JUNE 2017

Alice's Adventures in Volatility-land

Jeremy Evnine
John Raus

Athena Capital Advisors LLC
55 Old Bedford Road
Lincoln, MA 01773
781.274.9300
athenacapital.com

ATHENA
Capital Advisors

MULTI-ASSET CLASS PORTFOLIOS PROVIDE UNIQUE CHALLENGES REGARDING THE MEASUREMENT OF VOLATILITY.[1] *Why should we care? Managing portfolios according to Modern Portfolio Theory directs us to find investment portfolios with the highest level of return for a given level of risk. Volatility is the measure of that risk. As such, we need good estimates of volatility for strategic asset allocation, prospective portfolio management decisions of buying or selling one asset versus another, and meaningful performance attribution and benchmark comparisons.*

The first challenge occurs when calculating volatility for a portfolio of asset classes that have differing valuation periodicities. These periodicities include intraday for liquid exchange-traded securities, daily for liquid mutual funds, monthly for hedge funds and certain Limited Partnership (LP) vehicles, and quarterly for illiquid private investments. The second challenge is picking an appropriate look-back period, as volatility is influenced by events that occur during the period selected for estimation, making the length of look-back an important factor. For example, estimating volatility during a time period that includes the Global Financial Crisis of 2008 will almost certainly be higher than if excluding it. The third challenge presents itself when we have a limited number of observations resulting in a misleading estimate of volatility. An example would be a quarterly-valued fund with a single year track record, leaving us with only four data points with which to estimate volatility. Lastly, clients need to understand the volatility of their portfolios versus a benchmark, or "tracking error."

In the first part of this note, we examine some of these issues in greater detail. In the second, we discuss how we can minimize their impact in our volatility estimate. In the third, we discuss tracking error and its implications for portfolio management.

I. THE EFFECT OF PERIODICITY, LOOK-BACK, AND LIMITED DATA

Analyzing Periodicity and Look-back

Given a set of historical returns, the traditional approach to computing volatility is to compute standard deviation. We calculate this by subtracting the average of the series from each data point, which de-means the series, then we square each result, and finally we take the square root of the average of those squared results. If we apply this calculation to a specified time period then we will in effect end up calculating an estimate of the average volatility over the period covered.

The frequency of observation may have an effect on this calculation. For example, if a price tends to be mean-reverting, then annualizing a volatility calculated from daily returns can be expected to be higher than annualizing a volatility calculated from monthly returns. The industry standard calculation assumes that periodic volatility grows with the square root of the length of the period. The accuracy of this formula holds true if a price follows a "random walk" i.e. there is no discernable pattern between price today and the price yesterday, but is reduced in the presence of serial correlation (our mean-reverting example above).

To demonstrate the effects of varying periodicity and look-back, we now consider making the standard volatility calculation on a pair of indices—Energy

Equities and Developed Large Cap Equities—that reasonably represent sub-asset classes (SACs) present in the liquid part of an Athena portfolio. Our calculation will be computed as of the week ending July 1st, 2016 over periods of 1, 2, 3, 5 and 10 years (more precisely, 52, 104, 156, 260 and 520 weeks), using weekly, monthly and quarterly returns (more precisely, 1, 4 and 13 weeks). We obtain the following tables in Fig 1:

FIG 1: ANNUAL VOLATILITY

ENERGY EQUITIES

Periodicity (weeks)

Look-back (weeks)	1	4	13
52	31.8%	30.8%	45.5%
104	28.0%	31.2%	38.0%
156	24.7%	27.4%	32.8%
260	23.1%	24.3%	27.5%
520	27.5%	29.4%	32.6%

Source: Bloomberg[2]

DEVELOPED FOREIGN LARGE CAP EQUITIES

Periodicity (weeks)

Look-back (weeks)	1	4	13
52	17.2%	19.1%	11.6%
104	14.8%	15.7%	11.8%
156	13.8%	14.9%	12.5%
260	16.5%	17.7%	14.2%
520	20.6%	19.8%	20.0%

Source: Bloomberg[2]

In Figure 2 below, which highlights the magnitude change within each matrix by showing the ratio of volatility compared to the top left cell, we see for both SACs that the average volatility generally declines as we go from one to two to three years back, suggesting that recent volatility has been higher than volatility further back. But then average volatilities tick up: 10-year average volatility is higher than 5-year average volatility for both SACs. This is to be expected, since the 10-year look-back includes the 2008 Global Financial Crisis, whereas the 5-year look-back period does not. This shows that the period over which we compute average volatility can crucially affect the number that we obtain and use. We highlight extreme differences in green and red. (The opposite effects can also occur depending on the order of events of the specific examples chosen).

Regarding frequency of observations, we see that for Energy Equities, volatility based on quarterly returns (i.e. 13 weeks of data) is higher than volatility based on monthly returns (i.e. 4 weeks of data). In the case of Developed Foreign Large Cap Equities, volatilities increase slightly when we go from weekly to monthly observations (except in the case of the 10-year average volatility, which includes 2008), and decrease when we go from monthly to quarterly observations (except, again, for the 10-year average).

We can see that the magnitude of these differences can be non-trivial. In general, we should expect that more frequent observations result in more precise estimates. For example, we have 13 times as many observations for weekly returns than for quarterly returns. So, we might expect the estimated volatility to be $\sqrt{13} = 3.6$ times as accurate. In other words, to get the best measure of volatility, use the longest look-back period and highest frequency data possible.

FIG 2: VOLATILITY RATIOS COMPARED TO FIG 1'S TOP LEFT VALUE

ENERGY

Periodicity (weeks)

Look-back (weeks)	1	4	13
52	1.00	0.97	1.43
104	0.88	0.98	1.20
156	0.78	0.86	1.03
260	0.73	0.76	0.87
520	0.86	0.92	1.03

Source: Bloomberg[2]

DEVELOPED FOREIGN LARGE CAP

Periodicity (weeks)

Look-back (weeks)	1	4	13
52	1.00	1.11	0.67
104	0.86	0.91	0.69
156	0.80	0.86	0.72
260	0.96	1.03	0.82
520	1.20	1.15	1.16

Source: Bloomberg[2]

Problems with Limited Data

Note that in the case where we have a short look-back period and infrequent data (e.g. the top right corner in the above tables), we will obtain extremely unreliable estimates of volatility. Volatility measures that rely on short look-back periods or have few observations can be very misleading! If we have 52 weeks of history, and only sampling quarterly (13-week) data, we will have only 4 returns in our calculation. The volatility calculation based on this quarterly data would be less precise than if we were able to obtain a series of 52 weekly data points, or 520 data points if we had 10 years of weekly data. However, for illiquid assets weekly frequency is not obtainable.

In the case of illiquid assets, such as Private Equity and Private Real Estate, we may have only infrequent price observations, e.g. quarterly. There is nothing we can really do to increase the frequency of observations. But at any frequency, the prices we "observe" are frequently generated by the manager of the asset. For example, we may observe that a PE fund's value, which might be expected to correlate highly with a small cap index, fluctuates much less than the corresponding index. Does it really have a lower beta, or could the manager be marking the price to reduce reported price volatility? This would have the effect of making his fund look more attractive if the public markets drop and then recover, since the PE manager might show only a slight dip and a mild recovery. By the end, the price is realistic, but the intermediate price could have been manipulated or unrealistic. For such asset classes, we test for smoothing and de-smooth using an autoregressive model.

II. MINIMIZING THE IMPACT OF PERIODICITY, LOOK-BACK, LIMITED DATA

There are some clever techniques to get the most information out of short volatility data series. Now that we have gone over the issues with periodicity, look-back, and limited data, we will transition to how we can minimize their impact and get the best covariance (or correlation) matrix given the available historical data.

Data Periodicity

To minimize the issue caused by periodicity, our best approach is to use the most frequent historical data available between any two pairwise sub asset classes (SACs). If we have weekly returns for both SACs, we should use that. If we have only monthly, or quarterly, for both, then we would need to use that since it is the best we can do. But what if we have observations of different frequencies, say monthly for SAC "i" and only quarterly for SAC "j"? In that case, we have no choice but to cumulate the monthly returns for i into quarterly returns, and then compute the correlation between i and j using quarterly returns.[3] Note that we may have different look-backs for different SACs, or we might have missing observations for some SACs. For each pair of SACs, we should use as many good return couples as we have. To summarize—different pairs of SACs may use different pairs of returns to compute the correlations.

If we do this for all pairs, we can assemble these into a covariance or correlation matrix. But there is a problem. There are so-called triangle inequalities that must be satisfied by the elements of a correlation matrix.[4] If these fail, we may be able to construct a portfolio that has negative variance.[5] This would clearly indicate that we have a bad correlation matrix, as negative volatility is nonsensical for a portfolio of financial assets. To accommodate for this, we can adjust the correlation matrix using an algorithm due to Higham.[6] Alternatively, we can use a dimensionality reduction technique, such as Principle Components Analysis, to make the matrix more parsimonious and thus eliminate such anomalies.

Data Look-back

To minimize the issue caused by look-back, we can choose to incorporate a function that decays the weight of the observations with time. If we do this, older observations do not have as much impact on the covariance matrix as newer ones. Exponential decay is the usual method for doing this.[7] This approach has the benefit of better capturing the volatility characteristics of active managers who tend to turn over their portfolios since it gives a higher weight to a manager's most recent performance (and portfolio construction). In some cases we may either want to slow down the decay rate considerably, or not include the decay at all. For example, if we are modeling US Equity volatility over long time frames, we might want earlier data to have as much weight as more recent data to best capture the average volatility during the time period. To summarize—decaying the observations is entirely dependent on what volatility profile one is trying capture and understand.

Data Length

When we have limited data such as four quarterly returns, one thing we can do is to make an assumption about the mean return, and use that to de-mean the observations. With so few observations, the sample mean is likely to be a pretty bad estimate of the mean, and a value derived from other considerations might be just as good, or better. However, we still have only four observations. So another thing we might do is to define (again, from other considerations) a "Bayesian prior" value as an estimate for the volatility. This prior may be the volatility of the index of a manager's sub asset class benchmark. We can then compute a "Bayesian posterior" estimate by taking a weighted average of the prior volatility and the sample volatility.

The more observations we have, the less weight we give to the prior. In the case of only four observations, the prior will get a lot of weight. After two years (eight observations) it will get less weight. If we have five years of quarterly observations, the prior might get very little weight. We mention the manager's sub asset class benchmark in our example, but we might use observations from other portfolios, other managers, other markets, other time periods, etc. Using any reasonable prior, and a reasonable weighting scheme, will probably generate more reasonable results than simply using a sample volatility based on four or eight observations.

III. TRACKING ERROR

In this section we talk about how the issues involved in estimating volatility affect the volatility-based measurement of "tracking error." Tracking error measures active risk—how closely a portfolio tracks a benchmark—and is a major consideration in managing a portfolio. To calculate tracking error, we subtract a series of benchmark returns from a series of portfolio returns, then square each result, and finally take the square root of the average of those squared results. In the calculation of tracking error, the standard deviation of the data is calculated as in relation to a reference benchmark; in contrast to the basic volatility calculation discussed up to this point in the paper where the standard deviation is calculated only in relation to the data set's own mean.

If the goal were to minimize tracking error, we would passively invest in the benchmark by assembling a portfolio identical to the benchmark. The only tracking error (revealed as a difference in realized performance) would be from the fees associated with doing so (trading costs, fund expenses, etc.). However, most commonly, the goal is to outperform the benchmark—and by definition, you cannot outperform the benchmark if you own a portfolio identical to the benchmark. Therefore, to outperform the benchmark, we must take some degree of active risk by deviating from the benchmark's constituents. We could do so by allocating to the same underlying components in different weights than the benchmark, by investing in components not found in the benchmark, or a blend of the two. This is a double-edged sword in that the portfolio might underperform the benchmark during certain shorter periods of time, though the goal of taking active risk and incurring tracking error is to outperform the benchmark over longer periods of time. In other words, tracking error is a general assessment of the potential for a portfolio's performance to deviate from that of its reference benchmark as well as a notion of the investor's tolerance for performance outliers in up and down markets in pursuit of long term outperformance. Sometimes, tracking error is incurred to meet certain non-financial goals such as constraints on owning certain kinds of stocks. In either case, the investor has a need to know what kind of performance deviation from the benchmark that may occur over the short term in order to achieve the desired longer term objective.

Let us look at a case study in which we estimate the tracking error of a fund that takes active risk and compare it to a passive ETF. The active fund is the Fidelity® Magellan® Fund (FMAGX), the passive fund is the SPDR S&P 500 Trust ETF (SPY), and both are benchmarked to the S&P 500 index. It is critical that when we compute tracking error, we seek the best estimate possible. To do that, we must align the periodicity and look-back of both the active fund/passive ETF and the benchmark, which we do in the figure below where we estimate tracking error using weekly, monthly, and quarterly periodicity over 1, 2, 3, 5, and 10 years. We can see that various combinations of periodicity and look-back may result in a different estimate for both investments.

Starting with FMAGX, if we use a look-back period of one year, tracking error estimated from using 52 weeks of performance is lower than if we use 4 quarters of performance. This difference could be the result of noise from using both a short look-back period combined with limited data. It could also be effects from institutional investors concurrently adjusting allocations to FMAGX at quarter end, which tend to increase volatility of the fund versus the S&P 500 benchmark on those specific days.[8] In every case as we move from weekly to quarterly data, we see tracking error increase, as the weekly data would smooth the "outlier" volatility caused by quarter-end institutional portfolio adjustments. As such, our more precise estimate of tracking error in this case would involve using weekly periodicity.

We also see that as we extend the look-back period, excluding the prior year, tracking error increases. This could be that FMAGX is taking takes less active risk in recent periods than it did historically, or that recent equity correlations are higher than historical ones. Since FMAGX is an active manager, we want to focus on their current portfolio and investment strategy. Therefore its tracking error over the last 10 years is not as meaningful as its tracking error in the past few years, or even the past year, when analyzing its current portfolio. The most meaningful tracking error estimate is that which is calculated from its current portfolio, as this is our best expectation of what forward looking tracking error may be. Note that adding a decay function would supplement the accuracy of the tracking error estimate as it would most heavily weight the current portfolio's deviations vs the benchmark in the calculation.

Comparatively, the tracking error of the passive ETF to the S&P 500 is far lower than that of the active fund. The low tracking error is the result of zero active risk by the ETF, and is essentially the combination of management fees and the impact of trading liquidity before traders arbitrage it back to the underlying performance of the S&P 500. Choosing the periodicity and look-back In other words, tracking error is a general assessment of the potential for a portfolio's performance to deviate from that of its reference benchmark as well as a notion of the investor's tolerance for performance outliers in up and down markets in pursuit of long term outperformance. Sometimes, tracking error is incurred to meet certain non-financial goals such as constraints on owning certain kinds of stocks. In either case, the investor has a need to know what kind of performance deviation from the benchmark that may occur over the short term in order to achieve the desired longer term objective.

In this case it has little impact on the tracking error calculation with the exception of the weekly, 10 year look-back which includes the 2008 Global Financial Crisis in which trading liquidity had a higher-than-average impact on the ETF performance compared to the S&P 500.

ACTIVE FUND TRACKING ERROR: FMAGX TO S&P 500

Look-back (weeks) \ Periodicity (weeks)	1	4	13
52	3.6%	4.5%	4.7%
104	3.1%	3.7%	3.5%
156	3.2%	3.9%	3.6%
260	3.7%	4.2%	4.7%
520	5.8%	6.3%	6.2%

Source: Bloomberg

PASSIVE ETF TRACKING ERROR: SPY TO S&P 500

Look-back (weeks) \ Periodicity (weeks)	1	4	13
52	0.4%	0.2%	0.2%
104	0.4%	0.2%	0.1%
156	0.4%	0.2%	0.2%
260	0.4%	0.2%	0.1%
520	1.5%	0.6%	0.7%

Source: Bloomberg

An important aspect of tracking error applies to ESG[9] portfolios and Impact investing. These ESG investments constitute active investments versus broad market benchmarks, resulting in active risk to the portfolio. As such, part of the difference in the performance of an ESG portfolio is due to the impact factors selected compared to the benchmark. The more the portfolio emphasizes or selects Impact investments versus a benchmark, the greater the potential tracking error, and the more common it is to see out/underperformance versus the benchmark.

Conclusion

Multi-asset class portfolios provide unique challenges regarding the measurement of volatility. Chief among these are data periodicity, data look-back, and data length. We have tools to deal with these difficulties. For data periodicity we use the most frequent historical data available between any two investments to calculate variance and correlation. This may include weekly, monthly, and quarterly returns, which we then combine into a single covariance matrix used to estimate volatility. For data look-back we can choose to incorporate a function that decays the weight of observations with time. This decay function is especially helpful when modeling funds that tend to turn over their portfolios frequently. For instances where we have limited data length, we can introduce a Bayesian prior to approximate the volatility profile of the missing data. We can then blend this prior with the actual sample data to create a more reasonable output. Lastly, clients need to understand the volatility of their portfolios versus a benchmark, or "tracking error," as it has large implications in portfolio management decisions. Similar to our calculation of volatility, we use the best estimate of tracking error possible to make the most informed decisions

In summary, we use a combination of methods at Athena depending on the nature of the data and the intention of the analytic report in question. There is no single "right" answer on the best way to measure volatility. The lesson is simply to be aware of the possible pitfalls of naively interpreting any one single volatility number. Portfolio managers must understand the nuances of volatility and risk measurement discussed in this paper, including the potential optical illusions that can arise from unique periods in history or other data anomalies. This awareness better informs decision making and risk management of Athena client portfolios, especially in times of stress.

Endnotes:

[1] Strictly speaking, volatility cannot be "measured", since it is the inherently unobservable parameter of a process. The best we can do is to estimate it from a set of observations drawn from the process. Different sets of observations may yield different estimates. However, we use the word "measure", as it is prevalent in the industry.

[2] In our analysis we use total returns from the Bloomberg Energy Subindex and the MSCI EAFE to represent Energy and Developed Foreign Large Cap sub asset classes, respectively.

[3] One might wonder why we cannot do the reverse and simply disaggregate quarterly returns into three identical monthly returns. The reason we cannot do this is that it would artificially lower the volatility profile of the investment in question. The intra-quarter volatility of the newly created monthly returns—averaged into identical monthly returns—would be zero. This would lead to a host of volatility-related inaccuracies such as overstated risk-adjusted returns, understated risk contributions, etc.

[4] As an extreme example, if cor(A, B) = 0.92 and cor(B, C) = 0.88, then it obviously cannot be the case that cor(A, C) = -0.33. There is a limit as to how low cor(A, C) can be, given the other two correlations.

[5] If there exists no portfolio that generates a negative variance from the supplied covariance matrix, then the matrix is said to be Positive Definite (P.D.). Given our pairwise approach above, which we used to squeeze the maximum information out of disparate data, we have no such guarantee of being P.D.

[6] "Computing the Nearest Correlation Matrix – A Problem from Finance", Nicholas J. Higham (2001), MIMS EPrint: 2006.70, http://www.manchester.ac.uk/mims/eprints. It is worth noting that Higham developed this algorithm, which finds the nearest P.D. matrix to the given matrix, as a consulting project for an asset management firm.

[7] Note that equal-weighting observations back to a certain point and omitting older observations is a form of decay, where the decay function is a step function. Exponential decay has the advantage of smoothness.

[8] As mutual funds have to accommodate net subscriptions and redemptions, they need to buy/sell the underlying securities in the portfolio to meet these requests, resulting in changes to underlying prices and fund performance. These price changes are magnified when there is a large net subscription or redemption, causing larger than usual relative performance differentials versus a benchmark in cases when a fund holds different investments than the benchmark, or the same investments in different weights.

[9] Environmental, Social, and Governance.

Jeremy Evnine, PHD
Senior Advisor, Quantitative Research

Jeremy Evnine provides guidance to the Athena team in the areas of quantitative research and risk management. Jeremy is currently CEO of Evnine & Associates, Inc., a consulting and Investment Advisory firm engaged in quantitative strategies since 1992. From 1991 to 2003, Jeremy was also a partner in Iris Financial Engineering and Systems, a financial software firm specializing in providing high-end trading and risk systems to top-tier investment banks. He sold his interest in Iris in 2003. From 1984-1990, Jeremy was SVP in charge of research at WFIA (subsequently Barclays Global Investors, now BlackRock). In this capacity, he worked with such people as Fischer Black and Myron Scholes, Bill Sharpe, and Michael Brennan and Eduardo Schwartz. From 1980-1984, Jeremy was a consultant at Barra, where he developed the firm's option products.

Jeremy earned his B.Sc. in Mathematics at Manchester University in England, his M.Sc. in Pure Mathematics at the Hebrew University of Jerusalem, and his PhD in Operations Research and Finance at U.C. Berkeley. He has taught courses in finance at U.C. Berkeley, published articles in the financial literature on option pricing and tactical asset allocation, and lectured in the United States and abroad.

John Raus, CFA
Research Associate

John Raus focuses on Athena's Quantitative and Risk Analytics through development of proprietary portfolio and risk management models. He also leads the development of a wide range of analytics across the Research and Portfolio Management groups, as well as Athena's asset allocation framework.

Previously, John was the head trader for Ellington Management Group's long/short equity hedge fund, where his responsibilities included portfolio trading, idea generation, and maintenance of proprietary risk management systems. Prior to this, he worked in Ellington's Quantitative Strategies Group, where he researched and traded various medium and low frequency systematic strategies.

John earned a B.A. in Economics from The College of the Holy Cross. He is a CFA charterholder and a member of the Stamford CFA Society.

Disclosures

Athena Capital Advisors LLC ("Athena") prepared this document solely for the person to whom it has been given for informational and discussion purposes only. This document and the information contained herein are strictly confidential and may not be reproduced, distributed or communicated to any third party without the express written approval of Athena. Athena reserves the right at any time to amend or change the contents of this document without notice. The information and opinions herein reflect the views and opinions of Athena as of the date hereof and not as of any future date. All forecasts are speculative, subject to change at any time and may not come to pass due to economic and market conditions.

This document and the information contained shall not constitute an offer, solicitation or recommendation to sell or an offer to purchase any securities, investment products or investment advisory services. The material contained herein has not been based on a consideration of any individual client circumstances and is not investment advice, or should it be construed in any way as tax, accounting, legal or regulatory advice. An investment with Athena involves substantial risks and there can be no assurance that the investment objectives described herein will be achieved.

Athena believes that the research used in this presentation is based on accurate sources (including but not limited to economic and market data from various government and private sources and reputable external databases), but we have not independently verified those sources, and we therefore do not guarantee their accuracy. The opinions, projections and estimates contained herein reflect the views of Athena only and should not be construed as absolute statements and are subject to change without notice.

In considering the performance information contained herein, recipients should bear in mind that past and present performance is not necessarily indicative of future results, nor does it ensure that investors will not incur a loss with respect to their investment. Current performance may be higher or lower than the performance data quoted. Certain performance numbers in this presentation may be unaudited, preliminary and based on estimates. Final reported and audited performance numbers may vary considerably from these estimates due to many factors. Estimated gross (i.e., including any fees, expenses or taxes) and net (i.e., reflecting deduction of any fees, expenses or taxes) performance numbers could change materially as final performance figures and underlying investment costs and fees are determined and allocated. Certain information contained herein constitutes "forward-looking statements" which can be identified by the use of terms such as "may", "will", "should", "seek", "expect", "anticipate", "project", "estimate", "intend", "continue", "target" or "believe" (or the negatives thereof) or other variations thereon or comparable terminology. Due to various risks and uncertainties, actual events or results or actual performance may differ materially from those reflected or contemplated in such forward-looking statements. As a result, investors should not rely on such forward-looking statements in making their investment decisions. No representation or warranty is made as to future performance or such forward-looking statements.

Benchmarks are shown for illustrative purposes only and are provided for the purpose of making general market data available as a point of reference only. Such benchmarks may not be available for direct investment, may be unmanaged, assume reinvestment of income, do not reflect the impact of any trading commissions and costs, management or performance fees and have limitations when used for comparison or other purposes because they, among other reasons, may have different trading strategy, volatility, credit, or other material characteristics (such as limitations on the number and types of securities or instruments). No representation is made that any benchmark or index is an appropriate measure for comparison.

The investment examples contained herein are for informational and illustrative purposes only and should not be construed as a guarantee of actual or future performance results. Individual investment results may vary considerably based on various factors such as fees, expenses and the timing of capital contributions. To see specific performance results, gross and net, please contact Athena.

Any description of tax consequences set forth herein is not intended as a substitute for careful tax planning. Recipients of this material are advised to consult tax counsel for advice specifically related to any and all tax consequences of an investment made with or through Athena. The information provided herein is not intended to, nor does it specifically advise on, tax matters pertaining to federal, state, estate, local, foreign or other tax consequences of an investment. The recipient is solely responsible for all tax consequences with respect to any investment made with or through Athena.