Normal Return Gaps: Dispersion Illuminates Diversification

WILLIAM W. JENNINGS, THOMAS C. O’MALLEY, AND BRIAN C. PAYNE

ABSTRACT: Despite ever more sophisticated risk management and measurement, investment professionals have generally overlooked a simple but powerful measure of relative performance and portfolio diversification—the normal return gap. The authors develop a generalized specification of the expected difference in returns between two investments based on the folded normal distribution. Even highly correlated investments can have quite large expected return gaps. They then demonstrate the applicability of this dispersion to capital market forecasts, manager selection, performance evaluation, style tilts, sector bets, socially responsible investing, manager combinations, wash sale taxation, and rebalancing.

TOPICS: Performance measurement, wealth management, manager selection, ESG investing*

Persuaded by the benefits of diversification, investors incorporate new asset classes and subclasses into their portfolios. Correlation is typically the chief measure of diversification in doing so. However, investors are often later surprised by the actual relative performance of their new investments. These differences in relative performance are return gaps.

Even highly correlated investments can produce meaningful diversification; conversely, low correlations can produce small return gaps and, therefore, minimal diversification. Normal return gaps between investments are often the same magnitude as the expected returns of the underlying assets. Investors unaware of normal return gaps risk terminating worthy managers; establishing competitive “horse races” between managers is particularly unwise.

KEY FINDINGS
• Even highly correlated investments can produce meaningful diversification; conversely, low correlations can produce small return gaps and, therefore, minimal diversification.
• Normal return gaps between investments are often the same magnitude as the expected returns of the underlying assets.
• Investors unaware of normal return gaps risk terminating worthy managers; establishing competitive “horse races” between managers is particularly unwise.

Normal return gaps between investments are often the same magnitude as the expected returns of the underlying assets.

In investors unaware of normal return gaps risk terminating worthy managers; establishing competitive “horse races” between managers is particularly unwise.

The Journal of Wealth Management

Fall 2020
trade. Absent rebalancing trades, a portfolio’s return is merely its components’ returns weighted by their starting values. And without rebalancing, there is no benefit derived from the constituent assets’ interim imperfect correlation. Booth and Fama (1992) and Willenbrock (2011) show that the incremental diversification return arises solely from rebalancing, which comes from return gaps.

As an example of the size of the normal return gaps, we show that, although large-cap US stocks and small-cap US stocks have a 0.92 correlation under some widely available capital market assumptions, the expected value of their normal return gap is 7.60%. This expected return gap is as large as the expected return.

We contend that a realistic preview of these sizable expected return gaps between investments can help make asset mix and manager selection decisions more robust and sustainable. Consider an investor contemplating a new asset class. When the old and new investments have the same expected return, investors should expect returns to be as different as the normal return gap on average. As shown, these gaps can be substantial. Experiencing these gaps can make it seem that diversification has failed, even when inclusion of the new investment is a sound long-term decision. A realistic preview of this possibility mitigates the problem of “losing faith” in a diversification decision because of initial underperformance. It adds robustness to the diversification decision. Private wealth managers and investment consultants who properly calibrate expectations that even highly correlated investments can have meaningful return gaps build resilience into their clients’ portfolios and into their relationship with those clients.

Comparing two active managers is another instance where return gaps are useful. Here, there is no reference return, as when comparing a single manager to a benchmark or when comparing a new investment to an old one. We show the expected return gap is quite large, even for two exceptional managers—up to 9% to 14% annually in plausible equity scenarios. Investors hired the two managers because they believed them both skilled. Expected return gaps this large mean that investors risk wrongfully terminating one manager, even when she is, in fact, skilled. Foreknowledge of the surprisingly large expected dispersion between managers can mitigate this potential error.

While return gaps are not new to the literature, we believe we enhance and extend this research. Statman and Scheid (2005, 2006, 2008) make a compelling case that return gaps do a better job of characterizing diversification benefits than the standard approach using correlation. They persuasively assert most investors have a “faulty intuition about correlation” (Statman and Scheid, 2006, p. 25). Roll (2013) highlights other failings of correlation as a diversification measure while Menchero and Morozov (2011) show the importance of individual-asset volatility to diversification and return gaps. Solnik and Roulet (2000) use return gaps to calibrate correlation and diversification benefits instantaneously, without the need for a long time series. Our approach to return gaps is more general than these prior applications, correctly characterizes the benefits of diversification, and augments the case these other authors make for the relevance and usefulness of return gaps.

In the following, we develop formulas for the normal return gap. We illustrate that allowing the two underlying assets’ returns to differ in mean and in variance is crucial. We then apply our specification of the normal return gap to historical data, to investment consultant capital market return assumptions, and to rebalancing. We demonstrate the relevance of return-gap performance divergences to manager selection, manager combinations, style tilts, sector bets, socially responsible investing, calibrating uncertainty about return gaps, as well as wash sales and taxation. We conclude by highlighting the investment implications of our results.

### NORMAL RETURN GAPS

Start by assuming returns for two asset classes follow a normal distribution:

\[ X \sim N(\mu_X, \sigma_X^2) \]
\[ Y \sim N(\mu_Y, \sigma_Y^2) \]

While the true distribution of asset returns has been a subject of much research, normality is not completely baseless. For instance, our analysis of Vanguard asset class index funds cannot reject normality based on a Kolmogorov–Smirnov test over our 21-year time series (see the next section). The distribution of a linear combination of normal random variables is also normal. Defining \( \Delta = X - Y \) and letting \( \rho \) be the correlation between \( X \) and \( Y \), we have:

\[ \Delta \sim N(\mu_X - \mu_Y, \sigma_X^2 + \sigma_Y^2 - 2\sigma_X \sigma_Y \rho) \]

(Normal Return Gaps: Dispersion Illuminates Diversification)
Some of these Δ values will be negative, but we are concerned with the gap or distance between the returns. Here, we need the absolute value. With normal distributions, taking absolute values effectively “folds” the negative part of the distribution over zero; the distributional weight below zero is added (symmetrically) to the probability density function above zero. Exhibit 1 shows this relationship.

Because the folded normal distribution builds upon the pervasively used normal distribution, it offers an intuitive and relatively tractable approach to return gaps. However, we have not seen this approach applied to return gaps in the finance literature to date (although Hallerbach 2014 uses a related half-normal distribution to calibrate market timing). Leone, Nelson, and Nottingham (1961) and Tsagris, Beneki, and Hassani (2014) give the distribution of a folded normal distribution:

\[
\Delta_{\mu} = \sqrt{2 \pi} \left( \sqrt{\sigma_x^2 + \sigma_y^2 - 2 \sigma_x \sigma_y \rho} \right) e^{-\frac{(\mu_x - \mu_y)^2}{2(\sigma_x^2 + \sigma_y^2 - 2 \sigma_x \sigma_y \rho)}} \\
+ \{\mu_x - \mu_y\} \left[ 1 - 2 \Phi \left( \frac{-(\mu_x - \mu_y)}{\sqrt{\sigma_x^2 + \sigma_y^2 - 2 \sigma_x \sigma_y \rho}} \right) \right] \tag{5}
\]

Equation (5) is therefore the normal return gap. Notably, our general formula permits investments to vary in return mean and standard deviation, whereas prior return gap analyses assume identical return moments for the two assets.

Intuitively, Equation (5) shows the normal return gap is increasing in the expected returns difference. We can also see that higher correlations reduce the normal return gap, while higher volatilities increase the normal return gap.\(^1\)

\(^1\)Within the second line of Equation (5), the cumulative density function term in the square brackets scales the effect of the difference-in-means term in the curly braces. The difference-in-returns effect is larger at higher correlations. The main impact of volatility and correlation is through the first \(\rho\) in equation (5). The impact of the exponential term is small for typical inputs, particularly if \(\rho < 0.9\).
DISCUSSION

Exhibit 2 presents the normal return gap between two assets under three scenarios. Each scenario shows the relationship between the normal return gap and correlations, volatilities, and means. Panel A illustrates the simple case when the two assets have the same mean and same standard deviation. Intuitively, normal return gaps are higher with higher volatility and with lower correlation. Yet, counterintuitively, even at relatively high correlation levels (e.g., $\rho \approx 0.8$), large normal return gaps occur. For instance, after one year, two assets with $\rho = 0.8$ and 5% expected returns with 10% volatility have a normal return gap of 5.05%—greater than each individual asset’s expected return! Higher volatilities produce larger return gaps. High correlations ameliorate the impact but cannot fully offset high volatilities. As Statman and Scheid (2005, 2006, 2008) noted, for a given correlation, the return gap is linear in volatility. Doubling volatility doubles the return gap.

Panel B of Exhibit 2 demonstrates the importance of the difference in expected returns for the two asset classes. In Panel B, we use the same volatilities as Panel A. Moreover, the average return for the two asset classes in Panel B equals the average from Panel A, but here we assume a difference between the two asset classes. First and most obviously, a difference in expected means produces a larger normal return gap than the Panel A common means. Additionally, the difference in expected returns is more important at high correlation levels and at lower volatilities. Both low correlations and high volatilities dominate the importance of a difference in expected returns. That is, as correlations decrease, so does the importance of the difference in expected returns. The same holds for increasing volatilities. With high enough volatilities and low enough correlation, the expected return difference becomes much less relevant.

Panel C of Exhibit 2 isolates the impact of volatility differences. In Panel C, the average of the two assets’ volatilities is the same as the common volatility of Panels A and B (and the mean returns match as well). At high correlations, the normal return gap is meaningfully larger with different volatilities than with the common volatility of Panels A and B. If average standard deviation were a comprehensive measure, then Panel C should be identical to Panel A, yet we find all return gap values are dramatically higher. As before, normal return gaps are higher with higher volatility and with lower correlation. Large return gaps appear, even at relatively high correlations. The highest correlations reduce the gaps but do not eliminate them. However, Panel C shows that average standard deviation is an incomplete characterization of the risks—and risks drive much of the normal return gap.

<table>
<thead>
<tr>
<th>Panel A: Common Distribution ($\mu_x = \mu_y \text{ and } \sigma_x = \sigma_y$)</th>
<th>Panel B: Different Means ($\mu_x \neq \mu_y \text{ and } \sigma_x = \sigma_y$)</th>
<th>Panel C: Different Risks ($\mu_x = \mu_y \text{ and } \sigma_x \neq \sigma_y$)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Expected Return</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\mu$</td>
<td>5%</td>
<td>6%</td>
</tr>
<tr>
<td>$\mu$</td>
<td>5%</td>
<td>6%</td>
</tr>
<tr>
<td>$\mu</td>
<td>\Delta</td>
<td>$</td>
</tr>
<tr>
<td><strong>Standard Deviation</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\sigma_x$</td>
<td>10%</td>
<td>15%</td>
</tr>
<tr>
<td>$\sigma_x$</td>
<td>10%</td>
<td>15%</td>
</tr>
<tr>
<td><strong>Correlation</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.99</td>
<td>1.13%</td>
<td>1.69%</td>
</tr>
<tr>
<td>0.9</td>
<td>3.57%</td>
<td>5.35%</td>
</tr>
<tr>
<td>0.8</td>
<td>5.05%</td>
<td>7.57%</td>
</tr>
<tr>
<td>0.5</td>
<td>7.98%</td>
<td>11.97%</td>
</tr>
<tr>
<td>0</td>
<td>11.28%</td>
<td>16.93%</td>
</tr>
<tr>
<td>-0.2</td>
<td>12.36%</td>
<td>18.54%</td>
</tr>
<tr>
<td>0.99</td>
<td>2.10%</td>
<td>2.39%</td>
</tr>
<tr>
<td>0.9</td>
<td>3.92%</td>
<td>5.59%</td>
</tr>
<tr>
<td>0.8</td>
<td>5.30%</td>
<td>7.74%</td>
</tr>
<tr>
<td>0.5</td>
<td>8.14%</td>
<td>12.07%</td>
</tr>
<tr>
<td>0</td>
<td>11.40%</td>
<td>17.00%</td>
</tr>
<tr>
<td>-0.2</td>
<td>12.46%</td>
<td>18.61%</td>
</tr>
<tr>
<td>0.99</td>
<td>4.91%</td>
<td>5.07%</td>
</tr>
<tr>
<td>0.9</td>
<td>5.87%</td>
<td>7.10%</td>
</tr>
<tr>
<td>0.8</td>
<td>6.79%</td>
<td>8.83%</td>
</tr>
<tr>
<td>0.5</td>
<td>8.99%</td>
<td>12.67%</td>
</tr>
<tr>
<td>0</td>
<td>11.78%</td>
<td>17.26%</td>
</tr>
<tr>
<td>-0.2</td>
<td>12.73%</td>
<td>18.79%</td>
</tr>
</tbody>
</table>
Exhibit 3 graphs data from the first column of the three panels of Exhibit 2. Again, note that meaningful return gaps appear at high correlation levels, even in Panel A. At very low correlation levels, there are very large return gaps and the particular distributional assumptions matter less. At more typical correlation levels for diversifying asset classes in Panel C, say $\rho \in [0.4, 0.8]$, the distributional assumptions matter more—the lines are farther apart. Also note that the middle line, for the different means (using Panel B data from Exhibit 1), converges on the expected difference in means (here, 2%) as the correlations approaches unity ($\rho \to 1$), as should be expected. In sum, any differences in means, standard deviations, and correlations matter greatly when analyzing the diversification benefit of additional asset classes, as measured by normal return gaps.

Calibration to Historical Returns

We retrieved data for three core asset classes—US stocks, international stocks, and US investment-grade bonds—using returns for the corresponding Vanguard “total market” index mutual funds. The common data period begins in 1996 and extends 256 months, or over 21 years.

Panel A of Exhibit 4 shows summary statistics for the monthly return series. Using the mean, standard deviation, and correlation of the three underlying return series, we calculate the normal return gap using Equation (5). As Panel B of Exhibit 4 shows, this forecast value closely approximates the actual average return differential in these three cases. The method presented in this article differs from historical values anywhere from 7 to 21 basis points per month.2

Panel 4

Notes: These graphs plot the three panels of Exhibit 2 as separate lines. The solid line, dotted line with squares, and dashed line with triangles correspond to Panels A, B, and C of Exhibit 2. All data are from the leftmost column of Exhibit 2. Panel B of this exhibit highlights a portion of Panel A of this exhibit to show that meaningful return gaps remain, even with high correlations.

2Statman and Scheid (2005, 2006, 2008) had the important insight that correlations are misleading guides to diversification. They forecasted return gap is $2\sigma \sqrt{(1-\rho)} / 2$ where $\sigma$ clarifies that the standard deviation in their formulation is the average standard deviation of the two asset classes. We concur that risk is a key driver of return gaps. Our article can be construed as a mathematical refinement to their work; further, we extend their work by demonstrating the importance of return gap to a number of portfolio management problems.

In Exhibit 4, we should not use their average standard deviation specification because $\mu$ and $\sigma$ differ for the three assets. If attempted, their results differ from historical values by 35 to 55 basis points per month. Their results miss by 9% to 27%, whereas ours miss by 3% to 6%. In short, the normal return gap of our Equation (5) better calibrates investment reality.

In Appendix A, we demonstrate that their forecasted return gap is proportional to a specialized case of our Equation (5). Specifically, this proportionality holds when the expected returns and standard deviation are identical for the two assets under consideration.
Analyzing Capital Market Assumptions

Institutional investment consultants use capital market assumptions to forecast performance and design portfolios. These measures consist of return, risk, and correlation assumptions for a number of asset classes. Typically, these capital market forecasts cover 10-year to 20-year horizons. Their standard use is in setting a strategic asset allocation mix using mean–variance optimization.

While useful to financial professionals, the implications of these capital market assumptions are very likely nonintuitive to many of their clients. We believe quantifying normal return gaps make the portfolio impact of capital market assumptions substantially more understandable. Normal return gaps provide important context.

For illustrative purposes, we evaluate some J.P. Morgan (JPM) Long-Term Capital Market Return Assumptions. The JPM CMA is a widely available set of asset class risk, return, and correlation assumptions. It covers 45 asset classes and has been updated annually; see Shairp, Werley, and Feser (2014).

JPM forecasts, for example, a correlation of 0.92 between large-cap US and small-cap US stocks, which have an (expected return, standard deviation) distribution pair of (7.60%, 15.50%) and (8.81%, 21.50%), respectively. These similarities and high correlations
might seem to offer little prospect of diversification benefits. Yet, Example 1 of Exhibit 5 shows that the normal return gap between large-cap US and small-cap US stocks is fully 7.60%. This gap is mainly due to the high forecast volatility for the two asset classes. Again, to our knowledge, calculating this difference has previously been impossible based on the existing literature, which only quantifies the return gap for assets having identical risk and return. In other words, these two highly correlated asset classes that should move together with similar mean returns and similar risks will have an average performance differential of 7.60% in any given year.

Counterintuitively to us (and, we suspect, to others), we observe that the normal return gap is the same order of magnitude as the expected returns on the correlated asset classes. This gap is large.

To highlight this deficiency of correlation in illustrating diversification benefits and return gaps, consider two pairs of asset classes with nearly identical correlation assumptions. Example 2 of Exhibit 5 shows that JPM forecasts that US aggregate bonds and US long treasuries will have a 0.76 correlation; similarly, they forecast US large-cap stocks and emerging markets equity will have a 0.77 correlation. In the case of the bonds, the normal return gap is 8.42%, while with the stocks it is 14.44%, as we show. These similar correlations, 0.76 and 0.77, can provide widely divergent diversification benefits—71% different in this example. This gap occurs because of the stock pair’s higher volatility and higher difference in expected mean returns. A univariate focus on correlation can miss the full diversification benefit available when we account for other statistical properties of returns.

Exhibit 6 uses capital market assumptions to show the shortcomings of correlation as a measure of diversification. It shows the return gap for stocks, bonds, and 11 diversifying asset classes, using JPM capital market assumptions. The 11 chosen are typical diversifiers considered by institutional investors and private wealth managers—TIPS, high-yield bonds, EM bonds, small-cap stocks, EM equity, private equity, direct real estate, REITs, and three types of hedge funds. (Exhibit 6 includes the stock–bond correlation and the correlation of those with the 11 others, but not the intra-11 correlations.)

Correlation and the normal return gap are visibly quite different. They are also statistically significantly different. While lower correlations can possibly have

---

**EXHIBIT 5**

Evaluating Capital Market Return Assumptions

<table>
<thead>
<tr>
<th>Asset Class Pairs</th>
<th>Capital Market Assumptions</th>
<th>Calculated Normal Return Gap</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean</td>
<td>Standard Deviation</td>
</tr>
<tr>
<td>Example 1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>US Large Cap Stocks</td>
<td>7.60%</td>
<td>15.50%</td>
</tr>
<tr>
<td>US Small Cap Stocks</td>
<td>8.81%</td>
<td>21.50%</td>
</tr>
<tr>
<td>Example 2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>US Aggregate Bonds</td>
<td>4.08%</td>
<td>4.00%</td>
</tr>
<tr>
<td>US Long-term Treasuries</td>
<td>3.59%</td>
<td>13.25%</td>
</tr>
<tr>
<td>US Large Cap Stocks</td>
<td>7.60%</td>
<td>15.50%</td>
</tr>
<tr>
<td>Emerging Market Stocks</td>
<td>11.77%</td>
<td>26.50%</td>
</tr>
<tr>
<td>Example 3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>US Large Cap Stocks</td>
<td>7.60%</td>
<td>15.50%</td>
</tr>
<tr>
<td>US Aggregate Bonds</td>
<td>4.08%</td>
<td>4.00%</td>
</tr>
<tr>
<td>Example 4</td>
<td></td>
<td></td>
</tr>
<tr>
<td>US Large Cap Value Stocks</td>
<td>7.95%</td>
<td>16.25%</td>
</tr>
<tr>
<td>US Large Cap Growth Stocks</td>
<td>7.39%</td>
<td>15.75%</td>
</tr>
<tr>
<td>Example 5</td>
<td></td>
<td></td>
</tr>
<tr>
<td>US Large Cap Value Stocks</td>
<td>7.95%</td>
<td>16.25%</td>
</tr>
<tr>
<td>US Large Cap Stocks</td>
<td>7.60%</td>
<td>15.50%</td>
</tr>
</tbody>
</table>

Notes: Here, the examples contain pairs of asset classes’ capital market assumptions. Input data are from the JPM capital market assumptions; calculations are the authors’ own. Return gaps can be large, even with high correlations. Return gaps can be quite different, even with similar correlations.
Normal Return Gaps: Dispersion Illuminates Diversification

Fall 2020

8

EXHIBIT 6
Correlation can Mischaracterize Diversification Benefits

Notes: This figure plots correlation against the normal return gap for stocks, bonds, and 11 other diversifying asset classes using third-party capital market assumptions. The relationship between correlation and normal return gaps is weak.

wider return gaps (and thus more diversification benefit), this relationship comes with no guarantees. Indeed, the highest correlation produces a return gap larger than the return gap from one of the most negative correlations. Correlation alone mischaracterizes diversification.

Application to Rebalancing

Return gaps induce rebalancing. If an asset sufficiently outperforms the rest of the portfolio, then it will be sold when it reaches its trigger point, and the proceeds will be reinvested in the rest of the portfolio. And vice versa. Outperformance of asset \( i \) (which is a return gap) can come from either earning more or losing less. In both instances, it is the \( t_i \) trigger point for asset \( i \) that determines whether asset \( i \) needs rebalancing. Assuming the portfolio starts at the target allocations, \( w_i \) and \( w_j \equiv 1 - w_i \), then the following equation holds when the rebalancing trigger point, \( t_i \), is hit:

\[
\frac{w_i(1 + r_j)}{w_i(1 + r_i)} = \frac{w_j - t_i}{w_j + t_i}
\]  

(6)

The left-hand side of the equation shows the impact of relative returns, \( r_i \) and \( r_j \), on the ending values of asset \( i \) and the rest-of-the-portfolio \( j \). Both the numerator and denominator reflect starting weights adjusted for returns over the period. The right-hand side shows the rest-of-the-portfolio \( j \) being underweight by the trigger amount and asset-in-question \( i \) being overweight—the condition that will hold when the rebalancing trigger point is hit. Manipulating Equation (6), we obtain

\[
r_i - r_j = \left[ \frac{-t_i}{w_j(w_i + t_i)} \right](1 + r_i)
\]

(7)

The return gap necessary to induce rebalancing, \( r_i - r_j \), is a function of target weights, rebalancing-band triggers, and the level of returns. For example, a simple 60/40 stock/bond portfolio with ±5% rebalancing bands has a return gap sufficiently large to trigger rebalancing, if, when stocks return 30%, the return gap is 25%.

Note from Example 3 in Exhibit 5 that the normal return gap between (large-cap) stocks and bonds is 13.14% over one-year horizons, using the JPM assumptions. Because Equation (7) gives values larger than this amount for most values of \( r_i \), rebalancing is not triggered over one-year horizons on average. More precisely, evaluating Equation (4) for the variation in return gaps shows a 25% return gap is reached 23% of the time over annual horizons. ³

³Masters (2003) dichotomizes portfolio rebalancing into asset \( i \) and the rest-of-the-portfolio \( j \). We find this split a useful simplifying construct. Our intent is merely to highlight that normal return gaps are related to rebalancing, so we leave further development to others. There are actually two rebalancing-trigger formulas like Equation (6) with the other version reversing the ± signs on the right-hand side to reflect \( j > i \) outperformance instead of \( i > j \) outperformance. Return gaps trigger rebalancing whenever the minimum of these two values is hit.

Investors’ capital market assumptions about return, risk, and correlation could be rescaled to different time periods (instead of merely annual) using Levy and Gunthorpe (1993) scaling techniques to calibrate when return gaps should be expected to trigger rebalancing. These rescaled capital market assumptions could then be used in return gap Equation (5) and compared with Equation (7) to see when rebalancing should normally occur. We note that applying Levy and Gunthorpe (1993) adjustments leads to concave increases in return gaps.

Application to Active Management

The decision to use active management necessarily incurs additional risk. One way of calibrating the additional risk is with tracking error, or the standard deviation
of active returns. While active-management tracking error is perfectly uncorrelated with the benchmark, the active portfolio and the underlying benchmark are highly correlated. Even when tracking error might be low, the volatility of the active portfolio and of the benchmark could both be high, particularly for equity portfolios. Equation (5) tells us this joint-high-volatility situation creates the conditions for a large normal return gap.

Consider a low-tracking-error enhanced index product benchmarked to equities. If the benchmark has 20% volatility and the enhanced index has 3% tracking error, then application of the Pythagorean Theorem tells us the enhanced-index portfolio has 20.2% volatility. Moreover, some right-angle trigonometry tells us that the portfolio and benchmark have 0.989 correlation. This high correlation would not seem to offer much opportunity for a return gap, yet our Equation (5) tells us that the normal return gap is a nontrivial 2.39%. The gap results from high volatilities. (The 2.39% normal return gap is 79.8% of the 3% tracking error. This 79.8% comes from the \( \sqrt{2\pi - 1} \) of our Equation (5).)

Why do tracking error and the normal return gap differ? Tracking error is a standard deviation, where positive and negative differences offset in calculating the average difference. In contrast, normal return gaps focus on the absolute value of the differences and are measuring the unsigned difference. See Exhibit 1.

---

\(^4\) Tracking error is a widely used specification of active risk. See Roll (1992). Here, we assume the active management benchmark is correctly specified and the beta relative to this benchmark is one. (See Kaplan 2016 for mis-specified benchmarks.) This assumption creates two benefits: First, we can ignore the distinction between tracking error and residual risk, which assumes the benchmark-relative \( \beta \neq 1 \). Second, our approach accommodates both CAPM and multi-factor beta approaches.

\(^5\) The Pythagorean Theorem works because the active-management tracking error and benchmark are orthogonal. (We know from regression that \( \alpha \) and \( \beta \) are necessarily orthogonal; in risk space, then, alpha and beta risks—here, \( \sigma_\alpha \) and \( \sigma_\beta \)—are at right angles.) Tracking error, \( \sigma_t \), creates a correlation, \( \rho_{\nu} \), between the portfolio and its benchmark, as follows:

\[
\rho_{\nu} = \cos \left( \arctan \left( \frac{\sigma_t}{\sigma_p} \right) \right) = \cos \left( \frac{\sigma_t}{\sigma_p} \right)
\]

where \( \sigma_t \) is the volatility of the benchmark and \( \sigma_t^2 = \sigma_\alpha^2 + \sigma_\beta^2 \) is the variance of the actively managed portfolio. This trigonometric approach to correlation has a rich literature developed in Litterman (1996), Zerolis (1996), Singer, Terhaar, and Zerolis (1998), Herold and Maurer (2008), and Kaplan (2016).

Even when there is expected alpha, say 1% annually for the previously mentioned low-tracking-error enhanced equity index product, the normal return gap grows to only 2.53%. This means the bulk of the return gap is from high volatilities, not different means. The expected alpha causes only 0.14%, or 5%, of the expected total return gap.

When the tracking error is larger, so is the normal return gap. Consider a concentrated small-cap active fund. If the benchmark has 24% volatility and the fund has 10% tracking error, then the active portfolio has 26% volatility and the fund-benchmark correlation is 0.923. The normal return gap is 7.98%, even with a zero-alpha assumption.

With a typical promised alpha of 2% annually on a small-cap active fund, the normal return gap only grows to 8.14%. The bulk of the normal return gap between active managers and their benchmark is a function of volatility. Here, only a minuscule 2% of the normal return gap (1.97% = (8.14% – 7.98%)/8.14%) is due to expected alpha.

We have demonstrated with both a low-tracking-error enhanced equity index fund and a concentrated small-cap active fund that meaningful return gaps arise. In both cases, the normal return gap is chiefly a function of high volatility. High correlation is unimportant. The anticipated alpha is unimportant. What is important is recognizing we should expect non-trivial return gaps between the fund and the benchmark.

**Dispersion between two active managers.** Another scenario compares two active managers. Even when both are benchmarked to the same index, their performance return gap can be quite large. This is particularly true if the two managers are selected to complement each other. For instance, retirement plan sponsors often employ more than one active manager within an investment style to mitigate risk. Likewise, a number of subadvised mutual funds use a multi-manager approach (with multiple firms investing subportfolios), particularly in capacity-constrained spaces such as small-cap equities; examples include several Vanguard, Vantagepoint, Russell, Northern Trust, Bridge Builder, and Litman Gregory funds.

We move the (onerous) mathematics of the two-manager return gap to an appendix but include the qualitative conclusions here. In a range of typical active management scenarios shown in Appendix A, the normal return gap between two active managers is
the same order of magnitude as the expected return for each—9% to 14% annually in plausible equity scenarios. This difference is quite large.

These sizable between-manager normal return gaps still occur even if both managers have exceptional expected alphas, say $\alpha = 4\%$ annually. This result holds because Equation (5) for the normal return gap uses differences in means, not levels. The two exceptional alphas cancel each other. The important consequence of this between-manager return gap is that the appropriate normal expectation is that one good manager will look like a hero and the other good manager will look like a goat—even if both typically earn exceptional alphas.

That is, the investor should anticipate suffering regret about one manager even when both manager selections were wise because of their positive alpha. Waring, Whitney, Pirone, and Castille (2000) rightly note that investor risk aversion to this sort of return gap is likely higher than their asset allocation risk aversion. Investors unaware of normal return gaps will be unprepared for this eventuality and will likely terminate a worthy manager (as in Goyal and Wahal 2008). Investors prudently seeking complementary managers are particularly susceptible.

Moreover, plan sponsors should never set up a “horse race,” where multiple managers are hired with a plan aforethought to terminate the laggard or laggards. The mathematics of normal return gaps between active managers make it highly probable that large gaps will arise, even between equivalently good managers. As noted, the return gaps chiefly arise because of volatility, not skill. The manager horse race will only incur unnecessary transaction costs.

**Long–short investing.** Long–short investing can be viewed as a more extreme version of active management. (Conversely, long–only investing is merely a constrained version of long–short investing.) Long–short investing can enhance the efficiency of managers’ alpha-generating insights by allowing stronger, unconstrained expression of manager views (see Jacobs and Levy 1996; Grinold and Kahn 2000).

Normal return gaps offer insights on the likely gap between the long and short subportfolios. The normal return gap, Equation (5), of active managers against their benchmark gives the expected return differential of a long–short manager employing generic shorts, such as exchange-traded funds (ETFs) or index funds. Normal return gaps of the long and short subportfolios, properly scaled for leverage (see Roll 2013, p. 17), could provide insights about expected performance levels of the long–short portfolio; this performance could be calibrated against the sometimes-onerous fees associated with long–short investing (see Jennings and Payne 2016). Further, the variation of the normal gap, Equation (4), provides a means of calibrating the significance of any observed outperformance.

**Application to Style Tilts, Sector Bets, and Socially Responsible Investing**

Here, we consider common modifications to unconstrained core equity investing. Many investors divide their equity allocations along value–growth dimensions. Others favor particular industries. Some investors modify their investment criteria on social-values-based factors.

**Style tilts.** Fama and French (1993) show that the market has historically rewarded a value–tilt decision—there is a value premium. However, there is risk in over-weighting value stocks because of style risk—the risk that a favored investment style underperforms.

Normal return gaps help calibrate this style risk performance divergence. Example 4 of Exhibit 5 includes the JPM capital market assumptions for US large-cap value and US large-cap growth stocks. It shows that the normal return gap between value and growth investment styles is 4.81%, despite the return advantage JPM assumes for value stocks. This is much larger (an order of magnitude) than the 0.56% difference in their expected returns.

Even mild style tilts induce risk. Example 5 of Exhibit 5 shows the normal return gap for US large-cap stocks (perhaps the core investment) and US large-cap value stocks (the style tilt) is 3.64%, again using JPM capital market assumptions. So, a reallocation of 30% of the equity portfolio from core large-cap stocks to value stocks produces a normal return gap of 1.09% between the tilted and un-tilted portfolios. When contemplating the prospective return advantage of a factor tilt, some

---

A reallocation of 30% of the equity portfolio from core large-cap stocks to value stocks is equivalent to having a 65/35 value/growth split. (The 70% remaining in large-cap core stocks is split 50/50 value/growth, or 35/35 if summing to 70; adding 30/0 in the pure value portfolio yields the 65/35 value/growth split.) While many investors would not ordinarily consider a 65/35 value/growth split as particularly extreme because tilting portfolios toward sensible economic smart-beta factors is reasonable, our normal return gap math highlights a meaningful 1.09% return differential.
investors may find that much potential underperformance too daunting. 

**Sector bets.** Similarly, normal return gaps can help evaluate the risk of sector bets. For example, Jennings (2012) evaluates energy stocks as a separate portfolio allocation. Over the 1926–2009 period, he found energy stocks had an average return of 13.7% with 24.1% volatility and a correlation of 0.77 with the broad market, which had an average return of 11.5% with 21.0% volatility. These values indicate a normal return gap of 12.55%. Despite the relatively high correlation of energy stocks with the broad market, large return deviations are typical. A concentrated energy strategy is quite risky, despite the historical return advantage. (Recent events have borne this out.)

**Socially responsible investing.** As with investment style tilts, applying SRI (socially responsible investing) criteria induces risk (see Adler and Kritzman 2008). To provide a calibration of this social-screening risk, we used two Vanguard funds, the Vanguard 500 Index (representing a generic core equity approach) and the Vanguard FTSE Social Index (representing a socially screened core equity method) with our normal return gap approach. We examined historical periods ranging from 3 to 10 years to garner the mean, risk, and correlation values needed to calculate the normal return gap using Equation (5). It is worth observing that the “winner” of the two funds changed depending on the period selected, a point which underscores our thinking of normal return gaps as a risk measure. To reflect this uncertainty, we examined normal return gaps with both historical return differentials and with no differential. The normal return gap between the socially screened portfolio and the unscreened portfolio ranges from 1.35% to 3.09% a year.

When contemplating the virtues of socially responsible investing, some investors may find this degree of potential underperformance too daunting. For others, any adverse return gap is simply the price of being good.

**Abnormal Return Gap Distributions**

Because we make a distributional assumption, our approach offers insights on whether a particular realized return gap is typical or atypical. Substituting values from Equations (1) and (5) into Equation (4) gives the variation associated with the normal return gap estimate.

For example, Exhibit 4 with the historical asset class data shows that the standard deviation of the normal return gap for US and international stocks is 1.58%. Since the folded normal distribution has a specific variance, we can calibrate that return gaps greater than 5.17% (the 2.09% normal return gap + 1.96 × 1.58%) should be rare. Indeed, the historical data reveal they occur approximately 5% of the time, as expected. The distributional properties of our Equation (4) accurately parameterize the real world.

A private wealth manager or chief investment officer could use these abnormal return gap distributions to monitor their portfolio. Observed return gaps could be calibrated against the distribution to highlight how common or rare they were. A client’s *ex post* discomfort with a statistically common return gap suggests reduced risk-taking in the future whereas *ex post* discomfort with a statistically rare return gap might warrant counseling a client to have patience. Repeatedly exceeding the return gaps implied by the distribution suggests the investment analyst’s distributional assumptions were unsound.

**Application to Taxation and Wash Sales**

The “wash sale” is a construct in the US tax code on investing whereby capital losses are disallowed or delayed if an asset is sold and replaced with a “substantially identical” security. Similar rules apply in other jurisdictions. While there are a handful of Revenue Rulings that clarify “substantially identical,” the term is somewhat elusory and interpretable.

A number of commentators suggest statistical interpretations for “substantially identical” that rely on correlation coefficients or regression coefficients of determination (\(\rho\) or \(R^2\)). For example, a high correlation between the investment sold for a loss and its replacement suggests a wash sale.

Normal return gap mathematics contribute to this debate. As we have highlighted, meaningful performance differentials can arise from seemingly similar investments. Other statistical measures do not fully capture this expected difference in returns. While a 0.99 correlation suggests being “substantially identical,” a 2.4% expected annual return differential unambiguously does not. Recall that these were the correlation and return gap values associated with an enhanced index fund with 3% tracking error to a 20% volatility benchmark, described in the previous subsection Application to Active Management.
A default expectation of a 2.4% gap in returns is a meaningful difference. Selling an index fund for an enhanced index fund tracking the same benchmark sounds like a close substitute, but the evidence of a 2.4% expected return gap suggests otherwise. Normal return gaps highlight that this enhanced index fund does not meet the layperson definition of being “substantially identical” to its underlying benchmark, despite the 0.99 correlation.

**CONCLUSION AND INVESTMENT IMPLICATIONS**

Despite ever more sophisticated risk management and measurement, investment professionals have generally overlooked a simple but powerful measure of relative risk and portfolio diversification—the normal return gap. We developed a generalized specification of the expected dispersion between two investments based on the folded normal distribution. We showed that our approach improves upon prior techniques. We then demonstrated return gaps’ applicability to a range of real-world investing contexts.

Key practical observations about normal return gaps include the following:

- The difference in the expected return of the two asset classes affects the normal return gap. In contrast with prior published work, it does matter.
  - Bigger differences produce bigger normal return gaps.
  - The difference matters, not the level of the two expected returns.
  - The difference matters more at high correlations.
  - The difference matters more at low volatilities.
  - A proportional change in the difference induces a higher expected return gap than a similarly proportional change in volatilities. This effect attenuates at higher correlations.
- Higher volatilities produce bigger normal return gaps.
- Higher correlations ameliorate the impact of higher volatilities, but generally do not fully offset them.
- While counterintuitive, even highly correlated investments can produce meaningful diversification.
- Conversely, low correlations can produce small return gaps and, therefore, minimal diversification.
- Similar correlations can provide widely divergent diversification benefits, depending upon the associated volatility.
- Normal return gaps are often the same order of magnitude as the expected returns of the underlying assets.
- Normal return gaps themselves have a distribution, which is useful in performance analysis for calibrating the significance of any observed return gap, as well as the appropriate investor response.
- After-the-fact discomfort with a statistically common observed return gap suggests reduced risk-taking in the future.
- After-the-fact discomfort with a statistically unusual observed return gap suggests patience, forbearance, and perseverance.

- Return gap analysis helps calibrate rebalancing.
- Return gap analysis informs discussion of tax “wash sales.”
- Return gap analysis helps with manager structure and performance evaluation.
- Most of the performance gap between active managers is a function of volatility and correlation, not the difference in the managers’ expected alpha.
- Investors unaware of normal return gaps risk terminating worthy managers. Further, establishing competitive “horse races” between managers is unwise.

We believe many investors will find that normal return gaps are large relative to their expectations. In the various scenarios we consider—asset allocation, style tilts, sector bets, and manager selection—we believe a realistic preview of the expected gaps will make investment decisions more robust and sustainable. Forewarned is forearmed.

The normal return gap is an important addition to the risk budgeting toolkit. Like other analytical tools, it is only as useful as its inputs and assumptions. Our distributional assumption is an obvious limitation. While the normality premise is imprecise, it is also useful. Likewise, the presumption that correlation properly captures
co-movement may also be inexact and unstable. However, the fit to actual data in Exhibit 4 substantially counters these objections. Other researchers might consider extensions to this work that address alternate distributions.\(^7\) Another sensible extension would be to investigate how expected normal return gaps can be used in optimal portfolio construction, either as a substitute or complement to correlation.\(^8\)

In sum, normal return gaps are a useful and important tool for private wealth managers and chief investment officers. Return gaps better parameterize diversification than correlation and provide insights into diversification benefits beyond simply using correlation. We have demonstrated how properly calibrating the expected dispersion in returns between two investments helps with decisions about asset allocation, manager selection, manager structure, style tilts, rebalancing, and performance measurement.

**APPENDIX A**

**MATHEMATICAL APPENDIX**

Relation to Prior Work

Our history with Statman and Scheid (2005, 2006, 2008) is worth recounting. We found the 2008 article, the first of the three we read, incredibly compelling and of obvious practical importance to investors. We were also surprised that their findings were new to the literature. While struck by the elegance of their main finding (recapped in our Equation (A2)), we could not grasp the absence of a term with the difference in expected returns of the two assets. Statman and Scheid (2005, 2006, 2008), but we can show their return gap \(\sigma^{-2}(1-\rho)/2\) was a function of \(\pi\). (0.798 is \(\sqrt{2\pi^{-1}}\)). This led to the realization that normality was relevant, and the paper followed.

Our results differ from those presented in Statman and Scheid (2005, 2006, 2008), but we can show their return gap \(\sigma^{-2}(1-\rho)/2\) is related to a special case of our Equation (5) for the normal return gap. When the mean difference in expected returns for the two assets is zero, \(E(\mu_A - \mu_B) = 0\), then our normal return gap equation reduces:

\[
\mu_{g\,1} = \sqrt{2\pi^{-1}} \times (\sigma) e^{\frac{(\mu - \mu)}{(2\sigma)}} + (\mu) \left[ 1 - 2\Phi\left(-\frac{\mu}{\sigma}\right) \right]
\]

\[
\mu_{g\,2} = \sqrt{2\pi^{-1}} \times (\sigma) e^{\frac{(\mu - \mu)}{(2\sigma)}} + (\mu) \left[ 1 - 2\Phi\left(-\frac{\mu}{\sigma}\right) \right]
\]

\[
\mu_{g\,3} = \sqrt{2\pi^{-1}} \times (\sigma) e^{\frac{(\mu - \mu)}{(2\sigma)}} + (\mu) \left[ 1 - 2\Phi\left(-\frac{\mu}{\sigma}\right) \right]
\]

\(\sigma\) denotes the grand average \(\pm\) the two assets concerned. We respect them as scholars and followed their mathematical derivation; nonetheless, we did not find the term’s absence particularly intuitive. To persuade ourselves, we undertook a large-scale (one million pairs) Monte Carlo simulation with correlated random variables, which led to our discoveries. Our results confirmed their finding that correlation is a faulty diversification metric and that standard deviations matter as well, but they highlighted our intuition that the difference in expected returns of the two assets mattered as well. We were aided substantially when one coauthor’s high school son observed that the 0.798 ratio between the simulation and their Equation (A2) was a function of \(\pi\). (0.798 is \(\sqrt{2\pi^{-1}}\)). This led to the realization that normality was relevant, and the paper followed.

Algebraic manipulation allows showing that this Equation (A1) is proportional to the Statman and Scheid return gap. First, note that,

\[
\rho + 1 = 2 - (1 - \rho) = \left[ 2 \left[ 1 - \left(\frac{1 - \rho}{2}\right) \right] \right]
\]

We use the following equivalency. Now, manipulate the second radical from Equation (A1):

\[
\sqrt{\sigma_{xy}^2 + \sigma_{yx}^2 - 2\sigma_{xy}\rho} = \sqrt{\sigma_{xy}^2 + \sigma_{yx}^2 + \sigma_{yx}^2 - 2\sigma_{xy}\rho - 2\sigma_{yx}\rho}
\]

\[
\sqrt{\sigma_{xy}^2 + \sigma_{yx}^2 - 2\sigma_{xy}\rho} = \sqrt{(\sigma_{xy} + \sigma_{yx})^2 - 2\sigma_{xy}\rho - 2\sigma_{yx}\rho + (\rho + 1)}
\]

\[
\sqrt{\sigma_{xy}^2 + \sigma_{yx}^2 - 2\sigma_{xy}\rho} = \sqrt{(\sigma_{xy} + \sigma_{yx})^2 - 2\sigma_{xy}\rho - 2\sigma_{yx}\rho + 2\left[ 1 - (1 - \rho) \right]} \]

\[
\sqrt{\sigma_{xy}^2 + \sigma_{yx}^2 - 2\sigma_{xy}\rho} = \sqrt{(\sigma_{xy} + \sigma_{yx})^2 - 4\sigma_{xy}\rho + 2\sigma_{yx}\rho + 2\left[ 1 - (1 - \rho) \right]} \]

Now, under the assumption that \(\sigma_{xy} = \sigma_{yx} = \sigma\), we can reduce this equation to match the Statman and Scheid (2005, 2006, 2008) return gap:

\[
\sqrt{\sigma^2 + \sigma^2 - 2\sigma^2\rho} = \sqrt{(\sigma + \sigma)^2 - 2\sigma^2\rho}
\]
\[ \sqrt{\sigma_i^2 + \sigma_j^2 - 2\sigma_i\sigma_j\rho} = \sqrt{(\sigma_i^2 + \sigma_j^2) - 2\sigma_i\sigma_j\left(1 - \frac{(1-\rho)}{2}\right)} \]
\[ = 2\sigma \sqrt{\frac{(1-\rho)}{2}} \]  

Equation (A2) is the Statman and Scheid (2005, 2006, 2008) return gap. So, in the case of identical distributions, where both \( \mu_i = \mu_j \) and \( \sigma_i = \sigma_j \), our normal return gap Equation (5) is proportional to theirs:

\[ \mu_{\pi i} = \sqrt{2\pi^{-1} \left(2\sigma \sqrt{\frac{(1-\rho)}{2}}\right)} \]

by substitution of Equation (A2) into (A1).

That is, our normal return gap is \( \sqrt{2\pi^{-1}} \) times the Statman and Scheid return gap in the case of identical distributions. Their result is enhanced by our \( \sqrt{2\pi^{-1}} \) term. Exhibit 2, however, demonstrates that means and variances of the two asset return series do matter and should be incorporated. Overall, our Equation (5) result is more general.

Two Active Managers

In the body of the paper (“Application to Active Management” in the second section and particularly note 4), we discuss “triangle thinking” for a single active manager. Here, we expand on our application of this trigonometric approach to offer more intuition but do so without re-creating the literature; interested readers are referred to the references cited in note 4, particularly Kaplan (2016), which represents the fullest evolution of this methodology. If a single active manager requires “triangle thinking,” then two active managers require “tetrahedron thinking” where the four sides of the tetrahedron reflect triangle thinking about one manager in isolation, triangle thinking about the other manager, a base or end reflecting the two managers’ relative tracking error (and their correlation), and a third “long side” reflecting the two managers’ relative total risk. Thus, the four triangular surfaces of the tetrahedron and their three sides are as follows:

1. Manager 1 total-risk triangle: \( \sigma_{s1}, \sigma_{p1}, \rho, \)
2. Manager 2 total-risk triangle: \( \sigma_{s2}, \sigma_{p2}, \)
3. The smaller endpiece with the two managers’ relative tracking errors: \( \sigma_{i}, \sigma_{j}, \)
4. A triangle relating the two managers’ total risk: \( \sigma_{p1} \), \( \sigma_{p2} \), and side \( c \)

where \( \sigma_i \) is the benchmark risk for portfolio \( i \), \( \sigma_{pi} \) is the portfolio risk for portfolio \( i \), and \( \bar{\sigma}_i \) is manager \( i \)’s tracking error. This is shown in three dimensions of Exhibit A1. Exhibit A2 shows the “net” of the tetrahedron, where the different sides are “unfolded.” We will use the fact that two triangles—the endpiece and the triangle relating the two managers’ total risk—both have side \( c \).

Consider the base or end of the tetrahedron reflecting the two managers’ relative tracking errors, \( \bar{\sigma}_1 \) and \( \bar{\sigma}_2 \). Let the two managers be selected to complement each other with a correlation of their tracking errors, \( \rho_{12} \). By standard portfolio math, we know the risk of the manager combination:

\[ \sigma_i^2 = h_i^2 \bar{\sigma}_1^2 + h_i^2 \bar{\sigma}_2^2 + 2h_ih_j \bar{\sigma}_1 \bar{\sigma}_2 \rho_{12} \]
\[ \sigma_j^2 = h_j^2 \bar{\sigma}_1^2 + h_j^2 \bar{\sigma}_2^2 + 2h_ih_j \bar{\sigma}_1 \bar{\sigma}_2 [-\rho_{12}] \]

where \( h \) is the portfolio weight invested with manager \( i \).

Then, the third side of the base or end of the tetrahedron has length \( c \) by the law of cosines:

\[ c^2 = h_i^2 \bar{\sigma}_1^2 + h_j^2 \bar{\sigma}_2^2 - 2h_ih_j \bar{\sigma}_1 \bar{\sigma}_2 \cos(\angle c) \]  

(A3)

Make the plausible simplifying assumption that the two managers we are interested in comparing share a common benchmark, so \( \bar{\sigma}_1 = \bar{\sigma}_{s1} = \bar{\sigma}_{s2} \). Also note that, for calculating

**Exhibit A1**

Visualizing Portfolio Risks for Two Managers

![Two-Manager Risk Tetrahedron](image_url)

Notes: The various risks are shown on the tetrahedron. Manager 1’s total risk triangle is in the x-z plane. The dashed base reflects benchmark risk, the shorter vertical reflects active risk, and the diagonal reflects total risk. Manager 2’s corresponding total risk triangle is in the x-y plane. Their relative tracking errors are the small triangle to the right (here, with tracking-error correlation \( \rho_{12} = 0 \), making a right triangle) and the cantilever triangle relates the two managers’ total risk. The total correlation between the two managers’ total risk \( \rho_{12} \) is given by the cosine of the angle marked by the two solid lines coming from the origin.
E X H I B I T  A  2
Net for Visualizing Portfolio Risks for Two Managers

\[
\begin{align*}
\cos(\angle \alpha) &= -\rho_{12} = \frac{Q^2 - R^2 - S^2}{2RS} \\
&= \frac{(\sigma_1 + \sigma_2 - 2\sigma_1 \sigma_2 \rho_{12}) - (\sigma_{p1}^2) - (\sigma_{p2}^2)}{2\sqrt{\sigma_1^2 + \sigma_2^2} \sqrt{\sigma_{p1}^2 + \sigma_{p2}^2}} \\
&= \frac{\sigma_1 \sigma_2 \rho_{12} + \sigma_{p1}^2}{\sqrt{\sigma_1^2 + \sigma_2^2} \sqrt{\sigma_{p1}^2 + \sigma_{p2}^2}}
\end{align*}
\]

In the simplifying case, where the managers have the same tracking error, \( \sigma = \sigma_i \forall i \), we obtain Equation (A4):

\[
\rho_{12} = \frac{\sigma_{p1}^2 + \sigma_{p2}^2 \rho_{12}}{\sigma_1^2 + \sigma_2^2} \quad (A4)
\]

Note that Equation (A4) gives the correlation between the two managers’ total returns, \( \rho_{12} \), as a function of the correlation between the two managers’ active returns, \( \rho_{12} \). The denominator of Equation (A4) is simply the risk of the active portfolio. The numerator is the sum of the benchmark risk and the covariance of the two active returns. Also notice the manager correlation \( \rho_{12} \) becomes more influential on \( \rho_{12} \) the larger the tracking error variance \( \sigma_i^2 \) is relative to the benchmark risk \( \sigma_i^2 \).

Consider a pair of concentrated small-cap managers, each with 10% tracking error to a benchmark with 24% volatility. (In the body of the article, we showed these managers individually have a normal return gap of 7.98% relative to their benchmark.) We can then use Equation (A4) in Equation (5) to find the normal return gap between the two managers. In the extreme case when the two managers’ tracking errors are completely unrelated, with uncorrelated active risk, \( \rho_{12} = 0 \), Equation (A4) gives a total portfolio correlation \( \rho_{12} = 0.85 \). Here, the normal return gap between the two managers is 11.28%. The expected gap between the two managers is larger than the expected gap between their individual performance and the benchmark, which makes sense because there is an additional source of variation—the second manager’s active management.

If the two managers are selected to be complementary, the return gap will be larger. If, say the active risk is negatively correlated, \( \rho_{12} = -0.5 \), Equation (A4) gives a total portfolio correlation \( \rho_{12} = 0.78 \) and the normal return gap between the managers is 13.82%.

More typically, two managers’ alphas will have positive correlation.9 If the two managers’ tracking errors are somewhat correlated, say \( \rho_{12} = 0.35 \), then \( \rho_{12} = 0.90 \). In this case, the normal return gap is 9.10%.

Return gaps are particularly applicable to the two-manager case. Investors have two managers because they believe them both skilled. An anticipated alpha of 1% to 2% for each is vastly different in scale than the expected normal return gaps of 9% to 14% described in the preceding scenarios.

Our two-manager analysis takes the perspective of the chief investment officer or private wealth manager deciding manager structure and evaluating each manager’s relative performance. Someone buying a single multi-manager fund or pool could evaluate the pool as if it were a “single manager”;

\[9\text{See Tobler-Oswald (2008) for how the correlation of two managers' alphas, } \rho_{12}, \text{ might be negative. In contrast, see Litterman (2003, p. 179), for statistics showing the average pairwise correlation between managers is positive. In practice, alphas between managers in a particular investment arena are higher than one might expect. Survivorship increases correlations; underperformers die out, leaving more–positively correlated successful survivors. Investment categories may be broader than underlying factors or clusters of manager approaches, creating correlated clusters within the category. For a particular plan sponsor, a set of investment beliefs, manager selection policies, and/or behavioral finance biases may induce further correlation in their selected active investment managers.} \]
for them the chief advantage would be a risk reduction from uncorrelated ($\rho_{12} < 1$) managers.

We can then use the equations in this Appendix to find the normal return gap between the two managers. We confirmed this relationship with a large-scale (one million pairs) Monte Carlo simulation with correlated random variables.

**APPENDIX B**

**EXCEL TOOL**

We are aware that the normal return gap in Equation (5) may appear daunting. Our Excel software tool is available at www.williamjennings.com. It consists of a user-defined function in Excel that implements Equation (5) with five simple inputs. The syntax is as follows:

$$\text{JOMPToolReturnGap}(	ext{Correlation}, \text{Return}_1, \text{Return}_2, \text{Risk}_1, \text{Risk}_2)$$

Only five inputs are required to use this Excel function: the risk and return of each of the two asset classes being evaluated and their correlation. Exhibit 1 was created with this VBA-based user-defined Excel function. Additionally, we created a JOMPToolReturnGapRisk function, with the same five inputs in the same order, to implement variance Equation (4), measuring risk.

**ACKNOWLEDGMENTS**

The opinions included are those of the authors and not necessarily those of the US Air Force Academy, the US Air Force, or any other federal agency. We previewed the normal return gap Equation (5) in Jennings, O’Malley, and Payne (2016), which cited an earlier version of this paper, but that abbreviated presentation does not consider investment applications like manager selection, performance evaluation, SRI, manager combinations, wash sales, or rebalancing and does not include the mathematical derivation and relation to prior work. We appreciate the helpful insights of the editor, an anonymous reviewer, Steve Fraser, Winfried Hallerbach, Graham Jennings, Jesse Pietz, Kevin Terhaar, Moustafa Abu El Fadl, participants at the Eastern Finance Association and Academy of Financial Services conferences, and especially Barton Waring.

**REFERENCES**


ABSTRACT: Perhaps the most universally accepted precept of prudent investing is to diversify, yet this precept grossly oversimplifies the challenge of portfolio construction. Correlations, as typically measured over the full sample of returns, often belie an asset’s diversification properties in market environments when diversification is most needed. Moreover, upside diversification is undesirable. The authors first describe the mathematics of conditional correlations assuming returns.


To order reprints of this article, please contact David Rowe at d.rowe@pageantmedia.com or 646-891-2157.

**ADDITIONAL READING**

**The Myth of Diversification**

Davide B Chua, Mark Kritzman, and Sébastien Page

The Journal of Portfolio Management

https://jpm.pm-research.com/content/36/1/26

ABSTRACT: Perhaps the most universally accepted precept of prudent investing is to diversify, yet this precept grossly oversimplifies the challenge of portfolio construction. Correlations, as typically measured over the full sample of returns, often belie an asset’s diversification properties in market environments when diversification is most needed. Moreover, upside diversification is undesirable. The authors first describe the mathematics of conditional correlations assuming returns.
are normally distributed. Then they present empirical results across a wide variety of assets, which reveal that, unlike the theoretical conditional correlations, empirical correlations are significantly asymmetric. Finally, the authors show that a portfolio construction technique called full-scale optimization produces portfolios in which the component assets exhibit relatively lower correlations on the downside and higher correlations on the upside than mean-variance optimization portfolios.

Optimizing Manager Structure and Budgeting Manager Risk
M. Barton Waring, Duane Whitney, John Pirone, and Charles Castille
The Journal of Portfolio Management
https://jpm.pm-research.com/content/26/3/90

ABSTRACT: Selecting active managers and structuring their use has historically been more art than science. The authors explain why a manager structure can and should be approached as a problem of portfolio construction, a view that brings a whole set of tools to bear on the problem. This approach suggests optimizing to maximize expected active returns while controlling active risks. For a given risk budget, there is a single combination of the candidate managers who will best maximize the risk–adjusted expected active return of the portfolio. An optimized approach will also solve the active–passive or core–satellite allocation problem, a perennial area of indecision. The passive or core managers will naturally be held by the optimizer in inverse proportion to the active risk budget. This approach can be easily modified to process off–benchmark managers and to incorporate an optimal completion solution as part of the overall structure. This is a practical and robust manager structure solution appropriate to any investor using multiple managers or interested in the optimal active–passive mix.