markets are backwardated.

The formula below contains both the storage(u) and the convenience yield(y).

$$F_0 = S_0 e^{(r+u-y)T} \tag{2.4}$$

Another important feature to take into consideration when Analyzing Future time series are the Roll Dates. Expected dates on which the composition of the Benchmark Futures Contract is changed or "rolled" by selling the near month contract and buying the next month contract.

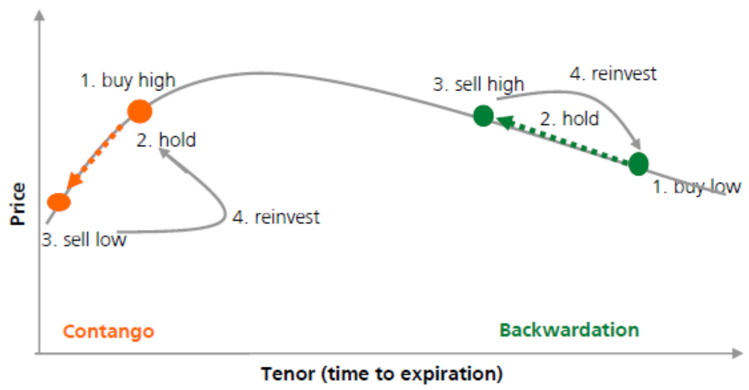
When a first-month contract matures and the second-month contract becomes the first-month contract, a commodity index specifies the “roll period”, the change occurs over five days, replacing the current contract in the index with a following contract. In this way, commodity indices provide returns comparable to passive long positions in listed commodity futures contracts.

This “rolling methods” actually lead to a phenomena returns that consists in the roll returns that investor have to face regularly.

$\boldsymbol{Roll\ Returns} = \ln(F_{t,1}) - \ln(F_{t,2}) < 0$ in Contango

$\boldsymbol{Roll\ Returns} = \ln(F_{t,1}) - \ln(F_{t,2}) > 0$ in Backwardation

The Table(2) below show the term structure of commodity futures prices in Contango (upward-sloping) and the term structure of commodity futures prices in Backwardation (downward-sloping)



2.1 Futures Sectorial

To analyze the time series it has been used sector specific returns referring our sample to 7 time series for different future commodities sector.
The Commodities Future Sectors Returns are the First Nearby daily returns, «First-nearby returns are log holding period returns, which equal the sum of the first-nearby roll return at the beginning of the holding period and the spot return of the first-nearby contract over the holding period, i.e., in between two roll dates».[5]
The time series used for this study goes over a period common to all sectors from 1979 the year in which Oil was introduced to 2015. Analysis of the distribution in Table (3).
Below the component of each sector.

Energy

Crude Oil(CL), Gasoline(HU), Heating Oil(HO)

Livestock

Feeder Cattle (FC), Live Cattle(LC), Lean Hogs(LH)

Agriculture 3

Coffe(KC), Orange Juice(JO) , Cocoa(CC)

Agriculture 4

Soybeans oil(BO), Soybeans meal(SM) , Soybeans(S)

Agriculture 5

Corn(C), Oats(O), wheat(W)

Agriculture 6

Cotton(CT), Lumber(LB)

Metals

Gold(GC), Silver(SI), Copper(HG)

2.2 First Generation Indices

A commodity index functions like an equity index, such as the S&P 500, in that

[5] Boons, Martijn, and Melissa Porras Prado. "Basis-Momentum in the Futures Curve and Volatility Risk."(2015)

its value is derived from the total value of a specified basket of commodities. Each commodity in the basket is assigned a specified weight. The weights are based on TDVT (Total Dollar Value Traded) calculated by using a simple average of the Total Dollar Value Traded for each individual commodity for the last five years within the Dow Jones Commodity Index.

The First generation Indices are Front month futures only, all the traditional indices are positioned on the same part of the forward curve, this broad-based commodity indexes representing unleveraged, long only positions in a basket of exchange-traded futures on commodities.

By far the largest two indices by market share are the S&P Goldman Sachs Commodity Index (GSCI) and the Dow-Jones UBS(DJCI).

Although SPGSCI it has been the first commodity Index launch back in 1991, Dow Jones UBS Commodity Index coincides with a period with a higher inflow in index tracking for this reasons I used DJ-UBSCI.
Dow Jones-UBS Commodity Index has been launched in 1999.
Commodity indices have been a popular vehicle for investors during the commodity markets expansion phase (2002-2005)[6]. However, the changing nature of the shape of the commodity term structure across the entire commodity world has meant that traditional commodity indices, which replicate front-month future contracts, have started experiencing a negative roll yield since 2005.
Observing that, first generation indices suffer from the pitfall of assuming that commodity futures markets are solely in backwardation. This situation whom was affecting the performance brought to the introduction of new benchmark to accomodates investor’s needs.

6 UBS Bloomberg CMCI Index (2014).

2.3 Second generation Indices

Long-only second generation indices, which attempt to minimize the harmful impact of contango on performance and use active long-only signals based on momentum or roll-yields, are found to outperform their first generation counterparts[7], due to also an Asymmetry at the Near End of the Curve (backwardated curves tend to be more linear
and contango markets less linear). Recognizing the term structure of commodity, the rolling mechanism of these indexes aimed at avoiding roll risks.

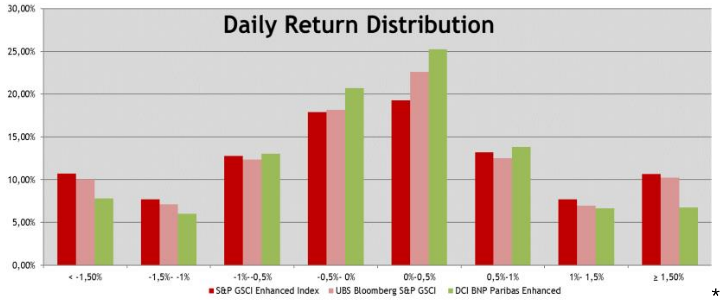
The more common strategy are:

Enhanced roll: Invest relatively liquid contract located in the mid to far end of the futures (S&P GSCI Enhanced Index).

Constant maturity: indices invest in a number of contract months across the futures curve, in order to achieve a targeted maturity. (UBS Bloomberg Constant Maturity Commodity Index).

Implied roll yield: Select the contracts with the maximum implied roll yield. (DCI BNP Paribas Enhanced Index).

Table(4) Daily Return Distribution for Second Generation Indices:



* Bloomberg data

To test my time series it will be used the UBS Bloomberg CMCI not only for being a benchmark in this market but also because as shown below allows me to have more comparable results due to a similar distribution compared to DJ-UBSCI and both DJ-UBSCI and UBS CMCI are not over weighted in energy commodity which is useful from a statistical point of view to avoid misleading results.

Table(5) Comparison Weights DJ-UBSCI and UBS CMCI:

7 Miffre, Joëlle. "Comparing first, second and third generation commodity indices." (2012).

		Dow Jones-UBS Commodity Index 2014		UBS Bloomberg CMCI 2014	
Commodity	Sector	Weight	Overall weight	Weight	Overall weight
WTI Light Crude Oil	Energy	31,8%	8,49%	38,2%	10,39%
Brent Crude Oil			6,51%		9,61%
Heating oil			3,72%		4,22%
RBOB Gasoline			3,62%		5,06%
Natural Gas			9,45%		8,89%
Copper*	Metals*	32,3%	7,51%	29,8%	12,37%
Zinc			2,31%		2,25%
Aluminium			4,72%		6,11%
Nickel			2,05%		2,20%
Lead			-		1,39%
Gold			11,53%		4,35%
Silver			4,14%		1,08%
Wheat*	Agriculture	30,8%	4,55%	28,0%	4,04%
Corn			7,20%		5,31%
Soybeans*			11%		9,59%
Sugar			4%		6,73%
Coffee			2%		1,02%
Cotton			2%		1,26%
Live Cattle	Livestock	5,1%	3%	4,1%	2,32%
Lean Hogs			2%		1,81%

The second-generation indexes have significantly outperformed the first-generation indexes, and this outperformance has continued in the recent, post-formation period. This since their recent introduction (2007) clearly increased the buying pressure, percentage of daily traded volume and open interest on Futures Market.

2.4 Third generation Indices

Launched around 2008, third generation have an active approach with strategies based on signals on momentum, basis and term structure.

As they accurately buy backwardated assets and short contangoed ones, are able to reduce overall volatility.

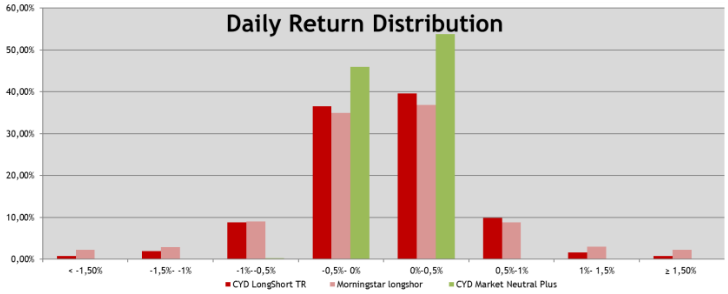
In addition they offer also Quantitative model based analysis based on the relationship between the returns of commodities futures and the state of inventories or hedging pressure position.

Moving to an always more active strategy a consideration that could be done is that from one point of view this alternative strategies do not increase pressure due to “announced” transaction typical of passive Indices.

One of the most representative Index is the Morningstar Long/Short

The Morningstar Long/Short Commodity Index is a fully collateralized commodity futures index that uses the momentum rule to determine each commodity is held long, short, or flat.

Table(6) Daily Return Distribution for Third Generation Indices:



* Bloomberg data

CHAPTER III

As previously introduced my data are comprehensive of the log return for each time series, and not the prices, for both the the Commodity Sectors and Indices. This choice will have a relevant impact on the the analysis of my data.

Time-series in many case have properties that make standard VAR analysis unsuitable:

Is important to estimate the VAR under hypothesis of Homoskedasticity and serial uncorrelated residuals.

In order to estimate a reliable VAR it is necessary to conduct an analysis on the model and verify if the residuals are uncorrelated and homoskedastic . To verify the absence of serial correlation I made the LM serial correlation test, while for the homoskedasticity of errors I used the White test.

All variables to be included in the VAR are required to be stationary in order to carry out joint significance tests on the lags of the variables.

Testing unit root, the results as I expected, the series are stationary.

We see from the augmented Dickey-Fuller function results that all the time-series are I(0) variables. We convincingly reject the augmented Dickey-Fuller null hypothesis of a unit root I(1) because our t-statistics for the series are more than the critical value at the 90% level.

Another important feature regards co-integration used for the case of two variables where we wish to test whether there is as an equilibrium relationship between the two series

I applied co-integrating augmented Dickey-Fuller test the series taken two at the time.

From the results of cadf we find that the time series analyzed two at a time are not cointegrated, again because the t-statistic does exceed the 90% critical value (in absolute value terms). We would conclude that an VEC model is not appropriate for these time-series. Results in Table(7) Table(8) Table(9) Table(10).

An important preliminary step in impulse response analysis is the selection of the

VAR lag order. impulse responses may depend critically on the lag order of the VAR model fitted to the data. These differences can be large enough to affect the substantive interpretation of VAR impulse response estimates[8]
Considering the big sample size for each series and the daily returns a simple and intuitive approach as Likelihood ratio tests is suitable for our study (LR test have been used by Sims (1980)).

As previously explained, Cholesky decomposition is one method of identifying the impulse– response functions in a VAR; thus, this method corresponds to a SVAR. There are several sets of constraints on $A_0$and B that are easily manipulated back to the Cholesky decomposition, and thus it makes the impulse response functions dependent on the ordering of the variables.
The order I choose is due to the fact that my goal is to obtain an index which contemporaneously affects the other variables.
A result that would be expectable is that the impulse responses are zero if one of the variables does not Granger-cause the other variables taken as a group.

3.1 Model estimation: Subsample1 (07/03/1979-05/01/1999)

The first subsample goes over a period common to all the commodities, since the oil has been introduced in 1979 and before the first Generation indices introduction.
In this period the flows of the institutional investors in terms of index tracking and other exchange traded products was not happened yet (the so called financialization). I do analyze this period to extract lead lag correlation between the sectors, without the buying pressure.
The LR test for our sample shows the following statistics. Table(11) p-value:

[8] Ivanov, Ventzislav, and Lutz Kilian. "A practitioner's guide to lag order selection for VAR impulse response analysis." (2005)

nlag	LR statistic	Probability
12 - 11	39.7587	0.8241
11 - 10	56.6619	0.2108
10 - 9	47.8137	0.5212
9 - 8	48.0320	0.5123
8 - 7	56.0752	0.2268
7 - 6	62.5548	0.0924
6 -5	53.1447	0.3176
5 - 4	50.4383	0.4164
4 - 3	**92.4024**	**0.0001762**
3 - 2	68.8967	0.03189

The results suggest that the lag length of 3 cannot be viewed as significantly degrading the fit of the model relative to a lag of 4,we find that at the 0.05 level. Also if we employ a 0.001 level of significance, we would conclude that the optimal lag length is 4, because the likelihood ratio tests reject lag 3 as significantly degrading the fit of the model at the 0.001 level of confidence

As explained in Chapter I, to extract the lead lag correlation using our VAR model, we focused on the Granger Causality test result.

The format of the output is such that the columns reflect the Granger causal impact of the column-variable on the row-variable. The table shows the p-value of the results with a cut-off at 0.05.

Table(12) p-value for Granger Causality Tests:

Variable	Energy	Livestock	Agri3	Agri4	Agri5	Agri6	Metals
Energy	0.00	NaN	NaN	NaN	NaN	NaN	NaN
Livestock	NaN	0.00	NaN	NaN	NaN	NaN	NaN
Agri3	NaN	NaN	0.00	NaN	NaN	NaN	NaN
Agri4	NaN	NaN	NaN	0.00	0.04	NaN	NaN
Agri5	NaN	NaN	NaN	0.00	0.00	NaN	NaN
Agri6	NaN	NaN	NaN	NaN	NaN	0.00	NaN
Metals	NaN	NaN	NaN	NaN	NaN	NaN	0.00

As expected, before the financialization there are no sectors that Granger Cause any other one, except for an interaction between Agri4 and Agri5 sectors.

3.2 Model estimation: Subsample 2 (20/01/1999-03/01/2007)

The LR test for our sample shows the following statistics.

Table(13) p-value:

nlag	LR statistic	Probability
12 - 11	60.4751	0.126
11 - 10	67.6164	0.04008
10 - 9	68.9772	0.03143
9 - 8	62.0087	0.1004
8 - 7	54.1638	0.2839
7 - 6	57.6570	0.1856
6 -5	32.9330	0.9621
5 - 4	**79.0159**	**0.004219**
4 - 3	51.1833	0.388
3 - 2	78.3449	0.004879

we find that at the 0.05 level, we might reject lag 10 in favour of lag11 as the optimal lag length. On the other hand, if we employ a 0.01 level of significance, we would conclude that the optimal lag length is 5, because the likelihood ratio tests reject lag 4 as significantly degrading the fit of the model at the 0.01 level of confidence

Granger Causal impact below.

Table(14) p-value for Granger Causality Tests:

Variable	Energy	Livestock	Agri3	Agri4	Agri5	Agri6	Metals
Energy	NaN	NaN	NaN	NaN	NaN	NaN	NaN
Livestock	NaN	NaN	NaN	NaN	NaN	NaN	0.00
Agri3	NaN	NaN	0.01	NaN	0.00	NaN	NaN
Agri4	NaN	NaN	NaN	NaN	0.01	NaN	NaN
Agri5	NaN	NaN	0.02	NaN	NaN	NaN	NaN
Agri6	NaN	NaN	NaN	NaN	0.01	0.00	0.02
Metals	NaN	NaN	NaN	NaN	NaN	NaN	0.00

It's easily observable how the past value do not Granger Cause anymore the actual value of the same time series implying a less "trend following" behavior. In addition Agri5, which includes wheat corn and oats the more traded in the Agri sector, leads the other Agri Commodities.

For the structural estimation I select a sample that comprise the recent years in order to evaluate the interactions between the Index and the Sectors, this differs from the sample used in the causation analysis which actually has to be compared with the subsample3

3.2.1 Subsample 2a (20/01/1999-27/02/2015)

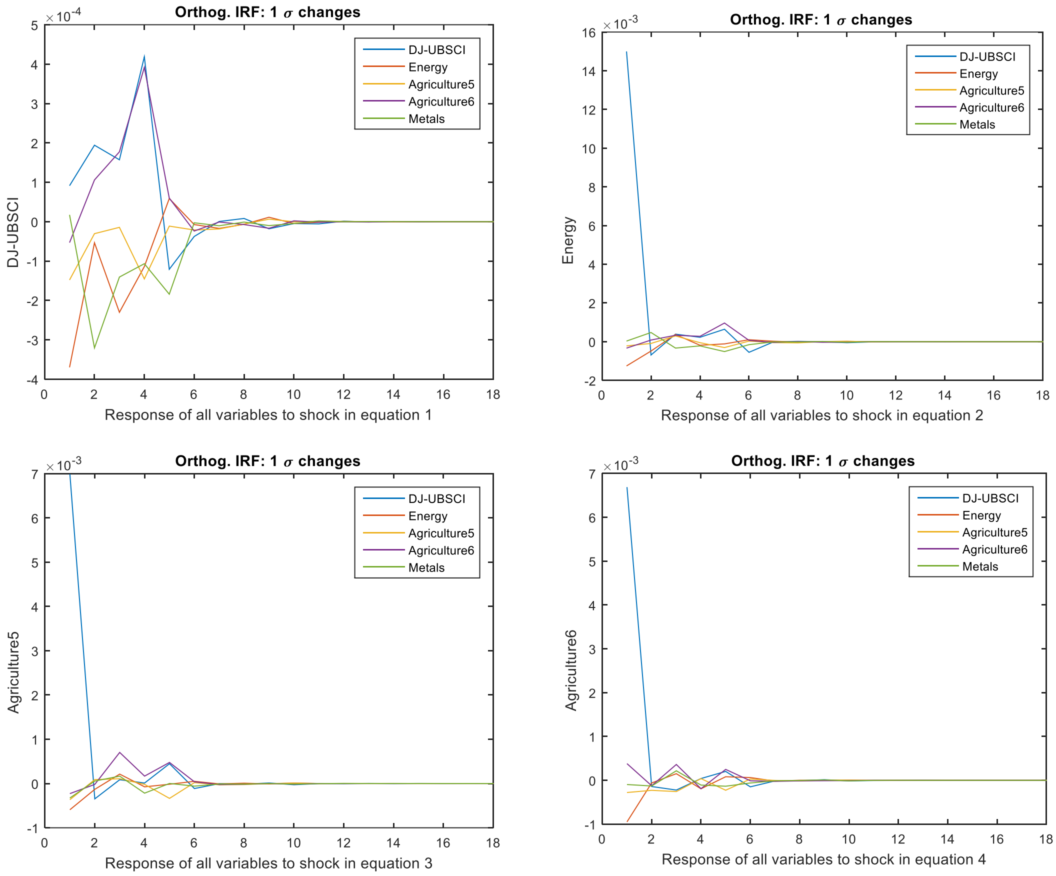
In this example, yt = (DJ-UBSCI, energy, agri5, agri6, metals),
We will impose the Cholesky restrictions on this system by applying equality constraints with the constraint matrices.
With these structural restrictions, we assume that the percentage change in DJ-UBSCI is not contemporaneously affected by the percentage changes in any of the selected sectors. We assume that the log returns of Energy is affected by contemporaneous changes in DJ-UBSCI but not the other three sector. We assume also that percentage changes in Agri5 are affected by contemporaneous changes in both DJ-UBSCI and Energy, but not Agri6 and Metals. Agri6 is affected by contemporaneous change in DJ-UBSCI , Energy, Agri5 but not Metals. Finally we assume that percentage change in Metals are affected by the percentage change of all the Variables.
Below Table(15) with the coefficients of the Lower Choleski Decomposition Matrix A, with the related Standard errors.

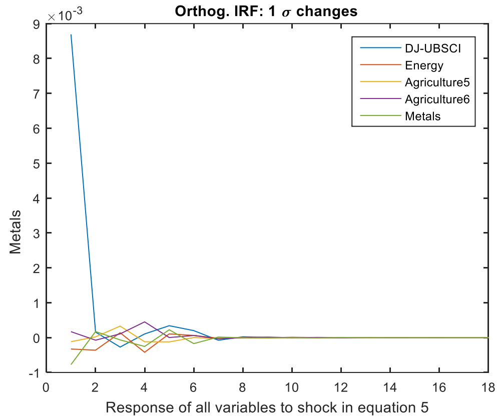
A	Djci	Energy	Agri5	Agri6	Metals	Standard errors	Djci	Energy	Agri5	Agri6	Metals
Djci	1	0	0	0	0	Djci	0	0	0	0	0
Energy	-0.0738	1	0	0	0	Energy	0.3326	0	0	0	0
Agri5	-0.0542	0.1563	1	0	0	Agri5	0.2920	0.2163	0	0	0
Agri6	-0.0067	0.1257	-0.4476	1	0	Agri6	0.2265	0.1703	0.1912	0	0
Metals	0.0075	0.1970	-0.0046	0.1137	1	Metals	0.2376	0.1815	0.2316	0.2588	0

Having estimated Cholesky decomposition for orthogonalised IRF we analyze the profile of the curves interpreting the response values in absolute terms and the speed of convergence towards equilibrium for the different series.

An important consideration is that the units are in absolute deviations from 0 trend line, not in percentage points and you can interpret the magnitude of the IRF as the size of the response per unit of the shock.

Table(16) Impulse Response Functions graphs:





As it would be expectable after the Granger Causality test results, few linkages between the series are established, the response to the shock are quite small, except for the Index which by construction is not contemporaneously affected by the other variables. In fact, the Index has a positive response to the shock since it is long only and it holds the front month futures by construction.

We can also observe that all the shock in this subsample tend to converge to zero after 6 lags. It would be an important characteristic to take in consideration when comparing to the other subsample.

Notable is in addition the similar reaction between the Index and the Agri6 sector.

3.3 Model estimation: Subsample 3 (01/01/2007-27/02/2015)

In the third subsample analysis, we focused on the period after the introduction of the second generation indexes, which is roughly also the subsequent period in which third generation was introduced. This period saw the most significant index investment for the commodities as an asset class.

The LR test for the sample shows the following statistics with an optimal lag length of 5

Table(17) p-value:

nlag	LR statistic	Probability
12 - 11	65.3348	0.05921
11 - 10	61.6693	0.1057
10 - 9	61.8489	0.1029
9 - 8	50.8866	0.3992
8 - 7	55.0408	0.2567
7 - 6	50.2291	0.4245
6 - 5	37.3816	0.8874
5 - 4	**84.0304**	**0.001363**
4 - 3	60.2733	0.1297
3 - 2	100.7393	1.931e-005

Table(18) p-value for Granger Causality Tests:

Variable	Energy	Livestock	Agri3	Agri4	Agri5	Agri6	Metals
Energy	0.05	0.01	NaN	NaN	0.01	NaN	0.00
Livestock	NaN	NaN	NaN	NaN	NaN	NaN	0.00
Agri3	NaN	NaN	NaN	NaN	0.04	NaN	NaN
Agri4	0.02	NaN	NaN	NaN	0.01	NaN	NaN
Agri5	0.02	NaN	NaN	NaN	NaN	0.01	NaN
Agri6	NaN	NaN	NaN	NaN	0.01	0.00	0.00
Metals	NaN	NaN	NaN	NaN	0.02	NaN	0.04

Analyzing again the interactions between the sectors, we can observe considerable difference in respect to both the previous period (subsample1, subsample2). We can observe how the log return of the Energy sector, the Agri5 sector and the Metals tend to Granger cause other sector.

These results are quite interesting since there is no reason why the past performance of some sectors can have any impact on the future performance of other ones.

3.3.1 Subsample 3a (01/01/2007-27/02/2015)

This subsample aim to analyze the Impulse Response Function under the assumption of orthogonal shocks. The subsample coincides with the launch in 2007 of the second generation Index UBS-Bloomberg CMCI.

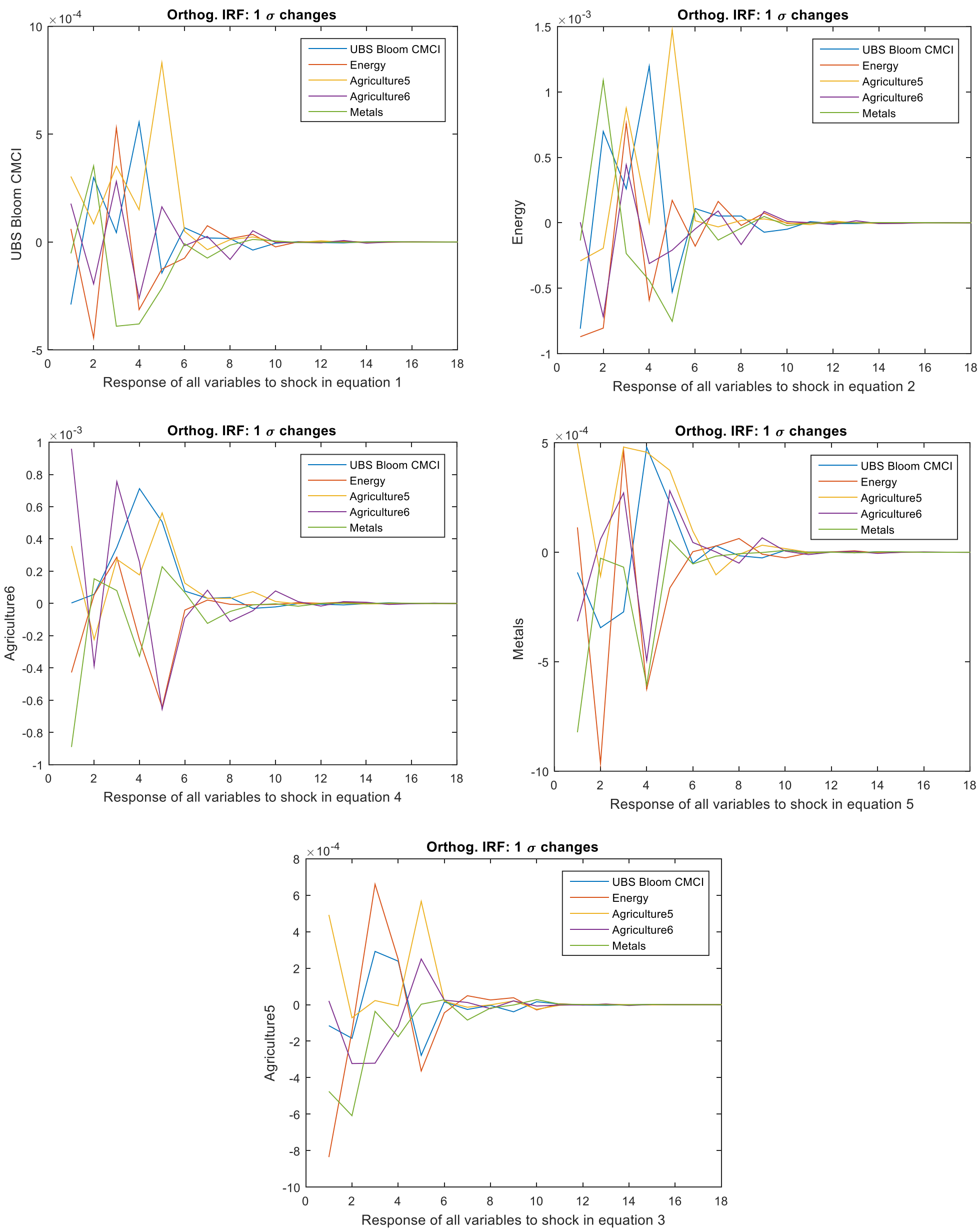
In this example, yt = (UBS-Blommberg CMCI, energy, agri5, agri6, metals).

We will impose the Cholesky restrictions on this system by applying equality constraints with the constraint matrices. Please refer to paragraph 3.2 for assumption regarding the constraints. Table(19) Cholesky Matrix coefficients and s.e.:

A	UBS Bloomberg CMCI	Energy	Agri5	Agri6	Metals	Standard errors	UBS Bloomberg CMCI	Energy	Agri5	Agri6	Metals
UBS Bloomberg CMCI	1	0	0	0	0	**UBS Bloomberg CMCI**	0	0	0	0	0
Energy	0.0404	1	0	0	0	**Energy**	0.0034	0	0	0	0
Agri5	-0.1623	-0.3332	1	0	0	**Agri5**	0.0024	0.0003	0	0	0
Agri6	-0.0385	-0.0768	-0.6016	1	0	**Agri6**	0.0019	0.0003	0.0004	0	0
Metals	-0.1187	-0.2764	-0.1825	-0.1280	1	**Metals**	0.0023	0.0004	0.0006	0.0006	0

Our orthagonalized IRF below, to notice that the absolute value of y-axis varies for the different shocks this could be misleading in the reading of the results.

Table(20) Impulse Response Functions graphs:

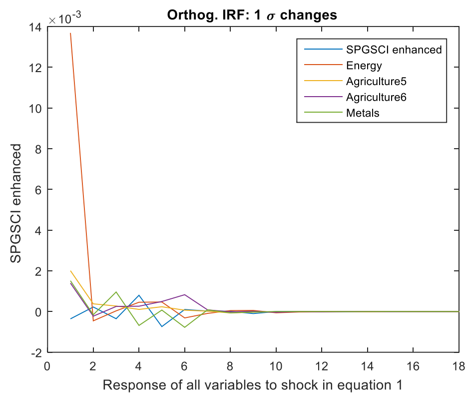


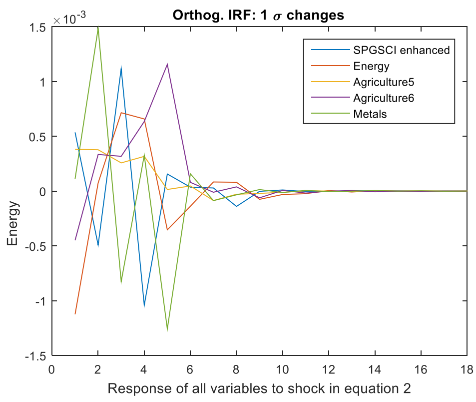
The first Notable difference from the previous subsample is that all the series tend to converge after 12 lags, double the time in which the time series of the subsample2a die down.

For the same subsample I run the same analysis for another second generation Indices (S&P GSCI Enhanced Index) which is over weighted in favor of the energy sector. Table(21)

The results as I Expected show a positive response in the energy sector to a shock in the Index equations.

Table(22) Impulse Response Functions graphs:





In both Indexes, shock for this period tend to converge after 12 lags differs from the first subsample and a unit variation shock in the energy sector is expecting to lead a 0,15% in metals.

3.3.2 Subsample 3b (01/02/2008-27/02/2015)

This subsample aim to analyze the Impulse Response Function under the assumption of orthogonal shocks.

The subsample 3b starts in 2008 considering that the third generation found the most consideration from the market after the downturn in 2008 commodities prices.

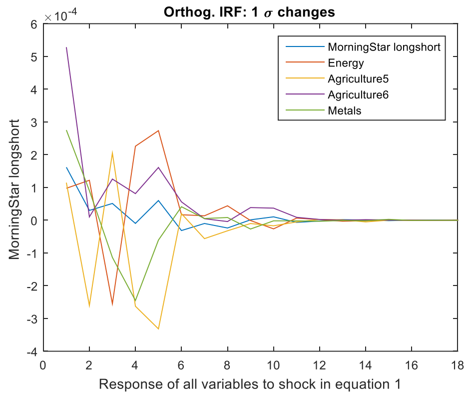
In this example, yt = (Morningstar long short, energy, agri5, agri6, metals), we will impose the Cholesky restrictions on this system by applying equality constraints with the constraint matrices. Please refer to paragraph 3.2 for assumption regarding the constraints.

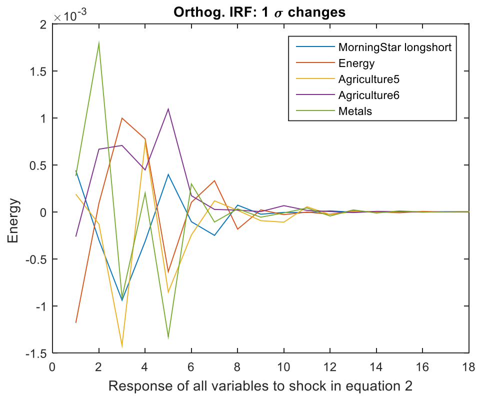
Table(23) Cholesky Matrix coefficients and s.e.:

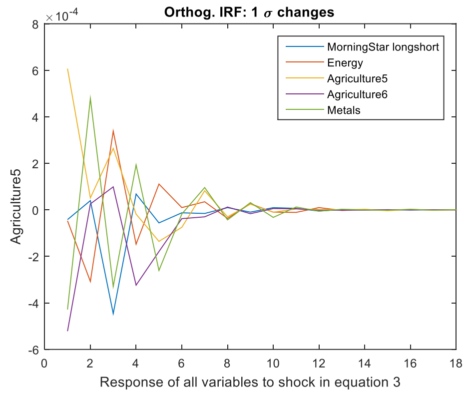
A	Morningstar long short	Energy	Agri5	Agri6	Metals	Standard errors	Morningstar long short	Energy	Agri5	Agri6	Metals
Morningstar long short	1	0	0	0	0	Morningstar long short	0	0	0	0	0
Energy	-0.5963	1	0	0	0	Energy	0.0015	0	0	0	0
Agri5	0.0865	-0.1109	1	0	0	Agri5	0.0006	0.0002	0	0	0
Agri6	-0.0006	-0.2651	-0.2026	1	0	Agri6	0.0011	0.0004	0.0010	0	0
Metals	-0.0037	-0.3162	-0.0441	-0.2314	1	Metals	0.0011	0.0004	0.0010	0.0005	0

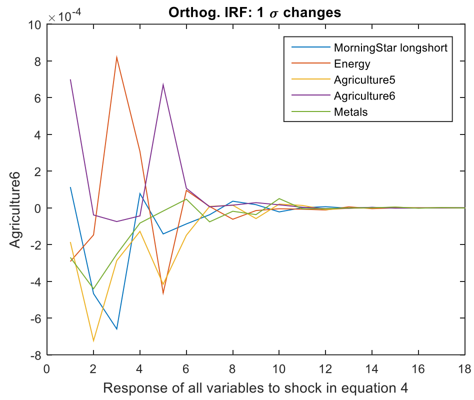
Similarly to the subsample 3a we can observe a convergence after 12 lags and a leading behaviour of the energy sector in respect to the metals one.

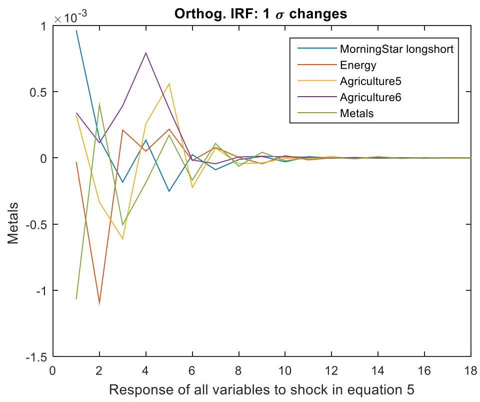
Table(24) Impulse Response Functions graphs:


Orthog. IRF: 1 σ changes
$\times 10^{-4}$
MorningStar longshort
MorningStar longshort
Energy
Agriculture5
Agriculture6
Metals
Response of all variables to shock in equation 1


Orthog. IRF: 1 σ changes
$\times 10^{-3}$
Energy
MorningStar longshort
Energy
Agriculture5
Agriculture6
Metals
Response of all variables to shock in equation 2


Orthog. IRF: 1 σ changes
$\times 10^{-4}$
Agriculture5
MorningStar longshort
Energy
Agriculture5
Agriculture6
Metals
Response of all variables to shock in equation 3


Orthog. IRF: 1 σ changes
$\times 10^{-4}$
Agriculture6
MorningStar longshort
Energy
Agriculture5
Agriculture6
Metals
Response of all variables to shock in equation 4


Orthog. IRF: 1 σ changes
$\times 10^{-3}$
Metals
MorningStar longshort
Energy
Agriculture5
Agriculture6
Metals
Response of all variables to shock in equation 5

CONCLUSIONS

Commodity price and forward curves are affected by many factors. This research does not pretend to analyze the interaction between external variables and the commodities future curves, but instead to see if the flows of passive/active investments affect the price level of the commodities in the different sectors and lead to a causality reaction between them. It is important to remember that causality simply implies chronological orders of movements of the series meanings that one variable appear to lead another on and not that a movement in one variable is considered a direct results of the other one. In addition, most traded assets tend to incorporate information into prices faster than others so they should lead, this consideration founds confirm on the Granger Causalities results of this study which show that over weighted assets like Energy Sectors tend to lead the others.

Considered so it’s hard to find a direct evidence that Index Tracking influenced the behavior of the first nearby returns, but there are signals that shows how the causality parameters changed over time.

This result can be caused by different factors, but from the described analysis it is evident how Institutional Investors tend to see commodities as a unique asset class and this increased the relationship between the variables.

This lead/lag relationship clearly increased after 2008, as distinctly evident from the Granger Causality between sector in subsample 3 and the tendency of the variables to incorporate shocks for longer lags.

I would like to bring as a support of my thesis a last feature about commodity price, Seasonality Levels for the active Contract of Corn before and after 2008. Table(25) and Table(26).

Is possible to observe in the graph how the price levels do not reflect anymore a seasonality behavior showed before 2008 and how probably the prices are not anymore a direct consequences of external factors such as macro-economic cycles, geopolitical events, inflation, weather and of course supply/demand.

Market dynamics are complex, and it is not easy to understand the interaction between the varied market participants and their motivations for trading.

But there is an emerging evidence that new players such as traders, broker-dealers and Funds are always more attracted from the development of the products in the Commodity Futures markets. This will have a key role for future understanding of the dynamics within this Market.

References

Amisano, Gianni, and Carlo Giannini. *Topics in structural VAR econometrics*. Springer Science & Business Media, 2012.

Basu, Devraj, and Joëlle Miffre. "Capturing the risk premium of commodity futures: The role of hedging pressure." *Journal of Banking & Finance* 37.7 (2013): 2652-2664.

BNP Paribas, "The DCI® BNP Paribas Enhanced (DCI®-B)" Index Methodology (April 12, 2007).

Boons, Martijn, and Melissa Porras Prado. "Basis-Momentum in the Futures Curve and Volatility Risk." *Available at SSRN 2587784* (2015).

Dunsby, Adam, and Kurt Nelson. "A Brief History Of Commodities Indexes." *Journal of Indexes* 13.3 (2010): 36-40.

Eduardo Rossi, "Impulse Response Functions" Università di Pavia.

Falkowski, Michał. "Financialization of commodities." *Contemporary Economics* 5.4 (2011): 4-17.

Gottschalk, Jan. *An introduction into the SVAR methodology: identification, interpretation and limitations of SVAR models*. No. 1072. Kiel Institute for the World Economy (IfW), 2001.

Herwartz, Helmut, and Helmut Lütkepohl. "Structural vector autoregressions with markov switching: Combining conventional with statistical identification of shocks." *Journal of Econometrics* 183.1 (2014): 104-116.

Hu, Conghui, and Wei Xiong. "The informational role of commodity futures prices." *Apres le Deluge: Finance and the Common Good after the Crisis, University of Chicago Press, http://www8. gsb. columbia. edu/rtfiles/finance/misc/Information_Commodity4_c. pdf* (2013).

Irwin, Scott H., and Dwight R. Sanders. "Financialization and structural change in commodity futures markets." *Journal of agricultural and applied economics* 44.03 (2012): 371-396.

Ivanov, Ventzislav, and Lutz Kilian. "A practitioner's guide to lag order selection for VAR impulse response analysis." *Studies in Nonlinear Dynamics & Econometrics*

9.1 (2005).

LeSage, James P. "Applied econometrics using MATLAB." *Manuscript, Dept. of Economics, University of Toronto* (1999): 154-159.

Marco Aiolfi, "A note for Svar function in Matlab" *Bocconi University and Banca Intesa, Milan*

Master, M. W., and A. K. White. "How Institutional Investors Are Driving up Food and Energy Prices." (2008).

Miffre, Joëlle. "Comparing first, second and third generation commodity indices." *Second and Third Generation Commodity Indices* (2012).

Sanders, Dwight R., and Scott H. Irwin. "Energy futures prices and commodity index investment: New evidence from firm-level position data." *Energy Economics* 46 (2014): S57-S68.

Sanders, Dwight R., and Scott H. Irwin. "Measuring index investment in commodity futures markets." *Energy Journal* 34.4 (2013): 105-27.

Sapienza Università di Roma, "Identificazione e VAR Strutturali" (2013) http://www.dipecodir.it/upload/file/Carlucci/Corso%20avanzato/Testo%2013.pdf (last view November 2016)

Scrocchia, Italo M. *Modelli strutturali in forma ridotta e modelli VAR*. No. 02-2005. Dipartimento di Scienze Economiche, Matematiche e Statistiche, Universita'di Foggia, 2005.

Seeking Alpha, "The New Generation of Diversified Commodity Indexes" http://www.seekingalpha.com/article/43586-the-new-generation-of-diversified-commodity-indexes (last view October 10, 2016)

Sims, Christopher A. "Macroeconomics and reality." *Econometrica: Journal of the Econometric Society* (1980): 1-48.

Stoll, Hans R., and Robert E. Whaley. "Commodity index investing and commodity futures prices." *Journal of Applied Finance (Formerly Financial Practice and Education)* 20.1 (2010).

Tang, Ke, and Wei Xiong. "Index Investment and Financialization of Commodities." *NBER Working Paper w16385* (2010).

UBS, “UBS Bloomberg CMCI Composite Index, State-of-the-Art Commodity Benchmark” Index Methodology (December 2012).

UBS, “UBS Bloomberg CMCI” Index Methodology (August 2014).

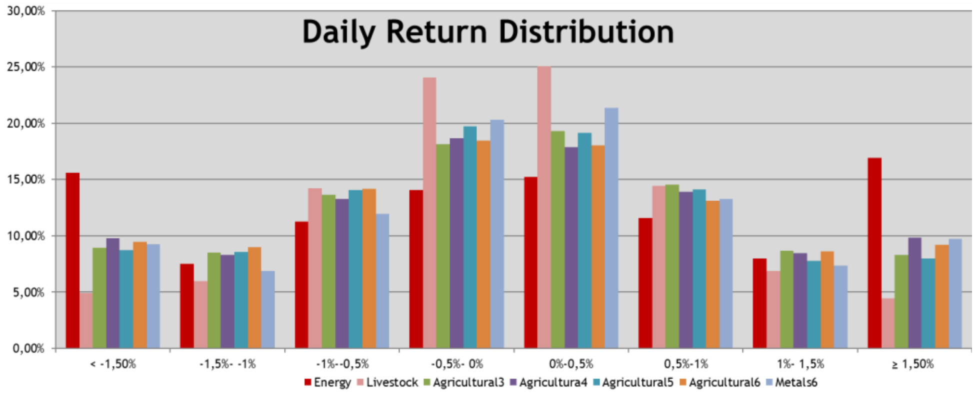
UBS, “UBS Bloomberg SPGSCI Constant Maturity, Combining the S&P GSCI composition and weighting with the innovative CMCI methodology” Index Methodology (January 2011).

UBS, “UBS Bloomberg SPGSCI Constant Maturity, CMCI with S&P GSCI components and weights” Index Methodology (March 2014).

Zivot, Eric, and Jiahui Wang. "Vector autoregressive models for multivariate time series." *Modeling Financial Time Series with S-PLUS®* (2006): 385-429.

Appendices

Table(3) First Nearby Daily Return 1979-2015:



Table(7) ADF e CADF Subsample 1:

Augmented DF test for unit root variable: energy

ADF t-statistic	# of lags	AR(1) estimate
-24.160433	8	0.004786
1% Crit Value	5% Crit Value	10% Crit Value
-3.458	-2.871	-2.594

Augmented DF test for unit root variable: agri5

ADF t-statistic	# of lags	AR(1) estimate
-21.565259	8	0.122236
1% Crit Value	5% Crit Value	10% Crit Value
-3.458	-2.871	-2.594

Augmented DF test for unit root variable: agri6

ADF t-statistic	# of lags	AR(1) estimate
-23.193125	8	0.062723
1% Crit Value	5% Crit Value	10% Crit Value
-3.458	-2.871	-2.594

Augmented DF test for unit root variable: metals

ADF t-statistic	# of lags	AR(1) estimate
-22.870175	8	0.100077
1% Crit Value	5% Crit Value	10% Crit Value
-3.458	-2.871	-2.594

Augmented DF test for co-integration variables: energy,agri5

CADF t-statistic	# of lags	AR(1) estimate
-24.17443512	8	-0.996109
1% Crit Value	5% Crit Value	10% Crit Value
-3.880	-3.359	-3.038

Augmented DF test for co-integration variables: energy,agri6

CADF t-statistic	# of lags	AR(1) estimate
-24.16601014	8	-0.994443
1% Crit Value	5% Crit Value	10% Crit Value
-3.880	-3.359	-3.038

Augmented DF test for co-integration variables: agri5 ,metals

CADF t-statistic	# of lags	AR(1) estimate
-21.40405755	8	-0.867392
1% Crit Value	5% Crit Value	10% Crit Value
-3.880	-3.359	-3.038

Augmented DF test for co-integration variables: energy,metals

CADF t-statistic	# of lags	AR(1) estimate
-24.03477768	8	-0.986337
1% Crit Value	5% Crit Value	10% Crit Value
-3.880	-3.359	-3.038

Augmented DF test for co-integration variables: agri5,agri6

CADF t-statistic	# of lags	AR(1) estimate
-21.71347809	8	-0.888096
1% Crit Value	5% Crit Value	10% Crit Value
-3.880	-3.359	-3.038

Augmented DF test for co-integration variables: agri6 ,metals

CADF t-statistic	# of lags	AR(1) estimate
-23.23957035	8	-0.946914
1% Crit Value	5% Crit Value	10% Crit Value
-3.880	-3.359	-3.038

Table(8) ADF e CADF subsample 2:

Augmented DF test for unit root variable: DJ-UBSCI

ADF t-statistic	# of lags	AR(1) estimate
-25.485411	5	0.058927
1% Crit Value	5% Crit Value	10% Crit Value
-3.458	-2.871	-2.594

Augmented DF test for co-integration variables: DJ-UBSCI,energy

CADF t-statistic	# of lags	AR(1) estimate
-25.58006446	5	-0.969225
1% Crit Value	5% Crit Value	10% Crit Value
-3.880	-3.359	-3.038

Augmented DF test for co-integration variables: DJ-UBSCI,agri5

CADF t-statistic	# of lags	AR(1) estimate
-25.55942801	5	-0.957359
1% Crit Value	5% Crit Value	10% Crit Value
-3.880	-3.359	-3.038

Augmented DF test for co-integration variables: DJ-UBSCI,agri6

CADF t-statistic	# of lags	AR(1) estimate
-25.51689803	5	-0.950002
1% Crit Value	5% Crit Value	10% Crit Value
-3.880	-3.359	-3.038

Augmented DF test for co-integration variables: DJ-UBSCI,metals

CADF t-statistic	# of lags	AR(1) estimate
-25.51689803	5	-0.950002
1% Crit Value	5% Crit Value	10% Crit Value
-3.880	-3.359	-3.038

Table(9) ADF e CADF subsample 3a:

Augmented DF test for unit root variable: UBSCMCI

ADF t-statistic	# of lags	AR(1) estimate
-18.207611	5	0.029194
1% Crit Value	5% Crit Value	10% Crit Value
-3.458	-2.871	-2.594

Augmented DF test for co-integration variables: UBSCMCI,metals

CADF t-statistic	# of lags	AR(1) estimate
-17.89004141	5	-0.960985
1% Crit Value	5% Crit Value	10% Crit Value
-3.880	-3.359	-3.038

Augmented DF test for co-integration variables: UBSCMCI,agri6

CADF t-statistic	# of lags	AR(1) estimate
-19.02485617	5	-1.025791
1% Crit Value	5% Crit Value	10% Crit Value
-3.880	-3.359	-3.038

Augmented DF test for co-integration variables: UBSCMCI,energy

CADF t-statistic	# of lags	AR(1) estimate
-19.19979511	5	-1.081375
1% Crit Value	5% Crit Value	10% Crit Value
-3.880	-3.359	-3.038

Augmented DF test for co-integration variables: UBSCMCI,agri5

CADF t-statistic	# of lags	AR(1) estimate
-19.20974993	5	-1.054053
1% Crit Value	5% Crit Value	10% Crit Value
-3.880	-3.359	-3.038

Table(10) ADF e CADF subsample 3b:

Augmented DF test for unit root variable: Morningstar

ADF t-statistic	# of lags	AR(1) estimate
-16.662828	5	0.058385
1% Crit Value	5% Crit Value	10% Crit Value
-3.458	-2.871	-2.594

Augmented DF test for co-integration variables: Morningstar,agri6

CADF t-statistic	# of lags	AR(1) estimate
-16.80226024	5	-0.952312
1% Crit Value	5% Crit Value	10% Crit Value
-3.880	-3.359	-3.038

Augmented DF test for co-integration variables: Morningstar,agri5

CADF t-statistic	# of lags	AR(1) estimate
-16.74757174	5	-0.946299
1% Crit Value	5% Crit Value	10% Crit Value
-3.880	-3.359	-3.038

Augmented DF test for co-integration variables: Morningstar,metals

CADF t-statistic	# of lags	AR(1) estimate
-16.73624699	5	-0.955663
1% Crit Value	5% Crit Value	10% Crit Value
-3.880	-3.359	-3.038

Augmented DF test for co-integration variables: Morningstar,energy

CADF t-statistic	# of lags	AR(1) estimate
-16.98648649	5	-0.964894
1% Crit Value	5% Crit Value	10% Crit Value
-3.880	-3.359	-3.038

Table(21) Weights based on TDVT for SPGSCI enhanced:

		SPGSCI enhanced Index	
Commodity	**Sector**	**Weight**	**Overall weight**
WTI Light Crude Oil	Energy	67,5%	35,04%
Brent Crude Oil			14,41%
Heating oil			4,60%
RBOB Gasoline			10,06%
Natural Gas			3,35%
Copper*	Metals*	11,6%	3,82%
Zinc			0,62%
Aluminium			2,41%
Nickel			0,81%
Lead			0,46%
Gold			3,00%
Silver			0,49%
Wheat*	Agriculture	16,5%	4,51%
Corn			4,01%
Soybeans			2,57%
Sugar			2,57%
Coffee			0,92%
Cocoa			0,30%
Cotton			1,65%
Live Cattle	Livestock	4,4%	3,05%
Lean Hogs			1,37%

Table(25) seasonality level price for Corn active contract from 1999 to 2007:

Table(26) seasonality level price for Corn active contract from 2008 to 2015: