NORTHFIELD | modeling financial markets worldwide

Simplified Investment Performance Evaluation

Dan diBartolomeo

Webinar October 2020

Motivation

- In 2018 Northfield did a complex consulting assignment for a large investment organization. The question at hand was how to fairly compensate portfolio management teams for their skill in managing the assets under their control.
  - The institution had put forward the simple concept that skill should be evaluated on "value added" or "risk adjusted returns".
  - This proved far more difficult than it sounds because the organization in question had a broad range of activities from running high turnover long/short strategies to owning long term illiquid assets like agricultural land.
  - The various groups within the organization had completely divergent views of what "risk adjusted returns" meant conceptually and how it should be evaluated.
- *What was needed was a universal metric that would applicable over all asset classes, strategies, and time horizons.*

Desirable Properties of a Performance Metric

- In formulating a metric to represent "risk adjusted return" there are numerous desirable properties.
  - Simple and transparent including statistical significance tests
  - Is fair in evaluation of investor and manager decisions
  - Closely related to investor economic utility, but without investor specific parameters.
  - Applicable to absolute returns and benchmark relative returns
  - Applicable to both active and passive strategies
  - Useful for evaluating both single period and multi-period performance
  - Applicable to all asset classes, both liquid and illiquid.
  - Minimal assumption dependence (e.g. returns are normally distributed)
  - Can be implemented by both investors and managers
- We need a universally applicable answer to "How much incremental return must I get to justify a choosing among alternative portfolios with different risk levels?"

The Current State of Performance Evaluation

- There are lots of metrics that are widely used to evaluate various situations in terms of investment performance
  - Return/Internal Rate of Return
  - Multiple of Investment
  - Excess return relative to specified market index benchmark
  - Excess return relative to the central tendency of the distribution of returns of similar investments (or peer funds).
  - Jensen alpha
  - Sharpe Ratio
  - Information Ratio / Covenant Information Ratio
  - Effective Information Coefficient (diBartolomeo, 2008)
  - Mean/Variance utility (Levy and Markowitz, 1979)

Let’s Consider Our Simplest Choices

- Return/Internal Rate of Return
  - Since investors investor to get returns there is a direct link to investor utility.
  - There is no form of risk adjustment
  - No context to consider statistical significance

- Multiple of Investment (MOI)
  - Often used in private equity
  - Equivalent to IRR is you assume the time horizon between investment (money in) and receipt of exit cash flows (money out) is a constant across all deals.
  - An obviously false premise, https://www.northinfo.com/documents/927.pdf

- Jensen Alpha
  - Return in excess of expectation from an asset pricing model (e.g. CAPM).
  - Very dependent on numerous theoretical assumptions.

Excessive Attention to Excess Returns

- Excess Return over a market index benchmark
  - Popular for evaluation of active management but moot for passive.
  - No consideration of absolute risk.
    - Excess return with greater than benchmark absolute risk is no different than the same return with less than benchmark absolute risk
  - Statistical significance can be considered only after many periods.

- Excess return relative to the central tendency of the distribution of returns of similar investments (or peer funds).
  - Usable for relative return but not absolute
  - There is an extensive literature showing industry practices to formulate peer groups can be "gamed"
    - diBartolomeo and Witkowski (1997)
    - Brown and Goetzman (1997)
    - Kim, Shukla and Thomas (2000)

Sharpe Ratio and Information Ratio

- Sharpe Ratio and Information Ratio
  - Captures both return and volatility and can be used ex-ante or ex-post
  - Assumes returns are normally distributed
    - Ledoit and Wolf (2008) extends to non-normal distribution
  - Statistical significance of comparisons is mathematically complex
    - Jobson and Korkie (1992)
  - Poor conformity to plausible investor utility
    - DeGroot and Plantinga (2001) shows that SR/IR conform to plausible investor utility only under assumptions of unlimited leverage.
    - Consider a manager with one basis point of alpha at zero tracking error. The IR is infinite but the value added for the investor is miniscule.
  - Covenant IR (Rudd and Clasing, 1982) is a lower boundary value for expected IR conditional on the Sharpe Ratio of the benchmark, and manager fees.
    - Put another way, can the active manager add more value after fees than just leveraging the passive index?

Effective Information Coefficient

- The EIC process (diBartolomeo, 2008) uses a risk model to infer manager alphas given the active weights of portfolio positions and relevant constraints.
  - Comparable to “IC * Transfer Coefficient” but the inference process allows EIC to be used by investors and fundamental managers.
  - IC * TC requires the manager’s input alphas to which only quant managers have access.
  - Since the returns on individual portfolio holdings are used rather than portfolio returns, sample size is largely achieving statistical significance in short time frames.
  - Usable only for active management, not passive.
  - Must be extended with ex-ante estimates of the cross-sectional dispersion of return of the securities in the universe in order to get into return units, similar to investor utility.

Mean-Variance Utility

- Levy and Markowitz (1979)
- Many investors find statistical variance an unintuitive measure and so prefer to think of the tradeoff between return and risk with the unit of risk being standard deviation
  - Or a scalar of standard deviation such as VaR or CVaR, see Cornish and Fisher (1937)

- As a performance evaluation metric, it is problematic because most investors are simply unable to numerically express their mean variance risk aversion ("lambda" or its percentage reciprocal which Northfield terms "RAP") with any confidence.
  – You get different evaluations of performance depending on the investor.
  – I know that if I had asked my grandmother about her "mean-variance risk aversion" she would have slapped me.

Simplified Investment Performance Evaluation

- Our proposed functional form for SIPE for *annualized* units is

$$\text{Risk Adjusted Return} = (R_p - R_f) - E(\sigma)/6$$

Where

R_p is the portfolio return for the observation period

R_f is the risk-free rate

(benchmark index return for the relative case)

(zero coupon riskless yield for a liability duration target)

E[] is the expectations operator

(you need a risk model but no time series history)

σ is the "volatility equivalent" risk measure

Where Did We Get the Scalar Constant of 6?

- In SIPE, we assume that the investor has chosen their current level of portfolio risk so as to bound a “worst case scenario”.
- From this boundary condition we can **infer** the mean variance risk tolerance of the investor as a *scalar function of the current portfolio expected return and volatility, as the investor has chosen this portfolio.*
- Algebraic simplification leads to the conclusion that for a wide range of situations and asset allocations, the optimal tradeoff parameter between *incremental expected return and incremental volatility is typically about one sixth.*
  - For example, a 2% increase in portfolio volatility is reasonably justified by a .33% (2/6) increase in expected return   The method works identically across absolute risk, tracking error, or any desired blend of the two risk measures.

Growth Optimal Investors

- Let's start with a "growth optimal" investor whose *only objective* is maximize their long term geometric mean return.   The usual objective function is

$$U = R - \lambda * \sigma^2$$

For a growth optimal investor who only cares about maximizing the future geometric mean return, λ =.5 or ½ assuming all units are in decimals. My preference is to remove the potential for $\lambda = 0$ so I'll rewrite this with the tradeoff parameter in the denominator and also convert to percentages

$$U = R - \sigma^2 / T$$

For the growth optimal case T = 200.

- The question is what is the appropriate <u>risk tolerance</u> T for a more risk averse investor, as almost all are.

The Maximum Loss Fraction

- Now let's work through a simple example, where our existing portfolio has R = 8 and σ = 12
  - Implicit in choosing a portfolio of a particular risk (e.g. 12%) is the idea that I don't want to put all my money at risk, just some of it (Wilcox's Discretionary Wealth Hypothesis, JPM, 2003). A reasonable expectation for the **maximum loss fraction** of the portfolio value would be something like

$M = (Z * \sigma) - R$, where Z is your choice Z-score of the worst case scenario (say something like 3.5)

$M = 3.5 * 12 - 8 = 34$

- So we're only willing to put 34% (.34) of the portfolio at risk (implicit T= 200) which means that the other part of the portfolio must be riskless (implicit T = 0)

$T = .34 * 200 + (1 - .34) * 0 = 68$

Getting To Approximately Six

- If we divide T through by σ, we obtain 68/12 = 5.67 (about six). So now we can express our objective as

$$U = R - \sigma^2/(5.67 * \sigma)$$

Which simplifies to

$$U = R - \mathbf{(1/5.67)} * \sigma$$

So our tradeoff between return and standard deviation is about 1/6. For a broad range of empirical cases, the 1/6 relationship holds rather nicely.

- For tracking error cases rather than absolute volatility, I tend to use Z = 3, R = 0, which **renders exactly 1/6.**

Discussion of Key Assumption

- The SIPE framework requires that the investor have a formal expectation for the risk of the investment they are undertaking.
  - This is the basis of evaluating when the performance was or was not satisfactory on a risk adjusted basis.
  - Did the investor or manager experience a good outcome knowing ***only what could reasonably have been known at the start*** of the observation period?
    - This removes the problem of hindsight biasing evaluations.
  - Investors and agent managers must agree on the risk expectation (as they do for benchmarks today).
  - Many institutional asset management contracts already comparable risk parameter targets or limits that require a risk forecast.

Broadening SIPE Application

- We propose three techniques to make SIPE universal across different types of securities, strategies and liquidity levels.
- All three are adjustments that take E[σ] from the expectation of volatility to a "volatility equivalent" value that compensates for differential aspects of investment activities.
  - Differences in liquidity between liquid and illiquid asset classes
  - Assets types (e.g. options) or dynamic strategies ("short vol") that require that the assessment of risk incorporate higher moments (skew and kurtosis)
  - The "Paradox of Active Management"
    - *While all active managers must believe they will outperform benchmarks, roughly half of all active managers must be **wrong** and underperform.*
    - *It is not mathematically possible for everyone to be above average.*

A Level Playing Field for Illiquid Assets

- There are really two ways to adjust σ to put liquid and illiquid assets on a level playing field in terms of risk adjusted returns as described by SIPE.
  - The best way is to formulate a portfolio of liquid assets that replicates the economic payoffs and risk exposures of the illiquid asset.
  - Since holding the liquid replicating portfolio provides the "option to get out", the risk of the replicating portfolio is the lower bound of the expected risk for the illiquid asset
  - A good example is provided for commercial real estate in https://www.northinfo.com/documents/813.pdf.
- For absolute risk, a less accurate way is to take asset class index return data (e.g. RE, PE) and adjust the historical index volatility for serial correlation.
  - Ex-ante volatility = Observed volatility * [(1+R)/(1-R)]^.5 where R is the first order autocorrelation coefficient
  - *This doesn't make sense for the benchmark relative case.* No two investors can hold a non-divisible illiquid asset. If I own a particular asset (e.g. an office building) you can't so there is no way to hold the index.

Incorporating Higher Moments

- Northfield risk reporting and optimization uses the method of Cornish and Fisher (1937) to convert a four moment return distribution to the economically equivalent two moment distribution.
  - If higher moments are present the *effective volatility* of each asset may be higher or lower.
  - Depending on degree of diversification of the portfolio and time horizon of the risk estimate you can calculate what portion of the increments (up or down) in volatility at the asset level work through to the portfolio level.
- For example, consider a portfolio with an expected return of 8% and a volatility of 10% under the normal distribution assumption.
  - Now we will introduce a 2% probability of a 90% loss (market crash) so there is a still a 98% probability of the original distribution.
  - This implies a large negative skew and positive excess kurtosis.
  - The volatility equivalent for being long this portfolio is about 33% and being short this portfolio is less than 6%

The Paradox of Active Management

- It is not possible for all active managers to achieve positive alpha, but they all believe it ex-ante. Somebody has to be wrong and that "risk of being wrong" is not accounted for in tracking error and hence IR.
- We will correct for this effect as follows:
  - Assume some probability p that the positive expected alpha is correct, which implies a probability of (1-p) that sign on the alpha is wrong.
  - We now have a bi-modal distribution with two modes centered on +alpha and –alpha, each of which has the same dispersion (tracking error)
  - The bi-modal distribution can be converted to a four-moment distribution with negative skew and positive kurtosis as a "mixture of normals".
- Use the Cornish Fisher method to convert the resultant four moment distribution to the "volatility equivalent".

“Paradox” Example

- Consider a hedge fund with historical return over the risk- free rate of 4.4% with a volatility of 2.8%, which gives a Sharpe ratio of 1.59.

- However, we will assume that there is a 60% chance that this good outcome came from skill (repeatable) and a 40% it was just luck.
  - After the mixture distribution algebra and the transformation of the higher moments into “volatility equivalent”, the alpha is just .88% and the expectation of the volatility equivalent is 6.46%.

- Using the SIPE metric the result is **negative** (.88 – 6.46/6) at about -.2% risk adjusted return.

Conclusions

- We have proposed SIPE as a universal metric for the evaluation of investment performance.
- SIPE is closely related to mean-variance utility as the functional form of "risk adjusted returns".
  - The key difference is that we infer mean/variance risk aversion from the ex-ante risk of the portfolio, so it is no longer investor dependent.
  - Use of the ex-ante risk expectation removes hindsight biases and the influence of rare, unusual conditions (e.g. pandemic)
  - Returns may be absolute, relative, or any blending of the two.
- We propose three analytical enhancements to the parameterization of SIPE that allows it to be widely applied:
  - SIPE works in evaluation of both passive and active strategies
  - Liquid and illiquid asset classes
  - Normally distributed returns or with material higher moments.