

Seven Quantitative Insights into Active Management

Ronald N. Kahn

Table of Contents

Introduction.....	1
Insight 1: Active Management Is Forecasting	3
Insight 2: Information Ratios Determine Value Added	7
An Empirical Addendum	10
Insight 3: The Fundamental Law of Active Management	11
Examples	12
Implications	12
Insight 4: Three-Part Alpha	15
Converting Raw Signals into Alphas	15
Providing Structure	17
Summary	18
Insight 5: Data Mining Is Easy	19
The Statistics of Coincidence	19
Investment Research	21
Four Guidelines for Backtesting Integrity	21
Conclusions	22
Insight 6: Implementation Subtracts Value	23
Value Added Versus Turnover	23
Relationship to Turnover	25
Insight	27
Insight 7: It's Hard to Distinguish Skill from Luck	29

Introduction

In past issues of BARRA's newsletters, we have presented seven quantitative insights into active management. We can summarize these seven insights by observing how they all fit into the process of active management. FIGURE 1 illustrates this process, which we define generally as efficiently utilizing superior information.

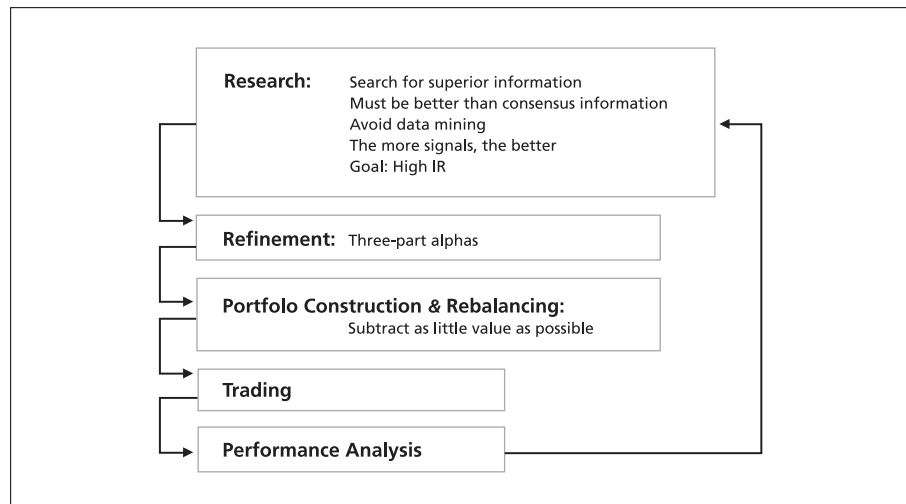


FIGURE 1

The *Research* stage involves the search for superior information. This information must be better than consensus information (*Insight 1: Active Management is Forecasting*). We can improve our chances for success by investigating many signals (*Insight 3: The Fundamental Law of Active Management*). At the same time, we need to avoid data mining (*Insight 5: Data Mining Is Easy*). The goal of this Research stage is a strategy with a high information ratio (*Insight 2: Information Ratios Determine Value Added*).

The *Refinement* stage takes our research signals and converts them to alphas by controlling for skill, volatility, and expectations (*Insight 4: Three-Part Alphas*).

The *Portfolio Construction and Rebalancing* stage and the *Trading* stage implement the strategy. Here the goal is to lose as little of the intrinsic strategy value as possible (*Insight 6: Implementation Subtracts Value*).

The *Performance Analysis* stage looks at results, identifying (imperfectly) what worked and what didn't work, in part as feedback to *Research (Insight 7: It's Hard to Distinguish Skill from Luck)*.

It is useful to note that while we have presented (and derived) these seven insights as "quantitative," they apply to all managers: fundamental, quantitative, top-down, bottom-up, equity, bond, and so forth.

Finally, let me point out that we have called this series "*Seven Quantitative Insights...*" and not "*The Seven Quantitative Insights into Active Management.*" This series has omitted some known insights. And, as a researcher, I will always claim that there remain more insights to uncover.

Active management combines art and engineering. The art involves finding valuable information about future returns. The engineering involves efficiently capturing that information in superior portfolios. By assuming that it is possible to find such valuable information, we can derive many important insights into the engineering of this process.

Over the next several *BARRA Newsletters*, I will outline seven insights which follow from this perspective. Richard Grinold and I discuss these points more comprehensively in our book *Active Portfolio Management*, and many of these items appear throughout the lore and literature of the profession.

Insight 1: Active Management Is Forecasting

By definition, active managers hold portfolios which differ from their benchmark, with the goal of returns in excess of that benchmark. We will assume a mean/variance framework. The manager's portfolio is optimal given his expected returns and risks. What happens if his expected returns match the consensus, i.e., $f = \beta \cdot f_B$?

Each asset's expected excess return is just the asset's beta with respect to the benchmark times the benchmark expected excess return. Well, if we input these consensus forecasts into our mean/variance optimizer, we will end up holding the benchmark! We show this situation in FIGURE 1.1.

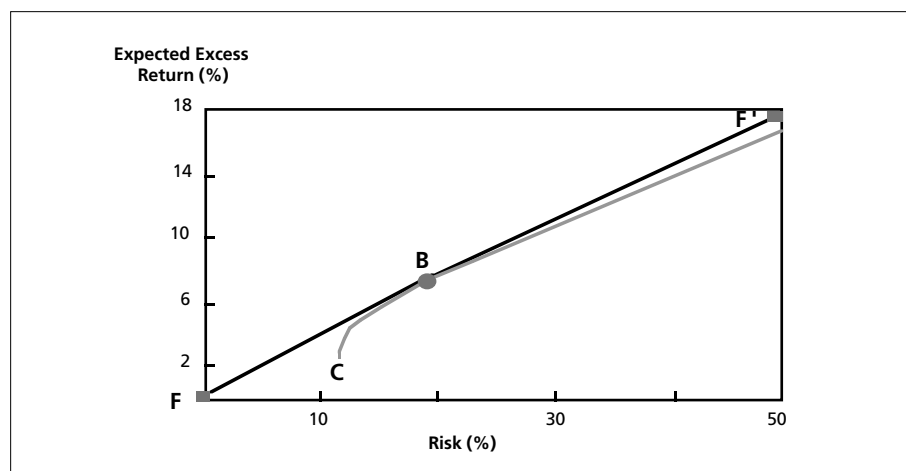


FIGURE 1.1

Portfolio B is the benchmark. Portfolio C is the minimum-risk fully invested portfolio. Portfolio F is the risk-free portfolio, and Portfolio F1 is a combination of B and F.

This is passive management, not active management. The benchmark is the portfolio with maximum Sharpe ratio. No matter what you think of the CAPM, it is the source of consensus expected returns for precisely this reason: with these consensus expected returns, we will hold the consensus portfolio: the benchmark. This “use” of the CAPM (really portfolio theory) is quite powerful, and it even works for any benchmark, not just “the market.”

What happens if our forecasts differ from the consensus, i.e., $f = \beta \cdot f_B + \alpha$? We now have asset alphas.

In this case, the benchmark is no longer efficient. Portfolio Q, which differs from B, is now the portfolio with maximum Sharpe ratio. Portfolio B is not on the efficient frontier. With these forecasts, we will hold an efficient portfolio not equal to B. We show this different situation in FIGURE 1.2.

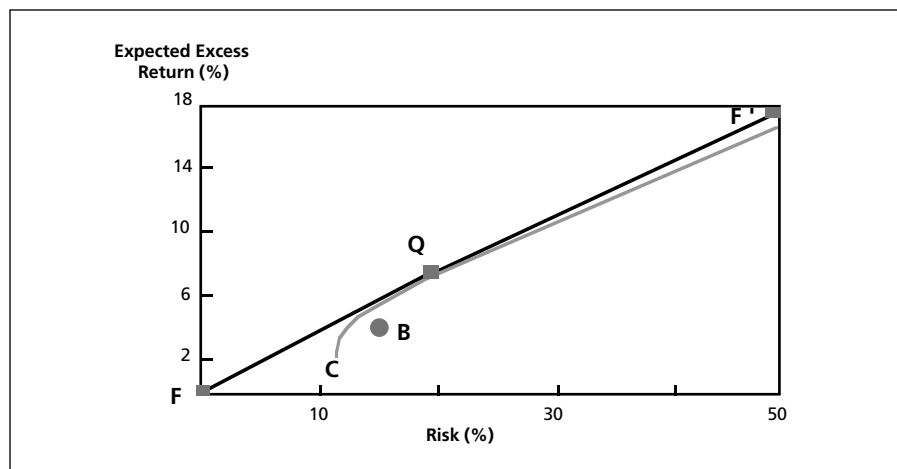


FIGURE 1.2

So only if your forecasts differ from the consensus forecasts will you hold a portfolio which differs from the benchmark. This works in the other direction too: your portfolio holdings imply expected returns (proportional to betas

with respect to your portfolio), and if your portfolio differs from the benchmark, your implied expected returns will differ from the consensus expected returns. Active management is forecasting.

Insight 2: Information Ratios Determine Value Added

Let's focus on the active manager's job: outperforming a benchmark. We can split the manager's return into two pieces — a component driven by the benchmark (the systematic return) and a component independent of the benchmark (the residual return):

$$r_p = \beta_p \cdot r_B + \theta_p \quad (2.1)$$

where r_p = portfolio excess return
 β_p = portfolio beta
 r_B = benchmark excess return
 θ_p = residual return

We will define the manager's alpha¹ as the expected (or later realized) residual return:

$$\alpha_p = E\{\theta_p\} \quad (2.2)$$

We also define residual risk ω :

$$\omega_p = STD\{\theta_p\} \quad (2.3)$$

Active managers build portfolios by trading off expected active return against risk.

We define the value added from active management as:

$$VA = \alpha - \lambda \cdot \omega^2 \quad (2.4)$$

where λ measures the investor's aversion to risk. The manager builds active portfolios to maximize this added value.

¹ Typically, active institutional managers choose $\beta_p = 1$, equating active returns and residual returns.

Individual preferences enter into the value added only in the ways in which individuals trade off residual return against risk. More risk-averse investors will demand more incremental return for each unit of risk.

The information ratio is the manager's ratio of residual return to risk:

$$IR = \frac{\alpha}{\omega} \quad (2.5)$$

We will consider this a fundamental constant defining the manager, assuming it is independent of time and the level of risk. A manager can deliver more residual return only by taking on more risk:

$$\alpha = IR \cdot \omega \quad (2.6)$$

This is exactly true in the absence of constraints. For example, if the manager overweights one position by 5 percent and underweights another by 3 percent, leading to a given forecast alpha, he can double both the alpha and the residual risk by increasing the overweighting to 10 percent and the underweighting to 6 percent.

We can use this "budget constraint" (EQUATION 6) to rewrite value added as:

$$VA = IR \cdot \omega - \lambda \cdot \omega^2 \quad (2.7)$$

FIGURE 2.1 shows graphically how value added depends on active risk.

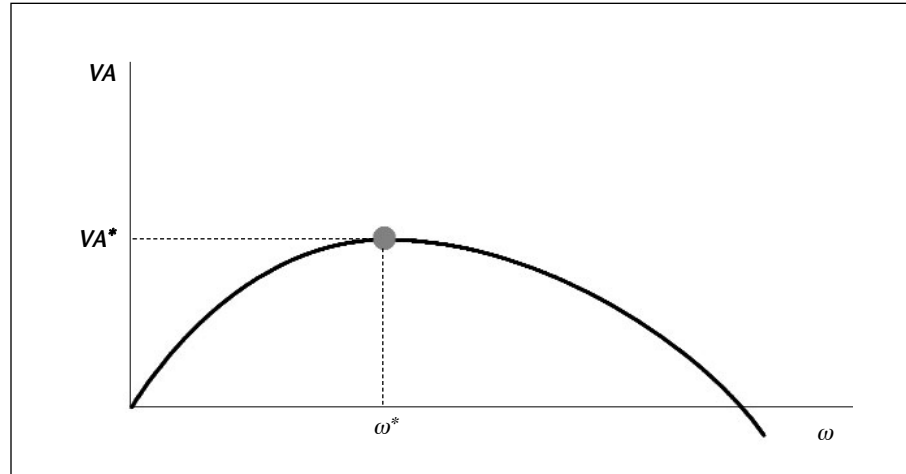


Figure 2.1

The active manager chooses the portfolio corresponding to the maximum point in FIGURE 2.1.

At this point:

$$\omega^* = \frac{IR}{2\lambda} \quad (2.8)$$

and

$$VA^* = \frac{(IR)^2}{4\lambda} \quad (2.9)$$

EQUATION 2.8 describes the optimal level of residual risk, ω^* . Optimal residual risk depends inversely on risk aversion, and directly on the information ratio. More risk-averse investors will choose lower levels of active risk. The higher the information ratio, the more residual risk a particular investor will tolerate.

Each investor's maximum utility depends, according to EQUATION 2.9, directly on the square of the information ratio and inversely on the risk aversion. This is the critical point. For any particular investor, for any particular amount of risk aversion, the most desirable manager will have the highest information ratio. So all investors, regardless of preferences, will agree that the highest information ratio can provide the most value. EQUATION 2.9 shows that information ratios determine value added.

An Empirical Addendum

Given the central role of information ratios, it is useful to know their typical values. Based on mutual fund research at BARRA, the typical before-expenses distribution of information ratios is roughly:

Percentile	IR
90	1.0
75	0.5
50	0.0
25	-0.5
10	-1.0

This holds for both equity and fixed income funds. A top quartile manager will have an information ratio of 0.5, adding 50 basis points of outperformance for every 100 basis points of active risk, before expenses.

Insight 3: The Fundamental Law of Active Management

In the previous insight we learned that the information ratio is the key to active management. Given that, how can we achieve high information ratios? Let's begin by looking at a relationship Richard Grinold has called "The Fundamental Law of Active Management." This law expresses the information ratio in terms of two other statistics, the information coefficient and the "breadth":

$$IR = IC \cdot \sqrt{BR} \quad (3.1)$$

where IR = information ratio

IC = information coefficient ("skill")

BR = independent bets per year ("breadth")

Now, expressing one quantity in terms of two others isn't always progress, but let's leave that thought for now.

The *information coefficient* is a measure of manager skill. In particular, it's the correlation of forecast and realized residual returns. It measures the manager's edge in choosing assets. The *breadth* measures the number of independent bets the manager takes per year. It measures diversification. We define breadth as bets *per year* because the information ratio is an annualized quantity.

According to the fundamental law, in order to achieve a high information ratio a manager must demonstrate an edge for every asset chosen and must diversify that edge over many separate assets.

In fact, the fundamental law of active management is an old result from gaming theory. Imagine a roulette wheel where players can bet on red or black. The casino has a small edge because two numbers, 0 and 00, are green. Through every spin of the roulette wheel, the casino maintains that same small edge. Now imagine that during the course of the year, players bet a total of \$10 million on this roulette wheel. Here are two possible scenarios: In the first, that \$10 million consists of five million spins of the wheel with \$2 bet for each spin. In the second scenario, the players all agree to pool resources and bet all \$10 million on one spin of the wheel. The casino's expected return is the same in both scenarios. However, it would clearly far prefer the first scenario from a reward-to-risk tradeoff.

Examples

Next consider two investment examples. First, imagine a stock picker with an information coefficient of 0.035, a small but reasonably impressive level of skill for the active management business. This manager follows 200 stocks per quarter, effectively taking 800 bets per year. The fundamental law implies an information ratio of 0.99 — indicative of a top decile manager.

Let's compare this example with the performance of a market timer. We'll assume that the market timer has a higher level of skill for every bet, with an information coefficient of 0.05. But this manager times the market by looking at broad macroeconomic trends and devises a new forecast once per quarter—taking four independent bets per year. In this case, the fundamental law implies an information ratio of just 0.10, slightly above the median for active managers. So a higher level of skill per bet does not necessarily translate into a higher information ratio. And given that it's easier and cheaper to increase breadth than skill, stock-picking strategies may have a higher chance for success than market-timing strategies.

Implications

The Fundamental Law of Active Management has several implications:

- *Given some skill, bet as often as possible.*
- *Combine models*, because breadth applies across models as well as assets. For example, an international equity manager can bet on countries, currencies, and individual stocks.
- *Don't market time.* Market-timing strategies are unlikely to generate high information ratios. While such strategies can generate very large returns in a particular year, they're heavily dependent on luck. On a risk-adjusted basis, the value added will be small. This will not surprise most institutional managers, who avoid market timing for just this reason.

- *Tactical asset allocation has a high skill hurdle.* This strategy lies somewhere between market timing and stock picking: it provides some opportunity for breadth, but not nearly the level available to stock pickers. Therefore, to generate an equivalent information ratio, the tactical asset allocator must demonstrate a higher level of skill.

In summary, information ratios, the key to active management, depend on both skill and breadth.

Insight 4: Three-Part Alpha

In the previous issue we learned that information ratios, as the key to active management, depend on both skill and breadth. In this issue we will learn the constituent parts of alphas, the fundamental inputs for active management.

Raw signals like analysts' earnings forecasts or broker buy/sell recommendations hopefully contain information useful in forecasting returns. But these raw data are not alphas: expected residual returns. They do not even necessarily have the units of return.

A basic forecasting formula governs the connection between these raw signals and alphas. This formula refines the raw signals into alphas by controlling for expectations, skill, and volatility. In many cases we can simplify this formula to a particularly intuitive form.

Converting Raw Signals into Alphas

The basic forecasting formula provides the best linear unbiased estimate of the residual return θ , given the raw signal g :

$$E\{\theta|g\} = E\{\theta\} + Cov\{\theta, g\} \cdot Var^{-1}\{g\} \cdot (g - E\{g\}) \quad (4.1)$$

According to EQUATION 4.1, the expectation of θ conditional on g equals the unconditional expectation of θ , plus a term dependent on the difference between the observed signal and its unconditional expectation. Reordering terms, we see that:

$$(E\{\theta|g\} - E\{\theta\}) = Cov\{\theta, g\} \cdot Var^{-1}\{g\} \cdot (g - E\{g\}) \quad (4.2)$$

So this formula controls for expectations. Only if the raw signal g differs from its unconditional expectation will the expectation of θ differ from its unconditional expectation.

This result is intuitive. If company earnings exactly match expectations, we do not expect the stock to move. That happens only when earnings do not match expectations.

Now let's simplify EQUATION 4.2 into a more intuitive form that reveals how alphas include controls for skill and volatility. First, the unconditional expected residual return is zero. In the absence of information, expected returns match their consensus (CAPM) values. Second, $E\{\theta|g\} - E\{\theta\}$ is the alpha, the expected residual return given signal g . Third, the covariance term includes a correlation term and two standard deviations.

Substituting:

$$\alpha = \text{Corr}\{\theta, g\} \cdot \text{STD}\{\theta\} \cdot \left(\frac{g - E\{g\}}{\text{STD}\{g\}} \right) \quad (4.3)$$

We commonly denote the correlation of raw signals and realizations as the information coefficient IC , and the standard deviation of the residual return as ω . We will refer to the standardized raw signal as the score S , because it has mean zero and standard deviation 1. It ranges roughly from -2 to +2. Hence:

$$\alpha = IC \cdot \omega \cdot S \quad (4.4)$$

So there are three parts to every alpha: an information coefficient, a volatility, and a score.

EQUATION 4.4 clearly shows how alphas control for skill and volatility. The information coefficient is a measure of skill. With no skill the IC is zero, and EQUATION 4.4 sets the alpha to zero, as it should. The greater the skill, the greater the alpha, other things equal.

Volatility serves two purposes in EQUATION 4.4. First, it provides the units of return. The *IC* and score are dimensionless. Second, it controls the alpha for volatility. For a given skill level, imagine two stocks with equal bullish scores of +1. We think both stocks will go up. EQUATION 4.4 says that the higher volatility stock will go up more. If both Pacific Gas & Electric and Netscape achieve earnings one standard deviation above expectations, then Netscape should rise more. Both stocks will rise, but the more volatile stock will rise more.

Providing Structure

Understanding the three constituent parts of an alpha can provide intuition. It can also provide structure in unstructured situations, where the connections between raw signals and alphas are unclear.

The ultimate example of an unstructured situation is the stock tip. Even in this case, EQUATION 4.4 can provide structure. Imagine that the stock in question has residual volatility of 20.%. Then TABLE 4.1 shows the range of possible alphas as a function of *IC* and score.

<i>IC</i>		VERY POSITIVE	VERY, VERY POSITIVE
Great	0.10	2%	4%
Good	0.05	1%	2%
Average	0.0	0%	0%

TABLE 4.1

Since stock tips are always presented as very, very positive (“I make only one or two recommendations a year; you are the first person I called...” etc.), converting from the tip to an alpha only requires estimating the tipper’s *IC*. Is Warren Buffet on the line, or someone you have never heard of?

For an institutional money manager, a more relevant example involves converting broker buy/sell recommendations into alphas. This common situation has relatively little structure, but understanding three-part alphas can help.

TABLE 4.2 shows an example, assuming that the broker has a good information coefficient of 0.05.

STOCK	ω	REC	SCORE	α
A	15%	Buy	+1	0.75%
B	20%	Buy	+1	1.00%
C	15%	Sell	-1	0.75%
D	30%	Sell	-1	1.50%
E	25%	Sell	-1	1.25%

TABLE 4.2

Our conversion from recommendations into scores is straightforward. Notice that Stocks A and B, both recommended, have different alphas. Stock B has the higher volatility: We expect it to go up more than Stock A.

Contrast this with simply giving every stock on the buy list an alpha of 1.0%. In an optimizer, the stocks would all have identical expected returns, so the optimal portfolio would be the minimum variance portfolio. The optimal portfolio would load up on the least volatile stocks. Even after controlling for volatility as in TABLE 4.2, an optimizer still favors low volatility,² but we have mitigated the effect.

Summary

In summary, to convert from raw signals to alphas requires controlling for expectations, skill, and volatility. Understanding this conversion provides insight and structure in many situations.

² If we ignore constraints and active return correlations, optimal active holdings are proportional to the ratio of alpha to active variance.

Insight 5: Data Mining Is Easy

Why is it that so many strategies look great in backtests and disappoint upon implementation? Backtesters always have 95% confidence in their results, so why are investors disappointed far more than 5% of the time? It turns out to be surprisingly easy to search through historical data and find patterns that don't really exist.

To understand why data mining is easy, we must first understand the statistics of coincidence. Let's begin with some non-investment examples. Then we will move on to investment research.

The Statistics of Coincidence

Several years ago Evelyn Adams won the New Jersey state lottery twice in four months. Newspapers put the odds of that happening at 17 trillion to 1, an incredibly improbable event. A few months later, two Harvard statisticians, Percy Diaconis and Frederick Mosteller, showed that a double win in the lottery is not a particularly improbable event. They estimated the odds at 30 to 1. What explains the enormous discrepancy in these two probabilities?

It turns out that the odds of Evelyn Adams winning the lottery twice are in fact 17 trillion to 1. But that result is presumably of interest only to her immediate family. The odds of someone, somewhere, winning two lotteries — given the millions of people entering lotteries every day — are only 30 to 1. If it wasn't Evelyn Adams, it could have been someone else.

Coincidences appear improbable only when viewed from a narrow perspective. When viewed from the correct (broad) perspective, coincidences are no longer so improbable. Let's consider another non-investment example: Norman Bloom, arguably the world's greatest data miner.

Norman died a few years ago in the midst of his quest to prove the existence of God through baseball statistics and the Dow Jones average. He argued that "BOTH INSTRUMENTS are in effect GREAT LABORATORY EXPERIMENTS

wherein GREAT AMOUNTS OF RECORDED DATA ARE COLLECTED, AND PUBLISHED” (capitalization Bloom’s). As but one example of thousands of his analyses of baseball, he argued that the fact that George Brett, the Kansas City third baseman, hit his third home run in the third game of the playoffs, to tie the score 3-3, could not be a coincidence — it must prove the existence of God. In the investment arena, he argued that the Dow’s 13 crossings of the 1,000 line in 1976 mirrored the 13 colonies which united in 1776 — which also could not be a coincidence. (He pointed out, too, that the 12th crossing occurred on his birthday, deftly combining message and messenger.) He never took into account the enormous volume of data — in fact, an entire New York Public Library’s worth — he searched through to find these coincidences. His focus was narrow, not broad.

The importance of perspective to understanding the statistics of coincidence was perhaps best summarized by, of all people, Marcel Proust — who often showed keen mathematical intuition:

The number of pawns on the human chessboard being less than the number of combinations that they are capable of forming, in a theater from which all the people we know and might have expected to find are absent, there turns up one whom we never imagined that we should see again and who appears so opportunely that the coincidence seems to us providential, although, no doubt, some other coincidence would have occurred in its stead had we not been in that place but in some other, where other desires would have been born and another old acquaintance forthcoming to help us satisfy them. (*The Guermantes Way, Cities of the Plain*, Volume 2 of translation of Marcel Proust’s *Remembrance of Things Past* [New York: Vintage Books, 1982], p. 178.)

Investment Research

Investment research involves exactly the same statistics and the same issues of perspective. The typical investment data mining example involves *t*-statistics gathered from backtesting strategies. The narrow perspective says: “After 19 false starts, this 20th investment strategy finally works. It has a *t*-statistic of 2.”

But the broad perspective on this situation is quite different. In fact, given 20 informationless strategies, the probability of finding at least one with a *t*-statistic of 2 is 64%. The narrow perspective substantially inflates our confidence in the results. When viewed from the proper perspective, confidence in the results lowers accordingly.

Four Guidelines for Backtesting Integrity

Given that data mining is easy, how can we safeguard against it? Here are four guidelines for data mining integrity:

- Intuition
- Restraint
- Sensibility
- Out-of-sample testing

The *intuition* guideline demands that researchers investigate only those strategies with some *ex ante* expectation of success. Investment research should never involve free-ranging searches for patterns without regard for intuition.

The *restraint* guideline attempts to minimize the number of strategies investigated — i.e., to keep the broad and narrow focus similar. In the best case, researchers decide *ex ante* exactly which strategies and variants they will investigate, run their tests, and look at the answers. They do not go back and continually refine their investigations.

The *sensibility* guideline deletes results that seem improbably successful. Observed *t*-statistics that are too large may signal database errors or an improper methodology rather than a new strategy.

The fourth guideline, out-of-sample testing, is the statistician's answer to the curse of data mining. Coincidences observed over one data set are quite unlikely to reoccur in another independent data set.

Conclusions

Many backtesting results are not foolproof demonstrations of strategy value but merely coincidence. Four backtesting guidelines can help avoid data mining.

Insight 6: Implementation Subtracts Value

Every promising investment strategy on paper loses value on implementation. Constraints imposed during portfolio construction, sometimes by the client, lower value added. Transaction costs lower value added. Insight, especially into the relationship between transaction costs and value added, can help minimize this loss.

Investors use constraints in portfolio construction for two reasons. Often they face external regulations — for example, lists of stocks they cannot own. Sometimes they use constraints to limit the impact of any flaws in their inputs: return and risk forecasts. These constraints lower realized value added relative to that of paper portfolios not subject to constraints. Much of this loss is either beyond the manager’s control or illusory (that is, constraints to counteract flawed inputs should increase value added).

Transaction costs lower value added, and the manager does have some control over these. In a utility function balancing expected returns, risk, and transaction costs, the transaction costs are particularly vexing because we incur them with certainty, in contrast to the expected returns we can only hope to achieve. There are two ways to control or reduce transaction costs: trade smarter (i.e., more cheaply trade by trade) or trade less. We will not deal here with approaches to reducing transaction costs trade by trade, though BARRA has pioneered methods here. Trading less doesn’t sound appealing on the surface, because less trading means acting on less of our superior information. But as this insight will show, we can trade substantially less without giving up much value added.

Value Added Versus Turnover

We want to understand the tradeoff between value added and turnover. Let’s take a minute to develop this relationship. If you want to avoid the technical details, skip down to EQUATION 6.16 and the insight.

First let's return to *Insight 2: Information Ratios Determine Value Added*. There we saw that the information ratio acted as a budget constraint: achieving higher alpha required taking on more risk:

$$\alpha = IR \cdot \omega \quad (6.1)$$

Defining value added as:

$$VA = \alpha - \lambda \omega^2 \quad (6.2)$$

$$= IR \cdot \omega - \lambda \omega^2 \quad (6.3)$$

leads to FIGURE 6.1, showing value added as a function of risk. FIGURE 6.1 is just a graph of EQUATION 6.3.

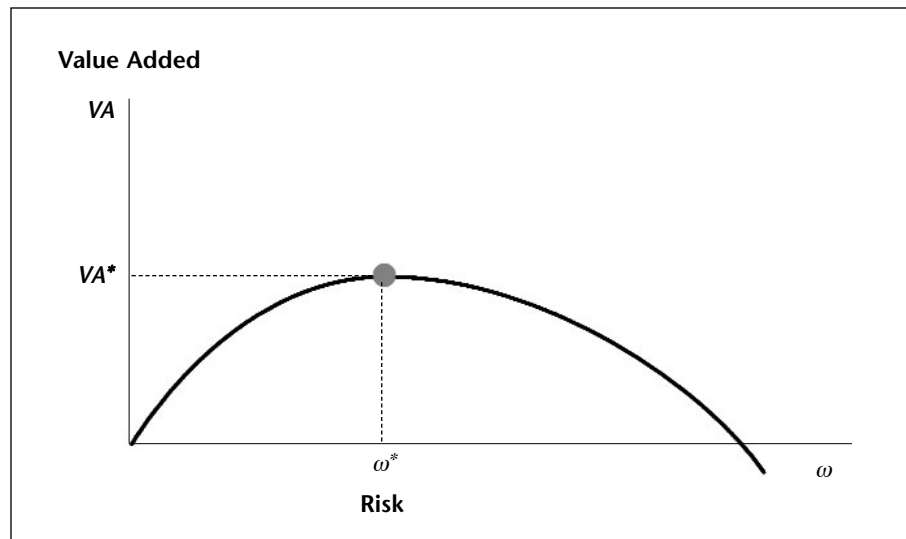


FIGURE 6.1

That previous insight also derived and discussed formulas for the optimal level of risk, ω^* , and the optimal value added, VA^* .

We simply reprint those results here:

$$\omega^* = \frac{IR}{2\lambda} \quad (6.4)$$

$$VA^* = \frac{(IR)^2}{4\lambda} \quad (6.5)$$

But we can rewrite EQUATION 6.4 as:

$$IR = 2\lambda\omega^* \quad (6.6)$$

and then substitute this into EQUATION 6.5, to find:

$$VA^* = \lambda\omega^{*2} \quad (6.7)$$

Substituting these back into EQUATION 6.3, we can rewrite value added as a function of risk as:

$$VA = VA^* \cdot \left\{ 2 \left(\frac{\omega}{\omega^*} \right) - \left(\frac{\omega}{\omega^*} \right)^2 \right\} \quad (6.8)$$

Relationship to Turnover

EQUATION 6.8 connects value added to risk. Can we connect risk to turnover? Imagine that the optimal solution involves a set of purchases and sales, Δn_i^* , with total turnover, TO^* , and optimal alpha, risk, and value added. Now the simplest strategy to cut turnover by a fraction x :

$$TO = x \cdot TO^* \quad (6.9)$$

is simply to reduce each trade by the same proportion.

Then:

$$\Delta n_i = x \cdot \Delta n_i^* \quad (6.10)$$

$$\omega = x \cdot \omega^* \quad (6.11)$$

$$\alpha = x \cdot \alpha^* \quad (6.12)$$

In this strategy:

$$\frac{\omega}{\omega^*} = \frac{TO}{TO^*} \quad (6.13)$$

and so:

$$VA = VA^* \cdot \left\{ 2 \left(\frac{TO}{TO^*} \right) - \left(\frac{TO}{TO^*} \right)^2 \right\} \quad (6.14)$$

In a more sophisticated strategy, reducing turnover from TO^* by conducting the most valuable trades first, we would expect:

$$\frac{TO}{TO^*} \leq \frac{\omega}{\omega^*} \quad (6.15)$$

and hence:

$$VA \geq VA^* \cdot \left\{ 2 \left(\frac{TO}{TO^*} \right) - \left(\frac{TO}{TO^*} \right)^2 \right\} \quad (6.16)$$

Insight

FIGURE 6.2 graphs this value added/turnover frontier. According to this result, we can achieve at least three-quarters of the value added with only half the turnover. And the key qualifier is “at least.” An extremely simple strategy, just reducing each trade by exactly the same fraction, can achieve three-quarters of the value added with half the turnover (the lighter line in FIGURE 6.2). You can do even better by distinguishing transaction costs between stocks and scheduling the most valuable trades first (the darker line in FIGURE 6.2).

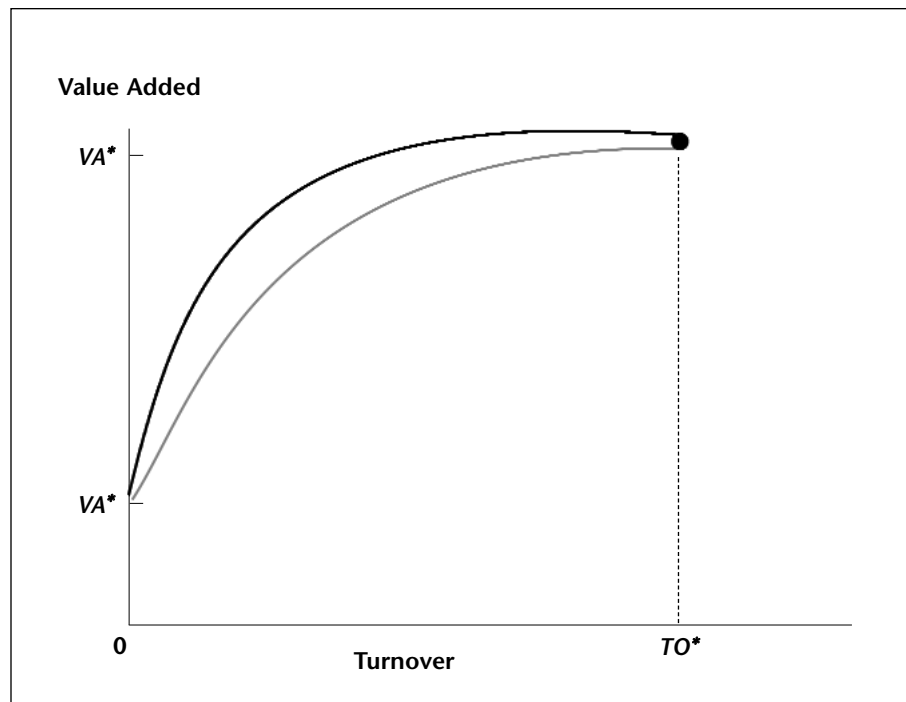


FIGURE 6.2

This insight has two practical implications. First, do not necessarily dismiss promising strategies with high turnover; you may be able to capture much of the value added with significantly less turnover. Second, transaction cost research is extremely valuable precisely for this reason: to distinguish trades and realize the most value at the least cost.

Insight 7: It's Hard to Distinguish Skill from Luck

Early in this series (Part 2) we learned that information ratios (IR) determine value added: they are the key statistic capturing manager performance. Here we will look at our ability to accurately measure information ratios.

To understand the measurement issues involved, we will need to analyze the standard error of an observed information ratio (IR). The standard error will help us to determine the statistical significance of the observed IR.

We start by writing the IR as:

$$IR = \frac{\alpha}{\omega} \frac{1}{\sqrt{\Delta t}}, \quad (7.1)$$

where we measure α and ω , the mean and standard deviation of active or residual returns, over time periods Δt , and the Δt in EQUATION 1 annualizes the IR. For example, we might calculate α and ω using monthly data, and the annualization multiplies the monthly ratio by $\sqrt{12}$.

Because we will assume that the statistical uncertainty in α dominates the statistical uncertainty in the IR,³ let's rewrite EQUATION 1 as:

$$IR = \left(\frac{1}{\omega \sqrt{\Delta t}} \right) \cdot \alpha. \quad (7.2)$$

Then, the standard error of the IR is:

$$SE[IR] \approx \left(\frac{1}{\omega \sqrt{\Delta t}} \right) \cdot SE[\alpha] \quad (7.3)$$

$$= \left(\frac{1}{\omega \sqrt{\Delta t}} \right) \cdot \frac{\omega}{\sqrt{N}}, \quad (7.4)$$

where N measures the number of observations underlying our measurements of α and we have used the classic result for the standard error of an estimated mean.

³ This is typically true unless the IR is extremely large (above 1.0) and/or the number of observations is very small. Accounting for the uncertainty in ω will only increase the standard error, so the analysis here presents the best case.

But we can simplify EQUATION 7.4 to find the result:

$$SE[IR] = \frac{1}{\sqrt{T}}, \quad (7.5)$$

where T measures the number of years of observation. Surprisingly, this standard error is independent of frequency: Given five years of data, the standard error of the IR is the same whether we use quarterly, monthly, or daily data.

Given the simple result in EQUATION 7.5, what can we say about the difficulty of measuring information ratios? We have seen that a top quartile manager has an IR of 0.5. How long must we observe such a manager to measure the IR with 95% confidence? We want the t -statistic, the ratio of the IR to its standard error, to exceed 2:

$$t = \frac{IR}{SE[IR]} = \frac{0.5}{\left(\frac{1}{\sqrt{T}}\right)} > 2 \quad (7.6)$$

or:

$$T > 16 \quad (7.7)$$

According to EQUATION 7.7, we require 16 years of data for that level of statistical confidence.⁴

But requiring so long a time series is quite problematic. Most active managers today probably do not have such lengthy track records. More importantly, managers and strategies may not retain their information ratios over such long periods. Superior information and great opportunities don't last forever.

⁴ Some purists may argue that we should invoke a one-sided test: What is the likelihood that random data ($IR = 0$) would generate such a large *positive* IR? We then require only $t > 1.65$. This one-sided test still demands 11 years of data.

So we find ourselves in a situation of some uncertainty, as captured in FIGURE 7.1. Here we classify managers along two dimensions — skill and luck: The “Blessed” possess both skill and luck. Their businesses thrive, and deservedly so. The “Doomed” lack skill and luck. We quickly and appropriately weed them out.

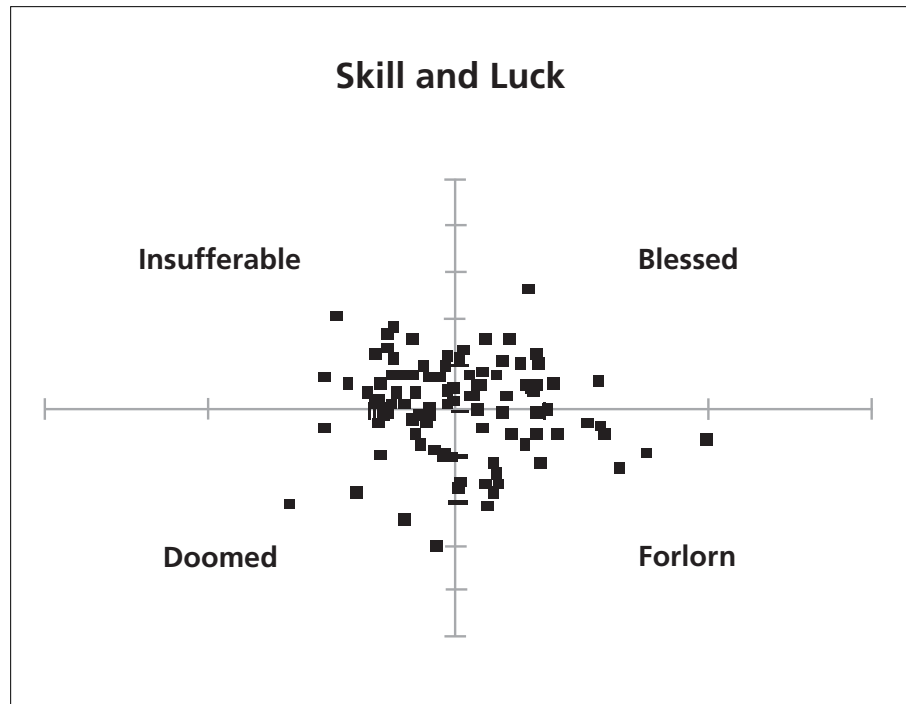


FIGURE 7.1

The other two categories are more problematic. The “Forlorn” possess skill but not luck. Their performance numbers do not convey their true level of skill, and they suffer for that. The final category is the “Insufferable.” They possess luck but no skill, and they thrive. Most managers I know can quickly cite examples of other managers they feel fall into this category.

The real insight at the end of this analysis is that successful active management depends on both skill and luck. Depending on your perspective, this can be good or bad news.