**How Predictable Are Precious Metal Returns?**

**Abstract**

This paper provides strong evidence of time-varying return predictability of three precious metals from January 1987 to September 2014. We use three variations of the variance ratio test, the nonlinear BDS test as well as the Hurst exponent to evaluate the time-varying return predictability of precious metals to reduce the risk of spurious results. Our full sample results report mixed findings where some tests indicate significant predictability while some suggest no predictability. However through a time-varying procedure, we show that each precious metal market goes through periods of significant predictability as well as periods of unpredictability. Therefore this finding suggests that return predictability does vary over time and is not a static, all-or-nothing condition and therefore is consistent with the Adaptive Market Hypothesis. We also show that platinum is the most predictable of the three precious metals and silver the least predictable, which may be of great to investors who include precious metals in their investment portfolios.

Keywords: Precious Metals; Predictability; Adaptive Market Hypothesis; Market Efficiency

*JEL* *classification:* G14; G15

**1. Introduction**

This paper extends the growing literature on precious metals by providing a detailed examination of the predictability of gold, silver, and platinum through a variety of testing procedures. We study the predictability of daily precious metals returns over the January 1987 to September 2014 period, as well as a rolling subsample to study how the degree of predictability changes over time, which also enables an examination of whether the Adaptive Market Hypothesis (AMH) is an appropriate model for the behaviour of precious metals returns. We show evidence of the time-varying predictability of each precious metal, which is consistent with the AMH and that silver is the most efficient of the three precious metals while platinum is the least efficient.

Our paper is motivated by the increased attention of precious metals in the finance literature over the last few years (for instance Aggarwal et al 2014; Batten et al 2014; Baur and Glover 2014; Lucey and Li 2014; Lucey et al 2014; Areal et al 2015; Aye et al 2015; Bampinas and Panagiotidis 2015; Baur and Glover 2015; Białkowski et al 2015; Bredin et al 2015)[[1]](#footnote-1). Given the increased attention of precious metals in the literature, there is a lack of detailed studies examining the predictability[[2]](#footnote-2) of the precious metals returns. This is an important area to study since the efficient market hypothesis (EMH) states that prices should reflect all available information, suggesting that returns are purely unpredictable from last prices (Fama 1970). However if precious metals returns are found to have predictive power, investors may be able to take advantage of these predictabilities, which in turn would be a violation of the EMH. Thus the predictability of precious metals is of great interest to academics and investors alike.

To accommodate the idea of a changing degree of market efficiency, Andrew Lo (2004), proposes a new version of the EMH derived from evolutionary principles. Lo argues that valuable insights can be derived from the biological perspective and calls for an evolutionary alternative to market efficiency. The paradigm is called the Adaptive Market Hypothesis (AMH) under which the EMH and market inefficiency can co-exist in an intellectually consistent manner. Lo (2005) states that individuals act in their own self-interest, but they make mistakes. They learn from these mistakes and adapt, and that competition drives adaptation and innovation. Finally evolution determines market dynamics. The AMH provides a number of practical implications within finance. Firstly, the risk premium varies over time according to the stock market environment and the demographics of investors in that environment. The second implication is that arbitrage opportunities do exist from time to time in the market. Thus from an evolutionary viewpoint, active liquid financial markets imply that profit opportunities must exist. However as they are exploited, they disappear. But new opportunities are continually being created as certain species/ traders die out and rather than move towards a higher degree of efficiency the AMH implies that complex market dynamics such as trends, panics, bubbles and crashes are continually witnessed in natural market ecologies. The third implication is that investment strategies are successful or unsuccessful, depending on the particular market environment. Contrary to the EMH, the AMH implies that investment strategies may decline for a time, and then return to profitability when environmental conditions become more conducive to such strategies. A consequence of this implication is that market efficiency is not an all-or-nothing condition, but is a characteristic that varies continuously over time and across markets. Lo (2005) argues that convergence to equilibrium is neither guaranteed nor likely to occur and that it is incorrect to assume that the market must move towards some ideal state of efficiency.

The AMH has been examined in some detail in the recent empirical literature with Lim (2007) studying eleven emerging and two developed markets through the portmanteau bicorrelation test. Using a rolling sample framework, the paper shows that market efficiency evolves over time in a way consistent with the AMH. Neely et al (2009) use daily exchange rate data to study the behaviour of technical trading rules and show that the returns in well studied rules have disappeared and declined in less-studied rules, consistent with the AMH. Kim et al (2011) implement an automatic variance ratio test as well as the automatic portmanteau test to examine the predictability of the DJIA over time. Using a rolling window framework, they find strong evidence of time-varying predictability that is driven by market conditions. Charles et al (2012) study the return predictability of major foreign exchange rates using daily and weekly rates from 1975 to 2009 and find a number of episodes of statistically significant return predictability. They also find that the returns predictability occurs from time to time depending on changing market conditions, consistent with the AMH. Urquhart and Hudson (2013) implement linear and nonlinear techniques to study the US, UK and Japanese markets using long run historical data and conclude that the AMH provides a better description of the behaviour of stock returns than the EMH. Hull and McGroarty (2014) study 22 emerging markets over a 16-year period through the Hurst-Mandelbrot-Wallis rescaled range as a measure of market efficiency and find strong evidence consistent with the AMH. Ghazani and Araghi (2014) investigate the daily returns of the Tehran stock exchange from 1999 to 2013 and show that the AMH gives an appropriate evolutional perspective of market efficiency. Further, Manahov and Hudson (2014) develop artificial stock markets using a special adaptive form of the STGP based learning algorithm and apply it to the FTSE100, S&P500 and Russell 3000. They show that the stock market dynamics are consistent with the AMH since the trader population behave in an efficient adaptive system evolving over time. Urquhart and McGroarty (2014) study four well-known calendar anomalies from 1900 to 2013 in the DJIA. Using subsample as well as rolling window analysis, they show that each calendar anomaly behaves in a manner consistent with the AMH and that some of the calendar anomalies are only present during certain market conditions. Also, Levich and Poti (2015) study the predictability in currency markets over the period 1972-2012 by constructing an upper bound on the explanatory power of predictive regressions of currency returns. They find that currency predictability exceeds this bound during recurring albeit short-lived episodes and that excess predictability is highest in the 1970s and tends to decrease over time but is still present in the final part of the sample period, thus providing evidence of the AMH. Furthermore, Urquhart et al (2015) study the well-known moving average technical trading rule and show its performance has weakened since the seminal study documenting its success, thus providing evidence of the AMH. They also show that if investors traded on anticipated signals, that is, they predict one-day-ahead signals, they can earn superior profits which is consistent with the learning and adaptation of the AMH.

In this study, we examine the predictability of gold, silver and platinum over time over the period January 1987 to September 2014 over the full sample period as well the time-varying level of predictability using a rolling window analysis. We use three different formulations of the variance ratio test, the nonlinear BDS test as well as the rescaled Hurst exponent to provide a detailed and robust overview of the predictability of these precious metals. Each of our testing procedures examines the level of predictability of returns in a different manner (more information in Section 3) and by employing a battery of tests we shall capture the main dynamics of the metals returns in several dimensions while also reducing the risk that a spurious result from one test may affect the conclusions. Therefore we contribute to the literature in several ways. Firstly, this paper provides a comprehensive examination of the linear and nonlinear predictability of precious metals returns through a number of statistical tests. There has been lack of studies examining the predictability of precious metals given their increased attention in recent years. Precious metals have been shown to be important components of investment portfolios mainly due to their hedge or safe haven properties (see for example (Baur and Lucey 2010; Baur and McDermott 2010; Hood and Malik 2013; Beckmann et al 2014; Mensi et al 2015). It has also been shown that precious metal markets are heavily influenced by economic and political conditions (see Batten et al 2010; Hood and Malik 2013; Areal et al 2015) and therefore it is very likely that the degree predictability will vary over time. Secondly, this study uses a range of tests, thus capturing the main dynamics of the precious metals returns in several dimensions, thus reducing the risk that a spurious result from one test may skew our conclusions. The BDS test also allows us to examine nonlinear predictability[[3]](#footnote-3) by filtering returns through an AR-GARCH model to remove any linear dependence due to conditional heteroscedasticity. Thirdly, we use rolling window analysis which gives a gradual description of the varying predictability over time which will not be distorted by arbitrary chosen subsample windows, which is the case with subsample analysis. Subsample analysis has the major issue of subsample selection bias, which can skew the results. Fourthly, this is only the second study to examine whether precious metal markets can be described by the AMH. Charles et al (2015) examine the predictability of precious metals through the automatic portmanteau and automatic variance ratio tests while we use three parametric and non-parametric variance ratio tests as well as the nonlinear BDS test. Nonlinear tests have gained much attention in the recent literature (see for example Lim and Hooy 2012; Urquhart and Hudson 2013) since it has been shown by Amini et al (2010) that a series can still exhibit strong nonlinear correlations even in the absence of linear dependence.

The remainder of the paper is organized as follows. The next section presents the literature review while Section 3 reports the methodology used in this study. Section 4 presents the data and Section 5 reports the empirical results. Finally Section 6 summarizes and provides conclusions.

**2. Literature Review**

The number of studies examining the efficiency of precious metals is growing, with Tschoegl (1980) testing the serial correlation in the gold market and modelling the changes as first-order Markov processes. The paper indicates some short-term dependence and that gold’s alpha is positive but insignificant. Solt and Swanson (1981) study the efficiency of gold and silver and find evidence of positive dependence in returns, but state that investors cannot exploit the dependency easily. Aggarwal and Soenen (1988) show that gold is not efficient, while Akgiray et al (1991) examine the time-series properties of gold and silver spot prices and find that both exhibit time dependence and GARCH effects. Lashgari (1992) studies the behaviour of gold and silver prices from January 1970s to December 1989 using daily, weekly and monthly time periods and show that information contained in the past prices of gold and silver does not allow one of predict next-period changes in prices in the short run. However, they show that longer-term predictions are possible and that gold exhibits a higher degree of dependency on past prices than silver. Cheung and Lai (1993) study the long memory behaviour of gold returns during the post-Bretton Woods period using the rescaled range Hurst technique and find that the long memory behaviour in gold returns is unstable and when major political events in the Middle East and the Hunts event in 1979 are omitted, little evidence of long memory is found. Smith (2002) tests for the random walk using the multiple variance ratio test and find that the twice-fixing prices do not conform to the random walk, while the closing price does follow a random walk.

Baur (2013) studies the autumn effect in monthly gold prices from 1980 to 2010 and find strong evidence of the effect and argue that it can be explained by hedging demand by investors in anticipation of the Halloween effect, wedding season gold jewellery demand in India and negative investor sentiment due to shorter daylight time. Arouri et al (2013) show that in the short and long run that future prices of precious metals do not constitute an unbiased predictor of future spot prices, suggesting that investors can design investment strategies based on past information to forecast future spot prices. Pierdzioch et al (2014) study whether publicly available information helps in real-time forecasting monthly excess returns on investing in gold. They show that using forecasts implied by the real-time forecasting approach to set up simple trading rules does not necessarily lead to a superior performance relative to a buy-and-hold strategy, implying that the gold market is informationally efficient with respect to the variables used in this study. Recently, Charles et al (2015) study the time-varying predictability of precious metals through the automatic portmanteau and variance ratio tests and show that return predictability of gold and silver does vary over time depending on the prevailing economic and political conditions. They also show that the predictability of gold and silver has been showing a downward trend, implying the degree of weak-form efficiency has been gradually improving. Also, Ntim et al (2015) study the weak-form efficiency of global gold markets by focussing on the random walks and martingale difference sequence hypotheses, by applying variance ratio tests of daily spot prices of 28 emerging and developed gold markets and January 1968 to August 2014. They find that while some of the markets are not weak-form efficient in respect of the random walk hypothesis but are for the martingale difference hypothesis, some are efficient for both hypotheses, and that some reject efficiency completely. They also show that the probability of rejecting weak-form efficiency is higher in emerging markets than developed ones and that greater changes in economic fundamentals are associated with lower levels of rejecting weak-form efficiency.

**3. Methodology**

In this paper, we study the dependency of stock returns from a statistical viewpoint, where persistent dependencies could potentially we exploited using a variety of trading strategies. If good estimates of future price levels are possible, conventional long/short investment strategies may be appropriate. Other moments of the underlying price distribution uncovered by nonlinear tests may also be predictable and this information can be potentially exploited using derivatives strategies.

*3.1. Variance Ratio*

Since the seminal work of Lo and MacKinlay (1988), the variance ratio (VR hereafter) test has emerged as a primary tool in examining whether stock returns are serially uncorrelated, with Hoque et al (2007) stating that it has become the most commonly used econometric tool for testing the random walk hypothesis. The VR test is based on the statistical property that if a stock price follows a random walk (and therefore are unpredictable), then the variance of the *k*-period return is equal to *k* times the variance of the one period return. Lo and MacKinlay (1988) provide a test for this hypothesis using the single VR test, denoted by VR(*k*). Let *rt* denote an asset return at time *t*, where *t* = 1,2,3….*T*. Then the variance ratio for *rt,* with holding period *k* is;

|  |  |  |
| --- | --- | --- |
|  |  | (1) |

Where  *= Variance(rt + rt-1 +….+ rt-k+1)* is the variance of *k*-period return. It can be rewritten as;

|  |  |  |
| --- | --- | --- |
|  |  | (2) |

Where is the autocorrelation of *rt* of order *j*. That is, the variance ratio is one plus a weighted sum of autocorrelation coefficients for the asset returns with positive and declining weights. The VR tests the null hypothesis that the variance ratio equals 1 for all *k*s since returns are serially uncorrelated with = 0. Alternatively, values for VR(*k*) greater than 1 imply positive serial correlations while values less than 1 imply negative serial correlations or mean reversion.

Lo and MacKinlay (1988) determined the asymptotic distribution of VR(*x*; *k*) by assuming that *k* is fixed when T → ∞. They showed that if *x*t is i.i.d., i.e. under the assumption of homoskedasticity, then under the null hypothesis that VR(*k*) = 1, the test statistic M1(*k*) is given by;

|  |  |  |
| --- | --- | --- |
|  |  | (3) |

which follows the standard normal distribution asymptotically. The asymptotic variance, , is given by;

|  |  |  |
| --- | --- | --- |
|  |  | (4) |

To accommodate the returns exhibiting conditional heteroscedasticity, Lo and MacKinlay (1988) proposed the heteroscedasticity robust test statistic M2(*k*);

|  |  |  |
| --- | --- | --- |
|  |  | (5) |

which follows the standard normal distribution asymptotically under the null hypothesis that VR(*k*) = 1, where;

|  |  |  |
| --- | --- | --- |
|  |  | (6) |

|  |  |  |
| --- | --- | --- |
|  |  | (7) |

The M2(*k*) test is applicable to returns of a price series and this study utilises M2(*k*) due to the heteroscedastic property of the returns series’ studied, as revealed in Table 1.

An issue with the traditional VR test is that under the random walk hypothesis, we must have VR(*k*) = 1 for all chosen values of *k.* As the null hypothesis of the VR test is rejected if any value of *k* rejects its own null hypothesis, implying that a sequential procedure of testing several *k* values leads to an oversized testing strategy (Borges 2010). To account for this Chow and Denning (1993) propose a multiple VR test where only the maximum absolute value of VR(*k*) in a set of *m* test statistics is considered. The Chow-Denning test statistics is defined as;

|  |  |  |
| --- | --- | --- |
|  |  | (8) |

And it follows the studentized maximum modulus (SMM) distribution with *m* and *T* degrees of freedom. One of the difficulties with the VR test is that the statistics are based on asymptotic theory making the statistical inference misleading in small samples (Richardson and Stock 1989). To overcome this problem, a wild bootstrap method for the VR test proposed by Kim (2006) is used for the Chow-Denning statistic.

Further, Wright (2000) proposes a non-parametric alternative to the conventional VR test using ranks and signs that overcome the problems of biased and right-skewed samples. These two tests can be more powerful than the traditional VR test since they have high power against a wide range of models displaying serial correlation, the signs-based test is exact even under conditional heteroscedasticity and the ranks-based test displays low-size distortion under heteroscedasticity. Given log returns as *y*t and r(*y*) be the rank of *y*(r) among (*y*1, … , *y*t) which, under the hypothesis that *y*t is i.i.d, is just a random permutation of the numbers 1,2, …, T, each with equal probability. Define the rank based VR tests R1 and R2 as (for *i* = 1 or 2);

|  |  |  |
| --- | --- | --- |
|  |  | (9) |

Where

|  |  |  |
| --- | --- | --- |
|  |  | (10) |

|  |  |  |
| --- | --- | --- |
|  |  | (11) |

Where φ-1 is the inverse of the standard normal cumulative distribution function. The test based on signs of the first difference is given by;

|  |  |  |
| --- | --- | --- |
|  |  | (12) |

Where *s*t = 2*u*(*y*t, 0) and *u*(*y*t, 0) is ½ if *y*t is positive and -1/2 otherwise. Under the assumption that *y*t is generated from martingale difference sequence with no drift, *s*t is an i.i.d. sequence with zero mean and unit variance and the critical values can be obtained by simulating its sampling distribution. Similar to the Chow-Denning statistic, we construct a joint variance ratio test for ranks and signs as proposed by Belaire-Franch and Contreras (2004).

*3.2. BDS Test*

The BDS test, proposed by Brock, Dechert and Scheinkman (1987), is a popular non-parametric test for serial dependence (or a nonlinear structure) in stock returns. The null hypothesis of this test is that the data generating processes are i.i.d., while the alternative hypothesis is “an indication that the model is misspecified” (Brock et al 1996). That is, Given a sample of i.i.d. observations, {*xt: t = 1,2, . . , n*}, Brock et al (1996) show;

|  |  |  |
| --- | --- | --- |
|  |  | (13) |

Where *Wm,n(ε)* is the BDS statistic, *n* is the sample size, *m* is the embedding dimension and the metric bound (*ε*) is the maximum difference between pairs of observations counted in computing the correlation integral. *Tm,n(ε)* measures the difference between the dispersion of the observed data series in a number of spaces with the dispersion that an i.i.d. process would generate in these spaces (*Cm,n(ε) – C1,n(ε)m*) and has an asymptotic normal distribution with zero mean and variance *V2m(ε).* The asymptotic distribution of the BDS test does not depend on the existence of higher-order unconditional moments. As Hsieh (1991) points out, structural changes in the data series can cause a rejection of the null hypothesis of i.i.d. on the basis of the BDS test. Thus it is rational to break up the sample period and examine subsamples separately. The choice of *ε* and *m* values can be problematic since too small a *ε* will capture too few points so we follow the common approach in the literature by setting *ε* as a proportion of the standard deviation of the data. With regards to *m*, we follow the literature by setting *m* from 2 to 5 due to the small sample properties of the BDS test degrade as *m* increases (Patterson and Ashley 2000).

Since the BDS test finds any dependency in returns, we need to remove the linear dependency in returns and use the BDS test on the residuals of the series to study if any nonlinear dependency is present. Therefore, we fit an AR(*p*)-GARCH(1,1) model to the data to remove all linear correlations and any dependence due to conditional heteroscedasticity. The specification of model is chosen where the standardised residuals are no longer correlated at lag 10 through the Ljung-Box *Q*-statistic, similar to Lim and Hooy (2013). The standard residuals are then examined through the BDS test to determine if any nonlinear dependency is present.

*3.3. Hurst Exponent*

We follow Cajueiro and Tabak (2004a; 2004b) in calculating a rolling rescaled range in order to analyse the predictability of markets over time. For each subsample, the rescaled range (R/S) exponent is calculated over time and is calculated as;

|  |  |  |
| --- | --- | --- |
|  |  | (14) |

where is the returns over period *j*, is the mean return, and is the standard deviation of returns given by;

|  |  |  |
| --- | --- | --- |
|  |  | (15) |

As Wang et al (2011) note, the greatest drawback of the rescaled method is the lack of natural significance test as the value of *H* will almost always deviate from 0.5. However Qian and Rasheed (2004) use Monte Carlo simulations to generate a range for which there is only weak evidence of persistence (0.5 < *H* < 0.65), whereas a *H* greater than 0.65 suggests strong evidence of persistence. Conversely, a Hurst exponent below 0.45 shows strong evidence for anti-persistence and mean reversion, while a Hurst exponent between 0.45 and 0.5 suggests weak evidence for anti-persistence. Therefore we suggest any Hurst exponent greater than 0.65 and less than 0.45 indicates significant predictability in the precious metals returns, similar to Hull and McGroarty (2014).

Therefore we employ a range of testing procedures to examine the predictability of precious metals. While the variance ratio tests examines whether the variance of the k-period return is equal to k times the variance of the one period return, the BDS tests examines the nonlinear dependence once the data has been filtered to remove all linear dependence. The Hurst exponent however examines the long-term memory of the price series to determine whether prices are persistent, anti-persistent or random. Therefore all three testing procedures examine different aspects of predictability and will therefore provide a comprehensive position of the predictability of precious metals.

*3.4. Time-Varying Predictability*

To obtain measures of time-varying predictability, we utilise the moving-subsample window of a fixed length over the grid of months similar to Kim et al (2011) and Urquhart and McGroarty (2014). We use a five-year window and calculate the test statistics using data from the first trading day in January 1987 to the last trading day of December 1991, and then move the window forward one-month to cover the period February 1987 to January 1992. We continue this process to the end of the data and obtain measures for predictability of returns up to September 2014, which generates 274 windows. We choose 5-yearly windows to provide enough observations to generate reliable results, while at the same time providing enough results to analyse how the precious metals level of predictability has behaved over time.

**4. Data**

The sample data consists of daily closing prices for gold, silver and platinum. The data period spans 5th January 1987 to 30th September 2014 and is chosen as the longest data period available, which generates 7236 return observations. The data is obtained from Thomson Financial Datastream and Figure 1 presents the time-plots of the three precious metals prices and Figure 2 depicts the returns of the three precious metals where returns are calculated by;

|  |  |  |
| --- | --- | --- |
|  |  | (16) |

where is the natural logarithm of the index at time *t* and is the natural logarithm of the index at time *t-1*. We can see from Figure 1 that all three precious metals volatility increases substantially since 2005 and that silver has increased in value the most. We can also see that in general, the precious metals have moved together over time except silver since 2010. The returns series’ shown in Figure 2 show strong evidence of volatility clustering at the same time during the sample period for each precious metal.

[Insert Figures 1 & 2 here]

Table 1 presents the descriptive statistics of the three precious metals returns. Silver has the highest mean and standard deviation, while platinum has the smallest mean and gold the smallest standard deviation. All series have a negative skewness and excess kurtosis, indicating a long left tail and that the distributions are leptokurtic. The Jarque-Bera statistics are significant at the 1% level indicating the non-normal nature of the precious metals returns. We examine the volatility of returns in a similar way to Batten and Lucey (2010) as we allow for an ARMA(2,2) process to accommodate any auto-correlated innovations and apply a GARCH(1,1) conditional variance specification, necessary given the volatility clustering clearly evident in Figure 2. We also allow for higher order ARMA terms and include them in our estimation if they are statistically significant. The results of this estimation are reported in Table 2 where the gold and silver returns require an ARMA(2,2) model while platinum requires an ARMA(3,3) model. All coefficients of the mean equation are statistically significant , with the exception of the AR(2) and MA(2) coefficients for silver, while we also find significant GARCH effects as both ARCH(1) and GARCH(1) terms are significant. To study the volatility graphically, we also we also plot the ARMA-GARCH volatility in Figure 3 where we can see that volatility for silver is much higher than that for gold and silver. There are also large spikes at certain periods for each precious metals, indicating that volatility is not constant and does vary over time providing a strong rationale to study the time-varying dynamics of precious metals.

[Insert Figure 3 & Tables 1-2 here]

**5. Empirical Results**

In this section, the empirical results of the tests given in the previous section are discussed and the rolling window estimation is conducted to determine how predictable the three precious metals have been over time.

*5.1. Full Sample Results*

Table 3 presents the full sample results of the tests for predictability of precious metals described in the previous section. The CD variance ratio test is conducted on the unfiltered data and shows that each of the series over the full sample fails to reject the null hypothesis of independence indicating returns are not predictable. This suggests that the prices of these precious metals from January 1987 to September 2014 follow a random walk. However, when we run the joint-rank and joint-sign variance ratio tests on the precious metals data, we find significant evidence of predictability in each of the precious metals. This suggests that precious metals are predictable and shows that different calculation methods of the variance ratio test can generate quite contrasting results[[4]](#footnote-4). The BDS test is conducted on the pre-whitened series, which was whitened using an AR(*q*)-GARCH(1,1) model to account for the volatility clustering in the data. The lag length of the autoregressive model is chosen to ensure that the Ljung-Box *Q*(10) statistic is insignificant at the 10% level. We can see that gold and silver both fail to reject the null hypothesis of independence in the series indicating no nonlinear predictability in gold and silver. However, the platinum result indicates significant evidence of nonlinear dependence in the returns at the 5% level of significance, indicating nonlinear predictability of platinum over the full sample period. The R/S Hurst exponent results are also reported in Table 2 where gold is found to have *H*-statistics of 0.605, indicating some evidence of persistence in returns. Silver however has a *H*-statistic of 0.694, indicating evidence of significant persistence of silver returns since the *H*-statistic is above 0.65. Platinum has a *H*-statistic of 0.557 suggesting some persistence in returns but not at a significant level.

Therefore, the full sample analysis of the predictability of the precious metals suggests no significant evidence in the precious metals according to the CD test while significant evidence of predictability according to the JR and JS statistics. Also there is some evidence of nonlinear predictability in platinum according to the BDS test and strong evidence of persistence in silver returns. To further examine the predictability of these precious metals, we examine how their level of predictability changes over time.

[Insert Table 2 here]

*5.2.* *Time-Varying Predictability*

Figure 4 presents the CD test p-values over time through the 60 month rolling window analysis for gold, silver and platinum. The statistical significance of the CD test is evaluated using the 95% confidence interval based on the wild bootstrap methodology described earlier. That is, if the p-value is less than or equal to 0.05, the p-value is deemed statistically significant and rejects the null hypothesis of independence and thus indicates predictability. We show that gold is generally deemed unpredictable over time, with only 4 subsamples from October 1999 to January 2000 indicating any level of significant predictability. All other subsample periods indicate no significant predictability indicating that gold is generally unpredictable according to the CD test. The silver results shown are very similar to the gold results, with only 1 statistically significant subsample (February 2004). This suggests that silver is generally unpredictable during this sample period, although the level of the CD p-value does varying substantially over time. However the platinum results clearly show a time-varying behaviour, with some periods generating statistically significant p-values and some periods generating insignificant p-values. At the beginning of the sample there are a number of subsample periods that report significant predictability, however there is no evidence of significant predictability from March 1995 to the end of the sample period. Therefore the CD results suggest the time-varying behaviour of predictability of platinum supporting the time-varying behaviour of predictability of the AMH but gold silver provide little of evidence of predictability and that the markets are quite efficient.

[Insert Figure 4 here]

The JR test p-values over time are shown in Figure 5, which are quite different to the CD p-values reported in Figure 4. Each precious metal experiences an initial period of significant predictability which then disappears at differing times. More specifically, gold experiences significant predictability up to and including August 1997. However after this point until September 2008, there is no evidence of significant predictability in gold returns. After September 2008, the p-values are generally low, fluctuating between significant and insignificant predictability indicating the varying behaviour of gold returns in the recent past. The silver results show significant predictability from the beginning of the sample until April 1999. There is a short period of no significant predictability until October 2001 that lasts until September 2005 after which, there is no evidence of significant predictability of silver returns. The platinum JR results show significant predictability from the beginning of the sample until March 2001. After this point, the JR statistic indicates no significant predictability until July 2007, which lasts until August 2008. After August 2008, there is no evidence of significant predictability in platinum returns. Therefore according to the JR test, each precious metal does go through periods of significant predictability and no predictability, which is consistent with the AMH.

[Insert Figure 5 here]

Figure 6 presents the JS variance ratio bootstrapped p-values over time. The gold returns document significant predictability from the start of the sample period until December 1997. From this point to the end of the sample period, there are no periods of significant predictability although the level of predictability does varying substantially over time. The silver results show no evidence of significant predictability in returns from the beginning of the sample until November 2004 where returns are found to be significantly predictable until August 2006. However after this period there is no evidence of predictability in silver returns. The platinum JS results reported in Figure 6 show significant predictability from December 1992 to May 1997, however after this point until August 2006 there is no evidence of significant predictability. From September 2006 to May 2010 there is again significant predictability in platinum returns, which disappears towards the end of the sample period. Therefore the JS returns suggests that each precious metal has time-varying predictability that is in agreement with the AMH.

[Insert Figure 6 here]

Figure 7 reports the time-varying average bootstrapped p-values of the BDS test from the four dimension sizes documented earlier through the 60-month rolling window analysis. Again the statistical significance of the test is evaluated using the 95% confidence interval based on the bootstrap methodology described earlier. The returns of each precious metals were initially whitened to remove any linear correlation or dependence due to conditional heteroscedasticity through an AR(*q*)-GARCH(1,1) model. The gold returns were whitened through an AR(2)-GARCH(1,1) model, silver returns through an AR(1)-GARCH(1,1) model and platinum through an AR(0)-GARCH(1,1) model respectively. The gold results show no level of significant nonlinear predictability up to June 1997, but significant nonlinear predictability during the July and August 1997 subsamples, and again significant nonlinear predictability from May 1999 to May 2004. The results also indicate no predictability from June 2004 to March 2008, however from April 2008 to the end of the sample there is significant evidence of nonlinear predictability in the gold market. The silver results show no significant evidence of nonlinear predictability except for one subsample period in October 2001. The level of nonlinear predictability does vary substantially over time but indicates that silver is generally unpredictable according to the BDS test. The platinum results show no level of significant nonlinear predictability up to November 1999. However from this point to May 2002, 21 of the 26 subsamples are statistically significant indicating the nonlinear predictability during this period. Nevertheless, after this period there is no significant nonlinear predictability until April 2003 where the level of nonlinear predictability fluctuates until April 2005 when there is a period up to October 2006 of no significant nonlinear predictability. From November 2006 to September 2011, the platinum market shows significant nonlinear predictability although this disappears after this period until the end of the sample period. Therefore the BDS test results suggest that gold and platinum do experience some periods of significant nonlinear dependence, consistent with the AMH. However silver experiences only one period of significant nonlinear dependence according to the BDS test, indicating that silver generally does not experience any nonlinear dependence and could be deemed efficient in this respect.

[Insert Figure 7 here]

The rolling window R/S Hurst exponent is reported in Figure 8 and shows that the majority (88.69%) of *H*-statistics are greater than 0.5 indicating persistence in gold returns. During the first half of the sample, the *H*-statistic for gold is generally higher than 0.5 indicating persistence in gold returns. However after March 2003, the *H*-statistic is generally lower than during the first half of the sample but higher than 0.5, indicating persistence in gold returns but not at the same magnitude as before. This suggests that gold returns are persistent but the level of persistence is decreasing over time. The R/S Hurst exponent for silver returns also shows the majority (81.02%) of the *H*-statistics are greater than 0.5 indicating more persistence than mean-reversion. Although the *H*-statistic does vary considerably over time for silver returns, there is no clear trend in the behaviour of the *H*-statistic over time. Similar to the gold and silver results, the majority (70.44%) of the *H*-statistics for platinum are greater than 0.5 indicating persistence in returns. There is no clear trend in the *H*-statistic over time although there is clear deviation in the value of the *H*-statistic indicating the time-varying behaviour of the dependence in platinum returns. Therefore the Hurst exponent also suggests the time-varying nature of precious metal returns.

[Insert Figure 8 here]

The global financial crisis from 2007 to 2009 was in interesting period for the global economy and a natural question is how this period affects our results, as Nguyen et al (2016) note that financial crises have a substantial negative impact on portfolio management and investment returns of different asset class. Further, Białkowski et al (2015) show that a model accounting for the financial crisis accurately tracks the gold price indicating that the gold market was affected by the financial crisis. Therefore we may expect our results will be affected by the financial crisis. By studying the results carefully, we can see that the parametric CD test indicates no significant predictability during this period. However, the nonparametric JS and JR tests for gold and platinum both show a noticeable move towards significant predictability. The BDS test for nonlinear dependence also shows a noticeable move towards significant predictability for gold and platinum while silver remains relatively unaffected by the financial crisis. The Hurst exponent test for long memory however shows no discernible change in pattern during the financial crisis. Therefore our results suggest that the financial crisis caused a shift towards inefficiency in returns in gold and platinum according to the nonparametric JS and JR tests and the nonlinear BDS test. This result for precious metals is consistent with evidence for equity markets, such as Boubaker and Sghaier (2015), who find that the dependence structure between US and four developed stock markets increases during the financial crisis.

*5.3. Predictability Ratios*

To determine the level of predictability the three precious markets, we calculate the percentage of windows that generate significant coefficients that indicate predictability. For the variance ratio tests and the BDS test we deem any p-value less than or equal to 0.05 indicates significant predictability. For the R/S Hurst exponent, we deem any *H*-statistic greater than 0.65 (less than 0.45) as significant evidence of persistence (mean-reversion) in precious metals returns[[5]](#footnote-5). Table 4 reports the predictability ratios and shows that according the CD variance ratio test, only 1.46%, 0.36% and 6.20% of gold, silver and platinum returns are predictable, respectively. However the JR and JS variance ratio tests both report more subsample periods that report significant predictability. For instance, the JR statistic reports 49.64% and 49.27% of silver and platinum subsample periods are predictable, while the JS statistic reports 26.28% of gold subsample periods are predictable. Once we remove all the linear correlations from the precious metals returns and examine the nonlinear dependence through the BDS test, we find that 49.27% and 34.31% of gold and platinum returns are predictable, compared to only 0.36% of silver. The R/S Hurst exponent shows that 14.60% and 11.68% of gold and silver returns are deemed to generate significant dependence, while 20.80% of platinum returns generate significant dependence. Therefore these precious metals experience more nonlinear predictability than linear predictability. Overall we find that silver is the least predictable of the precious metals, with only 14.01% of the subsample windows analysed generate significant predictability results while 24.04% of the windows generate significant predictability results for gold. Platinum is deemed the most predictable of the three series, with 29.34% of the windows studied generating significant predictability returns.

[Insert Table 4 here]

**6. Summary and Conclusions**

This paper studies the behaviour of gold, silver and platinum returns over the period January 1987 to September 2013 to examine whether these precious metals are predictable and whether the level of predictability varies over time in a way consistent with the AMH. We find mixed results over the full sample period, with the JR and JS tests suggesting significant predictability of all three precious metals while the BDS test and R/S Hurst exponent indicate predictability in platinum and silver respectively. However the CD test indicates no significant predictability over the full sample period. To determine how the level of predictability varies over time, we utilise a moving-subsample window of a fixed length over a grid of months. We use a five-year window and move the window forward one-month at time until the end of the sample period, generating 274 windows. We show that the level of predictability varies over time in each market for each testing procedure, with at least one window generating significant predictability through the sample period. This suggests strong evidence of the AMH and not the steady erosion of predictability over time that the EMH leads us to expect. For example, 33.58% of the windows suggest significant predictability of gold from the JR test, 8.03% of the windows suggest significant predictability of silver from the JS test and 34% of the windows indicate significant nonlinear predictability of platinum from the BDS test. We also show that the precious metals experience more nonlinear predictability than linear indicating that nonlinearity cannot be ignored. Therefore we find on average 14.01% of the windows indicate significant predictability of silver, while 24.04% of the windows indicate significant predictability of gold. However we find that platinum is the most predictability of the three precious metals, with on average 29.34% of the windows indicating significant predictability. Our results are similar to Charles et al (2015) in that the predictability of gold, silver and platinum varies over time in way consistent with the AMH. However our results differ from Charles et al (2015) who find a strong downward trend (and move towards efficiency) for gold and silver according to the automatic portmanteau test and automatic variance ratio test. We show that level of predictability has not decreased over time and that the precious metal markets have not moved towards efficiency. This demonstrates that different tests for predictability can generate different results, even when the different formulations of the same tests are studied[[6]](#footnote-6).

In summary, we find strong evidence of the time-varying behaviour of predictability of all three precious metals over time, even when the full sample results indicate no significant predictability. This suggests that the level of predictability of a market (or the level of market efficiency), should not be viewed as an all-or-nothing condition but should be examined over time since the level of predictability may not be constant, as suggested by the AMH. Therefore we find strong evidence of the AMH for the precious markets of gold, silver and platinum, while also finding that platinum is the most predictable of the three markets and silver the least predictable.

**References**

Aggarwal, R., Lucey, B. M., O’Connor, F. (2014). Rationality in precious metals forward markets: Evidence of behavioural deviations in the gold markets. *Journal of Multinational Financial Management*, 25, 110-130.

Aggarwal, R., Soenen, L. (1988). The nature and efficiency of the gold market. *Journal of Portfolio Management*, 14, 18-21.

Akgiray, V., Booth, G. G., Hatem, J., Mustafa, C. (1991). Conditional Dependence in Precious Metal Prices. *The Financial Review*, 26(3), 367-386.

Amini, S., Hudson, R., Keasey, K. (2010). Stock return predictability despite low autocorrelation. *Economics Letters*, 108(1), 101-103.

Areal, N., Oliveira, B., Sampaio, R. (2015). When times get tough, gold is golden. *The European Journal Finance*, 21(6), 507-526.

Arouri, M. E. H., Hammoudeh, S., Lahiani, A., Nguyen, D. K. (2013). On the short- and long-run efficiency of energy and precious metal markets. *Energy Economics*, 40, 832-844.

Aye, G., Gupta, R., Hammoudeh, S., Kim, W. J. (2015). Forecasting the Price of Gold Using Dynamic Model Averaging. *International Review of Financial Analysis*, forthcoming.

Bampinas, G., Panagiotidis, T. (2015). Are Gold and Silver a Hedge against Inflation? A Two Century Perspective. *International Review of Financial Analysis*, forthcoming.

Batten, J. A., Ciner, C., Lucey, B. M. (2010). The macroeconomic determinants of volatility in precious metals markets. *Resources Policy*, 35, 65-71.

Batten, J. A., Ciner, C., Lucey, B. M. (2014). On the Economic determinants of the gold-inflation relation. *Resources Policy,* 41, 101-108.

Batten, J. A., Lucey, B. M. (2010). Volatility in the gold futures market. *Applied Economics Letters*, 17:2, 187-190.

Baur, D. G. (2013). The autumn effect of gold. *Research in International Business and Finance*, 27, 1-11.

Baur, D. G., Glover, K. J. (2014). Heterogeneous expectations in the gold market: Specification and estimation. *Journal of Economic Dynamics and Control,* 40, 116-133.

Baur, D. G., Glover, K. J. (2015). Speculative trading in the gold market. *International Review of Financial Analysis*, 39, 63-71.

Baur, D. G., Lucey, B. M. (2010). Is gold a hedge or safe haven? An analysis of stocks, bonds and gold. *The Financial Review*, 45, 217-229.

Baur, D. G., McDermott, T. K. (2010). Is gold a safe haven? International evidence. *Journal of Banking and Finance*, 34, 1886-1898.

Beckmann, J., Berger, T., Czudaj, R. (2015). Does gold act as a hedge or a safe haven for stocks? A smooth transition approach. *Economic Modelling*, 48, 16-24.

Belaire-Franch, J., Contreras, D. (2004). Ranks and signs-based multiple variance ratio tests, Working Paper, University of Valencia, Valencia.

Białkowski, J., Bohl, M. T., Stephan, P. M., Wisniewski, T. P. (2015). The gold price in times of crisis. *International Review of Financial Analysis*, 41, 329-339.

Borges, M. R. (2010). Efficient market hypothesis in European stock markets. *The European Journal of Finance*, 16, 711-726.

Boubaker, H. and Sghaier, N. (2015). On the Dynamic Dependence between US and other Developed Stock Markets: An Extreme-value Time-varying Copula Approach. *Bankers, Markets & Investors Groupe Revue Banque*, 136-137, 80-93.

Bredin, D., Conlon, T., Poti, V. (2015). Does gold glitter in the long-run? Gold as a hedge and safe haven across time and investment horizon. *International Review of Financial Analysis*, forthcoming.

Brock, W. A., Dechert, W. D., Schieinkman, J. A., (1987). A test for independence based on correlation dimension, SSRI Working Paper No. 8702, Department of Economics, University of Wisconsin, Madison, WI.

Brock, W. A., Dechert, W. D., Schieinkman, J. A., LeBaron (1996). A test for independence based on correlation dimension. *Econometric Reviews*, 15, 197-235.

Cajueiro, D., Tabak, B. (2004a). Ranking efficiency for emerging markets. *Chaos, Solitons and Fractals*, 22, 349-352.

Cajueiro, D., Tabak, B. (2004b). The Hurst exponent over time: testing the assentation that emerging markets are becoming more efficient. *Physica A*, 336, 521-537.

Charles, A., Darné, O., Kim, J. H. (2012). Exchange-rate return predictability and the adaptive markets hypothesis: Evidence from major foreign exchange rates. *Journal of International Money and Finance*, 31, 1607-1626.

Charles, A., Darné, O., Kim, J. H. (2015). Will Precious Metals Since? A Market Efficiency Perspective. *International Review of Financial Analysis*, 41, 284-291.

Cheung, Y., Lai, K. S. (1993). Do gold market returns have long memory? *The Financial Review*, 28(2), 181-202.

Chow, K., Denning, K. (1993). A simple multiple variance ratio test. *Journal of Econometrics,* 58, 386-401.

Fama, E. F. (1970). Efficient Capital Markets: A Review of Theory and Empirical Work. *Journal of Finance*, 25(2), 383-417.

Ghazani, M. M., Araghi, M. K. (2014). Evaluation of the adaptive market hypothesis as an evolutionary perspective on market efficiency: Evidence from the Tehran stock exchange. *Research in International Business and Finance*, 32, pp. 50-59.

Hood, M., Malik, M. (2013). Is gold the best hedge and a safe haven under changing stock market volatility? *Review of Financial Economics*, 22, 47-52.

Hoque, H. A. A. B., Kim, J. H., Pyun, C. S. (2007). A comparison of variance ratio tests of random walk: a case of Asian emerging stock markets. *International Review of Economics and Finance*, 16, 488-502.

Hsieh, D. A. (1991). Chaos and nonlinear dynamics: Application to financial markets. *Journal of Finance*, 46, 488-502.

Hull, M., McGroarty, F. (2014). Do emerging markets become more efficient as they develop? Long memory persistence in equity indices*. Emerging Markets Review*, 18, 45-61.

Kim, J. (2006). Wild bootstrapping variance ratio tests. *Economics Letters*, 92, 38-43.

Kim, J. H., Shamsuddin, A., Lim, K. (2011). Stock return predictability and the adaptive markets hypothesis: Evidence from century-long U.S. data. *Journal of Empirical Finance*, 18, 868-879.

Lashgari, M. K. (1992). Information from successive changes in gold and silver prices during phases of business cycles. *Journal of Economics and Finance*, 16(2), 13-30.

Levich, R. M., Poti, V. (2015). Predictability and ‘good deals’ in currency markets. *International Journal of Forecasting*, 31(2), 454-472.

Lim, K-P. (2007) Ranking market efficiency for stock markets: A nonlinear perspective. *Physica A*, 376, pp. 445-454.

Lim, K-P., Hooy, C. (2013). Non-linear predictability in G7 stock index returns. *The Manchester School*, 81, 620-637.

Lo, A., MacKinlay, A. (1988). Stock market prices do not follow random walks: evidence from simple specification test. *The Review of Financial Studies,* 1, 41-66.

Lo, A.W. (2004). The Adaptive Markets Hypothesis. *Journal of Portfolio Management*, 30, pp. 15-29.

Lo, A. W. (2005). "Reconciling efficient markets with behavioural finance: the adaptive markets hypothesis." *Journal of Investment Consulting* 7(2): 21-44.

Lucey, B. M., Li, S. (2014). What precious metals act as safe havens, and when? Some US evidence. *Applied Economics Letters,* 22(1), 35-45.

Lucey, B. M., O’Connor, F. Larkin, C. (2014). Gold markets around the world – who spills over what, to whom and when? *Applied Economics Letters*, 21(13), 887-892.

Manahov, V., Hudson, R. (2014). A note on the relationship between market efficiency and adaptability – New evidence from artificial stock markets. *Expert Systems with Applications*. 41(16), pp. 7436-7454.

Mensi, W., Hammoudeh, S., Reboredo, J. C., Nguyen, D. K. (2015). Are Sharia stocks, gold and U.S. Treasury hedges and/or safe havens for oil-based GCC markets? *Emerging Markets Review*, 24, 101-121.

Neely, C. J., Weller, P. A., Ulrich, J. M. (2009). The adaptive market hypothesis: Evidence from the foreign exchange market. *Journal of Financial and Quantitative Analysis*, 44(2), 467-488.

Niemczak, K., Smith, G. (2013). Middle Eastern stock markets: absolute, evolving and relative efficiency. *Applied Financial Economics*, 23(3), 181-198.

Nguyen, C., Bhatti, M. I., Komorníková, M., and Komorník, J. (2016). Gold Price and Stock Markets Nexus under Mixed-Copula. *Economic Modelling,* forthcoming.

Ntim, C. G., English, J., Nwachukwu, J., Wang, Y. (2015). On the efficiency of global gold markets. *International Review of Financial Analysis*, 41, 218-236.

O’Connor, F. A., Lucey, B. M., Batten, J. A., Baur, D. G. (2015). The financial economics of gold – A survey. *International Review of Financial Analysis*, 41, 186-205.

Patterson, D. M., Ashley, R. A. (2000). A Nonlinear Time Series Workshop: a Toolkit for Detecting and Identifying Nonlinear Serial Dependence, Kluwer Academic, Boston, MA.

Pierdzioch, C., Risse, M., Rohloff, S. (2014). On the efficiency of the gold market: Results of a real-time forecasting approach. *International Review of Financial Analysis*, 32, 95-108.

Qian, B., Rasheed, K. (2004). Hurst exponent and financial market predictability. IASTED conference on Financial Engineering and Applications, FEA 2004, 203-209.

Richardson, M., Stock, J. H. (1989). Drawing inferences from statistics based on multiyear asset returns. *Journal of Financial Economics*, 25(2), 323-348.

Smith, G. (2002). London Gold Prices and Stock Prices in European and Japan. World Gold Council, London.

Smith, G. (2012). The changing and relative efficiency of European emerging stock markets. *The European Journal of Finance*, 18(8), 689-708.

Solt, M. E., Swanson, P. J. (1981). On the efficiency of the markets for gold and silver. *Journal of Business*, 54, 453-478.

Tschoegl, A., (1980). Efficiency in the gold market: A note. *Journal of Banking and Finance*, 4, 371-379.

Urquhart, A. (2014). The Euro and European stock market efficiency. *Applied Financial Economics*, 24(19), 1235-1248.

Urquhart, A., Gebka, B., Hudson, R. (2015). How exactly do markets adapt? Evidence from the moving average rule in three developed markets. *Journal of International Financial Markets, Institutions and Money,* 38, 127-147.

Urquhart, A., Hudson, R. (2013). Efficient or adaptive markets? Evidence from major stock markets using very long historic data. *International Review of Financial Analysis*, 28, 130-142.

Urquhart, A., McGroarty, F. (2014). Calendar effects, market conditions and the Adaptive Market Hypothesis: Evidence from long-run U.S. data. *International Review of Financial Analysis*, 35, 154-166.

Wang, Y., Wang, J., Gu, Y., Xinjue, W., Wang, J. (2011). FMH and its application in the Chinese stock market. *Beijing Institute of Tracking and Telecommunications Technology, 1, 349-355.*

Wright, J. (2000). Alternative variance-ratio tests using ranks and signs. *Journal of Business and Economic Statistics*, 18, 1-9.

1. For an excellent review of the literature on gold, see O’Connor et al (2015). [↑](#footnote-ref-1)
2. Linear and nonlinear predictability. [↑](#footnote-ref-2)
3. Which has often been ignored in the literature. [↑](#footnote-ref-3)
4. For examples of contrasting results from different variance ratio tests, see for example Smith (2012), Niemczak and Smith (2013) and Urquhart (2014). [↑](#footnote-ref-4)
5. Similar to Hull and McGroarty (2014). [↑](#footnote-ref-5)
6. Charles et al (2015) use the automatic variance ratio test while this paper uses the Chow-Denning, Joint-Rank and Joint-Sign variance ratio tests. [↑](#footnote-ref-6)