Informational efficiency of hard commodity futures prices

LAMPROS VELENTZAS

Supervisor: Dr. Sotiris K. Staikouras, Senior Lecturer in Banking and Finance, Faculty of Finance & Financial Econometrics

Cass Business School, City University, 106 Bunhill Row, London EC1Y 8TZ

This project is submitted as part of the requirements for the award of the M.Sc Finance and Investment on Sep 2019

The objective of the current study is to examine the market efficiency between the spot and futures prices of hard commodities. Cointegration theory is employed to test the assumption that the futures prices of hard commodities are a reliable predictor of the spot prices at the end of the contract. The results are consistent with this assumption which means that markets are efficient. In this study was also investigated the presence of time-varying risk premium under the general concept of market efficiency. The test showed that such a risk premium is present in the hard commodity market.

I. INTRODUCTION

The equilibrium price of a future contract is determined on the basis that there won't be any arbitrage opportunities between spot and future prices. By arbitrage we mean that a riskless profit gained is not possible through replication or synthetic creation of one of the products. At the presence of this equilibrium between spot and futures prices there is no form of riskless profit that can be resumed in that matter. Therefore, if market prices are possible to predict and the market is efficient itself, then there will be a mathematical relationship between futures of one period (in this study we are concerned about one month) and spot prices that will be observed after the end of the same period. On the other hand, if such a relationship doesn't exist, then the futures markets is said to be inefficient because they fail to consider relevant and significant information. Trading cost and unexpected daily economic, political and financial results are another two important variables that have an impact on this information. A hypothesis has been developed over the years, one that market participants are risk neutral (indifferent of risk) and that they use efficiently, reasonable and with unbiasedness all the available information. We are going to refer to this hypothesis as the "joint hypothesis" for the rest of the study. This translates to risk premium being zero (additionally benefit by taking extra risk) and arbitrage opportunities being non-existent. In reality though, price setting involves market inefficiency, which should be tested as well. The side effect is that when there is evidence of price anomalies, it's not easy to separate between market inefficiency flaws of the model. Therefore, the market efficiency is not testable under the strict definition of the term.

The current study concentrates on the relationship between spot and futures prices of hard commodities. What is examined is whether the futures can be an unbiased predictor of the spot prices. Previous empirical studies have concentrated on commodities (Bigman et al., 1983; French, 1986; Fama and French, 1987; Serletis and Scowcroft, 1991; Antoniou and Foster, 1994), the shipping market (Kavussanos and Nomikos, 1999), the live beef cattle market (Oellermann and Farris, 1985), the petroleum market (Serletis and Banack, 1990), the stock market (Kawaller et al., 1987; Chan, 1992) and the exchange rate market (Hansen and Hodrick, 1980;Huang, 1984; Hakkio and Rush, 1989; Lai and Lai, 1991; Copeland, 1993; Staikouras, 2002). From a practical point of view, the usefulness of the present study is that it provides traders and investors with understanding of how the financial markets operate as well as the potential to enhance investments when possible. Efficient markets, where the futures prices are an unbiassed predictor of the future spot prices, provide investors with signals that were once only available to well informed traders. In that sense they make them to maintain their faith and confidence that markets are efficient.

II. METHODOLOGY AND DATA

The efficient market hypothesis states that risk premium is zero and that investors are risk neutral. Under these assumptions the futures prices should be an unbiased predictor of the spot prices in the future. In other words, there shouldn't be an arbitrage opportunity. For this to take effect, the only way is for the future contracts to predict the spot prices in the future.

By following the same methodology with Serletis and Scowcroft (Serletis and Scowcroft, 1991), where the efficiency of the prices in soft commodities was checked, we decided not to work with prices but rather with the logarithms. Let S(t) be the natural logarithm of the spot price at time t and F(t,T) be the natural logarithm of the future price at time t with delivery at time T. Let also I(t) be the information available to market participants for taking an investment decision. If the efficient market hypothesis stands, then should be:

$$F(t,T) = E[S(T)/I(t)]$$

where: E[S(T)/I(t)] is the expected spot price at time T conditional of the given relevant information I(t).

It seems that we try to establish a positive linear relationship between F(t,T) and S(T) which means that increases in F(t,T) are accompanied by increases in S(T). It would therefore be an interest to determine to what extend the relationship can be described by the following equation:

$$F(t,T) = \alpha + \beta \cdot S(T)$$

To make the model more realistic, a random disturbance term, denoted by u, is added to the equation, thus:

$$F(t,T) = a + \beta \cdot S(t,T) + u(t,T)$$

The disturbance term can capture a number of features: a) we always leave out some determinants of F(t,T), b) there may be errors in the measurement of F(t,T) that cannot be determined c) random outside influences on the F(t,T) which we cannot model.

The hypothesis that needs to be tested now is the one that $\alpha = 0$ and $\beta = 1$.

The Data

We use daily observations from the Commodity Exchange, the London Metal Exchange and the New York Mercantile Exchange on spot and generic 1st futures from six hard commodities (gold, aluminium, nickel, silver, zinc and platinum). The sample period is the last 5 years ending in early Aug 2019. The futures contracts we use were the active 1st available future contract. Specifically:

Commodity	Spot/Cash (ticker)	Futures (ticker)
Aluminium	LMAHDY	LAQ19 (Aluminium active future)
Gold	XAUUSD: Currency	GC1: Com Generic 1st "GC" Futures
Nickel	LMNIDY LME	LNA (1st active future)
Silver	XAG	SI1 (generic 1st future)
Zinc	ZSDY	LXQ9 (active future)
Platinum	XPT	PL1 (active future)

For the study we followed a wide acceptable analysis policy: We collected spot and futures prices for each commodity at the contracts inception, we produced the spot price at the futures contract expiry S(T) and then we tested various relationships between these three prices.

The way that we processed the data was to take the future price and projected one month later and then process this price with the spot/cash one. The process was similar for Platinum except that we projected in two months' time (because the first active contract was in 2 months).

Time series concepts and definitions

One of the most important concepts addressed in this study is that one of stationarity. By understanding whether a time series is stationary or not is very important, as we are going to explain later. Some processes and their definitions will follow in order to understand deeply how we are going to work with our data.
A strictly stationary process is one where, for any t_i and for T=1,2, …

$$Fy_{t_1}, y_{t_2}, ... y_{t_T} = Fy_{t_1+k}, y_{t_2+k}, ... y_{t_T+k}$$

With F denoting the distribution function of a set of variables. In other words, a series is strictly stationary if the distribution of its values remains the same as time passes, or else that the random variable y has the same probability of falling within an interval at any time in the future as in the past.
On the other hand, a weakly stationary process needs to satisfy the following three criteria:

$$E(y_t) = \mu$$

$$E(y_t - \mu)(y_t - \mu) = \sigma^2 < \infty$$

$$E(y_{t_1} - \mu)(y_{t_2} - \mu) = \gamma_{t_2 - t_1} \forall t_1, t_2$$

In other words, the stationary process needs to have a constant mean μ, a constant variance σ^2 and a constant autocovariance.
A very important concept that we are going to address a lot in this paper is the white noise. By white noise in econometrics we refer to a process with no discernible structure. The definition of a white noise is:

$$E(y_t) = \mu$$

$$\text{var}(y_t) = \sigma^2$$

$$\gamma_{t-r} = \begin{Bmatrix} \sigma^2 & if & t = r \\ 0 & if & t \neq r \end{Bmatrix} \forall t_1, t_2$$

In other words, a white noise process has constant mean, constant variance and zero autocovariances. For clarification of the autocovariance part, it means that each observation in uncorrelated with all other values in the sequence.
There are several reasons why the concept of non-stationarity is important and why non-stationary data should be treated differently from stationary ones. We need to check for stationarity because: a) stationarity of a series can strongly influence its behaviour and properties b) use of non-stationary data can lead to spurious regression c) if the variables are not stationary then the standard assumptions for asymptotic analysis will not be valid.
We have two models in general that we use to characterise a time series as a non-stationary process, the random walk and the trend stationary process.
A random walk series is a series which has the following characteristics:

$$E(Y_t) \neq \mu$$
$$Var(Y_t) \neq \sigma^2$$
$$Co\,var(Y_t, Y_{t-i}) \neq \mathrm{c}$$

which essentially means the variable is not stationary (non-stationary). A random walk series can be represented as:

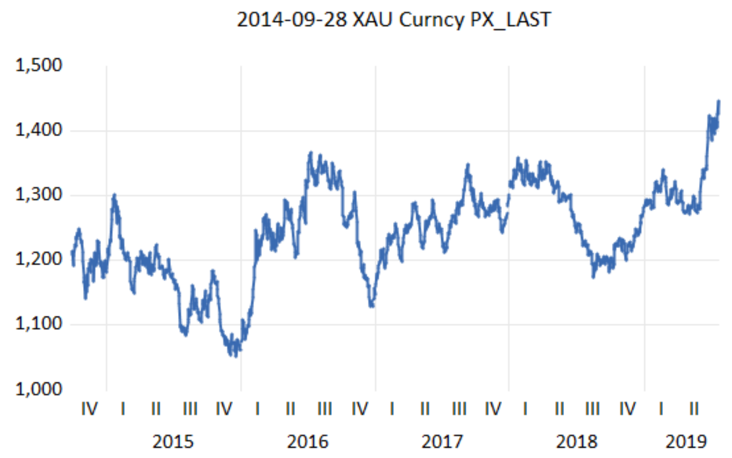
$$Y_t = Y_{t-1} + \varepsilon_t \quad \varepsilon_t \sim iid(0, \sigma^2)$$

The trend stationary process is a stationary process around a linear trend

$$y_t = a + \beta t + u_t$$

In this study we are mainly concerned with the random walk process.

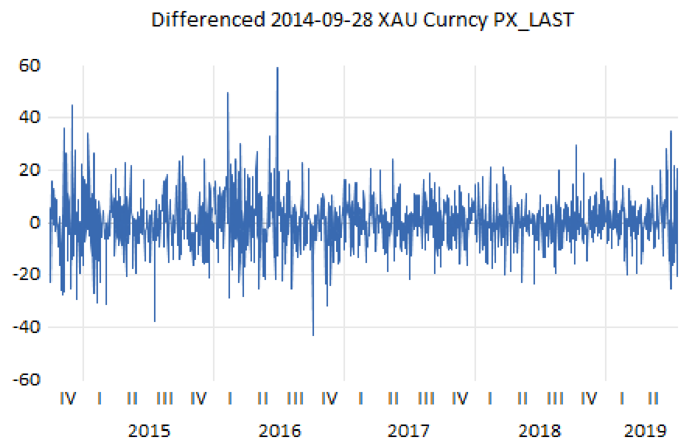
Time series plot for GOLD (used as an example) shows that the variable seems to be wondering around. This is an indication that maybe gold prices follow a random walk process. By assessing the autocorrelation function, it can be seen that the coefficients of the autocorrelation are all close to one and significant. Furthermore, from the correlogram of the series it can also be seen that the coefficients of autocorrelation die out very slowly, meaning that the series have a long memory.



Date: 08/10/19 Time: 18:22
Sample: 9/30/2014 7/19/2019
Included observations: 1248

Autocorrelation	Partial Correlation		AC	PAC	Q-Stat	Prob
		1	0.988	0.988	1221.2	0.000
		2	0.976	0.004	2414.6	0.000
		3	0.965	0.010	3581.3	0.000
		4	0.953	-0.023	4720.6	0.000
		5	0.941	-0.008	5832.8	0.000
		6	0.929	-0.040	6916.2	0.000
		7	0.916	-0.005	7971.4	0.000
		8	0.904	-0.012	8998.5	0.000
		9	0.891	-0.019	9997.0	0.000
		10	0.879	0.058	10971.	0.000
		11	0.868	0.018	11922.	0.000
		12	0.857	-0.016	12849.	0.000
		13	0.845	-0.024	13751.	0.000
		14	0.834	-0.016	14630.	0.000
		15	0.823	0.018	15486.	0.000
		16	0.811	-0.041	16318.	0.000
		17	0.799	0.003	17127.	0.000
		18	0.788	0.010	17913.	0.000
		19	0.775	-0.038	18677.	0.000
		20	0.763	-0.017	19416.	0.000
		21	0.751	0.028	20134.	0.000
		22	0.741	0.046	20832.	0.000
		23	0.733	0.082	21516.	0.000
		24	0.725	0.020	22185.	0.000

On the other hand, when we study the autocorrelation function of first differences, we find that the series vary around a constant mean, which indicate that the series are highly mean reverting. Also looking at the autocorrelation function and correlogram of the returns, we note that the series are not correlated. This in turn indicates that the return on gold is a random variable while gold prices is a random walk process.



Date: 08/10/19 Time: 18:32
Sample: 9/30/2014 7/19/2019
Included observations: 1247

Autocorrelation	Partial Correlation		AC	PAC	Q-Stat	Prob
		1	-0.043	-0.043	2.3442	0.126
		2	0.009	0.008	2.4569	0.293
		3	0.046	0.047	5.1372	0.162
		4	-0.006	-0.002	5.1778	0.270
		5	0.040	0.039	7.1865	0.207
		6	0.019	0.020	7.6323	0.266
		7	-0.017	-0.015	7.9744	0.335
		8	0.049	0.044	10.956	0.204
		9	-0.077	-0.075	18.478	0.030
		10	-0.035	-0.043	20.034	0.029
		11	-0.006	-0.014	20.086	0.044
		12	0.015	0.023	20.383	0.060
		13	0.011	0.014	20.544	0.082
		14	0.017	0.023	20.923	0.104
		15	0.006	0.014	20.969	0.138
		16	-0.014	-0.016	21.206	0.171
		17	-0.026	-0.025	22.044	0.183
		18	0.026	0.019	22.916	0.194
		19	0.019	0.016	23.372	0.221
		20	-0.014	-0.017	23.622	0.259
		21	-0.056	-0.058	27.630	0.151
		22	-0.074	-0.078	34.567	0.043
		23	-0.007	-0.009	34.625	0.057

If a non-stationary series (like the prices of GOLD earlier) must be differenced d times in order to become stationary, it is said that is integrated in order d and we write this as $y_t \sim I(d)$. What this means is that if we differentiate the series d times, the series will become I(0) or else stationary (equivalent to no unit roots). So, a I(0) series is a stationary process while a I(1) series contains one unit root. If a series is I(2) then it contains 2 unit roots and needs to be differenced twice to induce stationarity. Most of the financial data and financial series contain one unit root, which effectively means that almost all prices follow a random walk process but their returns (which is the prices differenced in one order) are stationary. As we are going to show later, all our commodity prices are integrated in the first order.

Unit root test

In order to test for unit roots in this study we use the traditional hypothesis testing. The fist work that was done in that context was from Dickey and Fuller with the basic objective to examine the null hypothesis that φ=1 in the following equation:

$$y_t = \varphi y_{t-1} + u_t$$

It said that the hypothesis testing is H_0: the series contains a unit root and H_1 : the series doesn't have a unit root (therefore is stationary)

Many researchers under the assumption that that the log levels of the variables are not stationary would carry out the analysis in terms of first differences on the variables. However, the first differencing practice to induce stationarity has been questioned by Engle and Granger (1987). They argued that this approach disregards potentially important equilibrium relationships among the levels of the series. They argued that if a set of variables are stationary in their first differences and they do not cointegrate, then the only valid relationship that can exist between them is in terms of their first differences. If however, they do cointegrate, then the model should be estimated in levels. Basically, what the above means is that modelling in first differences is only valid when the variables do not cointegrate, which is not the case here as we are going to demonstrate later.

Dickey-Fuller and Augmented Dickey-Fuller tests

Almost all the unit root tests in the literature are based on examining whether the series behave as a random walk. By definition, a random walk is a data generating process in which the current value of the series is equal to the value of the series last period plus an identically and independently distributed error term with zero mean and constant variance, $u_t \sim$ iid (0,σ).

$$y_t = y_{t-1} + u_t$$

The early work on testing for a unit root in time series was done by Dickey and Fuller (Fuller, 1976; Dickey and Fuller, 1979). The basic objective of the test is to examine the null hypothesis that φ = 1 in:

$$y_t = \varphi y_{t-1} + u_t$$

In practice the following regression is employed for ease of computation and interpretation:

$$\Delta y_t = \psi y_{t-1} + u_t \ (1)$$

so that a test of φ = 1 is equivalent to a test of ψ = 0 (and if ψ = 0 means that there is a unit root or else the data follow a RW process). If ψ is not significantly different from zero, then the series follows a random walk and is nonstationary (or else, has a unit root). On the contrary, if ψ is significantly different from zero, then the series is stationary. This means that the future value of the series will depend on their current value with a coefficient less than one, therefore the moments of the series are constant over time. Indeed, if the series in the above equation are not stationary then the statistical distributions of critical values will be different from the conventional ones. These critical values were calculated through Monte-Carlo simulations by Dickey and Fuller (1979).

A crucial point in using the Dickey-Fuller or any other unit root test is the specification of the regression model and inclusion of deterministic components. This is important since such deterministic components change the distributional properties of the unit root tests. This is because the data generating process assumed for unit root test in equation (1) is too restrictive, in fact, the model assumes that the mean of the dependent variable is zero, i.e. the series does not drift. It is also assumed that there is no trend in the data generating process. Dickey and Fuller (1981) relax these constrains by including different deterministic components in the model and reproducing the critical values for the more general model in equation (2)

$$\Delta y_t = \mu + \gamma t + \psi y_{t-1} + u_t \quad u_t = iid(0, \sigma^2) \quad (2)$$

Dickey and Fuller (1981) also suggest that testing the joint hypothesis that $\mu = \gamma = 0$ and $\psi = 0$ and $\mu = 0$ and $\psi=0$, through non-standard F test, Φ, for which they report the critical values, can be useful.

A problem that arises here is that when the true data generating process is not known, how one should decide to include deterministic terms in order to perform the test. Perron (1988) proposes a sequential testing approach, which starts with the most general test; that is, a model with intercept and trend. Then, insignificant terms are dropped one by one, using appropriate testing procedure shown in (2) and the critical values from (2), until the final model is obtained.

Clearly most of the economic series are not generated by a simple first order Autoregressive, AR(1), process (they might be generated by more complicated AR(p) or ARMA(p,q) processes) and using regression equation (2) for unit root tests will result in autocorrelated error term, while in the DF test the error terms are assumed to be identically and independently distributed with zero means and constant variance. This is a very strong assumption, which does not hold in most cases. There are two ways proposed in the literature to modify the standard Dickey-Fuller test. The first approach, suggested by Dickey and Fuller (1979) and Said and Dickey (1984), is a parametric approach which augments the test in order to make the residuals white noise, iid(0,σ). The second method, which is proposed by Philips and Perron (1988) is to apply some form of nonparametric corrections to the test statistics from equation (2).

In the former approach, Dickey and Fuller (1981) augmented the regression equation (2) by adding lagged dependent variables to the right hand side of the equation to obtain the Augmented Dickey-Fuller test (3)

$$\Delta y_t = \mu + \gamma t + \psi y_{t-1} + \sum_{\iota=1}^{k} \delta_\iota \Delta y_{t-i} + u_t \quad u_t = iid(0, \sigma^2) \qquad (3)$$

In addition, if the error terms in the DF test show a tendency to follow an ARMA(p,q) process, as Said and Dickey (1984) noticed, this type of augmentation (ADF) can be made in order to obtain white noise error terms. They suggested that the number of lag augmentation, k, should be large enough to whiten the errors.

Given the limitations of the Phillips - Perron approach, we always buttress it with the DF and the ADF test, and the results from these tests are presented in the following table (Table 1).

Table 1. *ADF test for stationarity on spot and futures prices*

Commodity	Observations	Mean	Standard deviation	ADF t-stat	MacKinnon (1996) one-sided p-values
Aluminium					
S(t)	1256	7.514205	0.120899	-2.078829	0.5566
F(t,T)		7.582172	0.097681	-1.778699	0.7147
Gold					
S(t)	1248	7.122806	0.059146	-2.690728	0.2406
F(t,T)		7.122544	0.059188	-2.769761	0.2089
Nickel					
S(t)	1256	9.360158	0.195954	-2.182366	0.4985
F(t,T)		9.39265	0.193143	-2.281711	0.4431
Platinum					
S(t)	1217	6.864012	0.120313	-3.001991	0.1320
F(t,T)		6.881997	0.13488	-3.512523	0.0384
Silver					
S(t)	1237	2.787836	0.082271	-3.013260	0.1289
F(t,T)		2.789758	0.085793	-3.343236	0.0599
Zinc					
S(t)	1256	7.790144	0.21116	-1.430589	0.8518
F(t,T)		7.765557	0.182507	-1.693416	0.7539

Notes:
test critical values

t-stat(1%)	t-stat(5%)	t-stat(10%)
-3.965488	-3.413451	-3.128767

On the basis of the ADF test, we are unable to reject the null of a unit root in any series. Therefore, all our data for spot and future prices follow a random walk process.
Since a unit root has been confirmed for all six commodities (for both spot and futures prices), the question is whether there exists some long run equilibrium relationship between them. Thus, we estimated the cointegration regression in the following equation:

$$S(T) = \hat{a} + \beta F(t,T) + \hat{u}(t,T) \Rightarrow$$

$$\hat{u}(t,T) = S(T) - \hat{a} - \beta F(t,T) \qquad (4)$$

By testing for the null hypothesis of no cointegration is like testing for a unit root in the regression residuals $\hat{u}(t,T)$. To make it clear, under the null hypothesis of no cointegration, $\hat{u}(t,T)$ will be integrated of order one. If on the other hand there is cointegration between spot and future prices, we expect the residuals of the above regression to be stationary.
The OLS estimates of (α, β) along with p-values are presented in the Table 2. The results from the ADF test for stationarity on the residuals of equation 4 are shown in Table 3.

Table 2. *Regression for spot and futures prices [S(T)=α+βF(t,T)+u(t,T)]*

Commodity	Variable	Coefficient	Std. Error	t-Statistic	Prob
Aluminium	c	-0.246074	0.149024	-1.651232	0.0989
	log(futures)	1.023490	0.019653	52.078200	0.0000
Gold	c	0.014757	0.010419	1.416333	0.1569
	log(futures)	0.997965	0.001463	682.234100	0.0000
Nickel	c	0.632926	0.108086	5.855788	0.0000
	log(futures)	0.929156	0.011505	80.760830	0.0000
Platinum	c	1.670703	0.093923	17.788000	0.0000
	log(futures)	0.754622	0.013645	55.303860	0.0000
Silver	c	0.782213	0.050401	15.519680	0.0000
	log(futures)	0.718924	0.018058	39.811860	0.0000
Zinc	c	-0.576470	0.092507	-6.231640	0.0000
	log(futures)	1.077400	0.011909	90.468010	0.0000

Table 3. *ADF test for stationarity on the residuals:* u(t,T)=S(T)-α-β·F(t,T)

Commodity	Observations	Mean	Standard deviation	ADF t-stat	MacKinnon (1996) one-sided p-values
Aluminium					
u(t,T)	1256	-0.24235	0.06798	-3.982674	0.0016
Gold					
u(t,T)	1248	0.01522	0.003056	-34.49674	0.0000
Nickel					
u(t,T)	1256	-0.00049	0.078680	-5.300568	0.0000
Platinum					
u(t,T)	1217	-0.01153	0.064152	-4.015340	0.0014
Silver					
u(t,T)	1237	0.0000791	0.054445	-5.252639	0.0000
Zinc					
u(t,T)	1256	0.003038	0.076968	-2.996602	0.0355

Notes:
test critical values

t-stat(1%)	t-stat(5%)	t-stat(10%)
-3.435348	-2.863635	-2.567935

Turning to the results, the OLS estimates of coefficients (α,β) as shown in Table 2 are generally close to (0,1). The p-values in Table 3 imply rejection of the Null hypothesis of unit root, therefore the residuals are stationary. Hence, we conclude that there is strong evidence of cointegration between daily spot and futures prices, so that S(T)-F(t,T) will be stationary quantities. Of course, these findings do not imply the absence of a time-varying risk premium.

Phillips and Perron test

As we mentioned earlier one of the major problems of DF and ADF test was the error terms in equation (2) which Dickey and Fuller overcome with the use of dependent variables. Phillips and Perron (1988) on the other hand, chose an alternative way to deal with the same problem. In an attempt to take into account the possible autocorrelation in the error terms, they applied a non-parametric correction to t test statistics in the regression (2).
Phillips and Perron mentioned that the fact that makes equation (2) biased because of the autocorrelation function, is because of the discrepancy between population variance and residual variance in the equation. The difference between consistent estimator of the variances of the population and sample is:

$$\sigma_\varepsilon^2 = \mathrm{T}^{-1}\sum_{t=1}^{T}\varepsilon_t^2$$

$$\sigma_p^2 = \mathrm{T}^{-1}\sum_{t=1}^{T}\varepsilon_t^2 + 2\mathrm{T}^{-1}\sum_{t=1}^{l}\sum_{t=j+1}^{T}\varepsilon_t\varepsilon_{t-j} \qquad (5)$$

where σ_ε^2 , σ_p^2 are the sample variance and population variance respectively, where l represents the lag truncation for the autocorrelation in the residuals. In equation (5), the term $2\mathrm{T}^{-1}\sum_{t=1}^{l}\sum_{t=j+1}^{T}\varepsilon_t\varepsilon_{t-j}$, is the difference between the sample and population residuals variances due to autocorrelation. In the absence of autocorrelation for the residuals, the above term will become zero. As a result, the above two equations will be equal and also the estimates in the first equation will be unbiased. In that sense, according to Phillips and Perron, the improved version of Dickey-Fuller test will be:

$$Z(\tau_\mu) = \frac{\sigma_\varepsilon}{\sigma_p}\cdot\tau_\mu - \frac{1}{2}\cdot\left(\sigma_p^2 - \sigma_\varepsilon^2\right)\cdot\left\{\sigma_p\cdot\left[T^2\cdot\sum_{t=2}^{T}\left(Y_{t-1}-\bar{Y}_{t-1}\right)\right]^{1/2}\right\}^{-1}$$

where τ_μ represents the t statistics for testing unit roots in DF test, equation (2). $Z(\tau_\mu)$ is the PP statistic for testing unit roots in the presence of residual autocorrelation. It concludes that the critical values for the PP test

are the same as the DF test, and also in the absence of autocorrelation in the error terms the t-statistic for both of the tests are also equal.
The PP test seems to accurately solve the problem of autocorrelation and time-dependent heteroscedasticity when it applies, but as Schwert (1987) showed, this test reject the Null hypothesis way to often when there is a moving average component. The ADF test on the other hand, has less restrictions when it comes to this. Therefore, we used the ADF instead of the PP process to test for unit roots in this study.

Criticism of Dickey-Fuller and Phillips-Perron type tests

Like all the statistical tools, these 2 unit root tests suffer from the problem of size and power. The most important element in these cases is the sample size, because all our data are in the form of a finite sample. This will often cause confusion because the distinction between a trend stationary process and a difference stationary process is not clear. Another problem which could significantly affect the results is the presence of structural breaks and mean-shift within the sample. If there is a permanent shift in the trend of a stationary process or a permanent change in the mean, the results could be spurious and misleading in deciding about the presence of a unit root.
Another very important criticism at unit root tests like the above is that they have low power if the process is stationary but with a root close to non-stationary boundary. If for example in equation (1) ψ=0.05 but significantly different from zero, technically the null hypothesis should be rejected. It has been therefore emphasized that unit roots tests are poor at deciding close to the boundaries whether a unit root is present, and this problem is enhanced if the sample size is small. The source of the problem is that the null hypothesis is either rejected or not rejected but it is never accepted. In that sense, failure to reject the null could be either because null is correct or because there is no sufficient evidence to reject it. One such a process of improving the result is to use KPSS test, which is available in E-Views, though this is beyond the course of this project and won't be addressed.

Cointegration methodology

Furthermore, with regards to the concept of cointegration that was addressed earlier (in the "Time Series" part), if we combine two variables of order 1, we expect the combination of them to be of order 1 as well in most cases. In general, if we combine 2 variables, one of $I(d_1)$ and another $I(d_2)$ with $d_2 > d_1$ we expect the combination of them to be $I(d_2)$ or else to have the order of the variable with the highest order of integration. So, typically a linear combination of I(1) variables will itself be I(1), but would be desirable to be I(0). In that case we state that the variables are cointegrated.
As a further illustration, consider the following regression model which contains the time series y_t and x_t which are I(1): $y_t = \alpha + \beta \cdot x_t + u_t$. For the estimating model, the regression function would be: $y_t = \hat{a} + \hat{\beta} x_t + \hat{u}_t$. We can easily derive from that equation: $y_t - \hat{a} + \hat{b} x_t = \hat{u}_t$. This linear combination of I(1) variable will be I(0), in other words stationary, if the variables are cointegrated. The rational behind is that many time series are non-stationary but "move together" over time – which imply that the two series are bound by some relationship in the long run. A cointegrating relationship may also be seen as a long-term or equilibrium phenomenon, since it is possible that cointegrating variables may deviate from their relationship in the short run, but their association would return in the long run.
Moreover, market forces that arising from no-arbitrage conditions suggest that there should be an equilibrium relationship between the series. If there were no cointegration, there would be no long run relationship binding the series together which means that all linear combinations of the series would be non-stationary and would have no constant mean that would be returned to frequently.
Since spot and future prices are for the same asset, it is expected (in some cases) for them to cointegrate since they will be affected by the same information. Hendry and Juselius (2000) note that if a set of series is cointegrated in levels, they will also be cointegrated in log levels.

Error correction models

Earlier in the study we mentioned that if the series will be non-stationary (as they are), the strategy that will follow would be to differentiate the series one time in order to induce stationarity. Additionally, we are going to use these differences in order to run the modelling process. In the case that the relationship between the variables is not important, this approach is completely valid. Although, when the relationship between the variables is important (as expected in the spot, futures regressions), this is not the best approach. The problem lies in the fact that first difference models have no long-run solutions, while at the same time there is a long-run relationship between the series. If for example we consider two series, y_t and x_t, which are both integrated of order 1, I(1) and we try to estimate the following model:

$$\Delta y_t = \beta \cdot \Delta x_t + u_t$$

Consider now the case where both variables have converged upon some long-term values and do change anymore. It will be that: $y_t = y_{t-1}$ and $x_t = x_{t-1}$. Hence: $\Delta y_t = 0$ and $\Delta x_t = 0$ and everything in the equation cancels. This model will have no long-run solution and cannot advise for a long-run relationship between x and y.

The above problem can be solved by models that use combinations of first differenced and lagged levels of cointegrated variables. If for example in the following equation:

$$\Delta y_t = \beta_1 \cdot \Delta x_t + \beta_2 \cdot (y_{t-1} - \gamma x_{t-1}) + \mathrm{u}_t$$

y_t and x_t are cointegrated (the cointegrated coefficient being γ) then the term $(y_{t-1} - \gamma x_{t-1})$ will be I(0) while the constituents are of the first order. In this case would be valid to use OLS for statistical inference. It is of course possible to have an intercept, either in the term $(y_{t-1} - \gamma x_{t-1})$ or even in the Δyt. Whether the constant is included or not could be determined on the basis of financial theory. The above model is known as “error correction model” while the term $(y_{t-1} - \gamma x_{t-1})$ is known as error correction term.

Error correction models have the following interpretation: y is determined to change between t − 1 and t as a result of changes in the values of the explanatory variable x, between t − 1 and t, and also to correct for any disequilibrium that existed in the previous period. Notice that the term $y_{t-1} - \gamma x_{t-1}$ (“error correction term”) is not at time t, but instead at time t-1. It would be better if it appeared without a lag, for this would imply that y changes between t − 1 and t. For the rest of the terms, γ defines the long-run relationship between x and y, β_1 describes the short-run relationship between changes in x and changes in y, β_2 describes the speed of adjustment back to equilibrium, and measures the proportion of last period’s equilibrium error that is corrected for.

Testing for cointegration in regression

The model we are going to use is: $y_t = a + \beta \cdot x_t + u_t$

It is essential to test the residuals u_t of the above equation to conclude whether they are stationary or have a unit root. By using the regression: $\Delta \hat{u}_t = \psi \cdot \hat{u}_{t-1} + v_t$, with v_t : iid error term, the ADF test can show the outcome. Additionally, because this is a test on the residuals, the critical values of the ADF test will change. Engle-Granger came up with a new set of critical values which they communicated and hence the new name of this test is Engle-Granger (EG) test. The reason for these new critical values is simply because the regression runs on the residuals rather than on the data (spot and future prices) themselves. The residuals are derived from specific coefficient estimates and therefore the sample error in those coefficients will change the distribution of the test statistic.

Another alternative for testing for stationarity in the residuals is to use the Durbin - Watson (DW) test statistic or the Phillips - Perron (PP) approach. When a DW test is applied to the residuals of a potential cointegrating regression, it is known as Cointegrating Regression Durbin Watson (CRDW). The Null hypothesis will be the presence of unit root in the errors, CRDW will be approximately zero and the Null hypothesis is rejected if the CRDW statistic is larger than the critical value.

The Null and alternative hypothesis for stationarity applied to the residuals of a cointegrating (or not) system are: $H_0 : \hat{u}_t \sim I(1) \quad and \quad H_1 : \hat{u}_t \sim I(0)$

According to the Null hypothesis (if not rejected) there will be a unit root in the residuals and under the alternative hypothesis the residuals will be stationary. Therefore, under the Null hypothesis, there is no possible combination of the non-stationary variables that provides a stationary outcome. Hence if this Null hypothesis is not rejected, there is no cointegration. If that's the case, the appropriate econometric strategy modelling will be to use the first differences of the variables. These models will exclude the presence of a long-run equilibrium solution, but this will be indifferent since no cointegration implies no long run relationship. On the other hand, if the null hypothesis is rejected, it would conclude that a linear combination between the non-stationary variables that gives a stationary outcome has been found. In this case the variables is said to be cointegrated. The appropriate econometric modelling strategy in this case will be to estimate an error correction model using the following method.

Engle-Granger approach

The modelling strategy which is going to be used for data that are non-stationary but they are cointegrated is the Engle – Granger approach which is described in the following paragraphs.

In the 1st step we make sure that all the individual variables are I(1). Then we can use OLS to estimate the cointegrating regression. It will not be possible to perform any inferences on the coefficient estimates in this regression. All that we'll do is to estimate the parameter values. We save the residuals of the cointegrating regression $\hat{u}_t$ and we test them to ensure that they are I(0).

In the 2nd step we use the residuals from the 1st step as one variable in the error correction model:

$$\Delta y_t = \beta_1 \cdot \Delta x_t + \beta_2 \cdot (\hat{u}_{t-1}) + v_t$$

Where $\hat{u}_{t-1} = y_{t-1} - \hat{\tau} x_{t-1}$. The stationary, linear combination of non-stationary variable is also known as the cointegrating vector. In this case the cointegrating vector would be $[1 - \hat{\tau}]$. Also, any linear transformation of this vector will also be a cointegrating vector. For example: $\hat{u}'_{t-1} = 2y_{t-1} - 2\hat{\tau} x_{t-1}$ will also has a unit root. In our case, in the previous equation: $y_t - \hat{a} - \hat{b} x_t = \hat{u}_t$ the cointegrating vector would be $\left[1 - \hat{a} - \hat{b}\right]$. It is now valid to perform inferences in the second stage regression concerning the parameters α and β since all variables in this regression are stationary.

There are some issues with the Engle – Granger method which will be addressed in the following:

Firstly, there is the usual problem with the limited amount of data in a sample, which has already been discussed and contributes to the lack of power in unit root tests as well as cointegration tests.

Secondly, there could be a simultaneous equations bias if the causality between y and x runs in both directions, but this single equation approach requires the researcher to normalise on one variable (i.e. to specify one variable as the dependent variable and the others as independent variables). The researcher is forced to treat y and x asymmetrically, even though there may have been no theoretical reason for doing so. A further issue is the following. Suppose that the following specification had been estimated as a potential cointegrating regression

$$y_t = \hat{a} + \hat{\beta} x_t + \hat{u}_t$$

What if instead the following equation was estimated?

$$x_t = \widehat{a'} + \widehat{\beta'} y_t + \widehat{u'}_t$$

If it is found that $u_t \sim I(0)$, does this imply automatically that $u'_t \sim I(0)$? The answer in theory is ‘yes’, but in practice different conclusions may be reached in finite samples. Also, if there is an error in the model specification at stage 1, this will be carried through to the cointegration test at stage 2, as a consequence of the sequential nature of the computation of the cointegration test statistic.

Thirdly, it’s not possible to perform any hypothesis tests about the actual cointegrating relationship estimated.

Lastly, there may be more than one cointegrating relationship.

Lead-lag long term relationships spot-futures

Under the EMH if markets are frictionless, changes in the logarithm of spot prices of commodities and changes in the logarithm of the corresponding futures prices, should be perfectly correlated and not cross-autocorrelated. In mathematics the above relationship is illustrated as follows:

$$\begin{aligned}
&corr(\Delta \log(\mathrm{f}_t), \Delta \log(\mathrm{s}_t)) \approx 1 \\
&corr(\Delta \log(\mathrm{f}_t), \Delta \log(\mathrm{s}_{t-k})) \approx 0 \quad \forall k > 0 \\
&corr(\Delta \log(\mathrm{f}_{t-j}), \Delta \log(\mathrm{s}_t)) \approx 0 \quad \forall j > 0
\end{aligned}$$

What these equations present is the following: According to the first one, changes in the spot prices and in the future prices are expected to take place simultaneously without any time lag. From the second equation we derive that changes in the current future prices should not be related to any change of the spot prices that took place in the past. From the third equation we derive that current changes in the spot prices should not be related to past changes in the future prices. When we write down changes in the logarithms, we also mean spot and futures returns.

For the case where the underlying asset is a hard commodity (which is not a consumption commodity but rather an investment commodity), the equilibrium relationship between spot and futures prices is given by:

$$F_t = S_t \cdot e^{r(\mathrm{T}-\mathrm{t})}$$

F_t is the futures price, S_t is the spot price at the same time, r is the continuously compounded risk-free rate of interest and T-t is the time to maturity of the futures contract. By taking logarithms from both sides of the equation we derive:

$$\log(F_t) = \log(S_t) + r(T-t)$$

From the above equation we conclude that the relationship between the logs of spot and futures prices should be one-to-one. Therefore, the basis (which is the difference between the futures and the spot prices) should be stationary. If that was not the case, hence if the basis was not stationary, arbitrage opportunities for traders would arise, which after full capitalization would bring the spot and futures prices back to the equilibrium prices and any future arbitrage opportunities would be eliminated.

By using simple regression and cointegration analysis we can prove that there shouldn’t be any lead-lag relationship between the logs of spot and futures prices and the only relationship between them should be a long-term one-to-one. We can see the results from the spot and futures prices for the aluminium. Our data consist of 1256 daily observations for the spot and futures prices (over a period of 5 year approximately). In order to form a statistically adequate model we should first seek whether the data can be considered stationary or not. The results of applying the ADF test on the logs of the spots and futures prices can be shown in the Table 4.

From the results is clear that the two log-price series (spot and futures) contain a unit root, while the returns are stationary. A statistically adequate model therefore should be one in the returns rather than in the prices. However, if we include only the first differences, we will find no long-run equilibrium solution. On the contrary, the theory suggests that spot and futures prices should have a long-run equilibrium solution. The solution is therefore to find out whether a cointegration relationship between spots and futures prices exists, which would mean that is valid to include level terms along with returns. We can do that by examining the residuals of the regression:

$$S(T) = \hat{a} + \beta F(t,T) + \hat{u}(t,T)$$

are stationary, using the ADF test, with u(t,T) being the error term. The coefficient values and the DF test are given in Table 5.

It’s clear that the residuals from the cointegrating regression can be considered stationary. Also, the estimated slope coefficient in the cointegrating regression takes on a value close to one, as predicted from the theory. It is not possible to test whether the true coefficient could be one, however, since there is no way in this framework to test hypotheses about the cointegrating relationship.

Table 4. *ADF test on log prices and returns for the Aluminum*

Null Hypothesis: LSPOT has a unit root
Exogenous: Constant, Linear Trend
Lag Length: 0 (Automatic - based on SIC, maxlag=22)

		t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic		-2.078829	0.5566
Test critical values:	1% level	-3.965375	
	5% level	-3.413396	
	10% level	-3.128734	

*MacKinnon (1996) one-sided p-values.

Null Hypothesis: LFUTURES has a unit root
Exogenous: Constant, Linear Trend
Lag Length: 0 (Automatic - based on SIC, maxlag=22)

		t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic		-1.778699	0.7147
Test critical values:	1% level	-3.965375	
	5% level	-3.413396	
	10% level	-3.128734	

*MacKinnon (1996) one-sided p-values.

Null Hypothesis: DLSPOT has a unit root
Exogenous: Constant, Linear Trend
Lag Length: 1 (Automatic - based on SIC, maxlag=22)

		t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic		-23.19977	0.0000
Test critical values:	1% level	-3.965387	
	5% level	-3.413402	
	10% level	-3.128737	

*MacKinnon (1996) one-sided p-values.

Null Hypothesis: DLFUTURES has a unit root
Exogenous: Constant, Linear Trend
Lag Length: 0 (Automatic - based on SIC, maxlag=22)

		t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic		-37.28556	0.0000
Test critical values:	1% level	-3.965381	
	5% level	-3.413399	
	10% level	-3.128736	

*MacKinnon (1996) one-sided p-values.

Table 5. *Estimated potentially cointegrating equation and test for cointegration for Aluminum*

Dependent Variable: LSPOT
Method: Least Squares
Date: 08/05/19 Time: 20:20
Sample: 8/01/2014 7/19/2019
Included observations: 1256

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-0.246074	0.149024	-1.651232	0.0989
LFUTURES	1.023490	0.019653	52.07820	0.0000

Null Hypothesis: REG2_LOGS_COINT has a unit root
Exogenous: Constant
Lag Length: 0 (Automatic - based on SIC, maxlag=22)

		t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic		-3.982674	0.0016
Test critical values:	1% level	-3.435348	
	5% level	-2.863635	
	10% level	-2.567935	

*MacKinnon (1996) one-sided p-values.

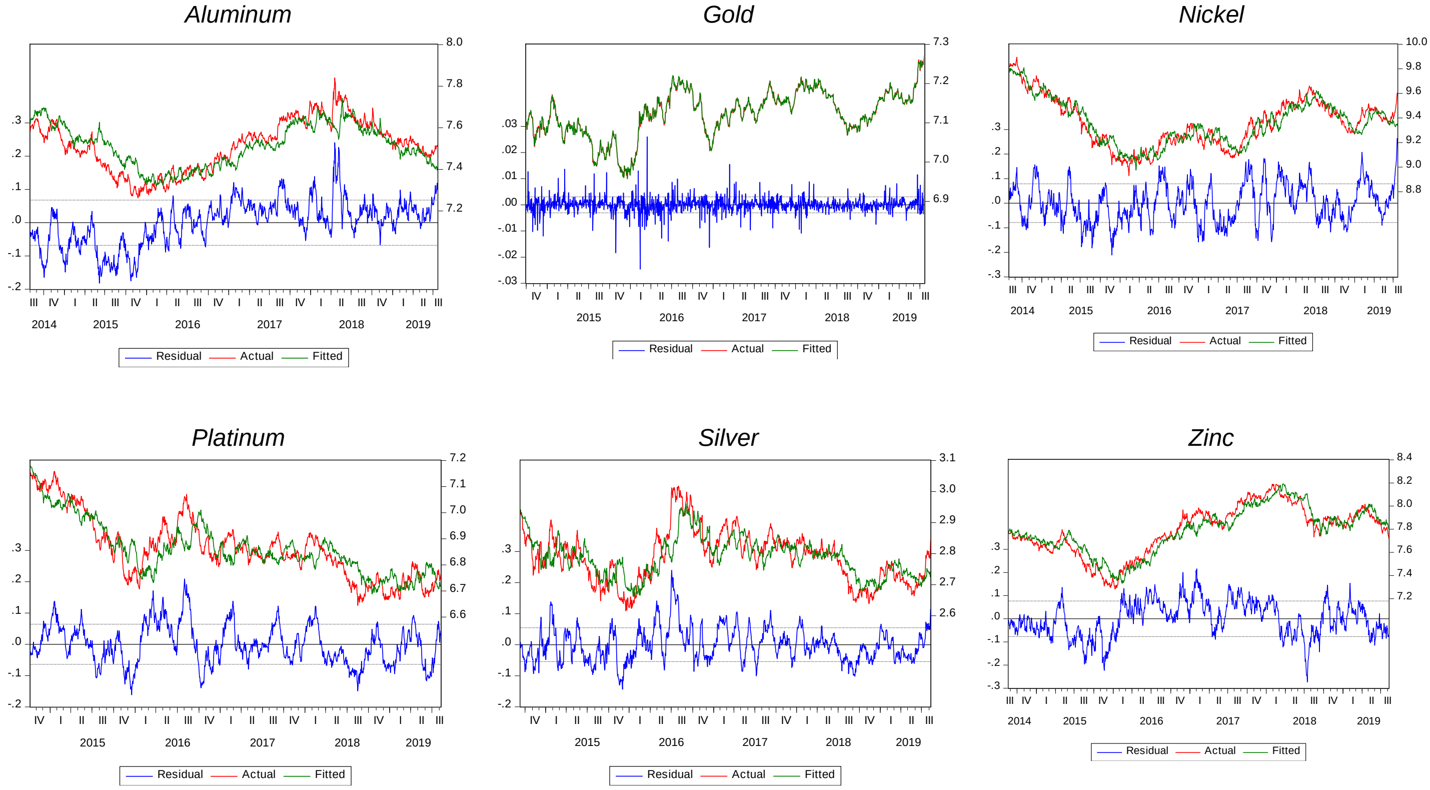
Testing for cointegration and modelling cointegrated systems using EViews

Now we will examine the spot and futures time series for the following 6 hard commodities: aluminum, gold, nickel, platinum, silver and zinc. We will examine whether there is a long term relationship between the two time series by testing for cointegration. This would mean that the two prices don't wander apart without any bound. We use the Engle - Granger approach in order to test for cointegration in the residuals of the regression of the spot and future prices. We first generated two new series, the log of the spot and the log of the futures prices for the data that we downloaded from Bloomberg. These two series were called "lspot" and "lfutures" respectively. We then run the regression:

"*lspot c lfutures*"

It is invalid to test for anything else than the coefficient values of this regression. By examining the residuals of this regression, we can see a plot of the levels of the residuals, which looks much more like a stationary series than the original spot prices (the blue line stands for the residuals and the red line stands for the spot prices at time T. It's obvious that the actual and fitted lines are very close together. The plot should appear in Table 6.

Table 6. *Actual, Fitted and residual plot to test for stationarity*



Following, we need to keep those residuals in a separate series to use later, therefore we generate the series:

"statresid=resid"

This is mandatory because the residuals as a time series is updated every time we run a regression, in order to include the residuals of the most recently run regression. We perform the ADF test on these residuals (on the "STATRESID". In this test, up to 22 lags are permitted, the Schwarz is the optimal criterion to select the best number of lags to take into consideration and a constant is employed in the regression in the levels of the series. The table 7 shows the results.

Table 7. *ADF test on the residuals for all six commodities*

Aluminum

Null Hypothesis: STATRESID has a unit root
Exogenous: Constant, Linear Trend
Lag Length: 22 (Automatic - based on SIC, maxlag=22)

		t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic		-11.07260	0.0000
Test critical values:	1% level	-3.965512	
	5% level	-3.413463	
	10% level	-3.128774	

*MacKinnon (1996) one-sided p-values.

Gold

Null Hypothesis: STATRESID has a unit root
Exogenous: Constant, Linear Trend
Lag Length: 0 (Automatic - based on SIC, maxlag=22)

		t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic		-35.27298	0.0000
Test critical values:	1% level	-3.965428	
	5% level	-3.413422	
	10% level	-3.128749	

*MacKinnon (1996) one-sided p-values.

Nickel

Null Hypothesis: STATRESID has a unit root
Exogenous: Constant, Linear Trend
Lag Length: 22 (Automatic - based on SIC, maxlag=22)

		t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic		-10.87898	0.0000
Test critical values:	1% level	-3.965512	
	5% level	-3.413463	
	10% level	-3.128774	

*MacKinnon (1996) one-sided p-values.

Platinum

Null Hypothesis: STATRESID has a unit root
Exogenous: Constant, Linear Trend
Lag Length: 0 (Automatic - based on SIC, maxlag=22)

		t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic		-35.00492	0.0000
Test critical values:	1% level	-3.965617	
	5% level	-3.413514	
	10% level	-3.128804	

*MacKinnon (1996) one-sided p-values.

Silver

Null Hypothesis: STATRESID has a unit root
Exogenous: Constant, Linear Trend
Lag Length: 0 (Automatic - based on SIC, maxlag=22)

		t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic		-34.16163	0.0000
Test critical values:	1% level	-3.965494	
	5% level	-3.413454	
	10% level	-3.128769	

*MacKinnon (1996) one-sided p-values.

Zinc

Null Hypothesis: STATRESID has a unit root
Exogenous: Constant, Linear Trend
Lag Length: 0 (Automatic - based on SIC, maxlag=22)

		t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic		-35.76492	0.0000
Test critical values:	1% level	-3.965512	
	5% level	-3.413463	
	10% level	-3.128774	

*MacKinnon (1996) one-sided p-values.

As we can see the test statistic is more negative than the critical values and the Null hypothesis of the presence of unit root in the residuals can be rejected. We would have to conclude therefore that the two series, spots and futures prices of the commodities are cointegrated. Since the series are cointegrated, an error correction model (ECM) could be estimated, as there is a linear combination of the spot and futures prices which would be stationary. The ECM model would be appropriate to catch the long – run relationship between the two series as well as the short – run one. Note that a model for the first differences would not be appropriate because of the cointegration factor. The ECM is estimated by running the regression:

“lspot c lfutures statresid(-1)”

Although the EG approach is very easy to use, one of the major disadvantages is that it can only estimate up to one cointegrating relationship between the variables. Because the study is related to spot and futures prices, there can only be one cointegrating relationship since there are only two variables in the system. In other situations however, if we had more variables, there could potentially be more than one linearly independent cointegrating relationship. In those cases would have been more appropriate to examine the cointegration with the Johansen technique.

Comments on lead-lag relationship

If the markets are frictionless and functioning efficiently, changes in the spot price of a financial asset and its corresponding futures price would be expected to be perfectly contemporaneously correlated and not to be cross-autocorrelated. Many academic studies, however, have documented that the futures market systematically ‘leads’ the spot market, reflecting news more quickly as a result of the fact that the commodity index is not a single entity. The latter implies that:

- Some components of the index are infrequently traded, implying that the observed index value contains ‘stale’ component prices
- It is more expensive to transact in the spot market and hence the spot market reacts more slowly to news
- Commodity market indices are recalculated only every minute so that new information takes longer to be reflected in the index.

Clearly, such spot market impediments cannot explain the inter-daily lead-lag relationships documented by Tse (1995). In any case, however, since it appears impossible to profit from these relationships, their existence is entirely consistent with the absence of arbitrage opportunities and is in accordance with modern definitions of the efficient markets hypothesis.

Long-memory models

It is widely acceptable these days that asset prices (and the logs of them) follow a random walk process. Its acceptable though that the returns of these assets do not possess a further unit root (they are stationary series) but note that this doesn't mean they are completely independent. It is possible and it has been proved to a certain extent, that data which have taken a distance apart have shown signs of dependence. Those series is said to have a long memory. In order to illustrate this phenomenon we use what is called "fractionally integrated model". Just for a clarification purposes, a series is said to be integrated of order d, when it becomes stationary if its differenced d times. But, in fractionally integrated framework, d is allowed to take any value and not only integer ones. This framework was applied to the estimation of ARMA models under which the autocorrelation function (ACF) will decline slower than an exponential decline. Therefore, the autocorrelation function dies significantly slower than an ARMA model with d=0. The long memory notion has been applied to GARCH models as well, where volatility was found to exhibit long time dependence. The new class of these models were named FIGARCH (fractionally integrated GARCH) and were developed by Granger and Engle 1993 and Bollerslev and Mikkelsen 1996.

III. RISK PREMIUM

The equation F(t,T) = E[S(T)/I(t)] relates the current futures price to the expected future spot price under the joint hypothesis of risk neutrality and rational expectations. If economic agents are risk averse however, then the future price may differ from the expected future spot price by a risk premium P. Thus, the former equation will be change to:

$$F(t,T) = E\left[S(\mathrm{T}) / \mathrm{I(t)}\right] + P(t)$$

Where at this point P(t) represents the bias of the futures price to forecast the future spot price.
By subtracting S(t) from both parts of the equation we end up with:

$$F(t,T) - \mathrm{S(t)} = E\{\left[S(\mathrm{T}) / \mathrm{I(t)}\right] - S(t)\} + P(t)$$

Where the F(t,T)-S(t) is called basis. The above equation implies that the basis can be split in two components: the premium P(t) and the change in the spot price E{[S(T)/I(t)]-S(t)}. Similarly, to the first study with regards to "testing for market efficiency", the regression that we are called to run is:

$$F(t,T) - \mathrm{S(t)} = \left[S(\mathrm{T}) - S(t)\right] + P(t) + \mathrm{u(t,T)}$$

In order to investigate the variability of risk premiums and expected spot-price changes and their covariability, we use Fama's model. Specifically, we split the above regression into two complementary regressions, F(t,T)-S(T) and S(T)-S(t), both observed at T, on F(t,T)-S(t), observed at t:

$$F(t,T) - \mathrm{S(T)} = a_1 + \beta_1\left[\mathrm{F(t,T)} - S(t)\right] + \mathrm{u_1(t,T)} \qquad (6)$$

$$\mathrm{S}(T) - \mathrm{S(t)} = a_2 + \beta_2\left[\mathrm{F(t,T)} - S(t)\right] + \mathrm{u_2(t,T)} \qquad (7)$$

According to Fama (1984) the regression slope coefficients are defined as follows:

$$\beta_1 = \frac{\mathrm{cov}\left[F(t,T) - S(\mathrm{T}), \mathrm{F(t,T)} - \mathrm{S(t)}\right]}{\mathrm{var}\left[F(t,T) - S(t)\right]}$$

$$\beta_2 = \frac{\mathrm{cov}\left[S(T) - S(\mathrm{t}), \mathrm{F(t,T)} - \mathrm{S(t)}\right]}{\mathrm{var}\left[F(t,T) - S(t)\right]}$$

Where covariance and variance are unconditional.
The previous 2 regression equations are dependent since the regressor is the same in both equations and the sum of the dependent variables is the stochastic regressor. Therefore, we are driven to the conclusion that $\alpha_1+\alpha_2$ should converge to zero and $\beta_1 + \beta_2$ should converge to 1. In other words, the above regressions contain identical information with regards to the variation of the premium and there is no need to be solved both.
Since F(t,T) - S(T) is the risk premium plus the random error term, estimates of the equations will show whether or not the risk premium of the basis has a variation that appears reliably in F(t,T) - S(T). Specifically,

when risk premium is constant over time, β_1 should be equal to zero and β_2 must be unity. Hence, the coefficients β_1 and β_2 describe the degree of variability in the components of the basis.
Although the above regressions allocate all basis variation to either risk premium, expected spot-price change or both β_1 would be equal to the proportion of the variance of the basis due to the variation of the risk premium only if the risk premium doesn't have any correlation with the change in the spot price. If this condition is satisfied, β_2 would be equal to the proportion of the variance of the basis due to the variance of the expected change in the spot price. The importance of analyzing both equations is that the difference the slope coefficients β_1 and β_2 indicated the relative size of variances of the premium and expected rate of change of the spot price.

$$\beta_1 - \beta_2 = \frac{\mathrm{var}\left[P(t)\right] - \mathrm{var}\left[E\left\{S(T)-S(t)\right\}\right]}{\mathrm{var}\left[F(t,T)-S(t)\right]}$$

Table 8 shows the σ for the premium, change and basis for each commodity. For all six commodities, basis standard deviation is low (in some cases significant lower) in general relative to the standard deviations of the premium and the change in the spot price. This indicates that is unlikely for the earlier shown regressions (6) and (7) to assign reliably basis volatility to premiums and change in the spot price volatility.

Table 8. *Test for stationarity in the premium, the change and the basis*

Commodity	Observations	Mean	Standard deviation	ADF t-stat	MacKinnon (1996) one-sided p-values
Aluminum					
Premium	1256	0.067967	0.068020	-5.238155	0.0001
Change	1256	-0.002168	0.049731	-6.090705	0.0000
Basis	1279	0.063690	0.048118	-3.370640	0.0558
Gold					
Premium	1248	-0.000262	0.003059	-34.513940	0.0000
Change	1236	0.002182	0.039651	-5.098453	0.0001
Basis	1259	-0.000362	0.006117	-30.649470	0.0000
Nickel					
Premium	1256	0.032492	0.079870	-5.626761	0.0000
Change	1256	-0.004836	0.084517	-5.607434	0.0000
Basis	1279	0.024983	0.015906	-2.082774	0.5544
Silver					
Premium	1237	0.001922	0.059546	-5.645492	0.0000
Change	1236	-0.003273	0.060113	-5.490173	0.0000
Basis	1259	-0.001458	0.010901	-27.441710	0.0000
Zinc					
Premium	1256	-0.024586	0.078254	-2.771270	0.2083
Change	1256	0.000471	0.068094	-5.741466	0.0000
Basis	1279	-0.023591	0.040702	-2.576397	0.2913
Platinum					
Premium	1217	0.017985	0.072186	-4.228439	0.0041
Change	1236	-0.009462	0.057436	-5.353491	0.0000
Basis	1259	0.000584	0.009194	-30.838750	0.0000

Notes:
Premium = F(t,T)-S(T) Change = S(T)-S(t) Basis = F(t,T)-S(t)
test critical values

t-stat(1%)	t-stat(5%)	t-stat(10%)
-3.965488	-3.413451	-3.128767

The only available hypothesis therefore is the one that premium, change in spot price and the basis are stationary processes. Indeed, after running the augmented Dickey-Fuller test we found out that all variables, apart from Zinc, are stationary. Hence, we are going to follow the procedure that we present next.
Table 9 shows the estimated regressions for the premium, the change in the spot price and the basis. Because the two regressions (one for the premium and one for the change) are complementary, we only show one standard error. We also show both of the intercepts α1 and α2 and both slopes β1 and β2 for the regressions

(6) and (7). Note that the coefficients of determination $R^2(1)$ and $R^2(2)$ are low because as we mentioned before the basis has significantly lower variation than the premium and the basis (look Table 8).

Table 9. *Regressions:* $F(t,T) - \mathrm{S(T)} = \hat{a}_1 + \hat{\beta}_1 \left[\mathrm{F(t,T)} - S(t)\right] + \hat{\mathrm{u}}_1(\mathrm{t,T})$ *(regression of the premium)*

$\mathrm{S}(T) - \mathrm{S(t)} = \hat{a}_2 + \hat{\beta}_2 \left[\mathrm{F(t,T)} - S(t)\right] + \hat{\mathrm{u}}_2(\mathrm{t,T})$ *(regression of the change)*

Commodity	α1	β1	α2	β2	S(α)	S(β)	R^2(1)	R^2(2)
Aluminum	0.003841	0.97417	-0.00384	0.02583	0.00236	0.02931	0.46832	0.00062
Gold	-0.002195	0.96304	0.00220	0.03697	0.00113	0.18406	0.02170	0.00003
Nickel	0.032627	-0.09395	-0.03263	1.09395	0.00444	0.14845	0.00032	0.04151
Silver	0.002727	0.62077	-0.00273	0.37923	0.00172	0.15545	0.01276	0.00480
Zinc	0.000000	1.01945	0.00000	-0.01945	0.00223	0.04697	0.27310	0.00014
Platinum	0.009281	1.30749	-0.00928	-0.30749	0.00164	0.17653	0.04256	0.00245

Notes: R^2(1) and R^2(2) are the coefficients of determination for the premium and change regressions. Because the two regressions are completely complementary it means that the standard errors of the estimated regressions coefficients are the same for the two regressions.

With regards to the coefficient estimates, we need to comment on some unexpected results and particularly on the estimated coefficients of Platinum. It was explained earlier that β1 contains the variance of the basis that is related to its premium component and β2 is the variance of the basis that is related to its change in the spot price component. The coefficients of the regression for Platinum although, cannot perceived that way because the coefficients estimate for the premium regression are almost always greater than one implying that their counterparts for the change in the spot price regression are almost always negative.
However, for the rest of the commodities, the fact that β1 and β2 are both positive and less than 1, reliably represents positive variance for the premium and for the change in the spot price. This means that the futures prices have reliable power in forecasting the spot prices in the future (at the end of the contract) and also that the futures prices contain a time-varying premium that appears reliably in the Premium $F(t,T) - \mathrm{S(T)}$.

IV. CONCLUSION

In this study we analysed six hard commodities' spot and futures prices (aluminum, gold, nickel, silver, zinc and platinum). With the powerfulness of the augmented Dickey-Fuller test, which is robust for a wide variety of serial correlation and time-dependent heteroscedasticity, we found solid evidence that the twelve time series are following a random walk process.

We used the cointegration theory to test for the joint hypothesis (risk neutrality and rational use of all available information), which is equivalent to the presence of market efficiency. The results were the presence of cointegration between the spot and futures prices (1 month futures specifically) in all six commodities. The cointegration is consistent with the market efficiency.

We also used Fama's regression approach (1984) to measure the information in the futures prices concerning with regards to time-varying risk premium and spot prices at the end of the futures contract. The results showed that there is variation in both the premium and the change in the spot prices components of the basis and also that the variance of the premium component is larger than corresponding variance of the change in the spot prices. It is concluded that the presence of time-varying risk premium makes the forecast of spot prices from the futures contracts less efficient.

REFERENCES

Antoniou, A. and Foster, A. J. (1994) Short-term and long-term efficiency in commodity spot and futures markets, Financial Markets, Institutions and Instruments, 3, 17–35.

Brooks C. (2014) Introductory econometrics for finance

Bigman, D., Goldfarb, D. and Schechtman, E. (1983) Futures market efficiency and the time content of the information set, The Journal of Futures Markets, 3, 321–34.

Bilson, J. F. O. (1981) The ‘speculative efficiency’ hypothesis, Journal of Business, 54, 435–51.

Dickey, D. A. and Fuller, W. A. (1979) Autoregressive timeseries with a unit root, Journal of the American Statistical Association, 74, 427–31.

Dickey, D. A. and Fuller, W. A. (1981) Likelihood ratio statistics for autoregressive time series with a unit root, Econometrica, 49, 1057-72

Engle, R. F. and Granger, C. W. J. (1987) Cointegration error correction: represenation, estimation and testing, Econometrica, 55, 251–76.

Fama, E. F. (1984) Forward and spot exchange rates, Journal of Monetary Economics, 14, 319–38.

Fama, E. F. and French, K. R. (1987) Commodity futures prices: some evidence on forecast power, premiums and the theory of storage, Journal of Business, 60, 55–73.

Fuller, W. A. (1976) Introduction to statistical time series, John Wiley and Sons, New York

French, K. R. (1986) Detecting spot price forecasts in futures prices, Journal of Business, 59, S39–S54.

Hakkio, C. S. and Rush M. (1989(Market efficiency and cointegration: an applicati, on to the sterling and Deutschemark exchange markets, Journal of International Money and Finance, 8, 75-88

Lai, K. S. and Lai, M. (1991) A cointegration test for market efficiency, The Journal of Futures Markets, 11, 567–75.

Marchese M. (2017) Quantitative methods for finance, Time series analysis, modelling and forecasting

Phillips, P. C. B. and Perron, P. (1988) Testing for a unit root in time series regression, Biometrica 75, 335-46

Serletis, A. and Banack, D. (1990) Market efficiency and cointegration: an application to petroleum markets, Review of Futures Markets, 9, 372–85.

Serletis, A. and Scowcroft, D. (1991) International efficiency of commodity futures prices, Applied Financial Economics, 1, 185–92.

Staikouras S (2004) The information content of interest rate futures and time-varying risk premia, Applied Financial Economics, 761-771