Abnormal Returns in Gold and Silver Exchange Traded Funds

Michael J. Naylor*, Udomsak Wongchoti and Chris Gianotti
School of Economics & Finance
College of Business
Massey University
Private Bag 11 222
Palmerston North
New Zealand.

* Corresponding author; M.J.Naylor@massey.ac.nz. T: +64-6-3509099 x7720

Acknowledgements
We wish to thank departmental colleagues for comments.

Version 1.0: 8th Feb 2011
Abnormal Returns in Gold and Silver Exchange Traded Funds

Abstract

Exchange Traded Funds (ETFs) are one of the fastest growing areas of investing and have significantly changed investor behavior, yet there is limited academic research on ETFs, with minimal on commodity based ETFs. This paper is the first to examine whether abnormal returns are available for Gold and Silver ETFS.

Gold and silver ETFs have attracted substantial investments, with $40.5 billion USD in gold (GLD) is and $5.22 billion in silver (SLV). Our study shows that while inefficiency is present in the gold and silver ETF markets for the period 2006-2009, when risk is factored the abnormal returns evaporate. However, use of simple filter trading rules does allow abnormal returns. GLD and SLV exhibit similar characteristics to the underlying physical assets and when ignoring risk are able to provide investors with greater profit than if investing in the market.

JEL classification: G10, G11, G12, G14

Keywords: Abnormal returns, exchange traded funds, commodities, filter trading,
1 Introduction

1.1 Overview

Exchange Traded Funds (ETFs) are one of the fastest growing areas of financial markets and have significantly changed how investors at both an institutional and retail level construct their investment portfolios investor behavior. The rapid growth raises the question as to whether it is due to superior characteristics or whether an abnormal return is present. However there is relatively little academic research on ETFs (Anderson, Born & Schnusenberg, 2008), and nearly no research on commodity ETFs, the fastest growing area. Very little is known about the dynamics of this market. This paper is the first to examine whether prices in commodity exchange traded funds in two main funds GLD (Gold) and SLV (silver) accurately reflect all the relevant information in the market, exhibit dependence and offer abnormal return on a risk adjusted basis. We also examine four smaller funds.

The study is also unique in that it applies existing knowledge of physical gold and silver prices/returns to examine whether the return properties and characteristics for gold and silver ETFs perform in the same way as the physical asset.

1.2 The Exchange Traded Fund Market

ETFs are similar to mutual funds, except that shares in an ETF can be bought and sold throughout the day and ETFs do not sell or redeem their individual shares at net asset value. Instead, financial institutions purchase and redeem ETF shares directly from the ETF in large blocks, called "creation units". ETFs generally provide the easy diversification, low expense ratios\(^1\), transparent portfolios and tax efficiency as index funds, while still maintaining all the features of ordinary stock, such as limit orders, short selling, and options. ETF funds also do not have to maintain a cash reserve for redemptions or pay brokerage expenses so do not charge a load.

ETF prices are set by markets and not based on day-end asset values. If there is strong investor demand for an ETF, its share price will rise above its net asset value per share, giving arbitrageurs an incentive to purchase additional creation units from the ETF and sell the component ETF shares in the open

\(^1\) Annual management fees are 0.3-.6%, vs mutual funds at 1-2% (Kelleher, 2010).
market. The additional supply of ETF shares increases the ETF's market capitalization and reduces the market price per share, generally eliminating the premium over net asset value. ETFs are dependent on the efficacy of the arbitrage mechanism in order for their share price to track net asset value.

This growth reflects the popularity of the ETF as an investment vehicle for both retail and institutional investors as an easy to use way to diversify across a broader range of securities. The value of total ETF assets under management at January 2010 was US$1,032B, up from US$710B in December 2008 (Fuhr, 2010). At the current growth rate ETFs will rapidly eclipse that of mutual funds. The fastest growing sector of ETFs has been commodities as the market is active and offers diversification. The first commodity ETF was established in 2002 and by January 2010 the GLD fund had grown to US$ 40.5B and SLV fund to US$5.22B (The largest is S&P500 fund SPY with US$78B). Both GLD and SLV physically own the commodity.

The impact of the rapid growth of commodity gold and silver ETFs is uncertain; the rapid growth in ETF assets can increase underlying demand, while supply is unchanged. Alternatively commodity ETFs may be substitutes for existing demand for the product. Brikner & Collins (2008) argue that growth in ETFs has boosted speculation in markets and creating unstable price fluctuations. The ETFs themselves are also speculators.

2 Literature on the Efficiency of the Physical and ETFs markets

Aggarwal & Soenen (1988) examines gold market efficiency vs the S&P 500. They argue that the strong movement in gold prices offer an opportunity for strong returns even risk adjusted, though there was significant positive skewness and daily changes was leptokurtic. However the inefficiencies diminished under the CAPM model and that when filter tests were applied the excess return likewise evaporated.

Solt & Swanson (1981) note the importance of including gold and silver in a risk return framework and note that gold and silver returns show a low correlation to the market. They thus conclude that gold and silver have a speculative rather than investment value.

Fama & Blume (1973) developed filter trading rules and examined the returns of 24 filters concluding that the evidence indicated trading rules do not consistently earn returns greater to those of a simple passive (buy and hold)
approach. The filter trading rules are $x\%$; if daily closing price (of gold or silver) increases $x\%$, buy and hold the commodity, until it’s price moves down at least $x\%$ from the subsequent high, at which time the investor should simultaneously sell and go short. The short position is maintained until the daily close price rises at least $x\%$ above the subsequent low at which time the investors buys.

De Jong & Rhee (2008) found that ETFs in general provide abnormal risk adjusted returns based on Fama and French 3 factor model exceed transaction costs. De Jong & Rhee do however distinguish commodity ETFs from other ETFs.

Buyuksahin, Haigh, and Robe (2008) examined the S&P Goldman Sachs Commodities Index (S&P GCSI) weekly return data and revealed no correlation between commodity and equity returns, with commodity returns being lower and no sustained pricing relationship over the long-run. Nor was evidence found of an increase in co-movement between equities and commodities during a period of extreme returns.

Engle & Sarkar (2006) found that where an arbitrage opportunity exists between the price of a given ETF and its Net Asset Value (NAV), this will be minor and quickly exploited eliminate any premiums/discounts. Miffre (2007) shows that higher than average returns can be achieved by investors taking long or short positions for a given level of risk when investing in iShares MSCI country funds ETF.

3 Methodology

3.1 Hypotheses

The objective of this study is to determine if by investing in ETFs GLD (gold) and SLV (silver), investors can achieve abnormal or above market returns due to inefficiencies. The study assumes gold and silver pricing are driven by fundamental supply and demand forces in the marketplace.

Time series analysis is used and based on a daily close price/return series derived from the NYSE Arca between 5th of May 2006 and 31st of December 2009. Within the study period the macro economy has experienced periods of growth, severe contraction and return to low/mild growth while maintaining extraordinarily high demand for gold and silver. It is thus an ideal period as for
ascertaining whether dependence actually exists in the series and returns are abnormal.

On this basis our hypothesis is that;

a) Returns between gold backed ETF GLD and silver backed ETF SLV are independent and inefficient, meaning that an abnormal return is possible if investing in GLD/SLV.

b) GLD and SLV provide investors with abnormal returns once adjusted for risk. Expressed as;

\[ (RGLD - R_t) = \alpha + \beta (R_{mt} - R_t) + \epsilon_t \]
\[ (RSLV - R_t) = \alpha + \beta (R_{mt} - R_t) + \epsilon_t \]

c) The application of filter trading rules to GLD and SLV will result in an evaporation of the abnormal return.

3.2 Data

GLD and SLV are the primary funds focused on in the analysis. However four additional funds on the NYSE Arca, the gold funds IAU, and DGL, and silver funds DBS and AGQ, were randomly chosen from twenty-one precious metals ETFs. The key criteria are the fund must trade gold or silver, be an ETF (not an ETN), use a passive strategy, offer minimum net assets of $50 million and have begun trading no later than December 2008. This criteria reduced the applicable funds for research from twenty-one to six. We also included the S&P 500 as a benchmark. The data has been sourced from the Centre for Research and Securities Prices (CRSP).

Table 1 about here

The time horizon is from 5 May 2006 to 31 December 2009 for the funds GLD and SLV. As the inception date of the four remaining ETFs IAU, DGL, DBS and AGQ varies between January 2005 and December 2008 the dataset for these funds is adjusted to the each funds applicable lifetime. Data will be based on daily closing prices of units/shares in each of the funds.
3.3 Methodology – Hypothesis a) Abnormal Returns

To prove hypothesis a) we must show that a dependency exists in series between prices. Conversely, if GLD and SLV were efficient, then our expectation is that today’s price is not reflective of yesterday’s price, rather it is reflective of today’s information and that there is no pattern or consistency so that pricing patterns follow a random walk, or that they are normally distributed.

The method for completing the test of hypothesis a) is to perform econometric tests across 923 observed price movements for GLD/SLV and between 261 to 925 observations for the remaining funds (IAU/AGL/DBS/AGY) when extending the test. For consistency with previous tests, the price series for GLD and SLV are converted to (natural) logarithms, or continuously compounded returns. The data set for this test is unfiltered as the use of logarithms reduces the affects of outliers.

Previous work in the study of efficiency in physical gold and silver pricing by Solt & Swanson (1981) identifies the occurrence of non-normality and non-stationarity in the means and variances of the price for distributions, concluding that (despite an element on non-stationarity) there exists a positive dependence in the series. Snedecor & Cochran (1967) likewise reject the presence of normality for a price series of gold.

Establishing stationarity is important for ensuring the significance of the series to predict future prices and will be tested using the autocorrelation function (ACF) and supported by the augmented Dickey-Fuller test (ADF). The series is expected to exhibit the following characteristics if stationary;

\[
\begin{align*}
\text{Mean:} & \quad E(Y)_t = \mu \\
\text{Variance:} & \quad \text{var}(Y)_t = E((Y_t - \mu)^2) = \sigma^2 \\
\text{Covariance:} & \quad Y_k = E[(Y_t - \mu)(Y_{t+k} - \mu)] \\
\text{where } Y_t & \text{ is the time series.}
\end{align*}
\]

Under the non-stationary process zero mean and constant variance is expected. A test of non-normality is then done using the Bera-Jarque method or to determine whether \( \mu_t \) are normally distributed. The test will support the outcome of the ADF unit root test.
3.4 Methodology – Hypothesis b) Risk Adjusted Returns

Hypothesis b) tests for risk-adjusted abnormal returns using the Sharp-Linter-Mossin model based on linear regression, otherwise known as the capital asset pricing model (CAPM). This follows Aggrawal & Soenen (1988) we use daily return series rather than convert into monthly returns. The market return is based on the S&P500 (ETF SPY) and the risk free rate is represented by the US 90 day Treasury bill rate. The beta coefficient ($\beta$) measures systematic risk. The market return ($R_m$) is represented by the ETF SPY, the broad based S&P 500. An aggregate of the market being the return on exchanges NYSE, AmEx, NASDAQ and Arca was also tested in the model to represent ($R_m$), the outcomes were comparable to that when using SPY. The CAPM is expressed as;

$$\text{CAPM:} \quad (RGLD - R_f)_t = \alpha + \beta (R_m - R_f)_t + \epsilon_t$$
$$\text{(3)}$$
$$\text{(RSLV} - R_f)_t = \alpha + \beta (R_m - R_f)_t + \epsilon_t$$

A consistent risk adjusted return by the applicable gold and silver ETF that exceeds the market will show the presence of abnormal returns for these investment vehicles. A Whites test (1980) is performed to ensure constant standard errors in the coefficients, or that homoskedasticity is present, thus proving the second assumption of the classic linear regression model (CLRM) holds;

$$\text{CLRM Second Assumption:}$$
$$\text{(4)}$$
$$\text{var}(\mu) = \sigma^2 \text{ and } < \infty$$

3.5 Methodology – Hypothesis c) Filtered Trading Rules

We then use filter trading rules to test for profitable trading strategies. Filter trading is a strategy determining by percentage movement of the stock price at which points in time investors should buy or sell stock, or in other words what is the entry and exit point of the trade, using 2% price movements as buy/sell signals. This was applied retrospectively to GLD and SLV to ascertain abnormal returns for the period December 2006 to December 2009. Testing is not extended to the smaller funds. The benchmark is the stock performance of a
passive buy and hold strategy rather than the market return or SPY. The starting point for the filer will follow whichever movement comes first, either upward or downward. Log returns are used and applied to filters at 3% and 5% in regression analysis against the market, following Solt & Swanson (1981).

4 Results

4.1 Preliminary Analysis

We plotted real gold and silver versus GLD and SLV to provide an indication of the relative performance of each investment over the same time horizon, using a 100 at 19 November 2004, and an end date at 31 January 2009.

Graph 1 about here

Graph 1 indicates that the movement between GLD and gold is quite close with very little separation between the two trend lines particularly since August 2008, and coinciding with a steady trend in appreciation of the asset.

Graph 2 about here

Graph 2 shows similar results for SLV and silver, with no significant points of deviation between the two return series. Comparing the time series of GLD versus the market, shows GLD has outperformed the market benchmark, represented by SPY especially since August 2007. A comparison of SLV versus the market (SPY) shows a smaller outperformance. SLVs performance began from a lower base relative to the market and comparative to GLD and was not as resilient when systematic risk increased. In fact SLV found support at a similar time to the market and began an upward trend with much larger oscillations.

Graph 3 about here

Graph 4 about here

4.2 Analysis and Results - Hypothesis a) Abnormal Returns, Stationary Testing

Graphical analysis provided a starting point to the analysis, offering a feel for the nature of the data and indication whether stationarity or autocorrelation is present before more formal tests are performed. The trend for the GLD
distribution offered no indication of a random walk with or without drift, or a process that is non-stationary. Rather the process for the GLD distribution is more consistent with white noise, or a stationary process that has a relatively even distribution over time. The same inference can be drawn for the SLV distribution.

The graphed series also indicated that negative autocorrelation may exist given the pattern where the residuals appear to cross the axis on a frequent basis that would be unusual for a random walk distribution. Detecting a autocorrelation relationship suggests that there is some inefficiency in the market and that successive price changes are not independent.

The Autocorrelation Function (ACF) is the first important formal test of the hypothesis and is in itself a relatively simple test for detecting dependence in the series. The test is expressed as;

\[ p_k = \frac{Y_k}{Y_0} = \frac{\text{covariance at lag } k}{\text{variance}} \]  

Plotting \( p_k \) against \( k \) will produce the correlogram although determining the appropriate lag length in the equation is important when performing the test and requires a trial and error process. Akaike and Schwarz information criterion has been chosen as the preferred method to assess the appropriate amount of lags.

Lags were tested from 5 to 35 at intervals of five and the model offering the lowest value in these criteria selected. The lowest value was achieved at 5 lags where it gave Akaike Schwarz information values at -5.49 and -5.45. 5 lags were then applied to both the GLD and SLV ACF and correlogram.

Figure 1 about here

The results for show GLD shows all lags are close to ‘0’, so GLD exhibits a typical stationary process. The AC range is between -.033 and .021, and the PAC range is the same, there is very little difference from zero.

Figure 2 about here
The correlogram for the SLV distribution shows AC and PAC values are insignificantly different from zero across all lags. SLV exhibits stationarity.

The analysis is now extended to the augmented Dickey-Fuller (ADF) test which is also known as the unit root test. The ADF test is suitable for larger sample consists of the following regression estimate;

$$\Delta Y_t = \beta_1 + \beta_{2t} + \delta Y_t - 1 + \sum_{i=1}^{m} (\alpha_i \Delta Y_t - i + \epsilon_i)$$

(6)

The ADF test is one sided and offers three critical values at 1%, 5% and 10% in comparison to a test statistic. When performing the test lags are applied to ensure that the error terms are uncorrelated and that unbiased estimates can be obtained preventing the test from to easily accepting whether a unit root is present.

*Figure 3 about here*

The test results for the GLD distribution are shown in Figure 3. The ADF for the GLD logarithm is performed with 5 lags applied to the dependent variable. The ADF test-statistic does not exceed the critical values at any level and has determined that the series is stationary; it does not have a unit root.

When extending the ADF test at intervals of five to include up to 35 lags the result in terms of the ADF statistic relative to the critical values was consistent which we would expect. The bulk of the coefficients produced at various lags are also shown to be negative implying a relationship exists between price changes.

The same tests are now performed on the series for SLV. The results suggest the null hypothesis should also be rejected. The series is stationary with zero mean.

*Figure 4 about here*

However, the SLV series is differentiated from the GLD series as even though the majority of coefficients are less than zero; the amount is fewer than those applicable to GLD as more lags are applied.

The ADF result is consistent with that obtained by Solt & Swanson (1981) that the rejection of a unit root in silver is slightly less obvious than that in gold, or
applied to this model that the rejection of a unit root in SLV is slightly less obvious than that evident in GLD. The implication is that not all available information is reflected in the prices for GLD and SLV as would the case if the market were truly efficient instead of weak form efficient.

When expanding the ACF, Correlogram and ADF test to funds IAU, DBS, DGL and AGQ, the tests produced comparable results to those obtained from GLD and SLV. The logarithmic distribution was found to possess no unit root. As with GLD and SLV the outcome did not change when increasing or decreasing lags. The findings held across all the four smaller funds although the number of observations included in the analysis was in most instances lower than those in the test of GLD and SLV, with AGQ having the least number of observations at 265. The null hypothesis tested in this case that the residuals have a unit root was rejected. This suggests that market for gold and silver ETFs is inefficient.

For the purpose of forecasting and predicating price a non-stationary process is of little value to investors. It only holds for the period for which it was studied.

4.3 Analysis and Results - Hypothesis a) Abnormal Returns, Normality Testing

To complete the analysis of hypothesis a) a test of normality is undertaken and applied to each fund. The null hypothesis tested in this case is that the residuals ($u_t$) are normally distributed; the alternate hypothesis is that they are non-normally distributed. Evidence of a non-normal distribution would identify dependence in the series, or that the series is weak form efficient. Such information is relevant to investors because it will influence buy and sell decision making as the investor attempts to predict when to enter or leave the market. As the data for all funds in the analysis has been transformed into logarithms and is based on a daily series, it has essentially had the effect of rescaling the data and pulling in extreme observations hence filters will not be applied to the test. It is also not necessary to filter in the cases of GLD, SLV and IAU due to the large amount of observations which encourages a normal distribution.

Table 2 about here
The summary statistics for each fund are available in Table 2 and include values for mean, median, skewness, kurtosis and the Bera-Jarque statistic. An examination of the summary statistics across all funds shows the mean and median for GLD to be the closest, at .0005 and .0008 respectively, suggesting there is a presence of normality in the series which is typical where the mean is close to zero. The mean and median for IAU, DGL and AGQ are also quite close and in some ways similar to GLD, indicating normality as there is not significant differences between each mean and median value. The variation between the mean and median values for IAU and DGL is both at .0004 whereas AGQ is at .004. Conversely, the mean and median for SLV are comparatively further apart at .0002 and .003, as is that for DBS at .0003 and 003. Both of these funds trade silver rather than gold and have an initial indication of non-normality.

The table also shows the Bera-Jarque statistic, and the measures of skewness and kurtosis. Aggarwal & Soenen (1988) found strong kurtosis for the gold price. We found strong positive kurtosis above the critical threshold of ‘3’ for all funds except AGQ., with a range between 8.310 for IAU at the lower end to 13.48 for DBS. The primary funds GLD and SLV offer kurtosis values of 7.5385 and 9.9616. The exception to the strongly positive kurtosis identified in the funds is silver fund AGQ which is below ‘3’, though this is based on fewer observations. These results suggest dependency exists in the series for GLD, SLV, IAU, DGL and DBS.

The Bera-Jarque is an asymptotic test based on OLS residuals. The test first calculates skewness then kurtosis testing whether the coefficients are jointly zero (Brooks, 2002). The norms of skewness and kurtosis are 0 and 3. The test statistic is given by;

\[ BJ = n \left( \frac{S^2}{6} + \frac{(K - 3)^2}{24} \right) \]

The Bera-Jarque output statistics are extremely high and very different from zero for all funds excluding AGQ. DBS has the highest Bera-Jarque value at 3608 while the lowest excluding AGQ is DBS at 642, comparatively GLS and SLV give values at 797 and 2026. This value of these statistics means the skewness
and kurtosis statistics is not jointly zero suggesting the null of normality should be rejected.

The p-value of the Bera-Jarque statistics is at the same time very low which we would expect; it is zero for all funds except AGQ. To not reject the null of normality the p-value should be larger than .05 at the 5% test level (Brooks, 2002). As the p-value is below .05 normality is conclusively rejected based on these outputs particularly given that the number of observations is large, at least for funds GLD, SLV, IAU and DBS. Comparatively the p-value for AGQ is greater than .05, meaning normality would not be rejected and fitting with the low Bera-Jarque value of 1.36.

Reviewing the residuals provides method to understand the shape of a distribution. The curve of the histograms under the formal Bera-Jarque test is representative of the frequency distribution and is consistent with that of a leptokurtic rather than a normal distribution. For GLD, SLV, IAU, DGL and DBS the distribution is shown to be high peaked and steep, so the distributions are not perfectly normally distributed. The curve for AGQ is conspicuous given it has a bell shaped curve more consistent with a normal distribution, although the number of observations in the series is somewhat fewer than those of other funds.

The existence of non-normality in the GLD and SLV series seems apparent; it establishes that GLD and SLV ETF prices behave like actual gold and silver prices giving some assurance around the funds ability to offer an abnormal return and consistency with prior studies into the distribution characteristics of the physical assets. The null is likewise rejected for funds IAU, DGL, DBS, but not for AGQ, although it seems logical that as the AGQ series extends from its current short base it will experience non-normality as the other funds have done.

4.4 Analysis and Results - Hypothesis b) Risk Adjusted Returns, CAPM

It is not enough to purely suggest abnormal returns are present in gold and without also considering the return within a risk return framework. The Capital Asset Pricing Model (CAPM) is the chosen tool with which to assess and explain the relationship as to whether GLD and SLV maintain the ability to offer investors an above market (expected) return relative to the risk of investing in the funds. The CAPM is a multivariate linear regression using least squares and
when performed in this analysis uses the logarithmic return series for GLD and SLV, both these series and the series of other funds used in the analysis are represented by \((R_t)\).

*Figure 5 about here*

The results for GLD are shown in Figure 5. The slope coefficient, which is the beta coefficient at 0.6095 is less than one, suggesting that the security moves less than the overall market and is not volatile, it is rather defensive in nature and suggests GLD would have some risk reducing properties in a portfolio of stocks. The intercept term at -0.0091 is not significantly different from zero. This means there is no significant risk adjusted abnormal return for GLD for the time period.

When applying sub-period analysis to GLD, the horizon of 1 May 2006 through 1 May 2007 are analyzed to assess whether in its relatively immature state if GLD demonstrated greater inefficiency. The beta coefficient over horizon is 0.5721, marginally less than the beta over the three years showing that GLD was slightly less risky over this period. The intercept is -0.0211, meaning that GLD did not offer a risk adjusted abnormal return for the period.

The horizon 1 March 2008 to 1 March 2009 is also tested, this will identify whether the rapidly increasing systemic risk produced an abnormal return if investing in GLD during the period of the financial crisis. The beta coefficient for the period showed a remarkable difference from the three years or one year immaturity period being 0.0491, meaning that GLD was highly defensive when market high risk was greatest.

This conclusion is similar to Aggarwal & Soenen (1988), though, it is important to point out that the asset, GLD in our case, has appreciated significantly during the period used in the analysis offering investors a strongly positive return on a simple passive or buy and hold strategy.

*Figure 6 about here*

The results SLV are in Figure 6 and are similar to those of GLD, with a low beta at 0.06503 and intercept at -0.0084. The same conclusions can be drawn that the return does obviously display the characteristics of inefficiencies when risk is
considered. The same quantity of daily logarithmic observations, 922 is used as that used in the GLD CAPM regression.

The SLV beta coefficient is less than one, evidence that the asset is not particularly volatile and is therefore risk reducing in nature within a portfolio. This would allow the investor to purchase other more risky assets with a higher beta and still maintain an acceptable level of chosen risk by including SLV (or GLD) to help balance the basket of assets. The model has a minor negative intercept coefficient at -0.0084 signifying that investors into SLV actually earned a rate less than the market requires. The same is true of course for GLD, and like GLD, SLV has experienced notable appreciation in underlying asset values during the GFC although not to the same extent.

When applying sub-period analysis to SLV, the horizon of 1 May 2006 through 1 May 2007 are analyzed to assess whether in its relatively immature state if SLV exhibited greater inefficiency. The beta coefficient over horizon is 0.4732, noticeably lower than the beta over the three years showing that SLV was a less risky investment over this period. The intercept is -0.0084, meaning that SLV did not offer a risk adjusted abnormal return for the period.

The horizon 1 March 2008 to 1 March 2009 is also tested, this will identify whether the significant increase in systemic risk produced an abnormal return if investing in SLV during this horizon. The beta coefficient for the period was much lower than the three years or one year immaturity period being 0.2844 and suggesting that SLV represented a low risk when market risk was greatest.

What is unexpected when comparing the regression results across all horizons tested is just how close the SLV intercepts and slope values are to GLD, especially given the differences in the supply and demand characteristics of the underlying assets. Another point of difference between the funds is the enormous disparity between market capitalizations, with GLD $35 billion USD larger.

As with GLD the standard errors under the SLV CAPM regression offer certainty around the model as they are reasonably small at .0277. $R^2$ at 0.2791 implies the model fits the data slightly better than that for GLD and that it explains the return variability. Nonetheless the generic outcome of the CAPM analysis is that market inefficiencies are not evident on a risk adjusted basis.
Applying the CAPM to IAU, DGL, DBS and AGQ offers a generally consistent result to that obtained when completing the CAPM regression for GLD and SLV, one of low market inefficiency.

The supporting regressions to some extent offer an additional element of sub-period analysis given the shorter listing time frames available for gold backed DGL and silver back DBS and AGQ. The shorter timeframes for these funds can be identified by the lesser amount of daily observations in the regressions at 389, 750 and 267 respectively. In terms of actual time horizon these observations reflect periods starting January 2007, June 2008, December 2008 and ending at 31 December 2009. The data in these sets is likewise converted in to a logarithm for consistency with the method applied to the GLD and SLV CAPM.

The beta coefficient of the IAU, DGL, DBS and AGQ beginning at lower end is in the range of 0.6053 applicable to the larger fund by capitalization IAU, and 0.6625 for the smallest fund by capitalization DBS. The intercept is in the range of -.0092 for IAU and .0024 for AGQ. With the exception of the AGQ intercept, the results are not overtly dissimilar from that produced in the regressions for the GLD and SLV. What is unusual is the very low beta applicable to gold fund DGL. At 0.1143 it is noticeably lower than that of the other funds and evidence of extraordinarily low relative volatility, in terms of market cap the fund is much smaller than GLD only $159.03 million USD.

The intercept is slightly positive on AGQ, whereas in all other tests the intercept is negative. This indicates AGQ investors earned slightly more than the risk associated with the investment would require whereas investors in the other funds earned slightly less. The value may be impacted by the fact that the fund is relatively immature.

In summary application of the CAPM regression to gold and silver ETF’s GLD, SLV, IAU DGL, DBS, essentially gave the same outcome as that drawn by Aggarwal & Soenen’s investigation into the physical asset, that the there is no indication of market inefficiency when adjusted for risk. In real terms this is interpreted as non existence of abnormal returns, meaning the null hypothesis of efficiency is accepted. What is interesting is that as investors have placed increasing value on investment in gold and silver which is reflected in the demand, the lack of a risk adjusted return is significant, it suggests investors...
are willing to accept a return at a level below what the market would expect as they seek to access the risk reducing properties.

4.5 Analysis and Results - Hypothesis b) Risk Adjusted Returns, Heteroskedasticity

We then tested for heteroskedasticity or whether the variance of the error terms is constant. Solt & Swanson (1988) discovered the presence of heteroskedasticity in physical gold and silver series.

Detecting heteroskedasticity though is usually not easy and often a matter of speculation (Gujarati & Porter, 2003). The price movements shown in Graphs 1 and 2 show strong movements over time, particularly for GLD, that suggests the series may actually be affected by homoskedasticity, that the errors are constant, the antithesis of heteroskedasticity. Graphical analysis though is not a sufficient test hence the application of Whites (1980) test as will be demonstrated. The use of Whites test differs from the approach taken by Solt & Swanson (1981) who employed the Goldfield-Quandt test.

As price/return variability is expected to be a factor throughout the past three years due to the economic climate and may affect the observations and create outliers, in order to perform the test analysis logarithms are taken of each series. Solt & Swanson experienced similar variability with the steeped change in gold and silver pricing throughout the 1970’s and also used logarithms. Outlying observations are a frequent driver of heteroskedasticity.

Whites (1980) test does not rely on the normality assumption and is a conservative test. It is expressed as the following auxiliary regression:

\[
\hat{U}_i = \alpha_1 + \alpha_2 X_2 + \alpha_3 X_3 + \alpha_4 X_4^2 + \alpha_5 X_5^2 + \alpha_6 X_6^2 + \epsilon_i
\]

Disproving heteroskedasticity improves the validity of the analysis by ensuring the OLS estimates give unbiased and consistent coefficient estimates (Brooks, 2002). However, as prior testing of physical gold and silver series behavior by Solt & Swanson encountered the presence of a non-constant finite variance, it is reasonable to expect the same phenomenon may occur in the ETF series given the equivalent underlying properties of the asset.
The test results derived from the application of Whites (1980) test are illustrated in Figures 7 and 8. Both outputs have a large critical value at 98.1 for GLD and 32.1 for SLV. The implication of this test output is that it is likely the two series possess a non-constant finite variance, especially as the given p-value is zero. Note that heteroskedasticity is still present despite the conversion of the series into logarithms. For GLD and SLV the null hypothesis of constant variance of the error terms is rejected.

IAU also evidences a very high critical value at 98.92 and a zero p-value, this suggests the presence of heteroskedasticity. The behavior of the series is essentially the same as that of GLD, which is not unexpected given the underlying properties, maturity and number of observations that are all consistent with GLD and that similar test result between the funds has been a regular trait in the prior analyses. DGL, DBS and AGQ offer far lower critical values that GLS, SLV and IAU, with the highest of these lower values being AGQ at 14.73 and lowest being DBS at 1.68. The result is interesting because DBS has a large number of observations at 750, which although less than GLD, SLV and IAU, would generally give an expectation that the fund would provide as similar outcome to those just mentioned yet it strongly suggests the presence of homoskedasticity, so the null hypothesis of homoskedasticity would be accepted for DBS.

Meanwhile DGL offers a critical value at 12.61 which would also suggest the null is accepted for this fund. What is obviously different of the DGL, DBS and AGQ series compared to GLD, SLV and IAU is that the p-values are above zero. Particularly, the p-value for DBS which at .4304 relative to the other funds DGL and AGQ at .0022 and .0006 is quite large. The p-value is important to note because it indicates that randomness could explain the output whereas there was not ability for randomness to explain the output for the funds.

The null hypothesis tested in the instance is that the variables have a constant finite variance (homoskedasticity); the alternate hypothesis is that the variance is non-constant (heteroskedasticity). Analysis of Whites test allows the null hypothesis to be accepted for the funds DGL, DBS and AGQ only and rejected for GLD, SLV and IAU. The funds accepting the null have far fewer observations.
than those applicable in Whites test for GLD, SLV and IAU, indicating that over shorter periods the presence of heteroskedasticity is not existent or at least it is less evident and more difficult to detect.

4.6 Analysis and Results – Hypothesis c) Filtered Trading Rules

Note that when interpreting the results it is important to remember the large increase in gold prices during the study period (2006-2009) and wider macroeconomic factors caused by the Global Financial Crisis, which would have increased systematic risk and induced a flight to gold and to a lesser extent silver as investors sought to diversify assets. By December 2009 gold had at the time reached new highs after earlier exceeding the $1,000 per ounce mark that once appeared unobtainable. The trend is common to that of the gold appreciation during the 1970’s ‘stagflation’ identified in Solt & Swanson’s (1981) work. The two periods are both times of increased liquidity in gold.

An investor employing a passive (buy and hold) strategy could have easily made large returns simply due to the wider factors mentioned. However we wil test whether there was an opportunity to generate an even greater return relative by active filter trading or momentum strategies.

Aggarwal & Soenen (1988) found that under Fama & Blume’s (1973) filters gold performed poorly and at no point did the return from filter trading exceed what the investor could have obtained via a passive strategy. Solt & Swanson (1988) found that on a gross basis applying filters to gold and silver likewise provides a poor result and that the investor would have been better to employ the buy and hold strategy. Solt & Swanson note that the result did not differ when managing for transaction costs and concluded there appeared to be no inefficiency to be exploited. Comparatively, Fama & Blume (1973) used 24 different filters finding that filters were inferior in almost all cases even when ignoring transaction costs.

We tested market inefficiency through filter trading at 3% and 5% levels. Given that a filter of 7%, 10% or up to 50% used by the other authors seems extreme for a daily series these will not be tested. It would be unusual for a stock to move 7%, 10% or 50% in one day. It has happened of course but the infrequency means the filter is unsuitable for a practical test of a daily series. Even over a month a movement 10% would equate to substantial market
movement, although Aggarwal & Soenen apply this filter to their monthly series as does Solt & Swanson to their weekly series.

Undoubtedly during the period under consideration the affects GFC have caused large downward and upward movements in values, nonetheless the study stresses that these are less important when based on a daily series and considers 3% and 5% more appropriate. The filters are applied to both upward and downward movements. The series is converted to logarithms excluding transaction costs which are ignored in the analysis. Note that as the funds are based on capital growth a dividend is not applicable in the analysis.

*Table 3 about here*

The results of filter trading applied to GLD and SLV are shown in Table 5. The start point is a ‘sell’ for both funds. The results show for GLD that at the 3% and 5% level filtered trading offers a better return than that applicable to buy and hold or passive investing. The same result is shown when filters are applied to SLV.

Our results differ from prior studies as filters marginally exceed the buy and hold strategy. The filtered return implies that an investor can generate a return greater than that available by having no trading strategy. The result is unique as it differs from that applicable to physical gold and silver returns when filters are implemented and is of value to investors seeking to trade the gold and silver ETF market.

5 Conclusion

Study has examined the series of three gold-backed and three silver-backed ETFs and tested their return performance against existing knowledge of physical gold and silver behavior. The primary focus of the study was the largest gold and silver funds by AUM being GLD and SLV with an extension of the tests applied to IAU, DGL, DBS and AGQ.

The tests evidenced a remarkable consistency between the return properties of the ETFs and the physical assets of gold and silver identified in prior studies. The log of GLD and SLV is shown to be consistent with a stationary, white noise, distribution, or that the market for GLD and SLV are inefficient and that prices do not reflect all available information. Non-normality was also proven for all funds except AGQ although some caution should be exercised given that the
funds are still maturing meaning inefficiency is not unexpected. This information is important to investors as it shows that an abnormal return can be achieved at today.

CAPM analysis revealed an abnormal return does not exist when considered in a risk-return framework. For investors this means that when risk is factored that cannot be assured to exceed the market’s risk adjusted performance.

Application of filter trading rules suggested a marginal above market return is possibly at the 3% and 5% filter level, implying that an investor using a trading strategy has a small opportunity to outperform a passive investor.

In summary the same fundamental behavior applicable to physical gold and silver returns also applies to GLD and SLV exchange traded fund prices/returns. Specifically, their price movements do not follow a random walk. However, we show that such inefficiency which was not exploitable on physical gold and silver in the past now provides an opportunity for abnormal returns through a simple filter trading rule. This finding is especially important in lights of David Swensen’s Yale model. Gold and silver EFTS may serve as the other asset class which is particularly worthwhile in the portfolio as they are inefficient (and thus profitable) and liquid, the best of both worlds. Our finding is also valuable to investors as the markets interest in gold and silver ETFs has made the metals tangible substitutes to other more traditional securities and is unlikely to abate in the near future. As the ETF vehicle and especially interest in gold continues to grow in an exponential manner undoubtedly more investigation into funds performance will be of interest.

References


~ 22 ~


Appendix

Tables:

Table 1: Commodity ETFs Used in the Data Set

Table 1 lists the selected funds and the benchmark SPY by; ticker, name, funds family, investment asset (either gold or silver), total net assets (USD), inception and the exchange the fund is traded on (note the funds are often not limited to trading on one exchange).

<table>
<thead>
<tr>
<th>Ticker</th>
<th>Name</th>
<th>Fund Family</th>
<th>Investment Asset</th>
<th>Total Net Assets USD</th>
<th>Inception Date</th>
<th>Exchange</th>
</tr>
</thead>
<tbody>
<tr>
<td>GLD</td>
<td>SPDR Gold Shares</td>
<td>SPDRs</td>
<td>Gold</td>
<td>$40.5bn</td>
<td>Nov 2004</td>
<td>NYSE Arca</td>
</tr>
<tr>
<td>IAU</td>
<td>Comex Gold Trust</td>
<td>iShares</td>
<td>Gold</td>
<td>$2.75bn</td>
<td>Jan 2005</td>
<td>NYSE Arca</td>
</tr>
<tr>
<td>DGL</td>
<td>DB Gold</td>
<td>PowerShares</td>
<td>Gold</td>
<td>$159.03m</td>
<td>Jan 2007</td>
<td>NYSE Arca</td>
</tr>
<tr>
<td>SLV</td>
<td>Silver Trust</td>
<td>iShares</td>
<td>Silver</td>
<td>$5.22bn</td>
<td>Apr 2006</td>
<td>NYSE Arca</td>
</tr>
<tr>
<td>DBS</td>
<td>DB Silver</td>
<td>PowerShares</td>
<td>Silver</td>
<td>$62.31m</td>
<td>June 2008</td>
<td>NYSE Arca</td>
</tr>
<tr>
<td>AGQ</td>
<td>Ultra Silver</td>
<td>ProShares</td>
<td>Silver</td>
<td>$171.23m</td>
<td>Dec 2008</td>
<td>NYSE Arca</td>
</tr>
</tbody>
</table>

Benchmark Equity ETF

<table>
<thead>
<tr>
<th>SPY</th>
<th>SPDR S&amp;P 500</th>
<th>SPDRs</th>
<th>Equities</th>
<th>Total Net Assets USD</th>
<th>Inception Date</th>
<th>Exchange</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>SPDR S&amp;P 500</td>
<td>SPDRs</td>
<td>Equities</td>
<td>$77.82bn</td>
<td>1993</td>
<td>NYSE Arca</td>
</tr>
</tbody>
</table>

(NYSE Arca, 2010)

Graph 1: GLD Return vs. Gold Return
Graph 2: SLV Return vs. Silver Return

Graph 3:

Graph 4:

Figure 1: GLD Correlogram
Figure 2: SLV Correlogram

<table>
<thead>
<tr>
<th>Autocorrelation</th>
<th>Partial Correlation</th>
<th>AC</th>
<th>PAC</th>
<th>Q-Stat</th>
<th>Prob</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.005</td>
<td>0.006</td>
<td>0.027</td>
<td>8.88</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>-0.053</td>
<td>-0.033</td>
<td>1.026</td>
<td>0.588</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>-0.005</td>
<td>-0.006</td>
<td>1.066</td>
<td>0.796</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>0.014</td>
<td>0.013</td>
<td>1.069</td>
<td>0.687</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>0.021</td>
<td>0.021</td>
<td>1.063</td>
<td>0.881</td>
<td></td>
</tr>
</tbody>
</table>

Figure 3: GLD ADF Test

<table>
<thead>
<tr>
<th>Variable</th>
<th>Coefficient</th>
<th>Std. Error</th>
<th>t-Statistic</th>
<th>P-Prob</th>
</tr>
</thead>
<tbody>
<tr>
<td>RGLD(1)</td>
<td>-1.05194</td>
<td>0.01882</td>
<td>-12.04852</td>
<td>0.0000</td>
</tr>
<tr>
<td>RGLD(2)</td>
<td>-0.25820</td>
<td>0.07478</td>
<td>-3.437244</td>
<td>0.0005</td>
</tr>
<tr>
<td>RGLD(3)</td>
<td>-0.08301</td>
<td>0.05701</td>
<td>-1.456132</td>
<td>0.1470</td>
</tr>
<tr>
<td>RGLD(4)</td>
<td>0.04109</td>
<td>0.04221</td>
<td>0.96969</td>
<td>0.3375</td>
</tr>
<tr>
<td>RGLD(5)</td>
<td>0.00020</td>
<td>0.00016</td>
<td>0.126232</td>
<td>0.9006</td>
</tr>
<tr>
<td>RGLD(6)</td>
<td>0.00005</td>
<td>0.00001</td>
<td>0.12686</td>
<td>0.9055</td>
</tr>
</tbody>
</table>

Augmented Dickey-Fuller Unit Root Test on RGLD

Null Hypothesis: RGLD has a unit root
Exogenous: Constant
Lag Length: 5 (Fixed)

t-Statistic | Prob
Augmented Dickey-Fuller test statistic | -12.04852 | 0.0000
Test critical values: 1% level | 3.437244 | 0.0000
5% level | -2.568384 | 0.0000
10% level


Figure 4: SLV ADF Test

<table>
<thead>
<tr>
<th>Variable</th>
<th>Coefficient</th>
<th>Std. Error</th>
<th>t-Statistic</th>
<th>P-Prob</th>
</tr>
</thead>
<tbody>
<tr>
<td>RGLD(1)</td>
<td>-1.05194</td>
<td>0.01882</td>
<td>-12.04852</td>
<td>0.0000</td>
</tr>
<tr>
<td>RGLD(2)</td>
<td>-0.25820</td>
<td>0.07478</td>
<td>-3.437244</td>
<td>0.0005</td>
</tr>
<tr>
<td>RGLD(3)</td>
<td>-0.08301</td>
<td>0.05701</td>
<td>-1.456132</td>
<td>0.1470</td>
</tr>
<tr>
<td>RGLD(4)</td>
<td>0.04109</td>
<td>0.04221</td>
<td>0.96969</td>
<td>0.3375</td>
</tr>
<tr>
<td>RGLD(5)</td>
<td>0.00020</td>
<td>0.00016</td>
<td>0.126232</td>
<td>0.9006</td>
</tr>
<tr>
<td>RGLD(6)</td>
<td>0.00005</td>
<td>0.00001</td>
<td>0.12686</td>
<td>0.9055</td>
</tr>
</tbody>
</table>

Re-reqired: 0.486895 Mean dependent var: -1.27605
Adjusted Re-reqired: 0.476893 S.D. dependent var: 0.521653
S.E. of regression: 0.16447 Akaike info criterion: -5.496151
Sum squared resid: 2.276325 Schwarz criterion: -5.468171
Log likelihood: 293.0244 F-statistic: 151.3983
Durbin-Watson stat: 2.201662 Prob(F-statistic): 0.000005

~ 26 ~
Table 2: Summary Statistics

<table>
<thead>
<tr>
<th>Variable</th>
<th>RGLD Mean</th>
<th>RIAU Mean</th>
<th>RDGL Mean</th>
<th>RDGS Mean</th>
<th>RAGS Mean</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
<td>0.0000589</td>
<td>0.0000234</td>
<td>0.000539</td>
<td>0.0000496</td>
<td>0.0000388</td>
</tr>
<tr>
<td>Median</td>
<td>0.0000762</td>
<td>0.0000257</td>
<td>0.000556</td>
<td>0.0000381</td>
<td>0.0000404</td>
</tr>
<tr>
<td>Maximum</td>
<td>0.1065974</td>
<td>0.1347032</td>
<td>0.115834</td>
<td>0.118026</td>
<td>0.154866</td>
</tr>
<tr>
<td>Minimum</td>
<td>-0.1777210</td>
<td>-0.195483</td>
<td>-0.086747</td>
<td>-0.096846</td>
<td>-0.222324</td>
</tr>
<tr>
<td>Std. Dev.</td>
<td>0.0150444</td>
<td>0.0255225</td>
<td>0.015431</td>
<td>0.019121</td>
<td>0.026372</td>
</tr>
<tr>
<td>Skewness</td>
<td>-0.186597</td>
<td>-0.151463</td>
<td>-0.071608</td>
<td>0.090070</td>
<td>-1.15432</td>
</tr>
</tbody>
</table>

Table 5: GLD CAPM
### Figure 6: SLV CAPM

<table>
<thead>
<tr>
<th>Variable</th>
<th>Coefficient</th>
<th>Std. Error</th>
<th>t-Statistic</th>
<th>Prob.</th>
</tr>
</thead>
<tbody>
<tr>
<td>C</td>
<td>-0.002028</td>
<td>0.000600</td>
<td>-5.04632</td>
<td>0.0000</td>
</tr>
<tr>
<td>RFSLV</td>
<td>0.885432</td>
<td>0.024927</td>
<td>36.0030</td>
<td>0.0000</td>
</tr>
</tbody>
</table>

- R-squared: 0.387768  
- Adjusted R-squared: 0.387114  
- S.E. of regression: 0.015899  
- Akaike info criterion: -4.931767  
- Sum squared resid: 0.364275  
- Schwarz criterion: -4.931767  
- Log likelihood: 2354.315  
- F-statistic: 639.6517  
- Durbin-Watson stat: 1.885965  
- Prob(F-statistic): 0.000000

### Figure 7: IAU CAPM

<table>
<thead>
<tr>
<th>Variable</th>
<th>Coefficient</th>
<th>Std. Error</th>
<th>t-Statistic</th>
<th>Prob.</th>
</tr>
</thead>
<tbody>
<tr>
<td>C</td>
<td>-0.002028</td>
<td>0.000600</td>
<td>-5.04632</td>
<td>0.0000</td>
</tr>
<tr>
<td>RFSLV</td>
<td>0.885432</td>
<td>0.024927</td>
<td>36.0030</td>
<td>0.0000</td>
</tr>
</tbody>
</table>

- R-squared: 0.279418  
- Adjusted R-squared: 0.276365  
- S.E. of regression: 0.022726  
- Akaike info criterion: -4.307536  
- Sum squared resid: 0.707213  
- Schwarz criterion: -4.307536  
- Log likelihood: 1594.471  
- F-statistic: 356.2620  
- Durbin-Watson stat: 1.895965  
- Prob(F-statistic): 0.000000

### Figure 8: DGL CAPM

<table>
<thead>
<tr>
<th>Variable</th>
<th>Coefficient</th>
<th>Std. Error</th>
<th>t-Statistic</th>
<th>Prob.</th>
</tr>
</thead>
<tbody>
<tr>
<td>C</td>
<td>-0.002028</td>
<td>0.000600</td>
<td>-5.04632</td>
<td>0.0000</td>
</tr>
<tr>
<td>RFSLV</td>
<td>0.885432</td>
<td>0.024927</td>
<td>36.0030</td>
<td>0.0000</td>
</tr>
</tbody>
</table>

- R-squared: 0.392935  
- Adjusted R-squared: 0.391662  
- S.E. of regression: 0.019865  
- Akaike info criterion: -4.952035  
- Sum squared resid: 0.367467  
- Schwarz criterion: -4.952035  
- Log likelihood: 2350.319  
- F-statistic: 504.0381  
- Durbin-Watson stat: 1.866934  
- Prob(F-statistic): 0.000000
Figure 9: DBS CAPM

Dependent Variable: RFDESL
Method: Least Squares
Date: 11/20/10  Time: 02:26
Included observations: 360 after adjusting endpoints

<table>
<thead>
<tr>
<th>Variable</th>
<th>Coefficient</th>
<th>Std. Error</th>
<th>t-Statistic</th>
<th>Prob.</th>
</tr>
</thead>
<tbody>
<tr>
<td>C</td>
<td>-0.0023546</td>
<td>0.001132</td>
<td>-3.1935</td>
<td>0.0019</td>
</tr>
<tr>
<td>RFSPY</td>
<td>0.114337</td>
<td>0.018456</td>
<td>6.1476</td>
<td>0.0000</td>
</tr>
</tbody>
</table>

R-squared  0.02126  Mean dependent var  -0.0023546
Adjusted R-squared 0.019540  S.D. dependent var  0.018456
S.E. of regression  0.01527  Akaike info criterion -5.782427
Sum squared resid  0.145022  Schwarz criterion -5.029040
Log likelihood  983.5361  F-statistic  0.950651
Durbin-Watson stat  1.783130 Prob(F-statistic)  0.000001

Figure 10: AGQ CAPM

Dependent Variable: RFAGQ
Method: Least Squares
Date: 11/20/10  Time: 02:11
Included observations: 267 after adjusting endpoints

<table>
<thead>
<tr>
<th>Variable</th>
<th>Coefficient</th>
<th>Std. Error</th>
<th>t-Statistic</th>
<th>Prob.</th>
</tr>
</thead>
<tbody>
<tr>
<td>C</td>
<td>-0.0023546</td>
<td>0.001132</td>
<td>-3.1935</td>
<td>0.0019</td>
</tr>
<tr>
<td>RFSPY</td>
<td>0.114337</td>
<td>0.018456</td>
<td>6.1476</td>
<td>0.0000</td>
</tr>
</tbody>
</table>

R-squared  0.02126  Mean dependent var  -0.0023546
Adjusted R-squared 0.019540  S.D. dependent var  0.018456
S.E. of regression  0.01527  Akaike info criterion -5.782427
Sum squared resid  0.145022  Schwarz criterion -5.029040
Log likelihood  983.5361  F-statistic  0.950651
Durbin-Watson stat  1.783130 Prob(F-statistic)  0.000001

Figure 7: Whites Test GLD
### Figure 8: Whites Test SLV

<table>
<thead>
<tr>
<th>Variable</th>
<th>Coefficient</th>
<th>Std. Error</th>
<th>t-Statistic</th>
<th>Prob.</th>
</tr>
</thead>
<tbody>
<tr>
<td>C</td>
<td>0.000226</td>
<td>4.738E-05</td>
<td>4.788E-01</td>
<td>0.00000</td>
</tr>
<tr>
<td>RFSPY</td>
<td>0.000384</td>
<td>0.000470</td>
<td>0.000350</td>
<td>0.00000</td>
</tr>
<tr>
<td>RFSPY2</td>
<td>0.000849</td>
<td>0.000719</td>
<td>10.512E-01</td>
<td>0.00000</td>
</tr>
</tbody>
</table>

R-squared | 0.107355 | Mean dependent var | 0.000350 |
Adjusted R-squared | 0.105412 | S.D. dependent var | 0.000350 |
S.E. of regression | 0.000362 | Akaike info criterion | -10.0001 |
S.E. of resid | 0.000387 | Schwarz criterion | -10.0001 |
Log likelihood | 50.7E-02 | F-statistic | 55.362E-05 |
Durbin-Watson stat | 1.807E-03 | Prob(F-statistic) | 0.00000 |

### Figure 13: IAU
Figure 14: DGL

White Heteroskedasticity Test:

F-statistic 52.80042  Probability 0.000000
OBS*R-squared 99.07467  Probability 0.000000

Test Equation:
Dependent Variable: RESID*2
Method: Least Squares
Date: 1/1/2010  Time: 03:54
Sample: 5/02/2008 11/1/2008
Included observations: 522

<table>
<thead>
<tr>
<th>Variable</th>
<th>Coefficient</th>
<th>Std. Error</th>
<th>t-Statistic</th>
<th>Prob.</th>
</tr>
</thead>
<tbody>
<tr>
<td>C</td>
<td>0.000233</td>
<td>4.88E-06</td>
<td>4.828990</td>
<td>0.0000</td>
</tr>
<tr>
<td>RFSFY</td>
<td>0.000526</td>
<td>0.001923</td>
<td>5.700536</td>
<td>0.0000</td>
</tr>
<tr>
<td>RFSFY+2</td>
<td>0.285451</td>
<td>0.027770</td>
<td>10.27911</td>
<td>0.0000</td>
</tr>
</tbody>
</table>

R-squared 0.103116  Mean dependent var 0.000000
Adjusted R-squared 0.101166  S.D. dependent var 0.001000
S.E. of regression 0.001923  Akaike info criterion -10.985629
S.D. squared resid 0.000237  Schwarz criterion -10.949658
Log likelihood 5057.936  F-statistic 52.80042
Durbin-Watson stat 1.703405  Prob(F-statistic) 0.000000

Figure 15: DBS

White Heteroskedasticity Test:

F-statistic 6.295903  Probability 0.002176
OBS*R-squared 12.16135  Probability 0.002237

Test Equation:
Dependent Variable: RESID*2
Method: Least Squares
Date: 1/1/2010  Time: 03:54
Sample: 5/02/2008 11/1/2008
Included observations: 522

<table>
<thead>
<tr>
<th>Variable</th>
<th>Coefficient</th>
<th>Std. Error</th>
<th>t-Statistic</th>
<th>Prob.</th>
</tr>
</thead>
<tbody>
<tr>
<td>C</td>
<td>0.000096</td>
<td>5.61E-06</td>
<td>5.369048</td>
<td>0.0000</td>
</tr>
<tr>
<td>RFSFY</td>
<td>0.003900</td>
<td>0.002041</td>
<td>1.821051</td>
<td>0.6555</td>
</tr>
<tr>
<td>RFSFY+2</td>
<td>0.003069</td>
<td>0.003530</td>
<td>2.737718</td>
<td>0.0085</td>
</tr>
</tbody>
</table>

R-squared 0.031263  Mean dependent var 0.000000
Adjusted R-squared 0.035244  S.D. dependent var 0.001029
S.E. of regression 0.003965  Akaike info criterion -10.576588
S.D. squared resid 0.003069  Schwarz criterion -10.449011
Log likelihood 2100.020  F-statistic 6.295903
Durbin-Watson stat 2.003151  Prob(F-statistic) 0.002176

Figure 16: AGQ
Histogram 1: RGLD

Histogram 2: RSLV

Histogram 3: RIAU

Histogram 4: RDGL
### Table 3: Filtered Trading Results

**Buy and Hold Vs. Filtered Trading Average Return**

<table>
<thead>
<tr>
<th></th>
<th>Buy and Hold</th>
<th>Filtered Trading</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
<td>0.000368</td>
<td>0.007650</td>
</tr>
<tr>
<td>Median</td>
<td>0.002894</td>
<td>0.004824</td>
</tr>
<tr>
<td>Maximum</td>
<td>0.115499</td>
<td>0.116572</td>
</tr>
<tr>
<td>Minimum</td>
<td>-0.225624</td>
<td>-0.116136</td>
</tr>
<tr>
<td>Skewness</td>
<td>-1.131352</td>
<td>-0.082971</td>
</tr>
<tr>
<td>Kurtosis</td>
<td>-0.348765</td>
<td>1.850862</td>
</tr>
<tr>
<td>Jarque-Bera</td>
<td>0.967613</td>
<td>0.968997</td>
</tr>
<tr>
<td>Probability</td>
<td>0.000000</td>
<td>0.000000</td>
</tr>
<tr>
<td>Fund</td>
<td>Period</td>
<td>3% Filter</td>
</tr>
<tr>
<td>------</td>
<td>------------------------</td>
<td>-----------</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Buy and</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Hold</td>
</tr>
<tr>
<td>GLD</td>
<td>1 May 2006 – 31 Dec 2009</td>
<td>-0.000118</td>
</tr>
<tr>
<td>SLV</td>
<td>1 May 2006 – 31 Dec 2009</td>
<td>0.0000</td>
</tr>
</tbody>
</table>