What Drives Gold Returns? A Decision Tree Analysis

Anastasios G. Malliaris
Loyola University Chicago, tmallia@luc.edu

Mary Malliaris
Loyola University Chicago

What drives gold returns? A decision tree analysis

A.G. Malliaris [a,*], Mary Malliaris [b]

[a] Departments of Economics & Finance, Loyola University Chicago, 1 East Pearson Street, Chicago, IL 60611, United States
[b] Information Systems & Operations Management, Loyola University Chicago, 1 East Pearson Street, Chicago, IL 60611, United States

A R T I C L E   I N F O

A B S T R A C T

The behavior of gold as an investment asset has been researched extensively. For the very long run, that is several decades, gold does not outperform equities. However, for shorter periods, gold responds to fears of inflation, stock market corrections, currency crises and financial instabilities very vigorously. In this paper we follow a decision tree methodology to investigate the behavior of gold prices using both traditional financial variables such as equity returns, equity volatility, oil prices, and the euro. We also use the new Cleveland Financial Stress Index to investigate its effectiveness in explaining changes in gold prices. We find that gold returns depend on different determinants across various regimes.

© 2015 Elsevier Inc. All rights reserved.

1. Introduction

In a recent paper, Barro and Misra (2013) compute the average real rate of price change for gold in the U.S. from 1836 to 2011 and find that it is 1.1% per year with a standard deviation of 13.1%. They suggest that gold's expected real rate of return, which includes an unobserved dividend yield, would be close to the risk-free rate of return. This low average rate of return and its high volatility suggest that a very long-term passive investment in gold does not appear to be very favorable and shorter term strategies that identify economic variables that drive volatilities may be more profitable.

In this paper we give a rapid description of the general characteristics of the gold market and then propose to investigate the behavior of gold prices across the most recent business cycle which

includes the Global Financial Crisis. We identify on the basis of a selective review of the literature several economic and financial variables that influence the price of gold and trace the significance of these variables across the recent business cycle. To find what drives the behavior of gold returns within a business cycle and across its various phases we employ a decision tree analysis.

2. Review of the literature

Erb and Harvey (2013) report that there are about 171,300 metric tons of gold above ground. Given that there are 32,150 troy ounces per metric ton with a price of about $1400 per ounce, this yields an overall value of about $7.7 trillion. More than half of all humanity's gold has been extracted the last 50 years. The annual global gold production is about 2,500 metric tons per year and this supply is relatively price inelastic. In contrast, the demand for gold is price elastic and is driven by jewelry, industrial use, investment strategies and by inventory management in official institutions (central banks and the IMF). Supply and demand fundamentals determine gold prices and such prices have fluctuated in the past forty-five years from about $35 an ounce in 1970 to a record high of $1900 in 2010 to a relative low of $1183 in late 2013 and back up to $1400 in July of 2014. How can such large fluctuations be explained?

There is a large literature that identifies economic variables that explain the theoretical grounds that connect these variables to the fluctuations in the price of gold. Most of these studies also perform empirical tests to confirm or reject the hypothesized relationships. In particular, the set of variables used to correlate with gold returns includes inflation as measured by the Consumer Price index, various energy prices, interest rates for both the short-term 3-month Treasury Bill and the longer-term 10-year Treasury Note, foreign currencies as described in the US Dollar trade weighted exchange rate, equities as measured by the S&P 500 Index and its volatility measure by VIX. Representative such studies are listed in both Erb and Harvey (2012) and also Aggarwal et al. (2015).

Aggarwal et al. (2015) offer a detailed analysis of the world metal markets and discuss in particular the behavior of gold prices. The efficiency of gold markets is presented in Canarella and Pollard (1986) and recently in Caminschi and Heaney (2013). Cheung and Lai (1983) investigate the long cycles of gold returns. Fama and French (1988) examine the behavior of metal prices over the business cycle, and as with equities, Aggarwal and Soenen (1988) report that gold returns are leptokurtic and negatively skewed. Ciner et al. (2013) investigate the return relations between major asset classes using data from both the US and the UK to examine time variation in conditional correlations to determine when these variables act as a hedge against each other. A worth mentioning finding of this study is that gold can be regarded as a safe haven against exchange rates in both countries.

Malliaris and Malliaris (2009) observe that while gold has been an important commodity for several centuries, oil's importance grew during the 20th century, and the euro has become important during the 21st. This paper analyzes the inter-relationships among the price behavior of gold, oil and the euro using a standard time series methodology and then employs neural networks to build a forecast for each of these three variables. The authors also compare the results of the neural network to those implied by the time series tests. The statistical evidence of time series analysis demonstrates that both short-term and long-term relationships exist between the three variables. Both the time series and neural network results indicate that the series move together though they identify slightly different relationships. The time series results imply that oil adjusts to gold, the euro and oil have equal influences on each other, and the weakest relationship is between gold and the euro. The neural network methodology indicates that oil impacts gold more than gold impacts oil. Also, oil's influence on the euro is found to be greater than the euro's effect on oil and lastly, that gold's impact on the euro is greater and faster than the euro's impact on gold. Mensi et al. (2013) obtain results that show significant transmission among the S&P 500 and commodity markets. In particular, these authors show that past shocks and volatility of the S&P 500 strongly influenced the oil and gold markets.

Malliaris and Malliaris (2013) investigate inter-relationships among the price behavior of oil, gold and the euro using time series and neural network methodologies. Traditionally gold is a leading indicator of future inflation. Both the demand and supply of oil as a key global commodity are impacted by inflationary expectations and such expectations determine current spot prices. Inflation influences both short and long-term interest rates that in turn influence the value of the

dollar measured in terms of the euro. Certain hypotheses are formulated in this paper and time series and neural network methodologies are employed to test these hypotheses. The authors find that the markets for oil, gold and the euro are efficient but have limited inter-relationships among themselves.

Pierdzioch et al. (2014) study whether the international business cycle, as measured in terms of the output gaps of the G7 countries, has out-of-sample predictive power for gold-price fluctuations. These authors concluded that the international business cycle has predictive power for gold-price fluctuations. Finally, Białkowski et al. (2014) examines whether an asset bubble exists in the gold market during the recent Global Financial Crisis period and show that a model accounting for the current European sovereign debt crisis accurately tracks the gold price observed in the market.

This rapid and selective review suggests that the behavior of gold returns has not been fully determined. This does not mean that we know only very little. Aggarwal et al. (2015) documents numerous findings about the efficiency of the gold market, the distribution of prices, the role of gold as a hedge or a safe haven and the great volatility of its return from 2% over very long investment horizons to extremely high returns over shorter periods, such as from 1970 to 1975. It is the purpose of this paper to investigate the behavior of gold returns during the recent financial period because financial instabilities add further complications to the behavior of gold returns. To do this we identify four regimes determined by the most recent business cycle and for each one of these we let a decision tree analysis identify the variables driving the corresponding returns.

3. Data and methodology

The NBER gives the dates of the crisis as beginning with a peak in December 2007 and continuing to the bottom in June 2009. [http://www.nber.org/cycles.html] This is a period of 388 days. Our data set encompasses this crisis period and buffers on either side. It runs from March 1, 2004 through October 20, 2014, and is divided into four parts, with beginning and ending dates, and total size, as shown in Table 1. From the period of the crisis top and bottom, we move backward and forward a little over 900 days, then we have an additional period, about the same size as the crisis at the very end. This table also gives the name of each data set. Before refers to the 919 days prior to the crisis period; During takes us from the peak to the bottom of the crisis; After moves from the bottom of the crisis to March of 2013; lastly, Post continues for another 384 days afterward.

Our data has six base variables, Gold, the Cleveland Financial Stress Indicator [CFSI], Cushing Oil, the S&P 500, the VIX, and the Euro. See the Appendix A for a listing of all inputs into the CFSI. Data was downloaded from the St. Louis FRED database (DEXUSEU for the Euro, GOLDAMGBD228NLBM for Gold, DCOILWTICO for Cushing Oil) and the Cleveland Federal Reserve for the CFSI (http://www.clevelandfed.org/research/data/financial_stress_index/index.cfm). Values of the S&P 500 and the VIX were sourced from Yahoo finance. Using these, we calculated derived variables on each series including the percent change from yesterday to today, and the direction the series moved from the close of yesterday to the close of today. We also created a Target variable by shifting the Gold direction one day to give us the movement of Gold tomorrow. Thus, we have 18 input variables and one target [Goldtp1], as show in Table 2.

The Charts in Fig. 1 shows the relationship of Gold with each of the other variables. Each of the charts has the data scaled between 0 and 1 over the entire set. We see that, in the scaled values, the relationship Gold has with each of the variables varies over time.

Table 1
Data set details.

Set name	Begin date	End date	Count
Before	1-March-04	30-November-07	919
During	3-December-07	30-June-09	388
After	1-July-09	28-March-13	914
Post	2-April-13	20-October-14	384

Table 2
List of variables used in the models.

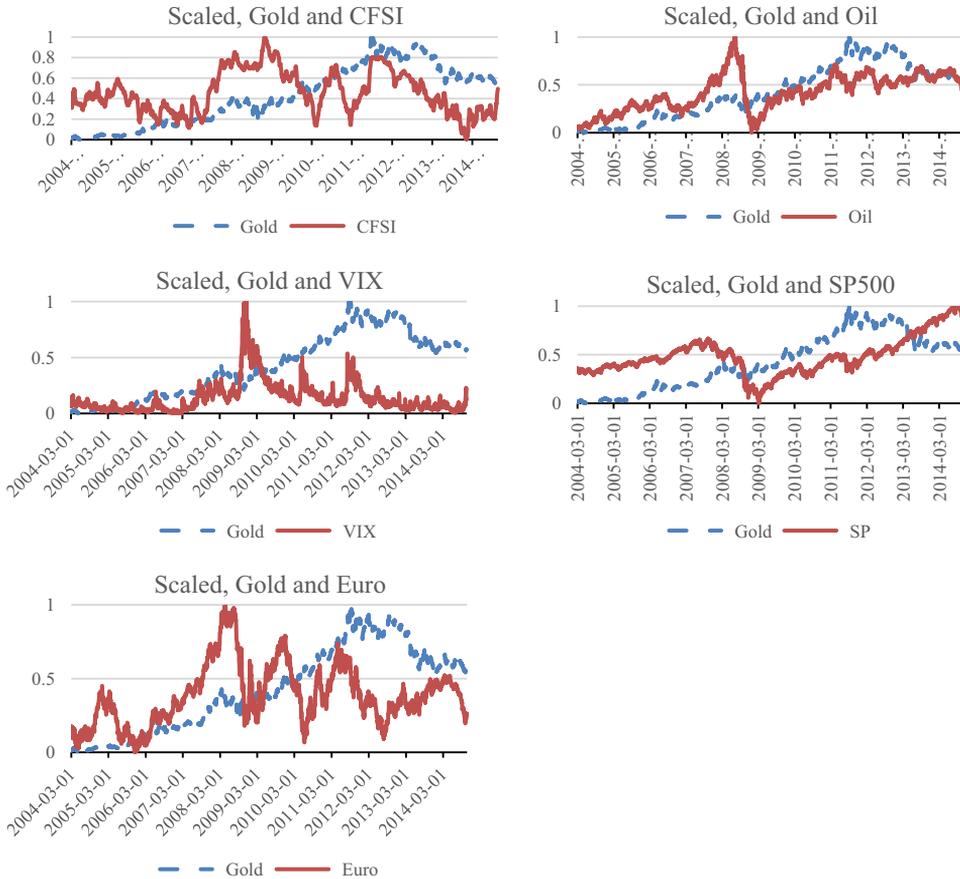| Gold	CFSI	Oil	SP	VIX	Euro
GChg	CFSIChg	Ochg	SPChg	VChg	Echg
GDir	CFSIDir	ODir	SPDir	Vdir	Edir
GDirtp1					

Fig. 1. Charts showing the relationship of Gold with each other variable.

In VIX, for example, prior to the crisis, the series intertwine and after a large peak in the VIX near the center, VIX drops to well below the scaled Gold values. Oil begins above Gold then plunges below and, after the crisis, stays close to, but below, Gold until early 2013 when these two again intertwine. The S&P 500 is above Gold for the initial period, then drops below where it remains until early 2013. At that point, the S&P climbs and stays above Gold. The Euro and the CFSI have more complicated relationships that cross over the scaled Gold prices numerous times. These shifts in relationships might indicate that a single model across the entire data set would not do as well as individual models might.

For each data set, we ran a C5.0 decision tree model using IBM's SPSS Modeler data mining software package. A decision tree is a classification model that, by examining rows in a training set, builds rules

that assign each row in the data set to a predefined class. These classes are the values of the target variable (Gold's direction tomorrow) which here consist of Up and Down. These models are based on Hunt's algorithm (Tan et al., 2006) and are derivatives of the original ID3 model (Han et al., 2012). The algorithm recursively partitions the data set into smaller sets, based on values of the input variables, which are more likely to have a single value of the target field. At each node of the tree, the variable is selected that gives the most pure (that is, most single-valued on the target variable) results in the following node. The splits are called branches, and each branch has a node that contains a subset of the initial data set. The tree stops growing either when all final nodes are single valued on the target, or when there is no remaining variable that can be applied to give a more pure split.

The results of the decision tree can be inspected in several ways. First, a matrix showing the counts of correct and incorrect classifications can be generated. Second, we can look at the variable importance chart. Last, we can inspect the shape of the tree and the rules. If the input rows on which the decision tree is trained are such that the rows do not give conflicting results (conflicting results occur when two rows are identical except for the value of the target) then the decision tree should be able to correctly assign the rows to their respective classes, and the matrix will show us higher percentages of correct classifications. The variable importance output shows the inputs that were used by the decision tree with their respective importance to the classification process. These weights are comparative and sum to 1. It is of interest to see, among multiple models, which variables move in and out of prominence. When looking at the shape of the tree, a deeper tree with more layers uses more variables in the classification process. A wider tree uses more splits of a single variable at a given level.

4. Results

A C5.0 decision tree was trained for each of the four data sets. Table 3 displays these results with the decision tree classifications for each data set in the columns. The left-most column of the table gives the actual direction that Gold moved on day $t + 1$. The interior cells contain the count and percentages for the intersection of the actual and decision tree directions. We see that, in the data set Before the crisis, that the decision tree classified 382 rows as Down for Gold tomorrow that did in fact move Down. The tree also classified 26 rows as Down that actually moved Up. Also in the Before data set, 35 rows were classified as Up that in fact moved Down, and 475 rows were classified as Up that did move Up the next day. The percentages of correct classifications for each of the data sets are shown in bold type. Decision tree classifications are correct over 90% of the time in 3 of the 4 sets. In the After set, correctness ranges from 86% to 89%. These results indicate that the decision trees were able to classify rows as indicating the Up or Down movement of Gold tomorrow in an accurate way.

The next step is to compare the importance of the specific variables to each of the models. Table 4 lists these relative importance values. For each model, the values sum to 100 and only those variables which were used have values assigned. Variables with values of 10 and over are shown as shaded and in bold font. We see that, in the Before model, much of the decision trees ability to classify correctly was due to the percent change in the Euro, the direction the Euro moved, and the percent change in Oil. During the crisis, the direction oil moved was most important, followed closely by the percent change in the CFSI and the Euro, and the price of Gold. After the crisis the CFSI and the change in the VIX dominate the variable importance, with the direction of the S&P also contributing. In the last period, we see the direction of the CFSI as most important with the percent change in Oil very close.

Table 3
Decision tree classification results.

GoldDirTp1		Before		During		After		Post	
		Down	Up	Down	Up	Down	Up	Down	Up
Down	Count	382	35	178	8	356	73	186	13
	Column%	**93.6**	6.8	**92.7**	4.1	**89.4**	14.1	**91.6**	7.2
Up	Count	26	476	14	188	42	443	17	168
	Column%	6.4	**93.2**	7.3	**95.9**	10.6	**85.9**	8.4	**92.8**

The percentages of correct classifications for each of the data sets are shown in bold type.

Table 4
Variable importance per model.

Variable	Before	During	After	Post
CFSI		9	**22**	**12**
CFSIChg		**15**		5
CFSIDir	2			**20**
Euro		9	6	1
Edir	**15**			
Echg	**56**	**13**	6	
Gold		**12**		6
GChg		3	2	
GDir		1		1
Oil		1	6	
Ochg	**13**			**19**
ODir		**18**	7	
SP		5		
SPChg				2
SPDir			**14**	13
VIX	9	4	6	1
VChg	5	9	**22**	5
Vdir		1	9	**15**

The VIX and the S&P directions and the values of the CFSI also contribute. While Table 4 shows the individual influence of each variable, Table 5 sums these values across each base variable and it derived variables. Thus CFSI in Table 5 represents the sum of CFSI, CFSIChg, and CFSIDir in Table 4.

Table 5 allows us to see more easily the impact of each of the base variables. We notice that the Cleveland Financial Stress Indicator is gaining more prominence while the Euro is dropping in impact. Gold had its largest influence during the crisis, and the pattern for Oil seems to repeat. The S&P shows a slighter increase, and the VIX peaked in the After period.

Finally, we have the shapes of the trees. Fig. 2 shows the shapes of each of the trees. The Before and After trees appear slightly more complex, and the During and Post trees have fewer layers. A less complex tree indicates that the algorithm was able to classify the data more quickly using fewer variables. However, the two sets on the right are smaller than those on the left, so this may account for the simpler structure.

Each shape is generated by the rules associated with the paths. For example, in the Before model, one shorter set of paths of the tree is shown below (see Fig. 3). A path reads from its beginning in the upper left and ends with either "=> Down" or "=> Up", indicating the direction that Gold will move tomorrow. The first path states "If Echg is less than or equal to .001 and less than or equal to -.008 and Oil is less than or equal to 61.89, then Gold will be Down tomorrow. This path branches on the value of Oil, and offers an alternative for Oil greater than 61.89 by adding the conditions that if GDir is Down and VIX is less than or equal to 16.03 then Gold will be Up tomorrow. These paths can be seen on the upper left area of the Before tree shape figure.

The total number of rules for each direction is shown in Table 6. We see that the smaller time periods generate fewer rules, but the number of rules is not proportional to the set size.

Table 5
Variable importance summed by category.

	Before	During	After	Post
CFSI	2	24	22	37
Euro	56	22	12	1
Gold	0	16	2	6
Oil	13	19	13	19
SP	0	5	14	15
VIX	14	14	37	20

Fig. 2. Tree structures.
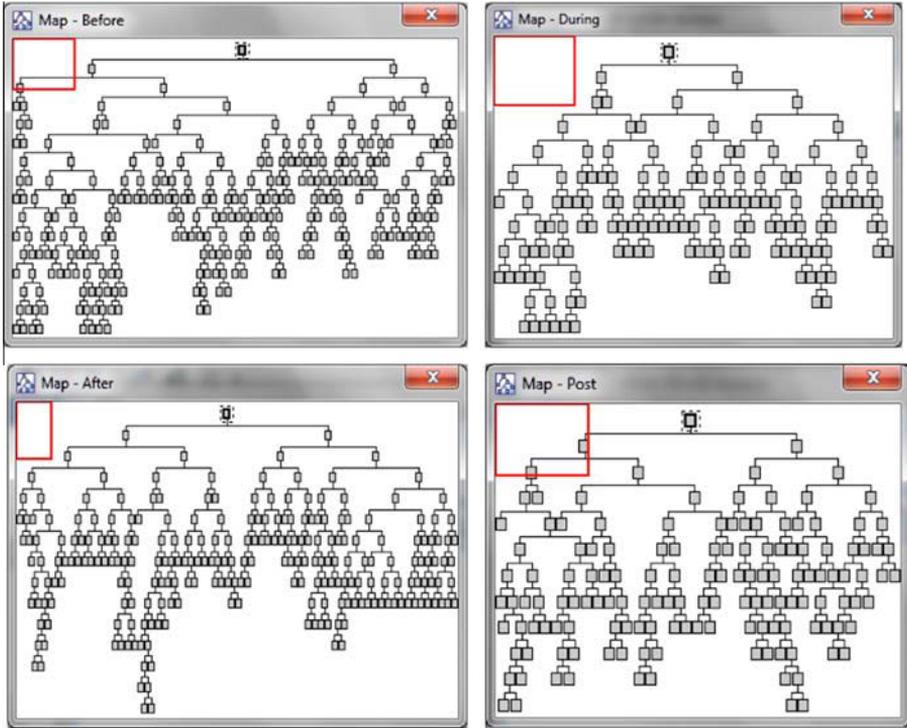
```
Echg <= 0.001
     Echg <= -0.008
          Oil <= 61.890 => Down
          Oil > 61.890
               GDir = Down
                         VIX <= 16.030 => Up
                         VIX > 16.030 => Down
               GDir = Up => Down
```

Fig. 3. One set of paths in the Before tree.

Table 6
Number of Down and Up Rules
in each Set.

	Down	Up
Before	68	68
During	36	36
After	62	61
Post	35	37

Overall then, we have four data sets, two larger and two smaller. For each time period, the decision tree built for that period is able to do a good job in classifying the direction that Gold will move the next day. However, we see that each decision tree values the input variables in different ways. These shifts over time are most apparent in Table 5 where the impact of each base variable and its derived counterparts are summed.

5. Conclusions

Gold has been described as the "golden constant" because it has behaved as a hedge against inflation in a very long investment horizon. So, real gold returns over, say, 100 years are zero. Real returns for equities over a similar very long investment horizon are much higher, around 7%. A fundamental question is: how did gold returns behave during the recent financial crisis when inflation was very low?

In this paper we use daily data of several variables from March 1, 2004 through October 20, 2014. The variables considered are: Gold, the Cleveland Financial Stress Indicator [CFSI], Cushing Oil, the S&P 500, the VIX, and the Euro. We use the National Bureau of Economic Research dates to identify 4 periods: before the crisis, the recession, the recovery and the post recovery described in Table 1. We employ a decision tree methodology and obtain very interesting results summarized in Table 5. In particular we find that the 3 most important variables for the BEFORE period are the Euro, VIX and Oil; for the DURING period are the CFSI, Oil and the Euro; for the AFTER period are the VIX, CFSI and the S&P 500 Index and for the POST period are the CFSI, VIX and Oil. The implication of our results is that gold has become an important asset whose return is driven by portfolio allocation strategies influenced by the regime changes in the economy.

Appendix A

Cleveland Financial Stress Index Components.

Markets	Component	Description
Credit	Covered Interest Spread	Measures uncertainty about government bond markets
	Corporate Bond Spread	Measures the broad perceptions of medium- to long-term risk in corporations of all sectors
	Liquidity Spread	Measures changes in the bid and ask prices on three-month Treasury bills, which reflects liquidity in financial markets
	Commercial Paper and T-bill Spread	Measures the short-term risk premium on financial companies' debt
	Treasury Yield Curve Spread	Measures the likelihood of recession because it captures long-term uncertainty and short-term liquidity
Equity	Stock Market Crashes	Measures the extent to which equity values in the S&P 500 financial Index have dropped over the previous year. It also captures expectations about the future of the banking industry
Foreign exchange	Weighted Dollar Crashes	Measures flight from the U.S. dollar toward a broad set of foreign currencies
Funding	Financial Beta	Measures the contribution of the banking sector to overall stock market volatility
	Bank Bond Spread	Measures the broad perceptions of medium- to long-term risk in banks issuing A-rated bonds

Appendix A (*continued*)

Markets	Component	Description
	Interbank Liquidity Spread	Measures the perception of counterparty risk in interbank lending
	Interbank Cost of Borrowing	Measures the degree of apprehension with which banks loan to one another
Real Estate	Commercial Real Estate	Measures the risk associated with investing in commercial real estate
	Residential Real Estate	Measures the risk associated with investing in residential real estate
Securitization	Residential MBS	Measures the ability of agencies to raise capital and relative riskiness of the securitized asset
	Commercial MBS	Measures the ability of originators to raise capital and relative riskiness of the securitized asset
	Asset-Backed Securities	Measures the ability of originators to raise capital and relative riskiness of the securitized asset

References

Aggarwal, Raj., Soenen, L., 1988. Nature and efficiency of the gold market. J. Port. Manage. 14, 18–21.

Aggarwal, Raj, Lucey, Brian, O'Connor, Fergal, 2015. World metal markets. In: Malliaris, A.G., Ziemba, W. (Eds.), World Scientific Handbook of Futures Markets. Hackensack, New Jersey, pp. 385–410.

Barro, Robert, Misra, Sanjay, 2013. Gold returns, NBER Working Paper No. 18759.

Białkowski, J., Bohl, M.T., Stephan, P.M., Wisniewski, T.P., 2014. The gold price in times of crisis. Int. Rev. Financ. Anal. http://dx.doi.org/10.1016/j.irfa.2014.07.001.

Caminschi, Andrew, Heaney, Richard, 2013. Fixing a leaky fixing: short-term market reactions to the London pm gold price fixing. J. Futures Markets. http://dx.doi.org/10.1002/fut.21636.

Canarella, G., Pollard, S.K., 1986. The efficiency of the London metal exchange: a test with overlapping and non-overlapping data. Journal of Banking & Finance 10, 575–593.

Cheung, Y., Lai, K., 1983. Do gold market returns have long memory? Financial Rev. 28, 181–202.

Ciner, C., Gurdgiev, C., Lucey, B.M., 2013. Hedges and safe havens: an examination of stocks, bonds, gold, oil and exchange rates. Int. Rev. Financial Anal. 29, 202–211.

Erb, Claude B., Harvey, Campbell R., 2013. The Golden Dilemma. Available at SSRN:.

Fama, Eugene, French, Kenneth, 1988. Business cycles and the behavior of metals prices. J. Finance 43, 1075–1093.

Han, J., Kamber, M., Pei, J., 2012. Data Mining Concepts and Techniques, third ed. Morgan Kaufmann Publishers, Waltham, MA.

Malliaris, A., Malliaris, M.E., 2009. Time series and neural networks comparison on gold, oil and the euro. In: Proceedings of the International Joint Conference of IEEE on Neural Networks, pp. 1961–1967.

Malliaris, A., Malliaris, M.E., 2013. Are oil, gold and the euro inter-related? Time series and neural network analysis. Rev. Quantitat. Finance Account. 40, 1–14.

Mensi, W., Beljid, M., Boubaker, A., Managi, S., 2013. Correlations and volatility spillovers across commodity and stock markets: linking energies, food, and gold. Economic Modelling 32, 15–22.

Pierdzioch, C., Risse, M., Rohloff, S., 2014. The international business cycle and gold-price fluctuations. Quart. Rev. Econ. Finance 54, 292–305.

Tan, P., Steinbach, M., Kumar, V., 2006. Introduction to Data Mining. Pearson Education Inc.