Do Risk Factors Eat Alphas?

JYH-HUEI LEE AND DAN STEFEK

JYH-HUEI LEE
is a vice president in Research at MSCI Barra in Berkeley, CA.
jyh-huei.lee@mscibarra.com

DAN STEFEK
is a managing director in Research at MSCI Barra in Berkeley, CA.
dan.stefek@mscibarra.com

Portfolio managers often use factor models to forecast risk and exceptional return, or alpha. Many use risk models based on one set of factors and alpha models based on another, overlapping set of factors. Risk factors are selected to explain portfolio volatility and alpha factors are chosen to forecast outperformance.

Portfolio optimization requires forecasts of both risk and alpha. The practice of using different models for risk and return in portfolio optimization, although widespread, has raised some concerns. Portfolio managers worry that discrepancies between risk and alpha factors may create unintended biases in optimized portfolios and has led some to wonder if it would be better to use a risk model that is more aligned with the alpha factors.

Little research has been published on this topic. Some researchers and practitioners have asserted that it is important to include the alpha factors in the risk model used in optimization.[1] However, not much has been written on the consequences of ignoring that prescription.

We analyze the ramifications of using different factor models of risk and alpha in portfolio optimization. Our results show the following:

1. Using different models for risk and alpha can lead to unwanted portfolio exposures and may hinder performance.
2. Aligning risk factors with alpha factors may improve the information ratio of optimized portfolios, even if doing so lowers the overall accuracy of risk forecasts.
3. Modifying a risk model may help remedy the problems previously described.

We begin by briefly reviewing risk and alpha models. Next, we demonstrate how differences between alpha and risk factors may lead to inadvertent bets in optimized portfolios and we rigorously analyze the root of the problem. Then using a simple model, we show how employing a consistent set of alpha and risk factors may improve the quality of the solution. Finally, we describe four approaches to surmounting these difficulties and explore their effectiveness on familiar optimization problems.

RISK AND ALPHA MODELS

Often a distinction is made between factor models used to forecast risk and those used to forecast return, or alpha. In theory, these models are fundamentally similar. Both are predicated on the belief that security returns are driven by pervasive influences in the market and that these factors account for both return and risk. There is no reason, in principle, that one factor model could not account for both risk and alpha.

In practice, portfolio managers often use different models for risk and alpha. Risk models typically include a comprehensive set of broad factors, such as style and industry, which explains the volatility and cross-sectional dispersion of security returns. Alpha models decompose return into benchmark and exceptional return, often using broad factors to account for the benchmark return as well as another set of factors—for example, momentum, earnings quality, and sentiment—which are designed to capture exceptional return.

Naturally, significant overlap may exist between risk and alpha factors. For example, the Barra U.S. equity risk model includes factors for momentum, earnings yield, value, and growth, which are also components of many managers' alphas. It is important to note, however, that the precise definitions and measures of these factors may differ materially between alpha and risk models.

Both risk and alpha models attribute return as follows:

$$r = X_R f_R + u_R \tag{1}$$

$$r = X_A f_A + u_A \tag{2}$$

where r is a vector of excess returns, $X_R(X_A)$ is a matrix representing the asset exposures to the risk (alpha) factors, $f_R(f_A)$ are the returns to the risk (alpha) factors, and $u_R(u_A)$ are the idiosyncratic, or specific, returns not explained by the factors.

Portfolio optimization requires an asset covariance matrix and a set of alphas. From Equation (1) we can write the covariance matrix as

$$\Sigma_R = X_R F_R X_R' + \Delta_R \tag{3}$$

where F_R is the covariance of the risk factors and Δ_R is the (diagonal) covariance matrix of the specific returns.

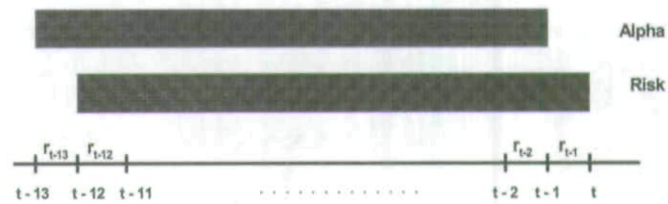
As is customary, the alphas are formed as a weighted combination of the alpha factor exposures,

$$\alpha = X_A w \tag{4}$$

WHAT COULD GO WRONG?

We now illustrate a problem that can arise when a manager uses different models of risk and alpha.

Consider the case of a portfolio manager who bets on 12-month price momentum, defined as the sum of the last 12 months' returns *lagged* one month from the date of usage. More precisely, the manager's alpha at the beginning of month t is $\alpha_t = r_{t-2} + r_{t-3} + \ldots r_{t-13}$, where r_t is the



return over month t. The risk model has a momentum factor that differs slightly from the alpha. Exposure to momentum is defined simply as the sum of the last 12 months' returns—without a lag. This is depicted as:

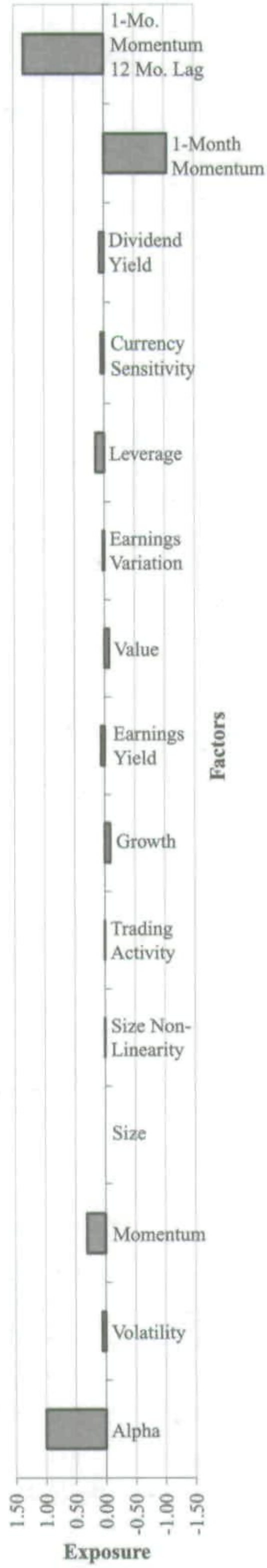
What happens when the manager optimizes the portfolio? To answer this, we perform an active, unconstrained optimization with the Barra U.S. equity risk model, modifying its momentum factor to match the description just given. The S&P 500 Index serves as both the investment universe and the benchmark.

Panel A of Exhibit 1 shows the optimal portfolio's active exposures to the risk model factors. It also shows the active exposure to 1-month momentum, defined as the return over the previous month, r_{t-1}, and to 1-month momentum lagged by 12 months, r_{t-13}. These measures reflect how well stocks did in the preceding month and 13 months ago. The alpha and all exposures are normalized (i.e., they have a mean of zero and a standard deviation of one) across the U.S. equity model's estimation universe.

The optimal solution takes a unit bet on the alpha.[2] Surprisingly, the portfolio's exposure to the risk model's momentum—which differs in composition by only two months—is just 30% of that bet on alpha. Even more surprising is that the portfolio takes a very large bet against stocks that outperformed in the previous month and a very large bet on stocks that outperformed 13 months ago. The chances are high that this is not what the portfolio manager has in mind.

Why does this happen? The root of the problem is the difference in how the alpha and risk models define momentum. The alpha model includes the return from

EXHIBIT 1A
Active Exposures



13 months ago, but the risk model does not. Thus, the optimizer sees return but no factor risk in the Month 13 exposure and loads up. The risk model, however, includes the return from one month ago, while the alpha model does not. In this case, the optimizer sees risk but no return, and thus underweights 1-month momentum.[3]

When the risk and alpha models agree on the definition of momentum, this problem disappears, as shown in Panel B of Exhibit 1. The results of the same optimization, using a risk model whose momentum matches the alpha, are illustrated in Panel B. To achieve this consistency, we simply make the momentum exposures in the risk model match the alphas, and leave the rest of the model alone.

A CLOSER LOOK

In the previous example, the optimizer exploited an inconsistency between risk and alpha factors, which resulted in inadvertent and unwanted bets. In this section, we provide a more general analysis of this problem.

To understand the interplay between risk and alpha factors in optimization, it is useful to decompose the manager's alpha into two parts: one part that is spanned by the risk exposures, α_R, and another that is orthogonal to them, $\alpha_{R_\perp}$. We can accomplish this by regressing the alphas against the risk exposures. The spanned alpha is the fit from the regression and the orthogonal alpha is the residual. This can be written as

$$\alpha = \underbrace{X_R (X_R' X_R)^{-1} X_R' \alpha}_{\alpha_R} + \underbrace{\left(I - X_R (X_R' X_R)^{-1} X_R'\right)\alpha}_{\alpha_{R_\perp}} \qquad (5)$$

A key point is that these components of alpha are viewed differently by the risk model. The spanned alpha is captured by the risk factors. A tilt in its direction incurs factor risk. In contrast, the orthogonal alpha is outside the risk factors, since $X_R' \alpha_{R_\perp} = 0$. Tilting the portfolio in this direction incurs no factor risk.

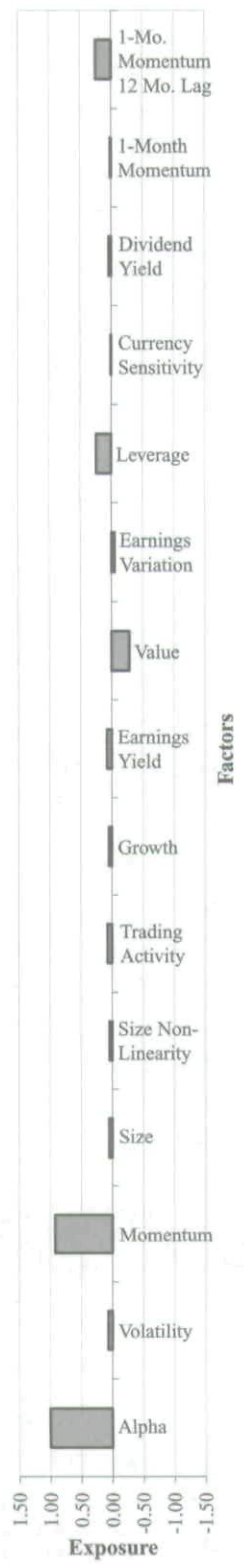


EXHIBIT 1B
Active Exposures after Replacing the Momentum Exposure

Now, suppose a manager solves the unconstrained, active optimization problem,

$$\text{Maximize } \alpha' h - \frac{\lambda}{2} h' \Sigma_R h \tag{6}$$

where h is the vector of active weights and λ is the risk aversion parameter.

The optimal portfolio is $h^* = \frac{1}{\lambda}\Sigma_R^{-1}\alpha$. We can rewrite this to highlight the role of the risk factors, as follows:

$$h^* = \frac{1}{\lambda\sigma_s^2}\cdot\alpha_{R_\perp} + \frac{1}{\lambda\sigma_s^2}\cdot\left(\text{I-}X_R(X_R' X_R + \sigma_s^2 F_R^{-1})^{-1} X_R'\right)\alpha_R \tag{7}$$

For simplicity, we assume that the risk model forecasts the same specific risk for all assets, σ_s.

The optimal solution is the sum of two terms. The first term is simply the orthogonal alpha, scaled to adjust for specific risk. The second term is the contribution of the alpha spanned by the risk exposures. This component of alpha is not represented as directly in the optimal solution. It is both scaled for specific risk and adjusted—twisted and shrunk—to mitigate the common factor risk that it bears. In a sense, the optimizer favors $\alpha_{R_\perp}$ over α_R!

To see this more clearly, assume there is only a single factor in a risk model with volatility σ_{f_R}. If we optimize over a universe of n assets, the solution is[4]

$$h^* = \frac{1}{\lambda\sigma_s^2}\cdot\alpha_{R_\perp} + \frac{1}{\lambda\sigma_s^2}\left(\frac{\sigma_s^2}{n\sigma_{f_R}^2 + \sigma_s^2}\right)\cdot\alpha_R \tag{8}$$

The optimal portfolio is simply a weighted sum of $\alpha_{R_\perp}$ and α_R. We see that $\alpha_{R_\perp}$ always has a greater weight than α_R. Furthermore, α_R's relative weight shrinks even more as the factor volatility increases and the number of assets increases.

By accentuating the component of alpha that is not captured in the risk model, an optimizer may produce unwanted portfolio exposures. Returning to our example of the momentum manager, much of the manager's alpha—the exposure to lagged momentum—is captured by the risk model, so α_R resembles α. In contrast, $\alpha_{R_\perp}$

reflects the difference between momentum as measured by the alpha and risk models. It is heavily weighted against 1-month momentum and toward outperformance that occurred 13 months ago. Since the optimizer favors $\alpha_{R\perp}$, the optimized portfolio takes these same concentrated bets.

UNEXPECTED, YES—BUT IS THIS BAD FOR YOU?

In this section, we investigate how using different models of risk and alpha affects the quality of an optimized portfolio. The problem is complex and the goal is insight. To that end, we analyze a simple case in which all asset returns are generated by a single factor. Here, the risk and alpha models are based on the same factor, but they measure it differently. In this context, alpha is just the expected return of the asset and the benchmark is cash.

Specifically, we assume that returns to a universe of n assets are driven by a single factor, f, as follows:

$$r = Xf + u \tag{9}$$

where $u \sim N(0, \sigma_s^2 I_n)$. The true asset covariance matrix is given by

$$\Sigma = \sigma_f^2 XX' + I_n \sigma_s^2 \tag{10}$$

where σ_f is the factor volatility and σ_s is the specific volatility which we assume is the same for all assets. All exposures are normalized.

We further assume that the manager uses separate single-factor models to capture risk and alpha. These models are imperfect, however, in that they estimate the exposures of the assets to the factors with error. Our measure of the accuracy of each model is the correlation between the model's exposures and the true exposures. We denote these ρ_A and ρ_R for the alpha and risk models, respectively.

Thus,

$$X_A = \rho_A X + \varepsilon_A \tag{11}$$

$$X_R = \rho_R X + \varepsilon_R \tag{12}$$

where the errors are distributed normally, $\varepsilon_A \sim N(0,(1-\rho_A^2)I_n)$ and $\varepsilon_R \sim N(0,(1-\rho_R^2)I_n)$, and they are uncorrelated to X, r, and to each other.

For simplicity, we ignore any estimation error in computing the factor returns, f_A, f_R, used for the alpha and the factor covariance matrix. Effectively, our calculations assume that the models are estimated using a very large universe over a long time.

The errors in exposures produce misestimates of alpha and risk. The true alphas (i.e., expected returns) are $XE(f)$, and the manager's alphas are noisy estimates of these, $\alpha = X_A E(f_A)$. Similarly, the risk model's forecasts of factor risk will, in general, be imperfect. We assume, however, that the risk model uses the true specific covariance matrix, so that its forecasts of specific risk are accurate.

Now, consider what happens when a manager optimizes a portfolio using different risk and alpha models. We use three risk models with different accuracies: $\rho_R = 1.0$ (perfect model), 0.9, and 0.8. With each, we perform a series of unconstrained optimizations with alphas of varying accuracies over our universe of stocks. To assess performance, we calculate the ex ante information ratios (IR) of the resulting portfolios using the true model.[5] In our examples, we set the maximum possible IR equal to one in order to facilitate comparisons.

Exhibit 2 shows the results for optimizations with 500 stocks. The true factor and specific annual volatilities are 1.5% and 30%, respectively.

As expected, the IR declines as the accuracy of the alpha decreases. What is striking, however, is that it is not always better to use a perfect risk model! Less accurate models work better when the alpha is sufficiently imperfect. Stranger yet, it is better to use an 80% accurate risk model than a 90% accurate risk model. What is going on?

To understand why this happens, we first decompose the manager's alpha into the part spanned by the risk exposures and the part orthogonal to them. In this case, we have

$$\alpha = \underbrace{(\rho_{A,R} X_R)E(f_A)}_{\alpha_R} + \underbrace{(X_A - \rho_{A,R} X_R)E(f_A)}_{\alpha_{R\perp}} \quad (13)$$

where $\rho_{A,R}$ is the correlation between X_A and X_R.

When the risk model is perfect, $\alpha_R = X\rho_A E(f_A)$ and $\alpha_{R\perp} = \varepsilon_A E(f_A)$. Thus, α_R is proportional to the true alpha, $XE(f)$, and $\alpha_{R\perp}$ is noise. Since the optimizer favors $\alpha_{R\perp}$ over α_R in building the optimal portfolio, it will favor noise over the true alpha! Paradoxically, the IR improves

EXHIBIT 2
Impact of the Risk Model on Performance

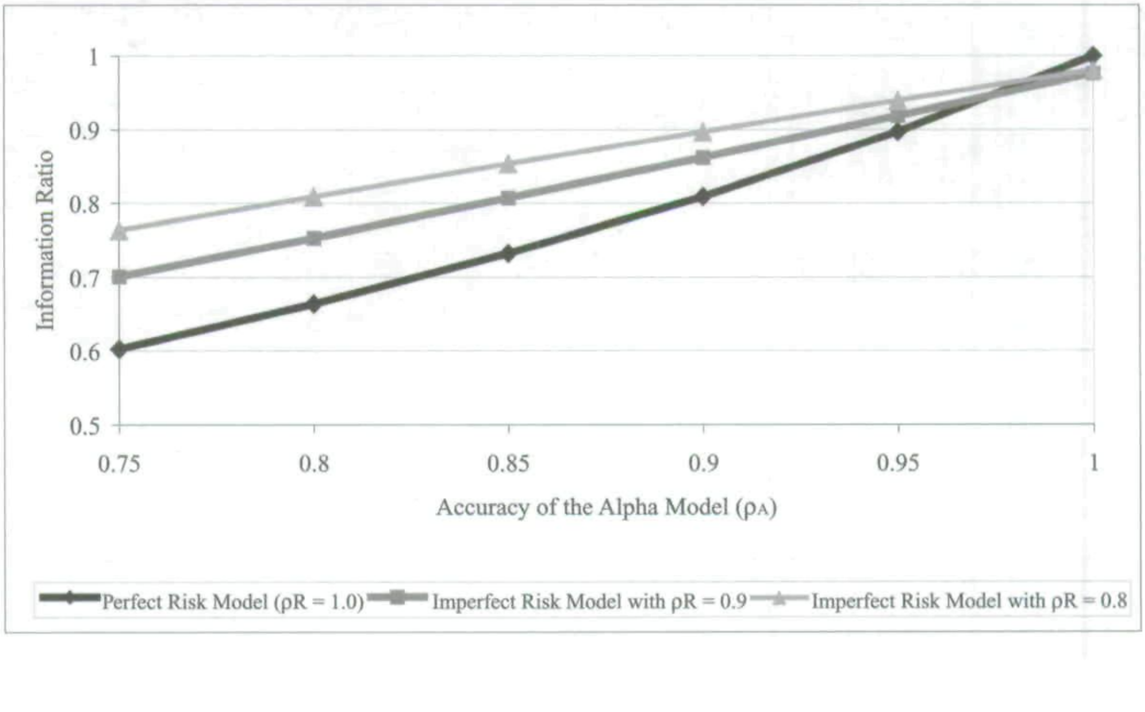
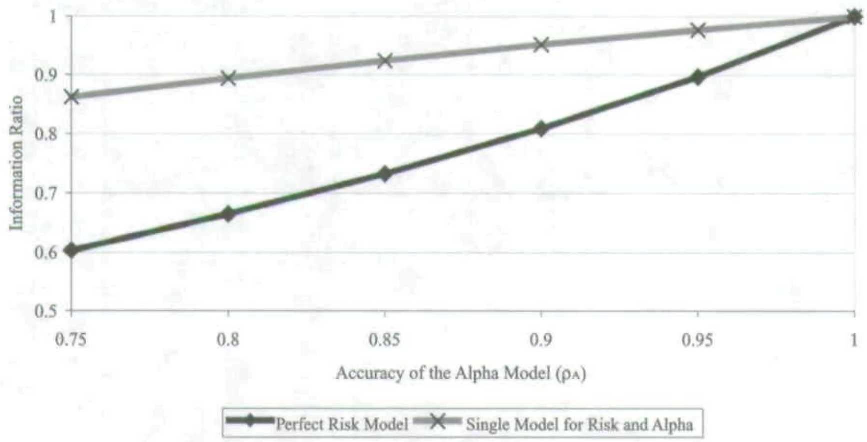


EXHIBIT 3
Aligning the Alpha and Risk Factor



as the risk model becomes less accurate, in part, because more of $\alpha_{R\perp}$ is then true alpha.[6]

What happens if we employ the same model for both risk and alpha? We can do this by using a covariance matrix based on the alpha model,

$$\Sigma_A = \sigma^2_{f_A} X_A X'_A + I_n \sigma^2_s \tag{14}$$

where $\sigma^2_{f_A}$ is the volatility of the alpha model's factor. This eliminates the opportunity for the optimizer to exploit discrepancies between risk and alpha factors.

As Exhibit 3 shows, using a single model for both risk and alpha improves performance across the board. In this case, there is no orthogonal alpha, so the optimal portfolio weights are proportional to the manager's alpha, as shown in Equation (13). The IR of the optimal portfolio is

$$IR = \frac{\rho_A E(f)}{\sqrt{(\rho_A \sigma_f)^2 + \frac{\sigma_s^2}{n}}} \tag{15}$$

Naturally, the IR falls as the accuracy of the alpha decreases. This occurs because a bet on the manager's alpha contains less true alpha, but the same amount of specific risk, as the accuracy worsens. When the factor volatility is large relative to average specific risk, however, the IR falls off more slowly—all else kept constant.

The amount to be gained by using one model depends, in part, on the strength of the factor. Exhibit 4 compares the performance achieved by a perfect risk model with that of a single model across a range of annual volatilities. The advantage is greatest when the factor volatility is strong and erodes as it weakens. Portfolios based on strong factors, such as momentum, may benefit more than those based on less volatile, fundamentals-based alpha factors.

Thus far, we have focused on the information ratio. Although using a single model improves the IR, does it do so at the expense of delivering a portfolio with the wrong risk level? Surprisingly, in our simple model the answer is no.

Exhibit 5 compares the accuracies of risk forecasts for portfolios formed using a perfect risk model, two imperfect risk models, and a single model for risk and alpha. As expected, the imperfect risk models underforecast the risk of the optimal portfolios that were built using them. The single model, however, accurately forecasts the risk of its optimal portfolio. Care must be exercised in

EXHIBIT 4
Factor Volatility Matters

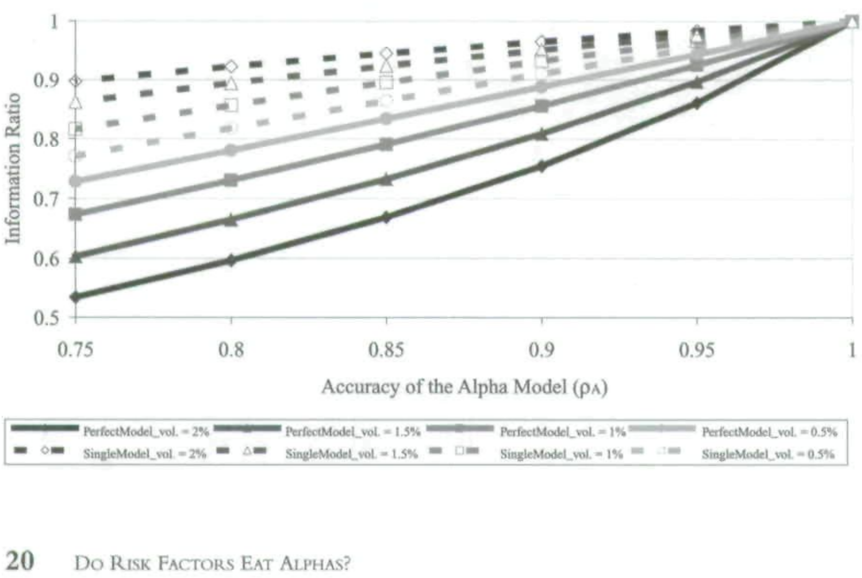
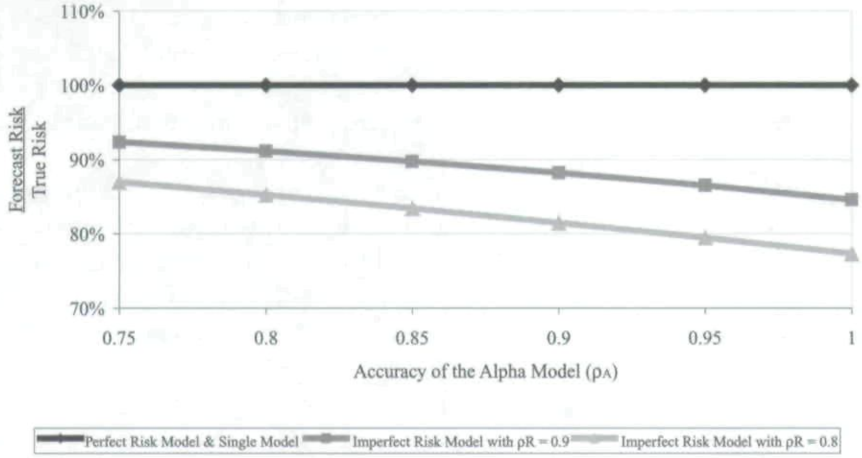


EXHIBIT 5

Risk Forecast Accuracy



interpreting these results. The single model gets the risk forecast right only for optimal portfolios containing all assets in the estimation universe. If we used a single model to optimize over a smaller number of assets, for example, the risk forecast would not necessarily be as good.

REMEDIES

In this section, we investigate approaches for remedying, or at least mitigating, the problems that arise when a manager uses different models for risk and alpha. Along the way we check whether insights from our analysis of the single-factor case apply to more realistic situations.

We focus on a fairly common situation in which a few of the risk model factors resemble the alpha factors; we call these the *related risk factors*. Our proposed remedies attempt to reduce the misalignment between the alpha factors and their risk factor counterparts. They are as follows:

1. The manager may simply *drop* the related risk factors from the risk model. This is accomplished by setting all asset exposures to those factors to zero in the manager's risk model.
2. The manager may alter the risk model by simply *substituting* the alphas for the related risk-factor exposures (i.e., swap X_{A_i} for the related X_{R_i}).
3. The manager may use a risk model that replaces the related risk factors with the alpha factors and retains all other risk factors. This requires *building a new risk model* that includes the alpha factors.
4. The manager may use his/her risk model to *emulate* the new risk model described in the previous point (3).

While the first three alternatives just listed are straightforward, the last requires further explanation. The idea is to approximate the covariance matrix of a risk model that is based on the manager's alpha factors and the retained risk factors. Let X_c denote the set of exposures to the alpha and retained risk factors. Our emulated risk model has the following form:

$$\Sigma = X_C F_C X_C' + \Delta_R \tag{16}$$

where Δ_R is from the manager's risk model and F_c is an approximation of the covariance matrix of the new model's factor returns.

We define F_c by using factor portfolios. A factor portfolio is a portfolio that has unit exposure to a given factor, no exposures to other factors, and minimum risk. For the factors of our new model, the asset weights of these portfolios are given by the rows of the matrix,

$$(X_C' W X_C)^{-1} X_C' W \tag{17}$$

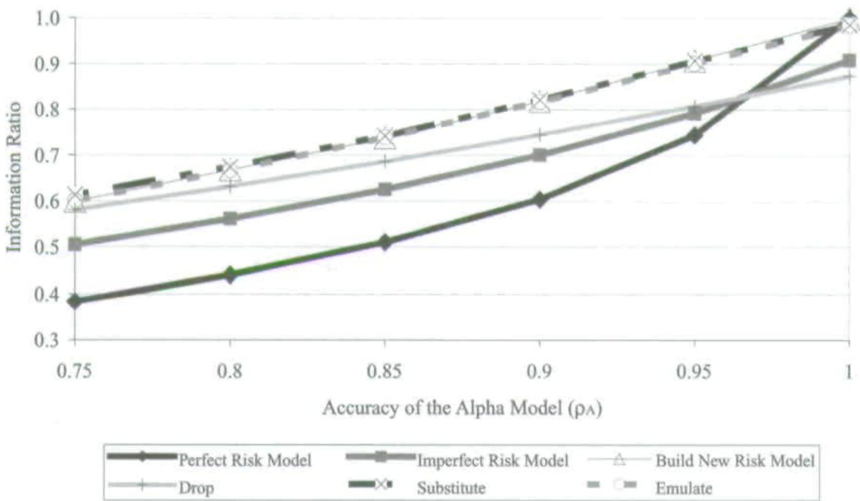
where W is a diagonal matrix that typically reflects each asset's specific volatility.[7]

We define F_c to be the factor covariance of these factor portfolios. Using the manager's risk model, this is

$$F_C = (X_C' W X_C)^{-1} X_C' W (X_R F_R X_R') W X_C (X_C' W X_C)^{-1} \tag{18}$$

To explore the effectiveness of these approaches, we apply the methodology used in the previous section to more realistic problems with several factors. We start by assuming that the Barra U.S. equity model is the true model of returns, that is, its factor structure and covariance matrix perfectly describe reality. The model contains 55 industries and 13 style factors, two of which are value and momentum. We assume that the true model's exposures do not change over time.[8]

EXHIBIT 6
Comparison of Remedies: Long-Short Portfolios



We use separate models for risk and alpha. The alpha model is based on value and momentum. The risk model employs the same factors as the U.S. equity model. Both models imperfectly estimate the exposures to value and momentum, and their errors are uncorrelated; the risk model perfectly knows all other factor exposures. As before, the risk model uses the true model of specific risk.

To keep things simple, the alpha model estimates momentum and value equally well, as does the risk model. This enables us to gauge the accuracy of each model with a single parameter reflecting the accuracies of its momentum and value exposures; we denote these accuracies ρ_A and ρ_R, respectively.

We run active optimizations using the S&P 500 Index as both the investment universe and the benchmark. We define the true alpha to be the equally weighted average of the Barra momentum and value exposures. The manager's alpha is the equally weighted average of the alpha model's estimate of value and momentum. In our examples, we assume that the manager's risk model is 90% accurate.

We first perform unconstrained optimizations, allowing the optimizer to both long and short securities. Exhibit 6 shows the average results of 100 simulations for each of the methods outlined, comparing them to what could be obtained by using a perfect risk model and the manager's imperfect risk model. In each simulation run, we generate different value and momentum exposures for the alpha and risk models, and reestimate them. We use the true model to compute the ex ante IRs of the optimized portfolios.

As in our single-factor study, we find that unless the alpha is sufficiently accurate the perfect risk model has the worst performance. Using an imperfect risk model or dropping risk factors helps in the case of less accurate alphas. All three of the other approaches—substitution, building a new risk model, and emulation—yield very similar and substantial improvements for a broad range of alphas. It is worth noting that substitution and emulation require significantly less effort, given that a risk model is already available.

The average accuracy of the active risk forecasts is shown in Exhibit 7. As expected, dropping risk factors or using an imperfect risk model leads to underestimates of risk. And the substitution approach overforecasts risk as the alpha degrades. This occurs because the manager's alpha carries less true factor risk as it becomes less accurate. The emulation approach underforecasts risk, especially

EXHIBIT 7
Risk Forecast Accuracy for Long-Short Portfolios

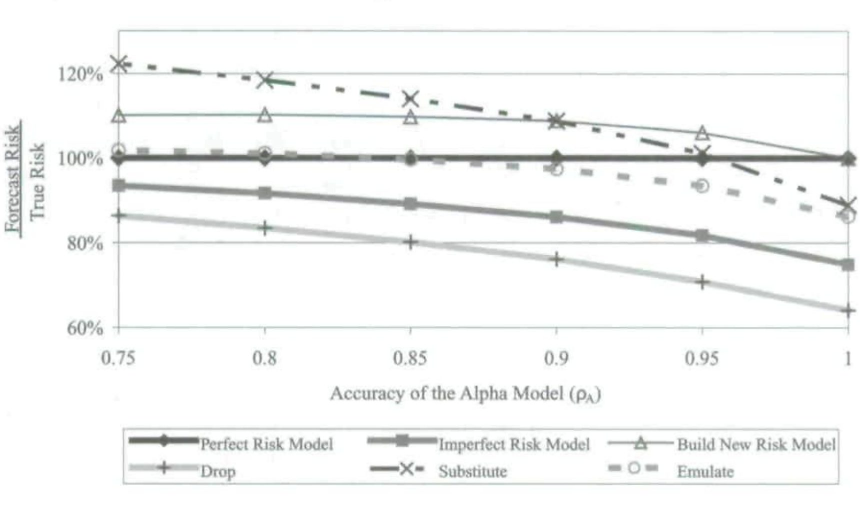
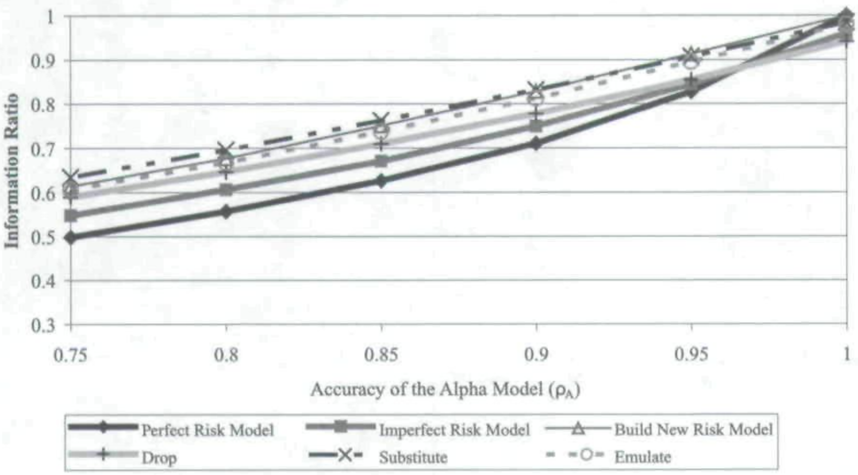


EXHIBIT 8A

Comparison of Remedies: Long-Only Portfolios with 1.5% Active Risk



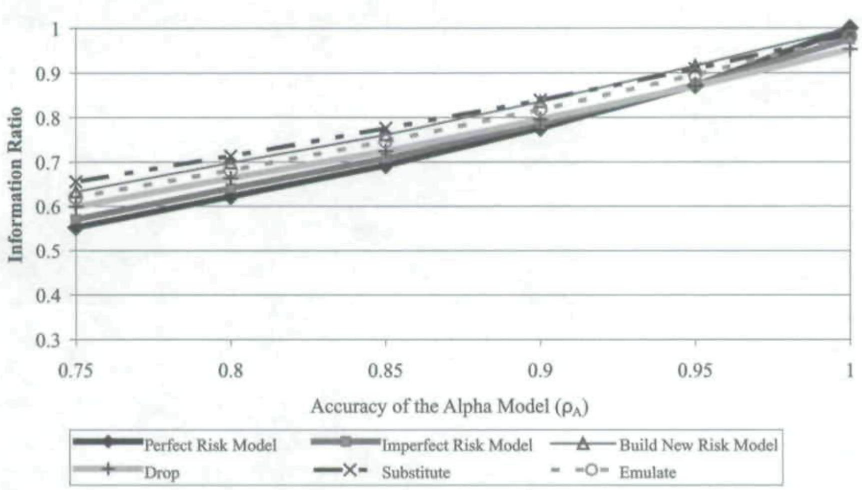
when the alpha is accurate, because the risk model on which the emulation is based is imperfect. Finally, unlike in our simple model, building the alpha into the risk model produces forecasts of risk that are too high.

Many managers are subject to long-only constraints. To see how they are affected by these issues, we rerun the optimizations, this time disallowing short positions. Exhibit 8 shows the results for optimal portfolios targeted at active risk levels of 1.5% and 3%.

The impact of the misalignment between risk and alpha factors is not as great for the long-only managers. The gap between the best- and worst-performing approaches narrows when we disallow shorting. Interestingly, at higher risk levels these differences become smaller still.

EXHIBIT 8B

Comparison of Remedies: Long-Only Portfolios with 3% Active Risk



SUMMARY

Potential pitfalls arise in portfolio optimization when one factor model is used to forecast risk and another is used to forecast alpha. Both models capture systematic sources of return and risk. We show that discrepancies between risk and alpha factors can create unintended bets in optimized portfolios that may hamper performance.

Aligning risk and alpha factors may improve the quality of optimized portfolios. We can achieve this by building a risk model that explicitly incorporates the alpha factors. Alternatively, we can emulate such a model using an existing risk model. In our problems, we find that both approaches improve the information ratio of optimized portfolios.

While we have provided a framework for understanding these issues, further study is needed to determine their impact on everyday portfolio optimization. We have made certain simplifications such as ignoring much of the noise in estimating the covariance matrix in order to concentrate on the main points. In practice, such factors might diminish some of the effects that we have observed.

Lastly, our studies have illustrated cases in which the risk and alpha models use the same factors, but measure exposures to the factors differently. The same analysis can be extended to include situations where the alpha model contains factors that are missing from the risk model.

ENDNOTES

The authors thank Miguel Alvarez, Vinod Chandrashekaran, Mark Ferrari, Neil Gilfedder, Lisa Goldberg, Mike Levinson, Jose Menchero, Guy Miller, Indrajit Mitra, Frank Nielsen, and Nicolo Torre for helpful

discussions and feedback on this article.

[1]From private communications; also see Jones [2007].

[2]We scaled the risk aversion to achieve this.

[3]The optimizer underweights 1-month momentum, rather than remaining neutral, in order to offset the risk of the overall bet on momentum.

[4]Here, we assume that the risk exposures and alpha exposures are normalized.

[5]The optimizations and calculations of returns and risks were done analytically assuming that $X'\varepsilon_A$, $X'\varepsilon_R$, and $\varepsilon_R'\varepsilon_A = 0$.

[6]More precisely, $\alpha_{R_\perp} = \rho_A(1-\rho_R^2)XE(f_A) + (\varepsilon_A - \rho_A \rho_R \varepsilon_R)E(f_A)$, where the first term is proportional to the true alpha and the second term is noise.

[7]In practice, W is defined in different ways. We simply set $W = I_n$. Our results are not materially affected by this choice.

[8]This assumption allows us to more easily construct the imperfect risk model and the risk model that includes the manager's factors and selected risk factors. As before, we ignore much of the estimation error in computing the factor returns.

REFERENCE

Jones, R., T. Lim, and P.J. Zangari. "The Black-Litterman Model for Structured Equity Portfolios." *Journal of Portfolio Management*, Winter 2007, pp. 24–33.