# Examination of Time-Variant Asset Correlations Using High-Frequency Data

Mingwei Lei

Professor George Tauchen, Faculty Advisor

Honors thesis submitted in partial fulfillment of the requirements for Graduation with

Distinction in Economics in Trinity College of Duke University

Duke University; Durham, North Carolina; 2012

Duke University is a community dedicated to scholarship, leadership, and service and to the principles of honesty, fairness, respect, and accountability. Citizens of this community commit to reflect upon and uphold these principles in all academic and non-academic endeavors, and to protect and promote a culture of integrity.

To uphold the Duke Community Standard:

- I will not lie, cheat, or steal in my academic endeavors;
- I will conduct myself honorably in all my endeavors; and
- I will act if the Standard is compromised.

## Abstract

Drawing motivation from the 2007-2009 global financial crises, this paper looks to further examine the potential time-variant nature of asset correlations. Specifically, high frequency price data and its accompanying tools are utilized to examine the relationship between asset correlations and market volatility. Through further analyses of this relationship using linear regressions, this paper presents some significant results that provide striking evidence for the time-variability of asset correlations. These findings have crucial implications for portfolio managers as well as risk management professionals alike, especially in the contest of diversification.

**Keywords**: asset correlations, market volatility, high-frequency data, financial crisis, time-variant correlations, time-variant volatility, diversification

**JEL Codes**: G, G1, G10, G11, G14

## **1. Introduction and Motivation**

In order to successfully measure and manage market risks, portfolio managers and risk management professionals scrutinize the variability and correlations of their portfolio assets. In finance, the correlation between two assets measures the extent to which they perform similarly to one another. For instance, equity stocks of companies in the same industry tend to be highly correlated; whereas two relatively unrelated securities such as stock of a bio-technology start-up and a gold futures contract usually have uncorrelated returns. Correlation measures are used in conjunction with expect returns to calculate the overall volatility of a collection of a portfolio. Since the true underlying volatility of an asset cannot be observed, portfolio managers and risk management professionals often utilize historical price data to estimate the asset correlations to make more informed financial decisions.

Some professionals and academics have conjectured that asset correlations change over time. For example, a former global risk manager of a major financial firm noted that, "during major market events, correlations change dramatically" (Bookstaber, 1997). If correlations do in fact change overtime, the use of historical data may provide misleading measures of the current level of asset correlations. This could potentially pose serious threats to portfolio selection methods as well as risk management techniques. As a result, the behavior of asset correlations has been an important ongoing concern in both academia and the financial industry.

Portfolio managers need to consider the behavior of asset correlations over time. It is considered good practice to reduce volatility in an investment portfolio by investing in a variety of assets whose values rise and fall independently of one another. This strategy is known as diversification. The fundamental premise behind diversification is that portfolio risk can be lowered via investments in a number of different assets with varying levels of risk, volatility, and returns. The amount of diversification in a given

portfolio is dependent on the correlation between assets that are in the portfolio. Effective asset allocation should reduce idiosyncratic risk, or the risk associated with owning an individual security. In order to achieve effective diversification, the assets held in the portfolio should not be highly correlated. Typically, there is a negative relationship between correlations of assets in a portfolio and the level of portfolio diversification. Since portfolio volatility is dependent on its level of diversification and diversification in turn is related with the correlations of the assets in the portfolio, it follows that portfolio volatility is dependent on the level of correlations between its assets. Thus, if asset correlations were to suddenly change, it could have a significant impact on overall portfolio volatility. Failure to account for this change may result in higher than expected losses for portfolios during certain market conditions, especially times of market turmoil. This is consistent with the events of the 2007-2008 financial crises. In this period, numerous seemingly unrelated assets together plummeted in value. Previously welldiversified pension and mutual funds lost their diversification. For example, before 2007, Vanguard Group proudly described their Target Retirement Funds as, "broadly diversified: Underlying funds invest in 6,000-plus U.S. stocks and bonds and 2,000-plus international stocks to help spread out risk". However, diversification seems to be lost between October, 2007 and March, 2009 as the Vanguard Target Retirement Fund lost 33% of its value (Schreiner, 2009).

Financial risk managers many also need to adjust their risk management techniques if correlations exhibit changes overtime. One way through which firms manage their exposure to market risk is the method of value at risk (VAR). For a given portfolio of assets, the N-day X-percent VAR specifies the loss amount that the portfolio is expected not to exceed in N-days with X-percent certainty. Firms and regulatory agencies often use this method to determine the amount of capital firms need to hold to absorb unexpected losses. In practice, risk managers often utilize only historical data from a relatively short time interval when calculating volatility and correlations for inputs

into VAR-based models. For example, one major banking company reported to use the variation of market prices in the most recent 264 days in its calculation of VAR (Chase Manhattan 1999). The use of a relatively short time period of input data for VAR calculation does have its desirable features and appeal. Since some financial market conditions tend to change over time, using data from the distant past may lead to erroneous and outdated depictions of the market. Nevertheless, if correlations are timevariant, using only the most recent data for VAR calculations may also have dangerous implications. Specifically, if market conditions rapidly and significantly deteriorate, the estimated asset correlation using data from the recent past of market stability could be far from the actual behaviors in the market, leading to a false sense of security. As a result, this assessment of market risk may overstate the amount of diversification currently present in the portfolio, misleading firms to take on excessive risk. Similarly, VAR may also understate the amount of diversification currently present in the portfolio during times of low market uncertainty, misleading firms to further diversify. The high cost associated with excessive diversification could impose a substantial strain on companies' bottom-line.

Because of the losses many firms incurred during the financial crisis, more professionals have come to realize the importance of a more complete understanding of the behavior of correlations. For example, Cassandra Toroian, the president and chief investment officer of Bell Rock Capital LLC, recently said in an interview, "It has become more important over the years, because the world is more interconnected, for people to understand on a macro level how things relate to each other... people look to use these kinds of [correlation analysis] because they're trying to smooth out the volatility in their investments" (WSJ). The growing need for a more thorough understanding of time-variant correlation movements among portfolio managers is a central motivation for the research presented in this paper.

### **Examination of Time-Variant Asset Correlations Using High-Frequency Data**

Due to the important implications of correlation time variability, there has been some previous research done on the topic. For example, in their 2002 paper, Campbell, Koedijk, and Kofman found evidence of elevated correlations in international equity returns during periods of low market returns (Campbell, Koedijk, Kofman 2002). In another paper published in 2002, Butler and Joaquin conjectured a systematic increase in asset correlations during bear markets compared to that during bull markets (Butler, Joaquin 2002). Although many previous research hint at the time-variability of asset correlations, none has produced results supporting its existence as convincing as those presented in this paper.

This paper presents some striking evidence for the time variability of asset correlations. This is accomplished through the use of high-frequency price data. Probabilists have long known about the benefits that high-frequency data provides. Specifically, high-frequency data generates very reliable measures of volatility and correlations. Anderson and Bollerslev also noted that the use of high-frequency price data improves measurements of volatility and correlations than using just daily closing data along (1998). Most previous research on the subject of time-variant correlations utilizes only daily closing data. This research uses high-frequency price data to produce much more convincing results.

Through the use of high-frequency data, this paper aims to examine the link between asset correlations and market volatility. The primary reason behind this exploration is the idea that if a relationship could be established between asset correlations and market volatility, which is a time-variant measure, then it is reasonable to conclude that asset correlations are also time-variant. Another motivation for exploring asset correlations in conjunction with market volatility is that portfolio diversification becomes more important in market conditions characterized by heightened volatility. As noted earlier, the level of diversification present in any portfolio is directly linked to the

6

correlations between assets in that portfolio. Thus, it is crucial to understand the behavior of asset correlations in relation to market volatility.

Before proceeding, a road map of the paper is presented here. First, in section 2, this paper presents some theoretical background on stochastic models of returns, volatility, correlation, as well as the benefits of using high-frequency data. Next, section 3 introduces the high-frequency data utilized in this research and discusses some of the concerns associated with using high-frequency data. To provide justification for using high-frequency data, section 3 also offers insight into microstructure noise reduction. Then, section 4 gives a brief discussion of Fisher Transformation, a statistical tool utilized in this research. The results of this research are presented in section 5, which offers strong evidence supporting a relationship between asset correlations and market volatility. Lastly, a discussion is given in section 6 on the results provided by this research.

## 2. Model

The foundation for estimating realized variance and covariance rests on the stochastic volatility model, defined by the differential equation:

$$dp(t) = \mu(t)dt + \sigma(t)dW(t)$$
(1)

where the change in a stock's log-price dp(t) is a function of the time-variant drift component  $\mu(t)dt$  and the time-variant stochastic component  $\sigma(t)dW(t)$ , in which dW(t) is a standard Brownian motion. This model treats the price, and thus the returns, of the underlying security as a random process. The geometric returns (r) of stocks can be calculated using the logarithmic price (p):

$$p_{t,j} = \log(P_{t,j}) \tag{2}$$

$$r_{t,j} = p_{t,j} - p_{t,j-1} \tag{3}$$

Here,  $P_{t,j}$  denotes the observed price of the underlying asset at time *j* in period *t*. The use of log-price here allows for convenient calculations of the percent change in stock prices.

To find a link between market volatility and asset returns correlation, appropriate methods must be applied on the given high-frequency price data to estimate the underlying volatility as well as correlation. A common method for approximating the underlying volatility of the returns of a given asset is to calculate the realized variance (Anderson and Bollerslev, 1998). The realized variance can be computed over some time period as the sum of the squared high-frequency geometric returns within that time period:

$$RV_t = \sum_{j=1}^{M} (r_{t,j})^2$$
(4)

where r is the log return and M is the total number of data samples in period t. Anderson and Bollerslev have noted that realized variance converges in frequency to the integrated variance plus a discrete jump component:

$$\lim_{M \to \infty} RV_t = \int_{t-1}^{t} \sigma^2(s) ds + \sum_{t-1 \le s \le t} \kappa^2(s)$$
(5)

Thus, assuming the magnitude of the jump component is relatively small, realized variance can be used as an effective estimate of the underlying volatility. The research presented in the paper utilizes realized standard deviation as the measure for volatility. Realized standard deviation is defined as:

$$RStd_t = \sqrt{RV_t} \tag{6}$$

where  $RV_t$  is the realized variance as defined above.

The underlying covariance of the returns of two given assets can be estimated using the realized covariance. The realized covariance between two assets, say A and B, can be calculated over some time period as the sum of the products of high-frequency geometric returns of those two assets:

$$RCov_{AB,t} = \sum_{j=1}^{M} r_{A,t,j} r_{B,t,j}$$

$$\tag{7}$$

Anderson and Bollerslev have noted that due to the convergence property of realized covariance, it could be used as an effective estimate of asset correlations when working with high-frequency data (1998). Realized covariance could then be used in conjunction with the realized variance of both assets A and B to calculate the realized correlation coefficient:

$$RCorr_{AB,t} = \frac{RCov_{AB,t}}{\sqrt{RV_{A,t}} \times \sqrt{RV_{B,t}}}$$
(8)

This research will utilize the realized correlation as a proxy measure for asset correlations. This is because realized correlation measures provide a scale that allows for convenient statistical comparisons across asset pairs.

It is important to note that these measures are applied to high-frequency data sets. The use of high-frequency data provides many benefits over alternative data sets. For instance, high-frequency data contains more information than daily data alone. This allows for more reliable realized measures of volatility and correlation. Moreover, the recent non-parametric approach using high-frequency data has also been found successful in estimating volatility and correlation. The framework for the integration of high-frequency intraday data into the estimations of realized variance and correlations are provided by Andersen, Bollerslev, Diebold, and Labys (2003).

## 3. Data

### 3.1 Raw Data

The research presented in this paper utilizes the price data for the S&P 500 (SPY), Bank of America (BAC), Goldman Sachs (GS), JPMorgan (JPM), Coca-Cola (HPQ), Wal-Mart (WMT), and Verizon (VZ) obtained from price-data.com, a commercial data provider. These particular companies were chosen for their market capitalization (i.e. more liquidity) and representation across different industries including financial, consumer conglomerate, and telecommunication. The data sets contain minute-by-minute stock prices from 9:35 am to 3:59 pm in the period between April 4, 1997 and January 7, 2009. The S&P 500 (SPY) was used as the market index for the purpose of this research. Stocks were randomly paired up for analysis: WMT and VZ, WMT and JPM, WMT and KO, BAC and GS, JPM and GS. Note that some of the chosen stocks were not publically traded by 1997, so its corresponding data set does not date all the way back to April 4<sup>th</sup>, 1997. For those stocks, only periods in which price data were available are considered in this research.

### 3.2 Market Microstructure Noise

Since the market often do not adjust fast enough to keep the spot prices exactly in line with the fundamental values as determined by common asset pricing models, our data sets on stock prices do not precisely reflect their theoretical counterparts. Instead, observed stock prices contain a microstructure noise element that could distort the estimations of volatility and correlation. Under ideal modeling conditions, the errors of estimating volatility with realized variance decreases as sampling frequency increases; however, this is not the case when working with market prices. As mentioned above, the observed asset prices  $(P_t)$  are not always equivalent to the fundamental prices of the stock  $(S_t)$ :

$$P_t \neq S_t (exactly) \tag{9}$$

This may be due to the inherent frictions in the market that may arise from various factors such as the lack of liquidity. According to asset pricing theory, the price of a stock at any given time should be equivalent to the sum of all discounted future cash flows associated with the stock. A well-known model that implements this widely accepted theory is the Gordon growth model,

$$S = \frac{D}{k-g} \tag{10}$$

where *S*, the underlying price of the asset, is equal to the expected future dividend payment (*D*) divided by the discount rate (*k*) as determined by the CAPM minus the expected perpetual growth rate (*g*) of the dividends. When the market determines the prices of different assets using the core Gordon growth model or a more sophisticated variation of this model, small changes to the input of the model could have a significant effect on the theoretical price output. For example, supposed a company is projected to issue a dividend of \$1.00 per share in the upcoming year. Its discount rate (*k*) and perpetual growth rate (*g*) are estimated to be 8 and 6 percent respectively. The underlying price of the asset as determined by the Gordon growth model would be:

$$S = \frac{D}{k-g} = \frac{\$1.00}{0.08 - 0.06} = \$50.00 \tag{11}$$

Now, suppose the market turns more risk adverse and changes its required rate of return for similar assets from 8 percent to 10 percent. The new underlying price of the asset according to the Gordon growth model is:

$$S = \frac{D}{k-g} = \frac{\$1.00}{0.10-0.06} = \$25.00 \tag{12}$$

This example epitomizes the phenomenon that a relatively small change in the market's view of input factors such as the discount rate could have a tremendous impact on the theoretical stock price of an asset. Moreover, determining the exact discount and growth rates of various assets are known to be notoriously difficult. Combining the two above observations, it directly follows that the market price of an asset at any given point in time is extremely susceptible to microstructure noise and may not reflect the underlying value of the asset.

Given the existence of microstructure noise, it makes sense to think of the log observed price  $p_t = log(P_t)$  as the natural log of the fundamental price of the stock  $(log(S_t))$  plus a microstructure noise component  $(\epsilon_t)$ :

$$p_t = \log(S_t) + \epsilon_t \tag{13}$$

Then, the observed geometric returns can then be expressed as:

$$p_{t+\delta} - p_t = \log(S_{t+\delta}) - \log(S_t) + \epsilon_{t+\delta} - \epsilon_t \tag{14}$$

Note that the size of the theoretical component  $(log(S_{t+\delta}) - log(S_t))$  is positively correlated with the size of the interval ( $\delta$ ) while both the microstructure noise component  $(\epsilon_{t+\delta} \text{ and } \epsilon_t)$  are independent of interval size. As the data is sampled more frequently, the magnitude of  $log(S_{t+\delta}) - log(S_t)$  decreases while the magnitude of  $\epsilon_{t+\delta} - \epsilon_t$  stays the same. Microstructure noise thus becomes more pronounced in the calculations as sampling frequency increases. As a result, many researchers have chosen to sample the data at intervals ranging from five to fifteen minutes in order to minimize the effects of microstructure noise (Zhang. 2005).

### **3.3 Signature Plots**

Due to the aforementioned distortions that microstructure noise could potentially cause, it is not suitable for the purposes of this research to use all the available data to estimate market standard deviation and asset returns correlation coefficients. Ideally, it would be best to choose a high sampling frequency in order to make use of as much available data as possible when making statistical inferences; however, due to the existence of microstructure noise, a choice of a data frequency that is excessively high could accentuate unwanted distortions that could significantly alter the results of this research. Thus, it is desirable to strike a balance and find an optimal sampling frequency. In order to find an optimal sampling frequency to use for the calculation of both realized market standard deviation as well as realized correlation coefficients, statistical methods were used to create volatility as well as correlation signature plots

A volatility signature plot shows the average calculated daily realized volatility as a function of sampling frequency. Because this research is only concerned with the market volatility, it is sufficient to create volatility signature plots for the market along. A volatility signature plot for SPY is shown in figure 1, with average daily variance on the y-axis and the number of minutes between each price data on the x-axis. In the absence of market microstructure noise, the average calculated daily realized variance is uncorrelated with the sampling frequency used in the calculation. Thus, the volatility signature plot should appear relatively horizontal (Andersen, Bollerslev, Diebold, and Labys, 1999). One would expect the presence of microstructure noise to create an upward distortion on calculated volatility, especially at high sampling frequencies, as market inefficiencies cause higher variations in stock prices. As shown in figure 1, the presence of microstructure noise does indeed create a significant upward distortion at high sampling frequencies, especially when working with data more dense than that at the 5minute sampling frequency.

#### **Examination of Time-Variant Asset Correlations Using High-Frequency Data**

Analogously, a correlation signature plot shows the average calculated daily realized correlation between the returns of two assets as a function of sampling frequency. Because the research examines the realized correlation coefficients of every asset pair chosen, correlation signature plots were created for each asset pair. However, for presentation purposes, only the correlation signature plot for Bank of America (BAC) and Goldman Sachs (GS) is included (figure 2). The average daily realized correlation coefficient is graphed on the y-axis while the number of minutes between each price data is shown on the x-axis. As is the case with volatility, average calculated asset correlations should be uncorrelated with sampling frequency in the absence of microstructure noise. In that case, the correlation signature plots would appear relatively horizontal. Given the existence of microstructure noise, one would expect a downward distortion on the correlation coefficient between positively correlated assets because market inefficiencies cause asset prices to have independent variability. This is indeed the case with all pairs of assets tested. Figure 2 exemplifies the downward distortion on asset returns correlations. The overall appearance of figure 2 generalizes to correlations signature plots of all asset pairs examined. The distortions caused by microstructure noise are especially significant when working with data more dense than that at the 5-minute sampling frequency.

The volatility and correlation signature plots both demonstrate the potential significant distortionary effects that market microstructure noise can cause. However, the distortions seem to be dramatically less significant beyond the 10-minute level. The 11-minute sampling frequency was chosen as the optimal to be used for statistical inference in this research. The 11 minute sampling frequency balances the trade-offs between sample size and noise distortions. Moreover, 11 divides event into 385, the number of minutes in a trading day. Thus the choice of 11 minute sampling frequency is also convenient when working with the data set.

### **3.4 Data Partitions**

To get a better sense of the fundamental relationship between market standard deviation and asset returns correlation coefficients examined in this research, it is desirable to explore the relationship over different length of periods. Thus, three partitions were used to make statistical inferences for each asset pair. The first partition separates the data into disjoint 1-day periods; the second partition separates the data into disjoint 20-day periods. For each partition, realized market standard deviation and realized asset-pair returns correlation coefficients were calculated over each separated period and matched up for analysis.

## 4. Fisher Transformation

This research utilizes a common method in statistics called the Fisher Transformation. Hypotheses about the value of the population correlation coefficient between variables X and Y can be tested using the Fisher transformation applied to the sample correlation coefficient ( $\rho$ ). The transformation is defined by:

$$\hat{\rho} = \frac{1}{2} \ln \frac{1+\rho}{1-\rho} \tag{15}$$

If (*X*, *Y*) has a bivariate normal distribution and the (*X<sub>i</sub>*, *Y<sub>i</sub>*) pairs used to form the sample correlation coefficient are independent, then  $\hat{\rho}$  is approximately distributed

$$N\left(\frac{1}{2}\ln\frac{1+\rho}{1-\rho},\frac{1}{\sqrt{n-3}}\right) \tag{16}$$

where n is the sample size. The Fisher transformation maps the correlation coefficient, which has a range of [-1, 1], to the entire real numbers set  $(-\infty, \infty)$ .

If one were to examine the relationship between two variables, say  $\rho$  and A, where  $\rho$  is distributed on [-1, 1] and A on  $(-\infty, \infty)$ , it would not be ideal to directly implement a linear regression analysis on the variables. This is because linear regression works the best when it is implemented on two variables that are both distributed on  $(-\infty, \infty)$ . It does not provide reliable results for a variable that is distributed between -1 and 1. In this case, since  $\rho$  is distributed on [-1, 1], a regression analysis should instead be done on the Fisher transformed  $\rho$ , call this  $\hat{\rho}$ , and A. Now, the regression becomes

$$\hat{\rho} = \frac{1}{2} \ln(\frac{1+\rho}{1-\rho}) = \hat{\beta}_0 + \hat{\beta}_1 \times A \tag{17}$$

where  $\hat{\beta}_0$  and  $\hat{\beta}_1$  are the regression coefficients. The Fisher transformation establishes the quasi-normality that is desired for regression analysis.

### **Examination of Time-Variant Asset Correlations Using High-Frequency Data**

If one uses the estimated regressions for forecasting or just plotting the regression line, the output values would be in units of the transformed variable,  $\hat{\rho}$ ; however, since the primary objective is to examine the relationship between  $\rho$  and A, one has to reverse the transformation on the predicted value of  $\hat{\rho}$ . Thus, the relationship becomes:

$$\rho = \frac{e^{2\hat{\rho}} - 1}{e^{2\hat{\rho}} + 1} = \frac{e^{2(\hat{\beta}_0 + \hat{\beta}_1 \mathbf{A})} - 1}{e^{2(\hat{\beta}_0 + \hat{\beta}_1 \mathbf{A})} + 1}$$
(18)

This process is exactly analogous to using the log of a variable as the dependent variable in a linear regression. The logarithmic function transforms a variable that is strictly positive to the entire real number set. Like the Fisher transformation, log transformation is often used to establish normality for linear regression analysis.

The Fisher transformation adjustment of asset correlations provides a scale that allows for direct implementation of linear regression analysis. For the purposes of the research presented in this paper, Fisher transformation is primarily implemented to transform the realized correlation coefficient, which is distributed on the interval [-1, 1], to a normally distributed random variable with range equal to the entire real number set. This new random variable is well-behaved and thus could be used for a more significant implementation of regression analysis.

## 5. Results

This section presents some striking evidence supporting the time variability of correlations. These results are presented through a myriad of plots and linear regression results. In order to effectively show the results, this section first offers a discussion of the process to create the plots using high-frequency data. Next, the results of different linear regression implementations are presented. Finally, this section concludes with a discussion of the results of linear regression analysis.

Numerous plots were created throughout the research process for various asset pairs using different partitions. The plots of the asset-pair Bank of America and Goldman Sachs using the 5-day sampling period length partition are provided in this paper for visual representation (figures 1 through 5). The choice to present this asset pair is random. Moreover, it should be noted that the correlation trends shown in the provided plots generalize to all data partitions as well as all other asset pairs examined in the research.

The first set of figures plot the realized correlation coefficients and realized market standard deviation to get an initial sense of the data. The realized correlation coefficient of the asset pair and the realized market standard deviation of the S&P 500 were calculated for each partition (1-day, 5-days, 20-days length sampling periods) using method described in the section 2. A scatter plot was generated for each partition with returns correlation coefficient on the vertical-axis and market returns standard deviation on the horizontal (in the case of figure 3, 5-day sampling period partition). Each data point on the graph plots the realized returns correlation of the two assets (in the case of figure 3, Bank of America and Goldman Sachs) and the realized market standard deviation for a particular sampling period. The graph seems to show a positive relationship between two variables, however, the relationship does not look linear.

#### **Examination of Time-Variant Asset Correlations Using High-Frequency Data**

Next, log-transformations were implemented on market standard deviation in an attempt to achieve normality for regression analysis. Because realized market standard deviation is always positive, it is not a good input for regression analysis. The logarithmic function maps the realized market standard deviation from  $[0, \infty)$  to  $(-\infty, \infty)$ . This transformation produces a near-normal distribution and provides a better scale for regression analysis. The plot of realized returns correlation versus log-transformed realized market standard deviation is provided in figure 4. Each data point on the graph plots the realized returns correlation coefficient and the log-transformed realized market standard deviation for a particular sampling period. The scatter plot still does not seem to suggest a completely linear relationship. This is because the realized correlation coefficients have not been normalized yet, but a regression analysis was nonetheless implemented. Specifically, the regression analysis was implemented on the realized asset returns correlation versus log-transformed realized asset returns correlation versus log-transformed realized asset returns correlation for a particular sampling period.

$$Corr_i = \hat{\beta}_0 + \hat{\beta}_1 \times ln(MktStd_i)$$
<sup>(19)</sup>

The results of the linear regression are shown in tables 1 through 3 and a graph of the linear regression for the asset pair Bank of America and Goldman Sachs using the 5 minute sampling frequency is included in figure 5.

Finally, the realized asset correlation coefficients are normalized through the Fisher transformations. Since correlation coefficients is distributed between -1 and 1, directly linear regression analysis may not be desirable. Instead, the Fisher transformation maps realized asset correlation coefficients from [-1, 1] to  $(-\infty, \infty)$ . The distribution of the new variable is near-normal, allowing for robust regression analysis. The plot of Fisher-transformed realized asset returns correlation versus log-transformed realized market standard deviation is provided in figure 5. Each data point on the graph plots the Fisher-transformed realized returns correlation coefficient and the log-transformed realized realized realized returns correlation coefficient and the log-transformed realized rea

market standard deviation for a particular sampling period. The scatter plot seems to suggest a linear relationship between the two variables. Moreover, both Fishertransformed realized asset correlation coefficients and log-transformed realized market standard deviation are near-normally distributed. Thus, linear regression analysis could implemented to effectively analyze their relationship.

Linear regressions of Fisher transformed realized asset returns correlation coefficients versus log-transformed realized market standard deviation were generated:

$$\frac{1}{2}\ln(\frac{1+Corr_i}{1-Corr_i}) = \hat{\beta}_0 + \hat{\beta}_1 \times \ln(MktStd_i)$$
(20)

The numerical results of the regressions are presented in tables 1 through 3 and a graph of the regression for the asset pair Bank of America and Goldman Sachs using the 5 minute sampling frequency is included in figure 5.

All of the regression coefficients and t-statistics are summarized in tables 1-3. The tstatistics suggests that the relationship between returns correlation and log market standard deviation, Fisher transformed returns correlation and market standard deviation, as well as Fisher transformed standard deviation and log market standard deviation are significant at virtually every significance level for every asset pair examined, suggesting the existence of a positive relationship exists between the returns correlation of stocks and market volatility.

Figure 5 presents striking evidence supporting the time variant nature of asset correlations. It is clear from figure 5 that there is a significant positive linear relationship between Fisher-transformed asset correlation coefficients and log-transformed market standard deviation. This suggests that there is also a significant relationship between asset correlation coefficients and market standard deviation. Although this relationship is not linear, it is still reasonable to infer that asset correlation changes as volatility in the

market changes. Since it is known that market volatility is time variant, one can conclude that asset correlations are time variant as well.

## 6. Conclusion and Discussions

The aim of this research is to explore the time-variant nature of volatility using highfrequency data. This is done through an empirical exploration of the relationship between asset correlations and market volatility using high-frequency data. The most notable difference between this research and previous research on the topic is the use of highfrequency data. High-frequency data generates very reliable measures of volatility and correlations which could be used to produce more accurate results. The results of this research present striking evidence supporting the time-variant nature of asset correlations (figure 5). Specially, the research results strongly support the existence of a statistically significant positive relationship between assets returns correlation and market volatility. The finding has crucial implications for portfolio managers as well as risk management practices.

In addition to the potential problems that the link between correlation and market volatility could cause for risk managers discussed in the introduction, the validity of stress testing and worst case scenario analysis could also be threatened by this phenomenon. These risk measures utilize recent price data to consider the possible risks that could be experience in a period of high volatility; however, the measures do not explicitly take into account the increase in asset correlations during period of market stress. These techniques could mislead the risk manager into a false sense of security when the portfolio is in fact not sufficiently diversified. Instead, risk managers should also consider information from historical periods of heightened volatility to form estimates of correlation conditioned on volatility. These conditional correlations could be more relevant then be used to evaluate the possible detrimental effects due to a period of high volatility. Risk managers should not consider the possible effects of a period of high volatility without also taking into account the effect of increased asset correlations.

### **Examination of Time-Variant Asset Correlations Using High-Frequency Data**

Finally portfolio managers must understand the relationship between volatility and correlations. Portfolio managers largely depend on diversification to hedge various potential market risks. The degrees to which managers diversify their portfolio mainly depend on the returns correlation of assets held in their portfolio. If diversification was done on the basis of correlations calculated with data from periods of low market volatility, it may not be sufficient in a bear market when market experience high volatility. Correlation breakdowns occur in periods of high market volatility, causing a decrease in the degree of diversification precisely when it is needed the most. Portfolio manager must be aware of this phenomenon and diversify with the link between volatility and correlation in mind.

	X-variable	Y-variable	β	t-stat	β <sub>o</sub>	t-stat	R <sup>2</sup>
BAC & GS	Corr	log(MktStd)	0.148	17.55	1.140	28.39	0.0955
	Fisher Corr	log(MktStd)	0.229	18.11	1.600	26.20	0.1200
JPM & GS	Corr	log(MktStd)	0.134	15.77	1.098	27.36	0.0835
	Fisher Corr	log(MktStd)	0.214	16.51	1.572	25.06	0.1058
WMT & JPM	Corr	log(MktStd)	0.1563	18.68	1.086	27.76	0.1019
	Fisher Corr	log(MktStd)	0.207	18.54	1.373	25.92	0.1151
WMT & KO	Corr	log(MktStd)	0.148	15.29	0.983	21.59	0.0851
	Fisher Corr	log(MktStd)	0.188	15.55	1.206	21.02	0.0975
WMT & VZ	Corr	log(MktStd)	0.195	22.66	1.243	30.42	0.1634
	Fisher Corr	log(MktStd)	0.247	21.97	1.533	28.30	0.1768

## Table 1- Regression Results (Sampling Period Length- 1 day)

#### KEY

Corr: Correlation Coefficient Fisher Corr: Fisher Transformed Correlation Coefficient MktStd: Market Standard Deviation log(MktStd): Log Market Standard Deviation

	X- variable	Y-variable	β	t-stat	β <sub>o</sub>	t-stat	R <sup>2</sup>
BAC & GS	Corr	log(MktStd)	0.1290	9.50	0.956	17.79	0.1131
	Fisher Corr	log(MktStd)	0.1878	10.01	1.248	16.52	0.1348
JPM & GS	Corr	log(MktStd)	0.1004	8.42	0.881	18.44	0.0860
	Fisher Corr	log(MktStd)	0.1537	8.85	1.161	10.45	0.1014
WMT & JPM	Corr	log(MktStd)	0.1233	10.17	0.832	17.82	0.1336
	Fisher Corr	log(MktStd)	0.1531	10.31	0.976	16.84	0.1431
WMT & KO	Corr	log(MktStd)	0.1242	8.78	0.768	13.97	0.1298
	Fisher Corr	log(MktStd)	0.1504	8.65	0.889	13.02	0.1439
WMT & VZ	Corr	log(MktStd)	0.1725	13.14	0.987	18.91	0.2462
	Fisher Corr	log(MktStd)	0.210	12.88	1.156	17.68	0.2679

## Table 2- Regression Results (Sampling Period Length- 5 days)

KEY

Corr: Correlation Coefficient Fisher Corr: Fisher Transformed Correlation Coefficient MktStd: Market Standard Deviation log(MktStd): Log Market Standard Deviation

	Х-	Y-variable	β	t-stat	β <sub>o</sub>	t-stat	R <sup>2</sup>
	variable						
BAC & GS	Corr	log(MktStd)	0.1314	5.74	0.869	11.64	0.1338
	Fisher Corr	log(MktStd)	0.186	6.10	1.098	10.80	0.1562
JPM & GS	Corr	log(MktStd)	0.0843	4.60	0.763	11.98	0.0858
	Fisher Corr	log(MktStd)	0.1316	5.03	0.979	10.67	0.1030
WMT & JPM	Corr	log(MktStd)	0.1229	6.69	0.735	12.59	0.2021
	Fisher Corr	log(MktStd)	0.1470	6.84	0.836	11.99	0.2084
WMT & KO	Corr	log(MktStd)	0.1240	5.90	0.672	9.92	0.2040
	Fisher Corr	log(MktStd)	0.1463	5.80	0.756	9.20	0.2149
WMT & VZ	Corr	log(MktStd)	0.1618	8.12	0.828	12.39	0.3351
	Fisher Corr	log(MktStd)	0.1947	8.06	0.951	11.70	0.3593

## Table 3- Regression Results (Sampling Period Length- 20 days)

KEY

Corr: Correlation Coefficient Fisher Corr: Fisher Transformed Correlation Coefficient MktStd: Market Standard Deviation log(MktStd): Log Market Standard Deviation

## **SPY Volatility Signature Plot**



## **BAC and GS Correlation Coefficient Signature**

## (Sampling Period Length- 5 day)



## **BAC and GS Returns Correlation Coefficient versus Market**

## **Standard Deviation**

## (Sampling Period Length- 5 Days)



Each data point represents the Bank of America and Goldman Sachs realized returns correlation coefficient and market (SPY) realized returns standard deviation calculated over a 5-day period. The market standard deviation is graphed on the x-axis while the correlation coefficient is graphed on the y-axis.

### **BAC and GS Returns Correlation Coefficient versus Log Market**

## **Standard Deviation**

## (Sampling Period Length- 5 days)



Each data point represents the Bank of America and Goldman Sachs realized returns correlation coefficient and log-transformed market (SPY) realized returns standard deviation calculated over a 5-day period. A linear regression ( $Corr_i = 0.956 + 0.1290 \times ln(MktStd_i)$ ) is also shown. The log-transformed market standard deviation is graphed on the x-axis while the correlation coefficient is graphed on the y-axis.

### BAC and GS Fisher-Transformed Returns Correlation versus Log-

## **Transformed Market Standard Deviation**

## (Sampling Period Length- 5 days)



Each data point represents the Bank of America and Goldman Sachs Fisher-transformed realized returns correlation coefficient and log-transformed market (SPY) realized returns standard deviation calculated over a 5-day period. A linear regression (*FisherCorr<sub>i</sub>* =  $1.248 + 0.1878 \times ln(MktStd_i)$ ) is also shown. The natural log-transformed market standard deviation is graphed on the x-axis while the Fisher-transformed correlation coefficient is graphed on the y-axis.

## References

Andersen, Torben and Tim Bollerslev (1998). Answering the Skeptics: Yes, Standard Volatility Models Do Provide Accurate Forecasts. *International Economic Review 39 no. 4*, 885-905.

Andersen, Torben, Tim Bollerslev, Francis Diebold, and Clara Vega (2003). Micro Effects of Macro Announcements: Real-Time Price Discovery in Foreign Exchange. *The American Economic Review 93*, 38-62.

Bookstaber, Richard (1997). Global Risk Management: Are We Missing the Point. *Journal of Portfolio Management*, pp. 102–107.

Butler, K.C. and Joaquin, D .C (2002). Are the gains from international portfolio diversification exaggerated? The influence of downside risk in bear market, *Journal of International Money and Finance*, pp. 981-1011

Campbell, Koedijk, and Kofman (2002). Financial Analysts Journal, Vol. 58, Iss. 1; p. 87

Chase Manhattan Corp. Annual Report 1998. New York, 1999.

Engle, R. (2002). Dynamic conditional correlation – A Simple Class of Multivariate GARCH Models. *Journal of Business and Economic Statistics*, 20, 339-350.

Groenen, P. J. F., and Franses, P. H. (2000): Visualizing time-varying correlations across stock markets. *Journal of Empirical Finance*, 7, 155-172.

Gupta, R, & Mollik, A.T. (2008). Volatility, time varying correlation and international portfolio diversification: an empirical study of australia and emerging markets. *International Research Journal of Finance and Economics*, 18<sup>th</sup> Issue

MARTE, J. (2010, October 4th). Why the math of correlation matters. *Wall Street Journal* 

Schreiner, R (2009, December 21<sup>st</sup>). Diversification is not enough. *Morningstar* 

Yang, S.Y. (2005). A DCC Analysis of International Stock Market Correlations: The Role of Japan on the Four Asian Tigers. *Applied Financial Economic Letters*, 1, 89-93.