Quantitative Hedge Fund Analysis and the Infinite Monkey Theorem

The Infinite Monkey Theorem states that a monkey banging randomly on a typewriter will eventually produce all the works of Shakespeare. This is technically true if allowed an infinite amount of time – every possible combination of letters would eventually be produced. But, how about a more realistic example – could 1,000 monkeys typing nonstop for 1,000 years produce just the opening six-word sentence of Shakespeare's *Romeo and Juliet* ("Two households, both alike in dignity,")? The line is only 38 characters long, so this would seem a trivial task to accomplish over ten centuries. But the odds of producing this sentence by chance alone on a 50-character keyboard are actually one in (50 multiplied by itself 38 times), or one in 36,379,788,070,917,129,516,601,562,500,000,000,000,000,000,000,000,000,000,000,000,000. Assuming the industrious monkeys randomly hit five keys a second 24 hours a day for 1,000 years, the probability of typing this one short sentence is so small as to be virtually impossible. In fact, there is only an infinitesimal chance that our tireless monkeys would output this sentence if they worked continuously over the entire estimated age of the universe (13.7 billion years)!

The lesson here is that humans have generally poor judgment when it comes to numbers and our intuitive perceptions may be refuted when mathematically scrutinized. Instead, we should understand and apply the proper formulas to reveal underlying characteristics and hidden relationships. Statistical analysis, in addition to dashing the hopes of any aspiring monkey-authors, is a key tool for guiding the selection and maintenance of your hedge-fund portfolio.

* * *

Quantitative analysis is one important part of the pre-investment process. While "past performance is not indicative of future results," alas none of us have a crystal ball and must rely on these past results in our evaluation. Hedge fund managers must have an edge or advantage, or an investment with them is no better than flipping a coin (actually worse, due to fees and expenses charged). Determining that edge is crucial, and just as crucial, determining if it will be sustainable in the future. Quantitative analysis gives us tools to help make these assessments.

The four steps in quantitatively evaluating a potential money manager are: 1) Data acquisition and preparation; 2) Individual analysis; 3) Comparative analysis versus peers and indices; 4) Portfolio analysis. The following will cover each step in detail, elaborating on the pros and cons of available analysis tools, with special focus on points to consider and pitfalls to avoid.

Step One: Data Acquisition and Preparation

Evaluating the Reported Performance Data

Before you begin, the appropriate data series must be obtained for your analysis. The manager may offer different vehicles (onshore/offshore), different share classes with differing terms, strategy variations through different products, etc. The correct starting point is of course the vehicle/product/class that matches your intended investment.

Some managers report a composite track record, which is a combination of various vehicles employing the same strategy, and may also include differing fee levels among various investors. Without knowing the exact composition and how it was weighted/calculated, the data is nearly impossible to interpret as it relates to the return you would have received at your fee level. Always check footnotes to performance tables, which may offer guidance on these points, and discuss thoroughly with the manager.

Shorter track records must be approached with caution. A longer performance history increases the chance that past results were not due to luck alone, and offers greater comfort that the manager has encountered a variety of macro environments and market conditions. A short seller may shine during a bear market, but fail spectacularly when the market turns bullish again: What is their edge over competitors and how will this manager fare after the bear market has ended?

Often, traders will leave a firm to strike out on their own. A short track record may be given more weight if the portfolio manager(s) traded a similar strategy at a predecessor firm. If you can obtain the prior-firm track record, and gain comfort that it is substantially similar to the current program being offered, this might enable you to consider a manager that would otherwise have missed your minimum evaluation criteria.

Audits, Reviews and GIPS Compliance

Audited returns are generally very reliable but unaudited returns are much more difficult to verify. Pooled private vehicles (such as limited partnerships) typically require at least an annual audit, per the terms of their legal and offering documents. Managed accounts, where the manager is given trading authority over the account of an investor, are not audited because the investor (account owner) has control and visibility of his or her own funds. If the manager has only offered managed accounts in the past, there will generally be no audits to confirm the performance record. The best course of action in this situation is to directly contact as many existing investors as possible, and ask each to independently verify that their own results comport with the manager's published track record (with an allowance for differing fee levels and entry points/high water marks).

Some managers may present performance that is "reviewed" but not "audited" and may even imply that a review is substantially similar to an audit. An audit is the highest level of assurance

that an accounting firm can provide. Audits confirm that the fund's financial statements are free from material misstatement and, as applicable, comply with Generally Accepted Accounting Principles (GAAP). Audits involve obtaining third-party confirmations, testing selected transactions and obtaining supporting documentation, and evaluating the management firm's internal controls and procedures. Reviews have a much more limited scope. The reviewer only expresses a reasonable basis for the accuracy of the financial statements. This "negative assurance" opinion is based solely on the fact that nothing has come to the reviewer's attention that would indicate otherwise. Reviewed financials are more reliable than non-reviewed, but are never a substitute for a proper audit by a recognized and peer-reviewed accounting firm.

"GIPS" compliance is another factor to consider. Investment managers may voluntarily choose to comply with Global Investment Performance Standards when reporting returns. GIPS is a set of standardized guidelines set forth by the CFA Institute that is designed to facilitate comparison between global investment options, improve transparency and help eliminate survivorship biases (as discussed below). It includes specifications on performance calculation, presentation and related disclosures, and is particularly useful when there are only managed accounts. Although participation is voluntary, third-party verification, if available, offers assurance that the manager has complied with all GIPS requirements. A verified, GIPS-compliant track record can eliminate much pre-analysis investigation and inquiry.

Track Record Modifications and Adjustments

It should never be assumed that a manager's results to date reflect that manager's current parameters and risk profile. Attention should always be given to any changes in a manager's style, risk, leverage, instruments or markets traded, fee levels, etc. over the historical track record. It is common that a money manager may take more risk in the early days, with a smaller amount of capital that is often internal/proprietary. Other aspects of the trading program may change over time due to manager experience, changing market environments and varying personnel in portfolio management and analyst positions.

If the manager has a ten-year track record but made a material change one year ago, should the prior nine years be eliminated from your analysis? Not necessarily. It may be possible to construct a pro forma data series, incorporating back-adjustments, that can then be used for your evaluation. If, for example, the manager now employs "1.5X" leverage versus earlier periods, you can simply multiply each monthly result from the first nine years by 1.5 (effectively increasing both gains and losses by 50%), to approximate the effect of the higher leverage. Manager fee changes can be similarly adjusted, though the process is a little more involved. Generally, expenses are deducted from a gross return, then the management fee is deducted, then the incentive fee (subject to high water mark and possible other adjustments) is deducted. To correctly account for fee changes, it is necessary to obtain monthly information for both gross returns and expenses; then a spreadsheet can be created that calculates the proper fee and expense deductions from the gross returns to obtain the net monthly returns. It is much easier for the manager to make these adjustments for you and they will usually be provided upon request; however the data should be carefully examined to ensure it approximates the

expected values (e.g., if management fee is now 1% higher than in prior years, adjusted returns should be roughly 1% less per year than the non-adjusted data).

Leverage, fee and expense adjustments are relatively straightforward, but other changes present much more of a challenge. A change in manager personnel, for one, is a qualitative assessment and cannot be quantified. Changes in markets or instruments traded, as further examples, will have to be evaluated on an individual basis to determine the materiality of their impact on the manager's track record.

"Past performance is not indicative of future results," but obtaining the proper data series, making adjustments as appropriate, and applying qualitative judgment to the data (how does the current environment relate to the manager's past experience?) gives one a strong base for making reasoned decisions.

Once the data series is ready, analysis can begin via your software of choice, such as Microsoft Excel, MATLAB or Stata. Developing an analysis template, such as an Excel spreadsheet, is enormously beneficial – returns can be quickly input and analyzed using a standardized set of statistics, allowing comparisons between alternatives using identical criteria.

Step Two: Individual Analysis

The first step in quantitative analysis is evaluation of the manager's individual performance and statistics.

CAGR and AAR

The single most important statistic to investors is historical returns. This is usually expressed as Compound Annual Growth Rate (CAGR). CAGR is the annual return that would compound to the manager's actual cumulative performance, if all annual returns were identical.

Some managers report performance using Average Annual Return (AAR) instead of CAGR. AAR is calculated by simply adding the annual returns for each year and then dividing by the number of years. This calculation can lead to misleading – even highly misleading – results. If you invest $1,000 and twelve months later it grows to $2,000, both CAGR and AAR would correctly report your first-year return of 100%. However, if you lose 50% in the second year, you're back to $1,000 (your original investment), for a two-year CAGR of zero percent. AAR adds the first-year return of 100% to the second-year return of -50% and divides by two, giving an average annual return of 25% (even though you made nothing at all on your investment). The problem with AAR is that it ignores compounding. The order of returns is important; higher returns are better in earlier years because that extra money then has more time to grow (compound) throughout the remaining years.

CAGR, while superior to AAR, can also obscure performance factors such as volatility and drawdowns. For example, one manager may produce a 15% return every year, give or take a couple percent, while another may be flat one year, up 30% the next, etc., but both compound to the same 15% CAGR. The second manager in this example has a higher *volatility* than the first.

Volatility

Volatility is a measure of dispersion or variation among returns. It is calculated using the standard deviation formula. A 1% monthly standard deviation means that about two-thirds of monthly returns in the period analyzed were plus or minus 1% from the average monthly return. So, if the average monthly return is 1.5%, two–thirds of the time (eight months out of 12), the actual monthly return was between 0.5% and 2.5%. Standard deviation is often stated as an annualized number. The monthly standard deviation can be converted to a yearly standard deviation by multiplying it by the square root of 12 (don't worry, we won't bore you with the math derivation). A 1.5% monthly standard deviation therefore annualizes to 5.2%. This means that, two-thirds of the time (eight years out of 12), the annual return was +/- 5.2% from the average annual return. Standard deviation of daily returns can be annualized by multiplying by the square root of 365, etc.

Note the use of the past tense in the preceding paragraph. Volatility and other statistics only describe the (past) data set being analyzed, but are frequently used to predict the probability of future events (in this case, investment returns). Confidence intervals are used to define these probability ranges. Assuming that a manager's returns are distributed in a normal, bell-shaped curve (not necessarily the case – see Return Distribution section below), one can measure the area under the bell curve to calculate a probability (refer to Figure 2). A statement can then be generated, such as "There is a 99% probability that Manager X's monthly return will be between -3% and 5%." This prediction is accurate if judged on the assumptions that there is a normal distribution and that future performance will have similar characteristics to past performance. It will fail if, for example, large monthly losses are more common than predicted by a symmetrical bell curve (i.e., a fatter "tail" on the left side). Another scenario that would invalidate the return forecast is a manager that takes more risk going forward (higher leverage, different trading instruments, larger positions), effectively expanding the range of likely returns (positive or negative). Finally, market conditions such as volatility may be very different than those contained in the original data set (track record).

Why is volatility important? All else equal, low volatility is always preferable to high volatility. If a theoretical manager had a volatility of zero (i.e., the exact same return every month), then you could enter or exit the investment, or add to it at any time with predictable results. Another manager with the same return but much higher volatility presents timing challenges. A string of loss months may be followed by outsized returns in the future, but what if you need to redeem your investment now? Making, or adding to, the investment presents a similar challenge; you risk adding at a temporary equity peak which may not be attained again for some time.

Drawdown

A drawdown is a peak-to-valley decline in the value of an investment (Figure 1). Any monthly loss is a drawdown, since it represents a decline, and can also be a part of a larger drawdown. A particular drawdown is considered over when a new equity high, or peak, is reached. "Maximum (or Max) drawdown" is the largest decline that has been historically experienced by a manager. Max drawdown presents a challenge when comparing investment options. The longer a manager's track record, the greater the chance of an ever-larger drawdown (in the same manner that a greater number of coin flips increases the probability of getting five tails in a row). Therefore, comparing max drawdown between managers with materially different operating histories is of limited utility. Another issue with max drawdown is that, unlike all other statistics covered herein, it only provides information regarding a single event, or point in time. Stock Trader A may have a track record that is only one month shorter than Stock Trader B, but if that one month is October 1987, then Trader B could be expected to have a much greater max drawdown.

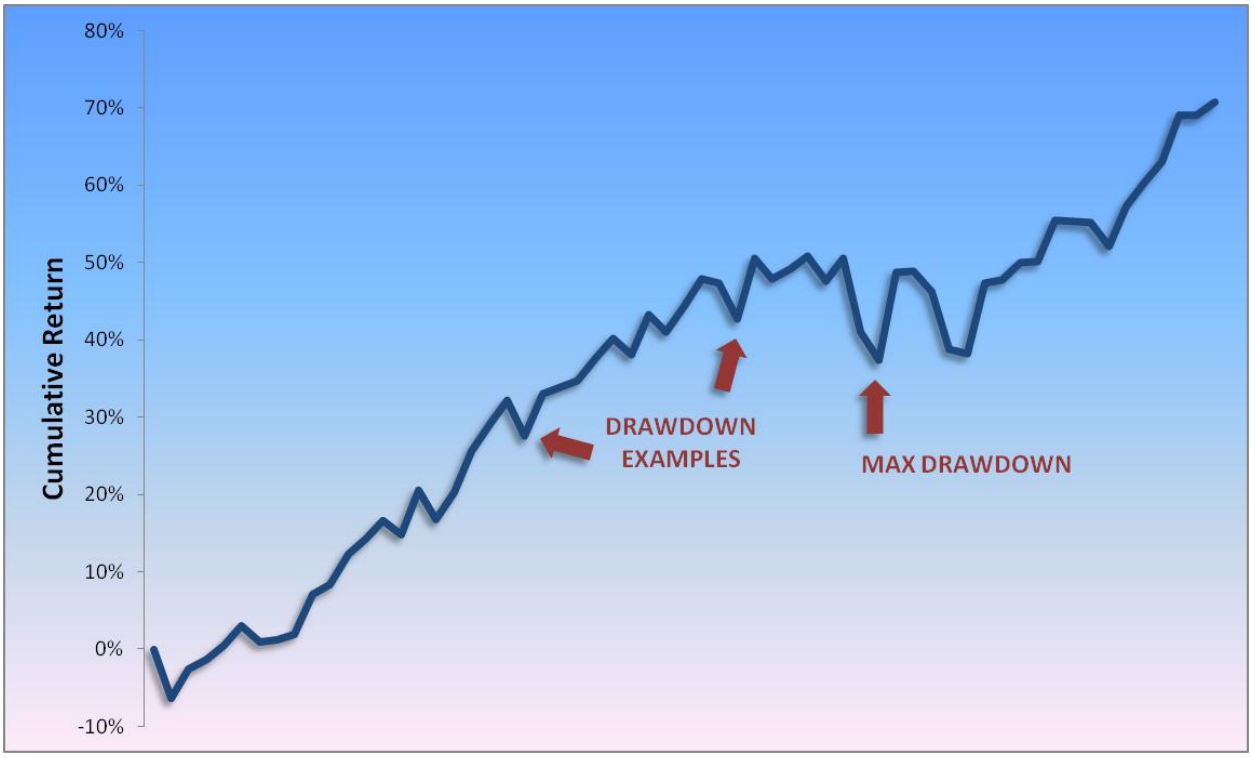

Figure 1: S&P 500 Index Drawdowns

A final point to note in regards to drawdown is that shorter analysis intervals tend to produce higher drawdowns (but never lower). Another way of stating this is that larger analysis intervals can mask, or even eliminate, a money manager's actual drawdowns. For example, a manager may have lost 5% in January, but when analyzing the weekly returns, was flat during the first two weeks of the month and down 8% in the third week before gaining 3% back in the

final week of the month. Similarly, reviewing each trading day within the third week will probably reveal a loss greater than 8% at some point. Shorter periods, especially trade-by-trade data, are much more representative of a manager's actual risk of loss, but keep the periods consistent when comparing two or more funds.

Rolling Period Analysis

It is useful to analyze different timeframes, not just the lifetime of the manager. Generally, more weight should be given to the most recent results.

"Percent Profitable Periods" is a helpful metric that analyzes a rolling time window of returns (three months, six months, etc.) to explore consistency of returns. A three-year track record contains only three annual returns (in this case, calendar-year), but a rolling twelve-month analysis contains 25 (January to December, February to January, etc.)! Yearend returns are not particularly significant when you could have invested at any point during the year, and could be misleading. Dividing the number of positive rolling periods by the total number of periods reveals the percent likelihood that you would have made a profit when investing at any random point in the past.

Knowing a manager's CAGR and their volatility is vital, but how do you compare multiple investment choices? Manager A has a higher CAGR than Manager B, but is more volatile – does Manager A's higher CAGR compensate for the higher volatility?

Risk/Reward Ratios

Sharpe, Sortino, Sterling, Calmar and Omega are among the popular ratios used in evaluating and comparing investments.

Sharpe is the most familiar, and is calculated by taking the average return minus a risk-free rate, divided by the standard deviation of the returns. This tells us what return (over a "risk-free" rate) the investment is producing per unit of risk. The risk-free rate is generally 90-day Treasury bills – as close to riskless as you can get outside of stuffing money into your mattress. Be sure to ascertain what risk-free rate is being used – which varies from manager to manager and can have a big impact on the calculated ratio. Sharpe allows risk/reward comparisons between two investments, using a single measure. Greater returns increase the numerator, resulting in a higher ratio. Greater volatility increases the denominator, resulting in a lower ratio. Therefore, a higher Sharpe Ratio is preferable to a lower one.

The main criticism of Sharpe is that is penalizes upside and downside volatility equally. Most investors would not complain if their returns varied widely from period to period (high upside volatility) but were never negative. The Sortino Ratio addresses this concern by replacing standard deviation in the Sharpe equation with standard deviation of negative returns only.

The Sterling and Calmar ratios also address the importance of capital loss to an investor, but replace downside volatility with drawdown measures.

The Omega Ratio is the newest and probably least known, but arguably the most powerful risk/reward measure. Omega measures the likelihood of achieving a target return. It is calculated as the probability-weighted gains above the target return divided by the probability-weighted losses below the target return. It takes into account both the probability of gains or losses and the size of those gains or losses. It does not directly penalize volatility, but considers the possibility of extreme losses. One of Omega's key advantages is that it does not assume the investment under consideration is *normally distributed*.

A normal distribution is the familiar bell-shaped curve (see Figure 2), where the highest probability occurs at the mean (average) tapering gradually and symmetrically on each side, with a smaller and smaller likelihood of achieving returns higher or lower than the average, and with the very thin "tails" at either side indicating a miniscule probability of extreme high or low values (returns). While the normal distribution is applicable to many real-world situations, investment returns do not necessarily fit this bell-shaped curve. In particular, and to the chagrin (and demise) of many an overleveraged investment manager, extreme losses are more likely to occur than the normal distribution suggests. Sharpe, Sortino and similar ratios assume a normal distribution, but Omega does not.

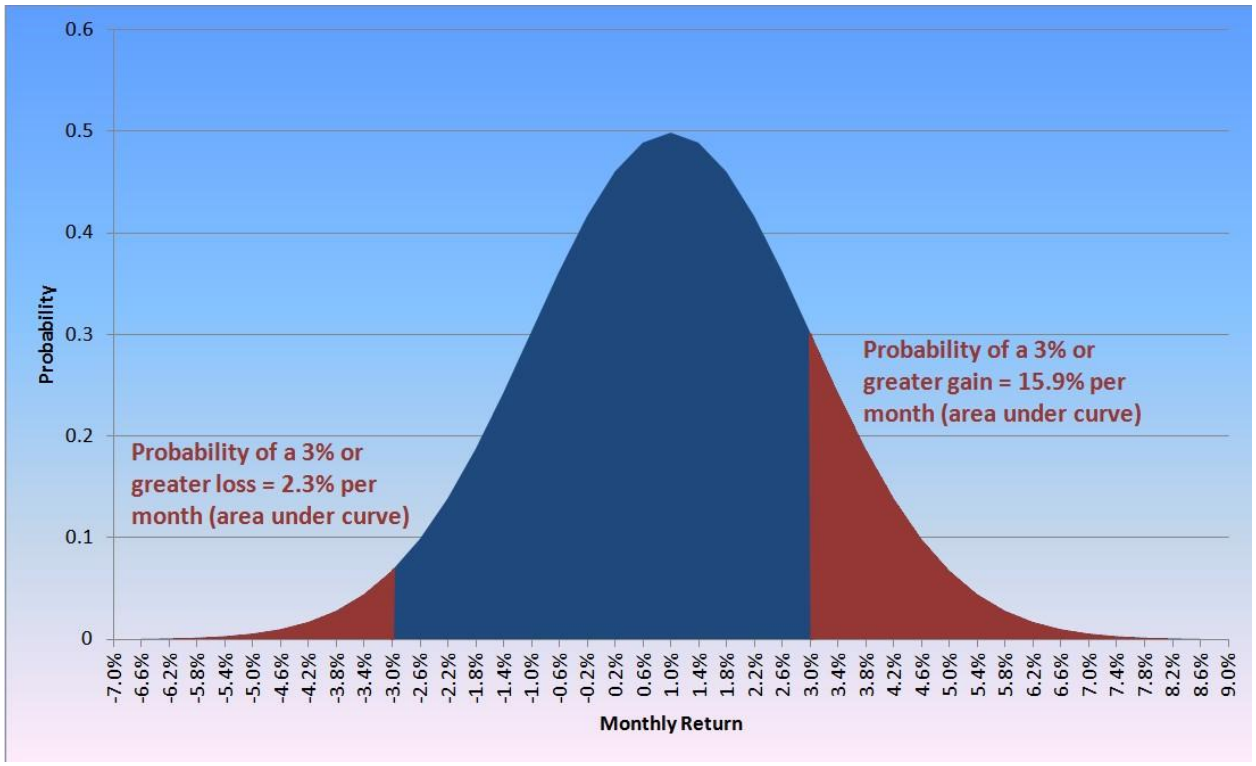

Figure 2: Hypothetical return distribution of a manager with a normal return distribution, 1% average monthly return and 2% standard deviation

Another advantage of Omega is that it incorporates an investor's target return – if you desire a 7% annual return, just plug 7% into the formula and invest in the alternative with the highest Omega ratio (again, all things being equal). If your target is 20%, you will have to accept more volatility and the Omega ratios will be lower than the previous example, but choosing the highest Omega ratio will maximize your chances of achieving that target return.

There is one important caveat with Omega: leveraging an investment with a positive cumulative return will always increase the Omega ratio, because the numerator will increase by a greater amount than the denominator. This makes intuitive sense – if we could look into the future and know that an investment would produce a positive return, and the cost of borrowing/leverage was low enough, we might leverage that investment by the maximum amount possible. So, caution is warranted when using Omega to compare similar investments with differing leverage.

For much more on risk/reward measures and especially the Omega Ratio, please see our whitepaper, "An Investor's Guide to the Risk Versus Return Conundrum," which is available at http://www.shinnecock.com/#articles.

Return Distribution

As stated above, most investment returns do not follow the normal distribution pattern. How the actual return distribution varies from the norm can be quantified.

Skewness measures the symmetry of the return distribution. Positive skew indicates the "tail" on the right (positive) side is longer or fatter than the left side. Negative skew indicates the tail on the left side is longer or fatter than the right side. The normal, bell-curve, distribution has a Skewness of zero because it is perfectly symmetrical. However, an asymmetric distribution can also have a Skewness of zero, if the differences on each side balance out (the left tail is very fat, but the right tail is very long). A very general inference from this measurement is that, for example, positive skew is preferable as it increases the likelihood of higher return values.

Kurtosis measures the shape of the return distribution. A normal distribution has a Kurtosis of 3; higher values indicate a more "peaked" distribution while lower values indicate a "flatter" distribution. A general inference from Kurtosis is that high values mean less variability of returns, because the values are more clustered around the peak (average).

Skewness and Kurtosis should not be applied as standalone measures, and interpretation is tricky. A more useful exercise may be to plot monthly return values on a chart and visually evaluate the resulting curve. Tightly clustered results with a few negative outliers, for example, would indicate mostly consistent returns but with the potential for significant down months.

Step Three: Comparative Analysis

The second type of analysis that should be performed on a money manager prospect is comparative analysis versus both peers and benchmark indices.

When considering adding a particular strategy or asset class to your portfolio, several promising managers can be compared side-by-side to highlight differences and rank the candidates by various criteria. Peer group comparisons are also useful for determining what edge, or Alpha, the manager is generating versus Alpha that may already be present in the strategy itself. Groups of similar managers can be reviewed concurrently to highlight the best performers. Looking at performance versus relevant indices is also important but must be evaluated in the context of the leverage or risk the manager is employing versus the index. Twice the index return may simply indicate twice the leverage, with no added value.

Alpha and Beta

Alpha and Beta are statistics that can be applied to the relationship between a manager under consideration to another data set such as an index, benchmark, or even another manager. For the following examples, we will assume that the manager is being compared to an index, and that a correlation exists between the two. Beta relates to volatility versus the index (whether the manager is more or less volatile) and Alpha is that portion of a manager's return that is independent of the index (that is, not explained by the up and down index movements). A higher Alpha is always preferable (per the explanation below), but Beta is not so clear-cut. All else equal, a low Beta (as with any volatility measure) is desired, but for practical purposes a higher return target may require acceptance of a higher Beta.

When comparing a manager to an index, one must consider which index is appropriate. Passive indices such as the S&P 500 measure the return of an asset class (more specifically, a formulized subset of an asset class). Active indices are comprised of actively-managed investment funds or programs, such as those published by BarclayHedge and Hedge Fund Research (HFR). HFR alone publishes over 60 different indices! Narrowing your focus to a specific geographic region, sector, strategy, sub-strategy , or sub-sub-strategy can make for a closer comparison to your manager under analysis, but there are some caveats. The narrower the focus, the smaller the number of managers comprising the index, making it potentially less reliable.

Also, be aware of "survivorship bias." Managers that close, or stop reporting due to very poor performance are removed from the index, leaving only the best performers. This distorts and skews the index returns to the positive side, and is potentially more pronounced in indices with fewer manager constituents. The best illustration of survivorship bias comes from World War II. Allied planes were being downed by enemy anti-aircraft guns, and therefore needed additional armor. But, armor is heavy and must be used sparingly. Planes returning from battle were examined and a statistical analysis was made of the parts of the plane that had the most bullet holes – so the extra plating could be added to those areas. But, mathematician Abraham

Wald, working for a classified program known as the Statistical Research Group (SRG),turned this idea on its head. If an airplane returned from battle, by definition any damage it had sustained was not enough to bring down the plane. Therefore, the extra armor should be added to the areas where there were *no* bullet holes. The Air Force was only examining the surviving planes, not the planes that had been downed by hits to the vulnerable cockpit or engine areas.

A Beta of 1 implies that the manager has a similar volatility to the index, that is, if the index rises 10% then the manager is expected to rise 10%. A Beta of 1.5 implies that the manager will rise 15%, and a Beta of negative 0.5 implies that the manager will fall 5%. If an investment is leveraged "2X" and everything else is the same, then the leveraged investment will always have a Beta of 2 versus the unleveraged investment.

Alpha represents the manager's excess (or deficit) return versus the index. In the leveraged example just above, the Beta is 2 but the Alpha is zero because there is no added value – leveraging simply doubles the returns, whether positive or negative. You would probably not pay fees to a manager who simply offered a 2X leverage version of the S&P 500 because of this absence of value (in other words, you can simply leverage it yourself). However, an annual Alpha of 4% indicates that a manager has outperformed the comparison index by an average of 4% per year – after any Beta adjustments. As a final example, the S&P 500 returns 10% in a year and the manager, who maintains a 200% (2X) exposure, returns 26%. 10% of the excess return is due to leverage and 6% of the excess return is due to Alpha. Beta provides valuable information, but it is positive Alpha that investors seek.

Step Four: Portfolio Analysis

The third and final type of quantitative manager analysis is evaluation of the manager within your existing portfolio.

A manager that appears excellent upon solo examination and even versus peers and benchmarks may offer only an insignificant improvement to your current portfolio performance and may even be a detriment. Conversely, a manager that does not appear outstanding when evaluated in isolation may sometimes have an unexpected, and favorable, portfolio impact.

Creating pro forma returns by including a manager prospect in your current portfolio is straightforward and provides a wealth of information. If 10% would be your allocation to the manager under consideration, for any given month in the past (for which returns exist for both the portfolio and the manager), use the formula (90% X A) + (10% X B), where A was the actual portfolio return and B is the manager return, to approximate what the return would have been with the 10% manager allocation. This new data series can then be examined using traditional statistical tools, and compared to the same statistics for the actual portfolio returns. A simple spreadsheet can compare the two data sets side-by-side, and easily calculate the percent change (better or worse) for any given statistic such as CAGR or maximum drawdown.

Correlation

"Correlation does not equal causation" is a phrase most of us have heard before; just because two number sets rise and fall together does not necessarily mean they are related to each other. Often this "spurious correlation" is caused by a third factor. For example, you could graph a positive correlation between lemonade stand sales and the average number of kids in the local community swimming pool. Lemonade and swimming have no connection, but both are positively correlated to temperature – higher temps equals more lemonade stands and more kids cooling off in the pool. Two data sets may also be mathematically correlated by chance alone without the involvement of any third factor. A popular humor website, Spurious Correlations, illustrates many examples of this, such as the Divorce Rate in Maine compared to Per Capital Consumption of Margarine.

The risk of spurious correlations may be reduced in a number of ways. Firstly, the more data you have, the greater the probability that any observed correlation is meaningful. The Spurious Correlations website graphs yearly data consisting of only about a dozen data points; the apparent relationships may evaporate when viewed on a monthly, weekly or daily basis. Secondly, common sense and qualitative evaluation is crucial. One can be pretty sure that lemonade does not cause swimming, or vice versa. If you can't understand the correlation, it may not be valid. There are further mathematical techniques to confirm or refute your correlation stats, such as de-trending: random data sets having a similar positive or negative trend over time can exhibit very high, and meaningless, correlation – removing the trend component can reveal the true non-correlation of these random up and down moves.

True portfolio diversification requires low correlation between asset classes and strategies. Therefore, the correlation of the manager prospect to your portfolio should always be investigated, subject of course to the caveats given above.

There are multiple types of correlation statistics that can be employed in your analysis; the following is a brief, simplified summary and is not meant to cover all aspects of this complex subject.

The Excel function CORREL calculates the Pearson Product-Moment Correlation Coefficient, or more simply, the "Pearson correlation" (PC). PC, in effect, makes a chart of the data points (one on the x-axis and the other on the y-axis); if January was a +1% month for your portfolio and a +2% month for the manager prospect, a dot is placed at these X,Y coordinates (1,2). Once all the dots are placed in this manner, PC draws a line that best fits the data points, that is, that minimizes the cumulative distances between the points and the line (see Figure 3). A positive-sloping line indicates a positive PC while a negative-sloping line indicates a negative PC. The slope of this line, a number between -1 and 1, represents the linear relationship between the two sets of data. A PC of 1 indicates a perfect correlation between the two; any set of data always has a PC of 1 when compared to itself. A PC of -1 indicates a perfect inverse relationship. Data sets that are positively correlated tend to move up and down at the same time. Data sets that are negatively correlated tend to move in opposite directions. The higher

(or lower) the PC, the stronger this relationship is. A PC near zero indicates little to no correlation between the data. Note that PC describes linear relationships only and will give no useful information about non-linear correlation (which would be a curved line in the example above).

Figure 3: Scatter plot of returns for S&P 500 Index and an equity manager, with linear regression line showing a positive slope (correlation)

Spearman's Rank Correlation (SRC) is another common correlation measure. SRC arranges the numbers from each data set in rank order (high to low), then performs a Pearson Correlation on the ordered data sets. If a manager's returns are always higher than your portfolio returns, by a consistent amount each month, then both PC and SRC will show a correlation of 1. If the manager's returns are always higher, but by differing amounts, SRC will still show a correlation of 1 and PC will show something less than 1 (but still positive). Like PC, SRC only measures linear relationships.

Should one use PC or SRC to calculate correlation between sets of investment or index returns? There are a number of factors to consider, but if your data looks close to a normal distribution, with no outliers (large positive or negative returns that are isolated and separate from the other data), PC may be best – otherwise, use SRC.

Of course, you are seeking to avoid high positive correlation to your portfolio when adding a new manager. High correlation will generally increase volatility. Negative correlation will lower volatility because the manager will tend to be up when the portfolio is down and down when the portfolio is up – this makes the total returns closer to zero and thus less volatile. Correlation can also be calculated between the manager prospect and any individual manager in your current portfolio, for more insights.

A manager that has positive returns every month is of course a welcome addition to almost any portfolio. But if the manager is flat overall when the portfolio is up, and positive overall when it is down, this would also be greatly beneficial, decreasing volatility and drawdowns while increasing returns by compounding money faster. Evaluation of a manager prospect's average return in the loss months for each existing manager and the portfolio as a whole can be quite revealing – downside correlation is a much more important factor than upside correlation. It is easy to create a new data series of monthly manager returns for only those months that the existing portfolio generated a loss (or a loss greater than a certain threshold, if you wish). This new series can then be used to calculate the average return for the manager in portfolio loss months (simply an average of all the returns), or the percent of time the manager was up during these portfolio losses (simply a count of the positive returns divided by a count of all returns in the series). This might reveal, for example, that a manager under consideration had a positive monthly return 75% of the time that the portfolio was down and that the manager's average return was 1.2% for those months (both encouraging findings).

Autocorrelation, or serial correlation, can be thought of as a manager's correlation to itself. More specifically, the correlation of one period's return to the next period's return. Managers with positive autocorrelation tend to follow good performance with more good performance and bad performance with more bad performance. Managers with negative autocorrelation tend to exhibit mean-reverting performance: losses are followed by gains, or vice-versa. This has importance for risk evaluation. Using only volatility to predict the likelihood of a drawdown of a certain size (depth) has been shown to produce inaccurate results. A 2012 study of world equity indices (which had positive autocorrelation in the study) and commodity trading advisors (CTAs, which had negative autocorrelation), illustrates the point. After equalizing differences in return levels and volatility, the expected maximum drawdown for equities was twice that of CTAs. This answered the question of why equity markets have demonstrated longer and deeper drawdowns historically than would be expected given their volatility.

When autocorrelation is non-zero, the formula given above for converting daily, weekly, monthly volatility to a yearly measure (multiplying by the square root of 12, for example to convert monthly to yearly) can produce inaccurate results. This is because the formula assumes that the probability of a gain/loss of a certain size during a particular period is independent of other periods. A non-zero autocorrelation tells us this isn't necessarily the case; a monthly loss, for example, may increase the chance of a loss the following month.

It would seem that timing one's investment (entries or exits) is an obvious use of autocorrelation statistics. But, the majority of published research seems to indicate otherwise.

Firstly, the sample size must be significant and because calculation errors decrease by the square root of the sample size, a hundred times more data is required to decrease the error rate by a factor of ten. A single manager's return history is often not of a meaningful size, the autocorrelation most likely does not significantly deviate from zero and can change at any time, particularly over short-term periods. Autocorrelation, however, can in some circumstances be utilized to give a truer picture of a manager's risk characteristics, particularly when combining long-term data from many managers into an index representative of the asset class and/or strategy.

Efficient Frontier and Modern Portfolio Theory

Modern Portfolio Theory (MPT) originated in a 1952 paper by Nobel Laureate Harry Markowitz. It hypothesizes that an investor can construct a portfolio that is optimized for maximum return at a given risk level.

The "efficient frontier" refers to the optimized portfolio that can, through the proper diversification, theoretically produce a higher return per unit of risk than a single investment. In other words, the portfolio is optimizing for the highest Sharpe Ratio (see Figure 4).

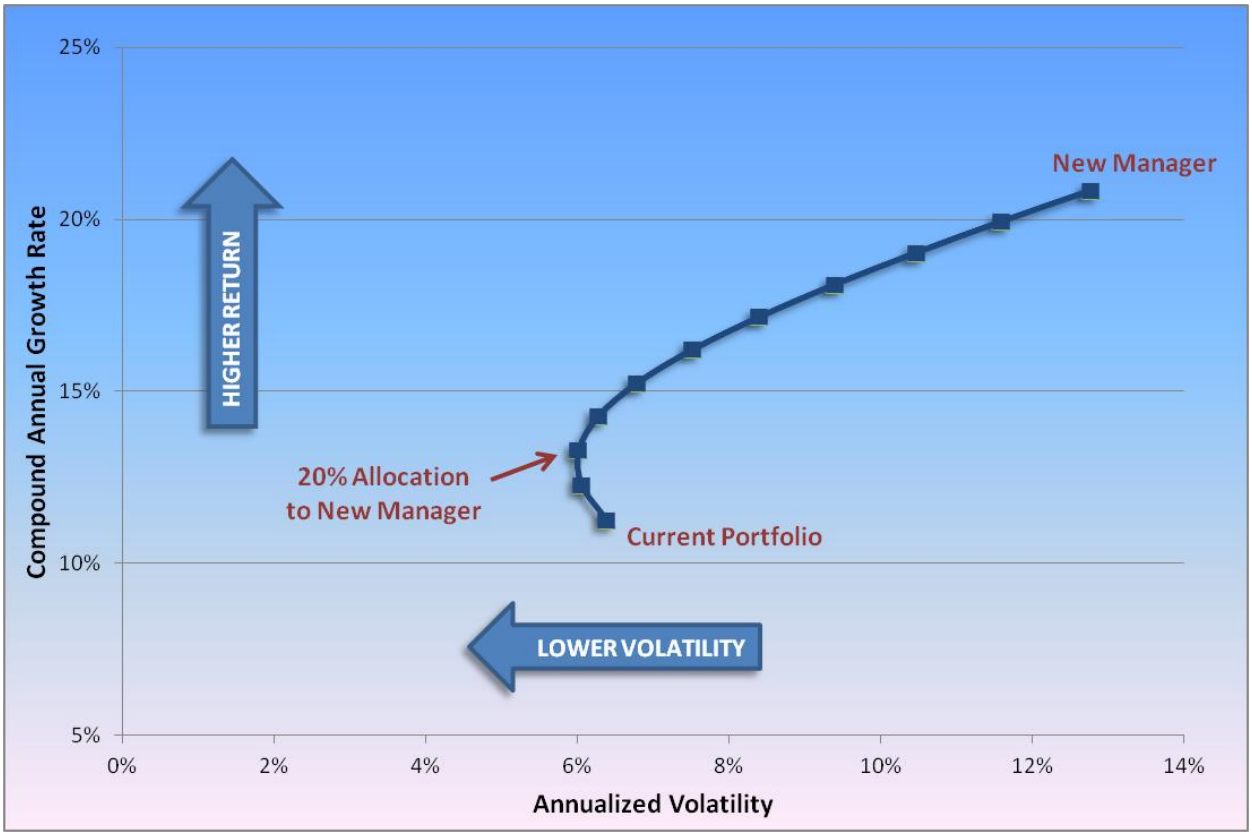

Figure 4: Effect of adding new manager to existing portfolio in 10% increments – 20% allocation optimizes risk/return ratio

The main criticism of MPT is that is assumes markets are efficient, investors act in a rational manner and market prices fairly reflect all information currently available. An opposing theory, Behavioral Portfolio Theory (BPT) focuses on investor psychology (particularly fear and greed) and states that markets are inefficient because investors do not behave rationally. Arguments for BPT (and against MPT) include irrational exuberance/market bubbles (hardly rational) and insider trading (i.e., information is not equally available to all investors). Another disputed aspect of MPT is the assumption of "homogenous expectations," that states investors will make the same decisions given the same data.

BPT posits that an informed investor can gain an advantage over other market participants by being aware of, and exploiting the psychological tendencies of the mob. BPT does not necessarily equate risk with volatility – a major deviation (no pun intended) from the MPT mindset. BPT allows for short-term volatility to enhance long-term returns. An obvious example is stocks versus bonds: Stocks usually outperform over long-term periods, but bond investors accept a much lower return in exchange for less volatility along the way.

Conclusions

Quantitative observations should be confirmed whenever possible by qualitative explanations. Why does the manager perform well in certain environments and not others? Why is the manager so uncorrelated to its peers? If you are unable to answer such questions, either through your own research or from the manager's input, it is a red flag and may increase the possibility of fraud. This analysis requires more than a surface examination. Two managers trading only stocks in the S&P 500 index can be legitimately uncorrelated – one may be a day trader while the other a buy-and-hold investor with a time horizon of months or years… or one may be a value trader (e.g., buy when the price is deemed low enough) and the other a momentum trader (e.g., go short as the price declines).

Once you've finished your analysis, selected the best manager and made your investment, your work is not over. Evaluation of all managers in your portfolio is important on a periodic and ongoing basis. Statistics are only valid for a point in time, and going forward you will of course have more data (returns) to assess (are the original conclusions still valid?), more alternatives to consider (is this manager still tops among its peers?) and a changing portfolio mix (is the manager still a good fit in my portfolio?)

"History cannot be told while it's happening," says author Duncan J. Watts; only after events have unfolded do we look back and ascribe meaning, causes and results. It may be "obvious" that a money manager is successful and talented when reviewing their past performance record, but getting to the truth requires careful application of the tools and techniques described above. History is only run once, so we cannot perform experiments to see how a manager would have performed under a different environment or at a different inception date. If a manager's track record is among the top 0.1% (one out of a thousand) of its peers, keep in mind that, on average, one out of a thousand managers will achieve this result by chance alone!

Even if one perfectly applies all mathematical concepts and encompasses all considerations when evaluating a potential manager, it does not guarantee success. You have simply increased the probability of success. The economy may change, a market edge may vanish or a portfolio manager may change their focus or risk tolerance. Though monkeys may never produce great literature, we live in an imperfect and random world, and improbable events do sometimes happen.

Quantitative analysis is time consuming, mentally challenging, and requires great care in evaluating the results. It must be combined with qualitative research and other due diligence to form a complete picture. Such efforts are, however, well worth the time and brainpower, and are an important tool to help protect and grow your capital.