On the Fundamental Law of Active Portfolio Management: *What Happens If Our Estimates Are Wrong?*

GUOFU ZHOU

GUOFU ZHOU
is a professor of finance
at Washington University
in St. Louis, MO.
zhou@wustl.edu

The fundamental law of active portfolio management was first proposed by Grinold [1989]. In its simplest and most intuitive form, the law says that the value-added of an active manager is proportional to the information ratio (IR) squared of the active portfolio, and that the IR is proportional to the information coefficient (IC) and the squareroot of the market breadth (BR). In short, the value-added is a product of the skill squared (measured by IC) and the times one plays the skill (measured by BR). The law has since received enormous attention for its key insights into portfolio strategy design and performance evaluation, and it continues to be developed and explored as a central research topic in the science of portfolio management. For example, Clarke, de Silva, and Thorley [2002, 2006] analyzed the law under various constraints, and Qian and Hua [2004] studied the impact of variability in IC. Quantitative portfolio management books, such as Grinold and Kahn [2000], Chincarini and Kim [2006], and Qian, Hua, and Sorensen [2007], all devote major chapters to the law and its applications in portfolio management. The law assumes that both alpha and the risk matrix are known; however, these inputs are unknown and must be estimated in practice. The question then is to what extent the estimated inputs affect the reliability of the law and its applications.

This article addresses this question of the law's reliability under uncertainties associated with using estimated alphas and an estimated risk matrix. We make two points. First, the estimation errors matter greatly in practice. Consider an imaginary case in which we know the true alphas. This is, of course, unrealistic because searching for alpha is the game every one is playing and so the estimation errors can obviously be large. Even in this hypothetically optimistic scenario, we find that the error in estimating the risk matrix alone can cause the value-added to be only a portion of what is claimed by the law, and that the proportional coefficient can easily approach zero if the sample size is small relative to the number of assets. The implication is clearly opposite to that of the law. Holding IC constant, the larger the BR, which is the number of assets, the larger the value-added. However, the larger the BR, the greater the difficulty in estimating the risk matrix, and the larger the estimation errors. Hence, there will be a trade-off between BR and the estimation errors in a given application. When BR is too large relative to the amount of data, the estimation errors can wipe out all the gains promised by the law. In general, when both alpha and the risk matrix have to be estimated, the estimation errors become larger, and the law becomes even less reliable. In short, the estimation errors add another layer of difficulty—on top of transaction costs, position limits, and

other market frictions—for an active manager who must outperform the real world benchmark index portfolio.

The second point we make is that the estimation errors can be mitigated to a certain degree. Theoretically, if the world is stationary and if we have an infinite amount of data, the true values of alpha and the risk matrix can be learned, and the estimation errors can be eliminated. However, this is clearly not the case in practice. Nevertheless, we suggest two feasible ways to reduce the impact of estimation errors. The first way is to scale the estimated optimal portfolio. In the absence of estimation errors, the scale is one. With estimation errors, the scale is optimally chosen to improve portfolio performance. The second way is to diversify away some of the unnecessary estimation risks by holding other portfolios in addition to the estimated optimal portfolio. A general framework for this will be provided after we discuss the estimation errors and assess their impact.

ESTIMATION ERRORS

Consider a managed portfolio with return R_p in excess of the risk-free rate. If R_B is the excess return on the benchmark portfolio, then we have

$$R_p = R_B + R_A \tag{1}$$

where R_A is the return on the active portfolio. Suppose that the benchmark consists of N risky assets. We can always decompose the excess return on the i-th asset into

$$R_i = \alpha_i + \beta_i R_B + \varepsilon_i \tag{2}$$

where α_i is the i-th asset's alpha, β_i is its beta, and ε_i is the residual with zero mean, conditional on available information. Mathematically, the return composition is simply a projection of R_i onto 1 and R_B. Then, $R_i - \beta_i R_B$ is the return without benchmark risk.

Denote α the N-vector formed by the alphas and Σ the covariance matrix of the excess asset returns after removing the benchmark risk. If the return decomposition is independently distributed over time, α and Σ will be constant parameters. If the return decomposition is stochastically evolving, they will be the parameters for the next period, which is what matters for the investment decision making in the standard mean–variance framework. Given the risk-aversion parameter γ, it is well known that the optimal active portfolio weights are

$$w = \frac{1}{\gamma}\Sigma^{-1}\alpha \tag{3}$$

and the value-added (risk-adjusted extra return relative to the benchmark) is

$$\mathrm{VA}^* = \frac{1}{2\gamma}\alpha'\Sigma^{-1}\alpha \tag{4}$$

and the IR, defined as the ratio of the expected active return to its standard deviation, is

$$\mathrm{IR}^* = \sqrt{\alpha'\Sigma^{-1}\alpha} \tag{5}$$

We see that the value-added is simply proportional to the square of IR*. However, this relation will break down in the presence of estimation errors. Note that the optimal weights are obtained without any constraints. There are, however, typically two constraints on the active portfolio. The first is that the sum of the weights is some constant in order to limit exposure. The second is that the weights are orthogonal to the betas (without the beta risk) to reflect a residual position. In this case, we can still obtain the same mathematical expression for the maximum IR by simply replacing the alpha and sigma matrix by those for only the first $N - 1$ assets, as shown in the appendix to Chapter 6 of Grinold and Kahn [2000] and, with more detail, in Appendix 2A of Chincarini and Kim [2006]. Because the conclusions for the unconstrained case carry through to the constrained case with simpler notation, we will focus only on the unconstrained case.

Economically, VA* is the maximum value-added, and IR* is the maximum IR obtainable. In other words, even with perfect foresight about both alpha and the risk matrix, a value-added no greater than VA and an information ratio no better than IR* can be obtained. However, neither the alphas nor the risk matrix are known in practice and have to be estimated from data. With estimated parameters, one can evaluate the estimated maximal VA and IR. But there is one obvious problem. If one inflates the alpha estimates by doubling their values, the IR will be doubled, and hence the value-added will be quadrupled. This is clearly incorrect. Therefore, one must use caution in estimating VA* and IR*.

Grinold [1989] took a novel approach in evaluating VA* and IR*, which is free of this obvious problem.

He related the alphas to forecasts of the returns and linked the IR to the ability to perform the forecast. In particular, he showed that

$$\mathrm{IR} = \mathrm{IC}\sqrt{N} \tag{6}$$

where IC is the information coefficient (skill) measured in terms of the correlation between the forecasts and the future returns, which are assumed constant across assets. Equation (6), which says that IR is linearly related to skill, is much more intuitive than the general formula (see Grinold and Kahn [2000, p.167]). If the manager doubles forecasting accuracy (IC), then IR doubles. If accuracy can be doubled at a given research cost, and if the cost is lower than the value-added, it should be doubled. Applying the same level of skill to a portfolio of 500 stocks will generate 10 times more value than applying it to 5 stocks. As a result, a small degree of predictability can potentially help an active manager make significantly greater returns than the benchmark.

While Grinold [1989] and others provided enormous insights into portfolio management via the fundamental law of active portfolio management, the usefulness of the law is subject to estimation errors. But this issue is largely ignored in the literature, although related issues are receiving increasing attention (see Fabozzi, Kolm, Pachamanova, and Focardi [2007]). For example, even in the most elegant derivation of Grinold [1989], the value-added was evaluated by treating it as if the estimated alphas were the true ones and as if we knew all other parameters. To understand the role played by the estimation errors, let $\hat{\alpha}$ and $\hat{\Sigma}$ be the estimated parameters, which could be the maximum likelihood estimators or some other estimators. Then, because the true weights are unknown, we can estimate the optimal active portfolio weights using Equation (3) to obtain

$$\hat{w} = \frac{1}{\gamma}\hat{\Sigma}^{-1}\hat{\alpha} \tag{7}$$

The important point is that, conditional on our estimates, the expected active return and variance (risk) should be evaluated by

$$E(R_A \mid \hat{w}) = \hat{w}'\alpha \tag{8}$$

and

$$\sigma^2(R_A \mid \hat{w}) = \hat{w}'\Sigma\hat{w} \tag{9}$$

where α and Σ are the the true parameters. These equations state that, no matter how we estimate our portfolio weights, the resulting portfolio return and risk should be judged by the market or by the true parameters, not by our estimates of them. Therefore, once we do not assume that we know the true parameters, the evaluation of $E(R_A)$ and $\sigma(R_A)$ becomes complex. Needless to say, we no longer have simple expressions for VA^* and IR^* as before. The question is how they can be determined using reasonable assumptions about the behavior of our estimates for α and Σ.

THE IMPACT

Although there is a 99.99% probability that the estimates will not equal the true parameters, they should be very close to them. By modeling stock returns, we can usually obtain unbiased estimates (at least asymptotically). For illustration and for simple analytical expressions, we assume that the excess asset returns are independent, identical, and normally distributed over time. Then, the standard regression analysis yields an estimator for the alphas that is normally distributed around the true alphas with covariance matrix $\tau\Sigma$, that is,

$$\hat{\alpha} \sim N(\alpha, \tau\Sigma) \tag{10}$$

where $\tau = 1/T$ and T is the sample size. In general, the estimator of the alphas may come from another source. The notation in Equation (10) allows a generic description of the accuracy of $\hat{\alpha}$. Given the normality assumption, the standard estimator of Σ is the sample covariance matrix, $\hat{\Sigma}$, of the excess asset returns after removing the benchmark risk. Then, $\hat{\Sigma}$ should follow a Wishart distribution around the true Σ. The question is how these estimates will change our earlier evaluation for VA or IR.

To understand the impact of the estimation errors, we consider three cases. The first case is a hypothetical one in which α is assumed known and Σ is estimated, which allows us to understand the effect of errors from estimating Σ alone. The second case is also a hypothetical one in which Σ is assumed known and α is estimated. The third case is a realistic one in which neither Σ nor α is known and both have to be estimated using $\hat{\Sigma}$ and $\hat{\alpha}$, respectively.

Consider the first case in which the true alphas are assumed known. Based on Kan and Zhou [2007] and the preceding equations, the expected active return and risk are

$$E(R_A) = \frac{T}{T-N-2}\frac{1}{\gamma}\alpha'\Sigma^{-1}\alpha \tag{11}$$

and

$$\sigma^2(R_A) = \frac{T^2(T-2)}{(T-N-1)(T-N-2)(T-N-4)}\frac{1}{\gamma^2} \times \alpha'\Sigma^{-1}\alpha \tag{12}$$

The terms involving T are caused by the estimation errors in Σ. The estimation errors not only make the expected active return higher, but also the risk higher regardless of the sample size, as long as $T > N + 4$. If the benchmark is the S&P 500 Index with $N = 500$, and if the data are monthly, the estimation errors will cause the expected active return to be 1.7 times higher than it should be and cause the risk to be 5 times the optimal level, even with 100 years of data. As a result, the VA becomes –1.62, a value destruction equal to 162% of VA*. In this case, if an active manager adopts the estimated portfolio strategy, he will, on average, experience sharp losses. Only when T is greater than 1,713, does the VA become positive. Of course, when the sample size T goes to infinity, the VA converges to the maximum VA*.

The fundamental law of active portfolio management suggests that, for any given level of skill, the higher the N, the greater the VA. In contrast, when the manager knows the true alphas, and his skill in predicting the alphas is perfect, the errors in estimating the risk matrix can still easily wipe out the value of skill, even if the standard risk estimator is using 100 years of monthly data. The intuition is that, when the sample size is not large enough relative to N, the estimated optimal portfolio weights based on the estimated risk matrix are so volatile they yield ex post a nonoptimal portfolio. Therefore, as N gets larger, if there are not enough data or efficient ways to estimate the risk matrix, no value is added in active management.[1] The message is that a larger N is not necessarily desirable in the presence of estimation errors.

Consider now the case in which the true Σ is assumed known. Again, based on Kan and Zhou [2007] and the preceding equations, the expected active return and risk are

$$E(R_A) = \frac{1}{\gamma}\alpha'\Sigma^{-1}\alpha \tag{13}$$

and

$$\sigma^2(R_A) = \frac{1}{\gamma^2}\alpha'\Sigma^{-1}\alpha + \frac{N\tau}{\gamma^2} \tag{14}$$

The uncertainty in estimating α has no impact on $E(R_A)$ as long as the estimator is unbiased, as in this case. However, the estimation errors affect $\sigma^2(R_A)$, increasing its value by the amount $N\tau/\gamma^2$. The greater the τ, the greater the increase. When the estimation errors decrease to zero, then $\sigma^2(R_A)$ reduces to the usual value. Nevertheless, whenever $\tau \neq 0$, VA is no longer proportional to the square of IR.

To obtain further insight, assume Σ is diagonal. Then, VA is simply given by

$$VA = \frac{1}{2\gamma}\sum_{i=1}^{N}\frac{\alpha_i^2 - \tau\sigma^2(R_i)}{\sigma^2(R_i)} \tag{15}$$

where $\sigma^2(R_i)$ is the variance of the i-th excess asset return. To ensure that the errors in estimating α_i will not be a drag on overall performance, $|\alpha_i| \geq \sqrt{\tau}\sigma(R_i)$. If an asset has an annual alpha of 2% and a standard deviation of 20%, then the estimation error, as measured by τ, should be smaller than 1/100 for the estimated alpha to add value. This says that if alpha is estimated from the regression, $T \geq 100$. But to obtain half of the value generated using true alpha, $T \geq 400$ is needed, and to achieve 90% of the maximum possible value, $T \geq 10{,}000$ is required!

And finally, consider the case in which both α and Σ are unknown. Again, based on Kan and Zhou [2007] and the preceding equations, the expected active return and risk are

$$E(R_A) = \frac{T}{T-N-2}\frac{1}{\gamma}\alpha'\Sigma^{-1}\alpha \tag{16}$$

and

$$\sigma^2(R_A) = \frac{T^2(T-2)}{(T-N-1)(T-N-2)(T-N-4)} \times \left[\frac{1}{\gamma^2}\alpha'\Sigma^{-1}\alpha + \frac{N\tau}{\gamma^2}\right] \tag{17}$$

It is interesting to observe that $E(R_A)$ is the same as in the first case. Despite uncertainty in both μ and Σ, $E(R_A)$ is still driven 100% by the uncertainty in Σ as long as $\hat{\mu}$ is unbiased. However, this is not true for $\sigma^2(R_A)$. The uncertainty in Σ contributes to the scalar terms involving T, and the uncertainty in μ adds the extra term, $N\tau/\gamma^2$. In practice, the estimator of $\sigma^2(R_A)$ is often computed based only on the middle term by replacing the associated parameters with their estimates. This ignores both the scalar and the extra terms, which can be large even when T is sizable. This explains the findings of Qian and Hua [2004], and others, that the realized $\sigma^2(R_A)$ is usually about 100% larger than predicted, a result of ignoring the estimation errors.

As previously demonstrated, when either the alphas or the risk matrix is unknown, the estimation errors can have a huge impact on the value-added and can easily wipe out all the potential values, if not properly guarded against. In the realistic case when both the alphas and the risk matrix are unknown, the errors in estimating them will jointly destroy the value of active portfolio management. The question is whether effective ways exist by which the impact of the estimation errors can be minimized.

THE REMEDIES

Because of the estimation errors, the estimated optimal portfolio must be different from the true optimal portfolio. The estimation errors can be quite large and thus the value-added can be substantially impacted. An important observation from our earlier impact analysis is that the errors disproportionately affect the expected active return and risk. In the utility function for evaluating the valued-added, however, the estimated expected active return is weighted linearly and the estimated risk is weighted quadratically. Hence, a suitable scaling may be used to adjust the level of desired expected active return and active risk in such a way that the utility is maximized. This is the first method that we propose as a remedy.

Instead of using the estimated optimal portfolio weights based on Equation (3), we consider its scalar

$$\tilde{w} = \frac{k}{\gamma}\hat{\Sigma}^{-1}\hat{\alpha} \tag{18}$$

where k is a constant scalar. Plugging this into the expected utility, we know that optimal k is

$$k^* = \left[\frac{(T-N-1)(T-N-4)}{T(T-2)}\right]\left(\frac{\theta^2}{\theta^2+N\tau}\right) \tag{19}$$

where $\theta^2 = \alpha'\Sigma^{-1}\alpha$ is an unknown parameter, but can be very accurately estimated by the procedure proposed by Kan and Zhou [2007]. It is clear that the scalar is less than one. Intuitively, the estimation errors make risky investments riskier, and hence we invest less aggressively than otherwise. Quantitatively, the estimation errors make the active risk increase relatively more than the increase in the expected active return as judged by the utility, and so the estimated optimal portfolio weights should be scaled back.

When τ is small or when μ is reasonably accurate, the value-added can be well approximated by

$$\text{VA} = \frac{T-N-1}{T-2}\text{VA}^* \tag{20}$$

where VA^* is the maximum value-added without estimation errors. To see the improvement of the valued-added by scaling, consider a sample size of $T = 1{,}714$. Even assuming we know the true alphas, the value-added by using the estimated optimal portfolio is only 0.11% of VA^*. In contrast, with the optimal scaling, the value-added is about 70.85% of the VA^*. This improvement is clearly of substantial economic importance.[2]

Our second method for minimizing the impact of the estimation errors is to diversify away some of the unnecessary estimation risk by holding other portfolios as well as the estimated optimal portfolio. Let $\hat{w}_1, \ldots, \hat{w}_q$ be q estimated portfolios. For example, $\hat{w}_1$ can be the earlier optimal estimated portfolio, $\hat{w}_2$ can be the global minimal-variance portfolio, and $\hat{w}_3$ can be the equal-weighted portfolio, and so on. In practice, for example, the score-weighted portfolios or portfolios of different managers can also be used. The objective is to choose the optimal combination to maximize utility. Because the utility is quadratic, the solution to the combination coefficients can be obtained analytically so that the optimal combination is feasible to construct.

For example, Kan and Zhou [2007] examined a combination of the estimated optimal portfolio and the global minimal-variance portfolio,

$$\hat{w} = \hat{w}(k_1, k_2) = \frac{1}{\gamma}(k_1\hat{\Sigma}^{-1}\hat{\alpha} + k_2\hat{\Sigma}^{-1}1_N) \tag{21}$$

where k_1 and k_2 are constants to be chosen optimally to maximize utility. They can be solved explicitly as

$$k_1 = c_0 \left(\frac{\psi^2}{\psi^2 + N\tau} \right), \quad k_2 = c_0 \left(\frac{N\tau}{\psi^2 + N\tau} \right) \mu_g \quad (22)$$

where ψ^2 is the squared slope of the asymptote to the ex ante minimum-variance frontier, μ_g is the expected excess return of the ex ante global minimum-variance portfolio, and c_0 is a constant function of the sample size and the number of assets. Theoretically, the combination will always outperform the alternative. Moreover, it can be shown that the combination method, although somewhat complex, will generally do better than the scaling method. More examples and their simulation analysis can be found in Kan and Zhou [2007], Kan and Smith [forthcoming], and Tu and Zhou [2008].

The analysis of the estimation errors in this article is conducted in the classical statistical framework. An alternative approach is to use a Bayesian set-up in which the model parameters are treated as random variables and hence are subject to uncertainty. The advantage of the Bayesian framework is that it allows the incorporation of prior information into the decision process. Useful priors may not be found easily, however, and the analysis is less intuitive, numerically more complex, and demanding. Examples of such studies can be found in Pástor [2000], Pástor and Stambaugh [2000], and Tu and Zhou [2004].

CONCLUSION

The fundamental law of active portfolio management pioneered by Grinold [1989] provides profound insights on the value-creation process of managed funds. However, a key weakness of the law and its various extensions is that they ignore the estimation risk associated with the parameter inputs of the law. We show that the estimation errors have a substantial impact on the value-added of an actively managed portfolio, and they can easily destroy all the value promised by the law if they are not dealt with carefully. To improve the chance of active managers beating benchmark indices, we propose two methods—scaling and diversification—which can be used to significantly minimize the impact of the estimation errors.[3]

ENDNOTES

The author is grateful to Campbell Harvey, Ronald Khan, Edward Qian, Jack Strauss, Jun Tu, and especially Raymond Kan for helpful discussion and comments.

[1]When N is large, say 500, one may use a K-factor risk model to estimate the risk matrix instead of using $\hat{\Sigma}$. But if the loadings are estimated from factor regressions, this makes no difference. Only when they are estimated under the nonlinear factor constraints via numerical optimization can the resulting risk matrix be different. The larger the N, the more likely the model is incorrectly specified. Then, the constrained estimator is biased even asymptotically, but $\hat{\Sigma}$ is always consistent under stationarity; its effective dimensionality, difficult to derive analytically, can be anywhere between K and N. In practice, K can be close to 100. Even if $K = 50$, it would still take at least 15 years of monthly data for the estimated strategy not to lose money. Since a risk model is known to often underestimate the risk by as much as 100%, the ex post VA decreases dramatically and may even diminish to zero since the initial IR is usually far less than one. The point is that, no matter how one estimates the risk matrix, the estimation errors are important. Our use of $\hat{\Sigma}$ makes the illustration analytical and easier to understand. Our conclusions and solution methods are applicable not only to $\hat{\Sigma}$, but also to other estimators of the risk matrix.

[2]The optimal scalar is solved analytically due to the simplifying normality assumption. The analytical solution is useful to illustrate the ideas, methods, and their implications. If one imposes a complex model for the asset returns, such as a multivariate GARCH process, simulation methods can be used to numerically determine the optimal scalar. Therefore the utility maximization problem can be solved with a single variable and thus be readily implemented by many practitioners.

[3]Theoretically, the law uses a linear strategy (in the alphas) that is only conditionally optimal. Zhou [forthcoming] provides the unconditionally optimal nonlinear strategy.

REFERENCES

Chincarini, Ludwig B., and Daehwan Kim. *Quantitative Equity Portfolio Management*. New York: McGraw-Hill, 2006.

Clark, Roger, Harindra de Silva, and Stephen Thorley. "Portfolio Constraints and the Fundamental Law of Active Management." *Financial Analysts Journal*, Vol. 58, No. 5 (2002), pp. 48–66.

——. "The Fundamental Law of Active Management." *Journal of Investment Management*, Vol. 4, No. 3 (2006), pp. 54–72.

Fabozzi, Frank J., Petter N. Kolm, Dessislava Pachamanova, and Sergio M. Focardi. *Robust Portfolio Optimization and Management*. New York: Wiley, 2007.

Grinold, Richard C. "The Fundamental Law of Active Management." *The Journal of Portfolio Management*, Vol. 15, No. 3 (1989), pp. 30–37.

Grinold, Richard C., and Ronald N. Kahn. *Active Portfolio Management*, 2nd ed. New York: McGraw-Hill, 2000.

Kan, Raymond, and Daniel R. Smith. "The Distribution of the Sample Minimum-Variance Frontier." *Management Science* (forthcoming).

Kan, Raymond, and Guofu Zhou. "Optimal Portfolio Choice with Parameter Uncertainty." *Journal of Financial and Quantitative Analysis*, Vol. 42, No. 3 (2007), pp. 621–656.

Pástor Ľuboš. "Portfolio Selection and Asset Pricing Models." *Journal of Finance*, Vol. 55, No. 1 (2000), pp. 179–223.

Pástor Ľuboš, and Robert F. Stambaugh. "Comparing Asset Pricing Models: An Investment Perspective." *Journal of Financial Economics*, Vol. 56, No. 3 (2000), pp. 335–381.

Qian, Edward, and Ronald Hua. "Active Risk and Information Ratio." *Journal of Investment Management*, Vol. 2, No. 3 (2004), pp. 20–34.

Qian Edward, Ronald Hua, and Eric Sorensen. *Quantitative Equity Portfolio Management: Modern Techniques and Applications*. New York: Chapman & Hall, 2007.

Sorensen, Eric, Edward Qian, Robert Schoen, and Ronald Hua. "Multiple Alpha Sources and Active Management." *Journal of Portfolio Management*, Vol. 30, No. 2 (2004), pp. 39–45.

Tu, Jun, and Guofu Zhou. "Data-generating Process Uncertainty: What Difference Does It Make in Portfolio Decisions?" *Journal of Financial Economics*, Vol. 72, No. 2 (2004), pp. 385–421.

——. "Being Naive about Naive Diversification: How to Improve the 1/*N* Rule?" Working paper, Washington University, 2008.

Zhou, Guofu. "On the Fundamental Law of Active Portfolio Management: How to Make Conditional Investments Unconditionally Optimal." *Journal of Portfolio Management* (forthcoming).

To order reprints of this article, please contact Dewey Palmieri at dpalmieri@iijournals.com or 212-224-3675