When Factor Timing Makes Sense

Farzin Barekat*

April 17, 2019

Abstract

In this paper we investigate when it makes sense for portfolio managers to implement factor timing in their quantitative investing, that is we outline necessary circumstances under which the benefits of factor timing (measured by the improvement in Sharpe ratio and the skewness of the returns) outweighs the challenges associated with development and implementation of factor timing. In particular, we mathematically show that factor timing for a single strategy does not yield substantial improvements unless either (1) the Sharpe ratio of the strategy is orders of magnitude different across states, or (2) the signal used for factor timing can accurately predict when the strategy will deliver negative returns (and the portfolio manager is willing to go short the strategy at that time). On the other hand, using simulation, we provide evidence that for a multi-factor portfolio (containing more than 10 factors) allocating risk based on instantaneous correlations between the factors at the beginning of each time period improves performance above the passive approach of allocating risk based on the long-term factor correlations.

1 Introduction

In the world of portfolio management, there is a constant desire to improve upon existing strategies. One popular belief among practitioners is that factor timing, or developing signals to predict a strategy's return, can significantly enhance performance of a strategy. In this paper, we show that this belief is typically wrong: we demonstrate that unless the signal can flag situations when (1) the Sharpe ratio of the strategy is orders of magnitude different across states, or (2) the strategy is expected to deliver negative returns (and the investor is willing to go short the strategy at that time), the improvement from actively changing the risk allocated to a single strategy is marginal, despite perfect foresight on the Sharpe ratio of that strategy over each time period. Therefore, rather than trying to enhance the performance of a strategy by developing signals that predict its Sharpe ratio over each time period, the portfolio manager (PM) is better off investing her resources to develop new, low or negatively correlated strategies and include them with the current strategy.

*Assistant Adjunct Professor at UCLA Math Department. Former Senior Vice President at PIMCO.

However, if the PM is running a multi-factor portfolio, we provide evidence that it might be beneficial to do factor timing: allocating risk based on instantaneous correlations between the factors at the beginning of each time period helps improve performance above the passive approach of allocating risk based on the long-term factor correlations. The improvement becomes more significant with the number of factors in the portfolio, as well as with the variation in pairwise correlations among the factors. The benefits of factor timing for a multi-factor portfolio become apparent when it consists of more than circa 10 factors. Moreover, note that in the context of multi-factor portfolio, factor timing mitigates concentration of risk in particular sector/asset class.

The results of this paper have practical consequences for portfolio managers. To provide a concrete example consider the following situation: Suppose a PM has established that the spread between two securities (with a given hedge ratio) exhibits mean reversion. Therefore, she decides to go short the spread whenever its current z-score (calculated over a rolling or a expanding window) is above say +1. Alternatively, she might consider to go short when the z-score is above +1, but double the size of the position whenever the z-score becomes above +2 (because when z-score becomes +2 her conviction in the position is doubled). Based on the results of this paper, the latter strategy is only marginally better than the former strategy. This is even before taking into account that the latter strategy will incur higher transaction costs than the former strategy, because it has higher turnover.

2 Factor Timing for a Single Strategy

2.1 General results

In this section we present the inefficacy of factor timing for a single strategy, except in certain specific scenarios that we outline below. In particular, we establish a theoretical upper bound on the amount of improvement in the long-run ex-post Sharpe ratio that an investor can achieve even when she has perfect foresight on the Sharpe of the strategy at the beginning of each time interval.

Suppose the investor is running a strategy and she is allowed to change the risk allocated to the strategy only at a specific time t_i, where $t_i = i\Delta t$ and Δt is a constant time period (e.g. $\Delta t = 1/52$ for weekly re-balancing). For $1 \leq i \leq T$, denote the return and the prevailing Sharpe ratio and the ex-ante Sharpe ratio of the strategy for time interval of $[t_{i-1}, t_i]$ by $r(i)$, $S(i)$, and $\hat{S}(i)$ respectively. Note that $\hat{S}(i)$ is known at time t_{i-1}, whereas, $r(i)$ becomes known at time t_i. The next theorem provides an upper bound on the ex-post Sharpe that the investor can achieve over the entire time period (and not just a particular time interval) using any risk allocation policy as long as the strategy is run over a long time.

Theorem 2.1 *Suppose the Sharpe ratio, $S(i)$, for $1 \leq i \leq T$, is sampled independently from a fix distribution μ, and the risk allocated to the strategy at the beginning of each time period is a function of $S(i)$. Furthermore, assume the associated return is normally distributed and is independent for each time period. Then, for a sufficiently large T, the* *(ex-post) Sharpe ratio of the strategy over the time period* $[0, t_T]$ *is bounded above by*

$$\sqrt{\mathbb{E}_\mu[S^2]}, \tag{1}$$

where $\mathbb{E}_\mu[\cdot]$ *denotes the expectation with respect to the distribution* μ*.*

Proof: See appendix A.

Throughout this section, we refer to an investor who maintains the same amount of risk in the strategy at all times as the *passive investor*[1]. On the other hand, an investor who has perfect foresight on the Sharpe ratio of the strategy during each interval (i.e. $\hat{S}(i) = S(i)$ for $1 \leq i \leq T$) and implements the optimal risk allocation (i.e. risk allocation is proportional to the ex-ante Sharpe ratio) is referred to as the *clairvoyant investor*[2].

Theorem 2.2 *Under the assumptions of theorem 2.1, over a long time period, the Sharpe ratio of the passive investor is slightly less than*

$$SR_{passive} = \mathbb{E}_\mu[S]. \tag{2}$$

On the other hand, over a long time period, the clairvoyant investor achieves Sharpe ratio slightly less than

$$SR_{clairvoyant} = \sqrt{\mathbb{E}_\mu[S^2]}. \tag{3}$$

Proof: See appendix A.

It is easy to show that equation (3) is greater than or equal to (2) by using the Cauchy–Schwarz inequality.

Some might argue that the purpose of factor timing for a single strategy is not necessarily to improve its Sharpe ratio but to improve other aspects of the strategy's returns, for example improving skewness of its returns. Here skewness of the returns refers to the following quantity

$$\frac{m_3(r(t))}{(\mathrm{Var}[r(t)])^{3/2}},$$

where $m_3(r(t))$ denotes the third central moment of the returns. The next theorem provides estimates on the skewness of returns for a single strategy with and without factor timing.

Theorem 2.3 *Under the assumptions of theorem 2.1, over a long time period, the skewness of the returns for the passive investor is*

$$SK_{passive} = \frac{m_3(S)}{(1 + \Delta t\,\mathrm{Var}[S])^{3/2}} \Delta t^{3/2}. \tag{4}$$

On the other hand, over a long time period, the skewness of the returns for the clairvoyant investor is

$$SK_{clairvoyant} = \frac{3\,\mathrm{Var}[S^2] + \Delta t\ m_3(S^2)}{(\mathbb{E}[S^2] + \Delta t\,\mathrm{Var}[S^2])^{3/2}} \Delta t^{1/2}. \tag{5}$$

[1] By "passive", we mean the investor is not readjusting risk allocation to the strategy at the beginning of each sub-interval.

[2] Note that the "clairvoyant" investor has perfect foresight on the Sharpe ratio of the strategy, not its return, over each time period.

Proof: See appendix A.

Remark 2.4 *Note that in theorem 2.3,* $\mathbb{E}[\cdot]$, $\mathrm{Var}[\cdot]$, *and* $m_3(\cdot)$ *are taken with respect to distribution* μ. *To avoid cluttering the notation, we have dropped* μ *from equations* (4) *and* (5).

In the next subsection, using a simple toy example, we will investigate the scenarios under which the improvement of $SR_{clairvoyant}$ over $SR_{passive}$ justifies the challenges of actively adjusting the risk allocated to the strategy.

2.2 Simple example of two state Sharpe

This subsection uses a simple example to elucidate the ideas presented in the previous subsection. Suppose the Sharpe of a strategy over each time interval is given by the following two-state random variable

$$S = \begin{cases} S_A & \text{when state } A \text{ occurs with probability } p_A, \\ S_B & \text{when state } B \text{ occurs with probability } p_B = 1 - p_A. \end{cases} \tag{6}$$

Next, suppose an investor is able to create a signal that perfectly predicts the state of the Sharpe for each time interval, and decides to allocate k_A and k_B units of risk to states A and B respectively. By equation (A.7), over the long run, the investor achieves the Sharpe ratio slightly lower than the following:

$$\frac{\mathbb{E}_\mu[kS]}{\sqrt{\mathbb{E}_\mu[k^2]}} = \frac{p_A k_A S_A + p_B k_B S_B}{\sqrt{p_A k_A^2 + p_B k_B^2}}. \tag{7}$$

The plot in figure 1 shows the ex-post Sharpe ratio (7) for different values of k_B/k_A when $p_A = p_B = 0.5$, $S_A = 1$, and $S_B = 2$. As mentioned in theorem 2.2, the Sharpe ratio is maximized when the investor allocates double the risk to state B than state A (because in this example, we have set $S_B/S_A = 2$).

Figure 2 shows the Sharpe ratios of the passive and clairvoyant investors for different values of S_B. Here we assume again that $p_A = p_B = 0.5$ and $S_A = 1$. The Sharpe ratios for these investors are given by equations (2) and (3), which in the context of this example simplifies to

$$SR_{passive} = p_A S_A + p_B S_B \qquad \text{and} \qquad SR_{clairvoyant} = \sqrt{p_A S_A^2 + p_B S_B^2}.$$

Of course, the clairvoyant investor achieves higher Sharpe ratio than the passive investor; however, **the amount of improvement gained is underwhelming considering the challenges that the clairvoyant investor faces (e.g. developing an accurate timing signal, dynamically changing the size of the strategy, related transaction costs as well as costs associated with technology and infrastructure).**

Figure 3 exhibits improvement in Sharpe ratio of the clairvoyant investor over the passive investor (dashed-blue line over solid-red line in figure 2). As can be seen, even

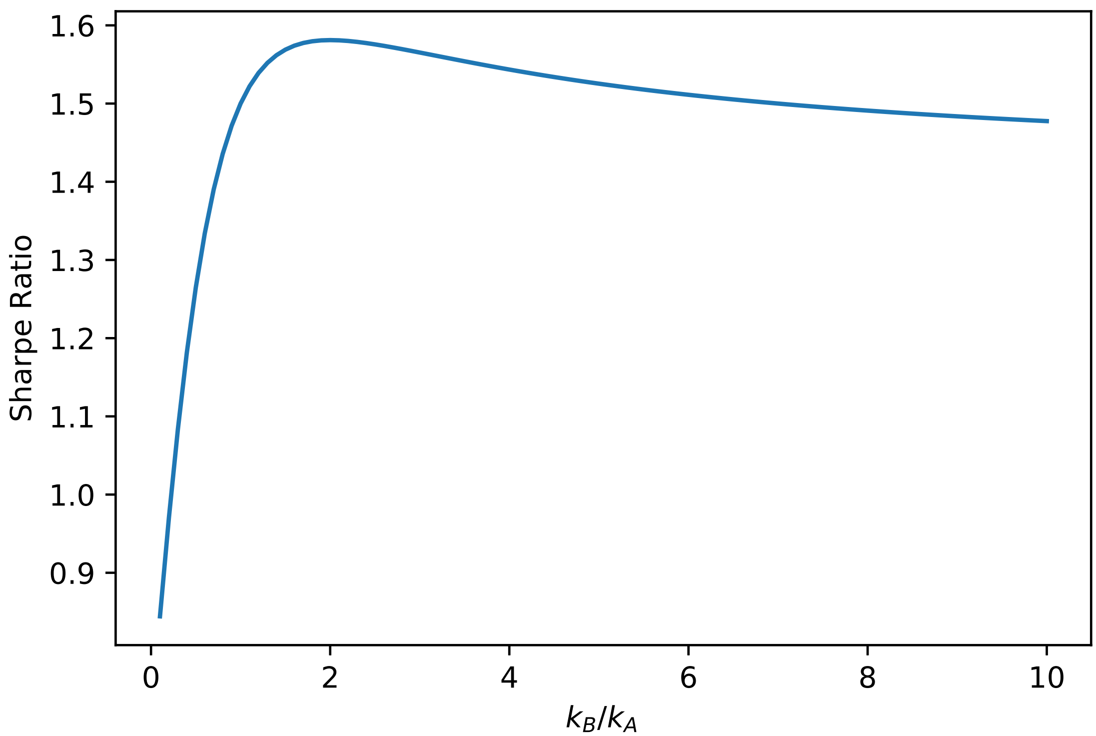

Figure 1: Investor's ex-post Sharpe ratio over long run for different values of k_B/k_A. Here $p_A = p_B = 0.5$, $S_A = 1$, and $S_B = 2$.

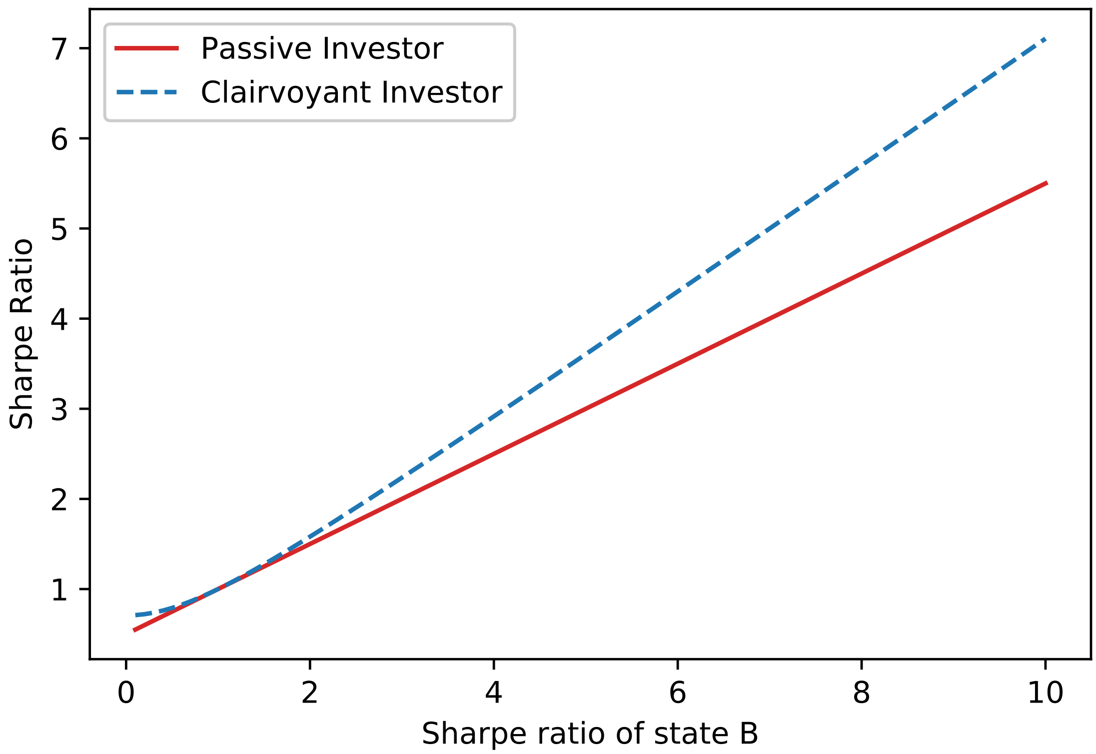

Figure 2: Ex-post Sharpe ratio of the passive and clairvoyant investor over a long time for different values of S_B. The passive investor maintains a constant amount of risk in the strategy at all time. The clairvoyant investor has perfect foresight on the state of Sharpe ratio at the beginning of each time period and implements optimal risk allocation. Here $p_A = p_B = 0.5$ and $S_A = 1$.

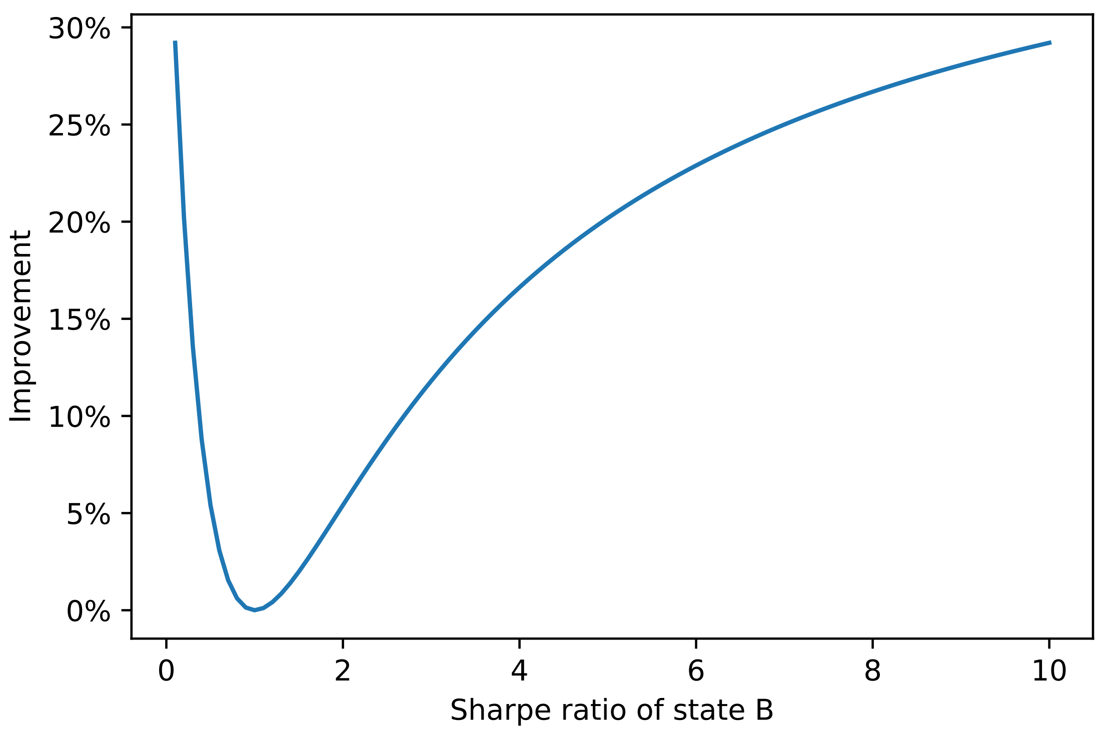

Figure 3: Long-term Sharpe ratio improvement of the clairvoyant investor over the passive investor for different values of S_B (i.e. $SR_{clairvoyant}/SR_{passive} - 1$). The passive investor maintains a constant amount of risk in the strategy at all time. The clairvoyant investor has perfect foresight on the state of Sharpe ratio at the beginning of each time period and implements optimal risk allocation. Here $p_A = p_B = 0.5$ and $S_A = 1$.

when S_B is ten times greater than S_A, the improvement that the clairvoyant investor achieves is below 30%.

Figure 4 complements Figure 3 by considering two additional cases when $p_B = 0.1$ and $p_B = 0.25$. Note that we only need to consider cases where $p_B \leq 0.5$ because by symmetry without loss of generality we may assume that state A represents the state with higher probability.

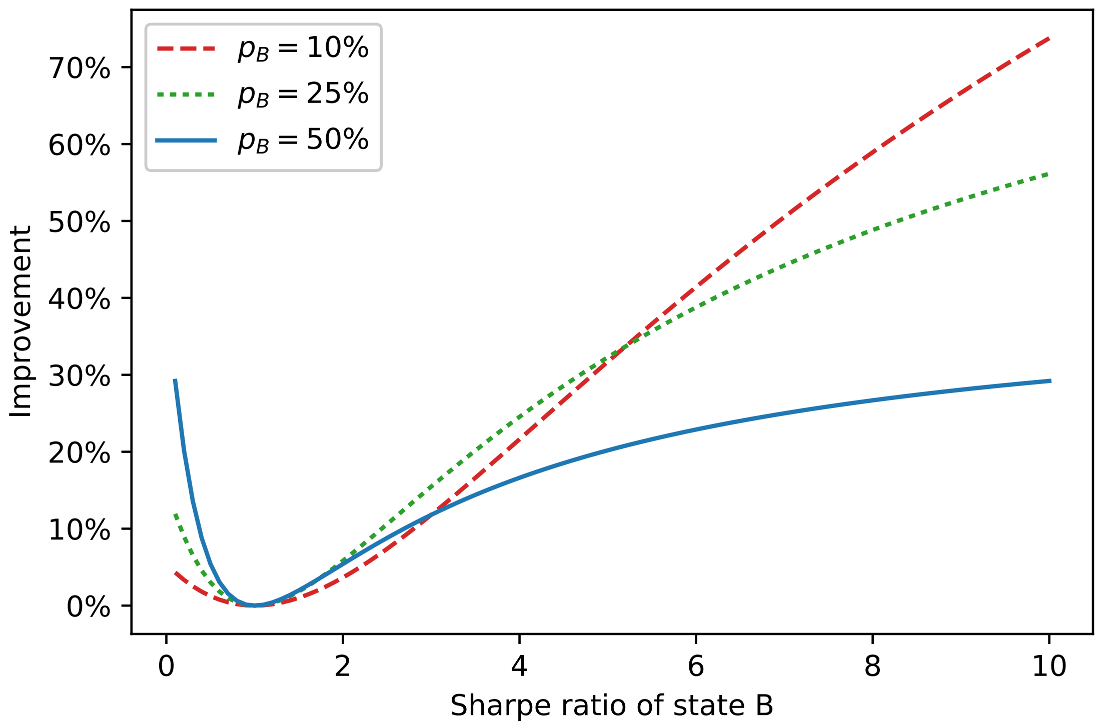

Figure 4: Long-term Sharpe ratio improvement of the clairvoyant investor over the passive for different values of S_B (i.e. $SR_{clairvoyant}/SR_{passive} - 1$). The passive investor maintains a constant amount of risk in the strategy at all time. The clairvoyant investor has perfect foresight on the state of Sharpe ratio at the beginning of each time period and implements optimal risk allocation. Here $S_A = 1$ and the plots for three cases $p_B = 0.1$, $p_B = 0.25$, and $p_B = 0.5$ are shown.

Typically, a strategy that has low overall Sharpe ratio $SR_{passive}$, say below 0.5, would not be implemented, let alone be considered for factor timing. For that reason, in figure 5 we hold $SR_{passive}$ constant at 0.5 and show the improvement achieved from factor timing. Unsurprisingly, it shows that **if the signal can accurately predict huge draw downs in the strategy and the investor is willing to go short the strategy in those cases, then the improvement from factor timing can be significant.**

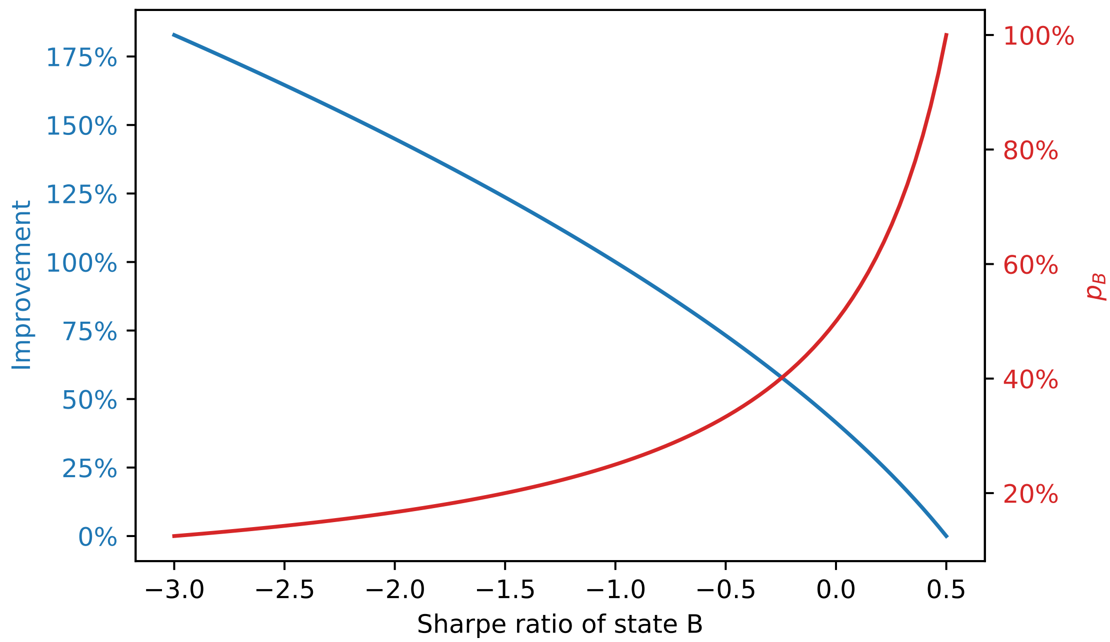

Figure 5: The left-axis shows the long-term Sharpe ratio improvement of the clairvoyant investor over the passive for different values of S_B (i.e. $SR_{clairvoyant}/SR_{passive} - 1$). The passive investor maintains a constant amount of risk in the strategy at all time. The clairvoyant investor has perfect foresight on the state of Sharpe ratio at the beginning of each time period and implements optimal risk allocation. Here $S_A = 1$ and p_B (shown on the right-axis) is chosen so that the Sharpe ratio of the passive investor $SR_{passive}$ is equal to 0.5.

Finally, some might argue that even though factor timing for a single strategy might not improve the Sharpe ratio it has other benefits: for example it would improve the skewness of the strategy's returns. Using the results of theorem 2.3, figure 6 plots the difference in the long-term skewness of the weekly returns of the clairvoyant and passive investors. **As it can be seen there is an improvement to the skewness of the returns, but it is not substantial unless S_B is at least 4x the value of S_A.** To provide some context on the magnitude of the skewness metric, note that the skewness of the weekly returns for S&P500 and AAPL from 2000 to 2019 is around -0.5 and -1, respectively.

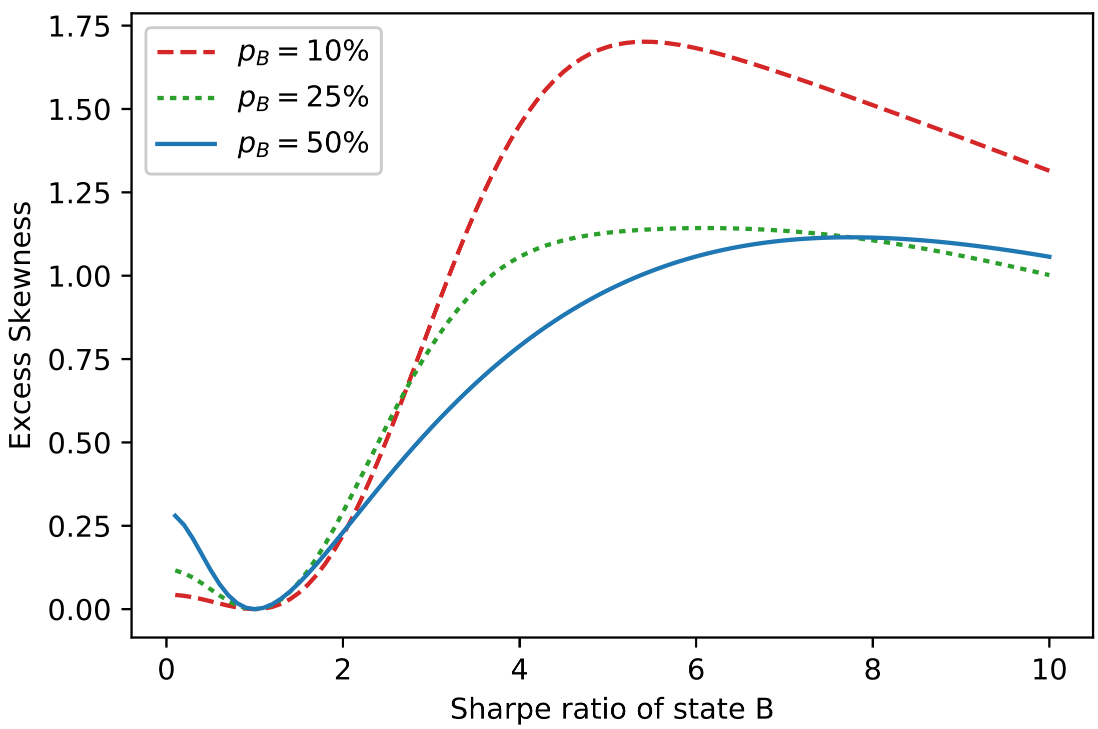

Figure 6: Difference in the long-term skewness of the returns of the clairvoyant and passive investors for different values of S_B (i.e. $SK_{clairvoyant} - SK_{passive}$). The passive investor maintains a constant amount of risk in the strategy at all time. The clairvoyant investor has perfect foresight on the state of Sharpe ratio at the beginning of each time period and implements optimal risk allocation. Here $\Delta t = 1/52$ and $S_A = 1$. The plots for three cases $p_B = 0.1$, $p_B = 0.25$, and $p_B = 0.5$ are shown.

3 Factor timing for multi-factor portfolio

In this section we provide evidence that if the investor is running a multi-factor portfolio, then it might be beneficial to do factor timing: actively and periodically adjust risk allocation across factors. Using simulation, we show that when the portfolio consists of many factors (e.g. around 10 factors or above; see figure 7), risk allocation based on instantaneous pair-wise correlations between the factors for each time interval improves performance above and beyond the passive approach of allocating risk based on long-term correlations between the factors. This magnitude of improvement grows with the number of factors in the portfolio.

Suppose the investor is running a portfolio of N factors. For simplicity of exposition assume each factor has an Sharpe ratio of 1. Furthermore, suppose over a long time period, any two factors are uncorrelated. More precisely, let $\lambda(N,\rho)$ denote the uniform distribution over the space of $N \times N$ correlation matrices whose off-diagonal entries are in $[-\rho,\rho]$ [3]. Assume the correlation matrices $\Omega(i)$ of the N factors for time period $(t_{i-1},t_i]$, with $1 \leq i \leq T$, are sampled independently from λ. For information on generating random correlation matrices and their properties see for example [2] and [3].

In this context, because the factors are uncorrelated over a long time period and have equal Sharpe ratios, the passive investor allocates equal amount of risk to each factor and does not change it over time (i.e. risk-parity allocation). Therefore, the ex-post Sharpe ratio of the passive investor after a long time has passed is given by

$$SR^{passive}_{N,\rho} = \sqrt{N}. \tag{8}$$

On the other hand, suppose the clairvoyant investor has perfect foresight on the value of the correlation matrix for each time interval and optimally allocates risk to factors at the beginning of each interval based on the prevailing correlation matrix. Explicitly, for $1 \leq i \leq T$, let $\vec{k}(i)$ denote the N-dimensional vector of risk allocated to factors at the beginning of time period $(t_{i-1},t_i]$. The clairvoyant investor determines $\vec{k}(i)$ by solving for the allocation that maximizes the ex-ante Sharpe ratio of her portfolio for each time period, that is,

$$\vec{k}(i) = \underset{\vec{k}}{\operatorname{argmax}} \frac{\vec{k}'\mathbb{1}}{\sqrt{\vec{k}'\Omega(i)\vec{k}}} \qquad \text{subject to} \qquad \vec{k} \geq 0. \tag{9}$$

Remark 3.1 *The constraint in* (9) *indicates that the investor does not go short any of the factors. This is to avoid corner situations where for some correlation matrices the optimal allocation is to be long (with a very high leverage) a group of factors and be short with (very high leverage) the remaining factors, which is not practical. If we do not impose the constraint in* (9)*, the Sharpe ratio improvement of the clairvoyant investor over the passive investor would be exorbitant but not realistic.*

Now, the following theorem provides an expression for the long-term ex-post Sharpe ratio of the multi-factor portfolio when the risk allocated to factors at the beginning of each time period is a function of the prevailing correlation matrix between the factors. In particular, ex-post Sharpe ratio of the clairvoyant investor over a long time period is given as a corollary of the next theorem.

[3]For the sake of notation, we use λ in place of $\lambda(N,\rho)$ when it is clear from the context.

Theorem 3.2 *Using the above notation, suppose the correlation matrices* $\Omega(i)$*, for* $1 \leq i \leq T$*, are sampled independently from distribution* λ*, and the risk allocated to factors at the beginning of each time period is a function of* $\Omega(i)$*. Furthermore, assume the returns of the factors over each time period are jointly normally distributed and independent of the returns over the other time periods. Then, for sufficiently large* T*, the (ex-post) Sharpe ratio of the portfolio over the time period* $[0, t_T]$ *is given by*

$$\frac{\mathbb{E}_\lambda[\vec{k}'\mathbb{1}]}{\sqrt{\mathbb{E}_\lambda[\vec{k}'\Omega\vec{k}] + \Delta t\, \mathrm{Var}_\lambda[\vec{k}'\mathbb{1}]}}, \tag{10}$$

where $\vec{k} = \vec{k}(\Omega)$ *denote the* N*-dimensional vector of risk allocated to factors based on correlation matrix* Ω*.*

Proof: See appendix A.

Corollary 3.3 *Using the above assumptions, the ex-post Sharpe ratio of the clairvoyant investor over long time period is given by*

$$SR^{clairvoyant}_{N,\rho} = \frac{\mathbb{E}_\lambda[\vec{k}'\mathbb{1}]}{\sqrt{\mathbb{E}_\lambda[\vec{k}'\Omega\vec{k}] + \Delta t\, \mathrm{Var}_\lambda[\vec{k}'\mathbb{1}]}}, \tag{11}$$

where $\vec{k} = \vec{k}(\Omega)$ *is the solution to the optimization problem* (9) *for given correlation matrix* Ω*.*

Figure 7 shows the improvement of $SR^{clairvoyant}_{N,\rho}$ over $SR^{passive}_{N,\rho}$ for several values of N and ρ. To calculate expression (11), we use Monte Carlo simulation, where the acceptance-rejection method is used to independently sample correlation matrix Ω according to distribution λ (see, for example, [1] for information on the acceptance-rejection method). As it can be seen, for small ρ the improvement is negligible: this is intuitive, because when ρ is small, the correlation matrix for each time period is close to identity; therefore, portfolio allocation of the clairvoyant investor becomes very similar to the passive investor. On the other hand, for moderate values of ρ (say circa 0.3-0.4), the performance improvement that the clairvoyant investor achieves over the passive investor grows quickly with N. Moreover, figure 7 indicates that for typical values of ρ, when the number of factors is less than 8, the improvement is below 30%. Therefore, benefits of factor timing for a multi-factor portfolio become apparent when it consists of many factors.

Acknowledgement

The author benefited greatly from insightful conversations with Nic Johnson, Masoud Sharif, and Hashim Zaman for this project.

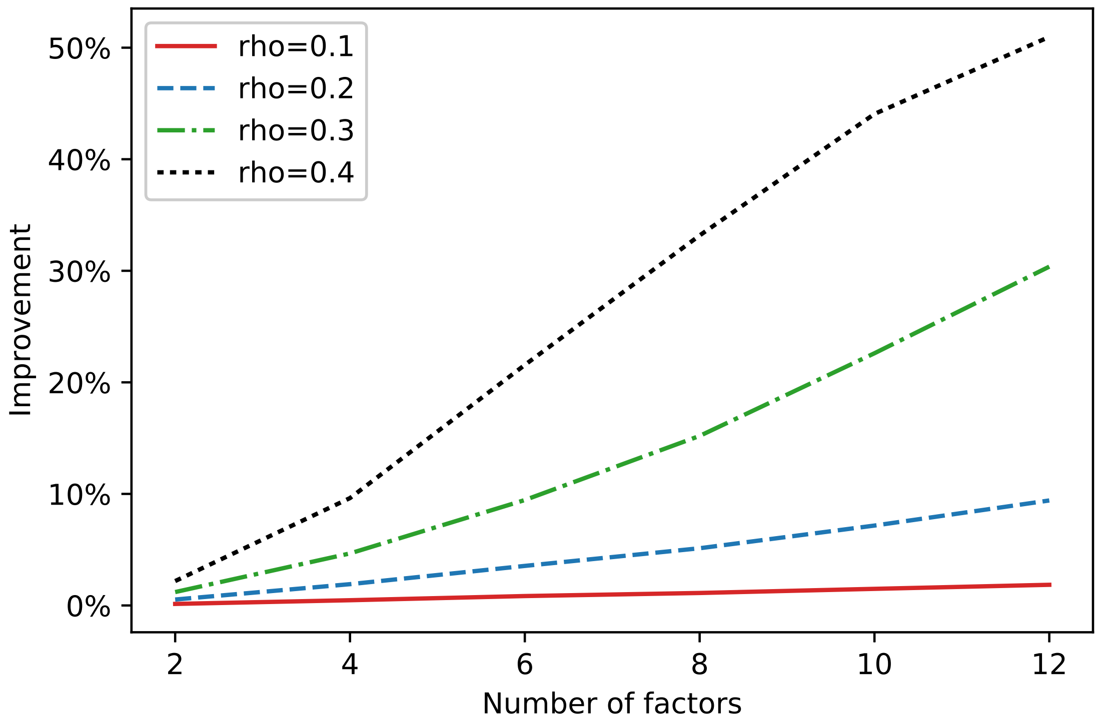

Figure 7: Improvement in the long-term Sharpe ratio of the multi-factor portfolio for the clairvoyant investor over the passive investor (i.e. $SR_{N,\rho}^{clairvoyant}/SR_{N,\rho}^{passive} - 1$). Here the clairvoyant investor dynamically adjusts risk allocation to factors based on instantaneous correlation matrix; whereas, the passive investor maintains constant risk allocation to factors over time. Here $\Delta t = 1/52$ and Monte Carlo simulation is used to calculate expression (11) where correlation matrix Ω is sampled uniformly from the space of $N \times N$ correlation matrices whose off-diagonal entries are in $[-\rho, \rho]$.

A Derivation of Theorems

Proof of Theorem 2.1: Let $k(i)$, for $1 \le i \le T$, denote the amount of risk that the investor allocates to the strategy at the beginning of time period $(t_{i-1}, t_i]$. Given the assumption of the theorem, the strategy's return over time period $(t_{i-1}, t_i]$ (conditioned on the period's Sharpe $S(i)$) is given by normal distribution

$$r(i)|S(i) \sim N(k(i)S(i)\Delta t, k(i)^2\Delta t).$$

With a slight abuse of the notation, write the above expression as

$$r|S \sim N(kS\Delta t, k^2\Delta t), \tag{A.1}$$

where k is determined by the value of S.

In particular,

$$\mathbb{E}[r|S] = kS\Delta t \qquad \text{and} \qquad \mathrm{Var}[r|S] = k^2\Delta t. \tag{A.2}$$

Let ν denote the empirical distribution for $\{t_i\}_{i=1}^{i=T}$. Observe that over the time period $[0, t_T]$, the annualized average return is given by

$$\frac{1}{\Delta t}\sum_{i=1}^{T} r(i)/T = \frac{1}{\Delta t}\,\mathbb{E}_\nu[r], \tag{A.3}$$

and the annualized standard deviation is given by

$$\sqrt{\frac{1}{\Delta t}\sum_{i=1}^{T}(r(i)-\bar{r})^2/T} = \frac{1}{\sqrt{\Delta t}}\sqrt{\mathrm{Var}_\nu[r]}. \tag{A.4}$$

Denote the distribution of S by μ, and taking expectation with respect to μ by $\mathbb{E}_\mu[\cdot]$. Note that by the law of large numbers, when T is sufficiently large, $\mathbb{E}_{\nu|\mu}[r|S]$ and $\mathrm{Var}_{\nu|\mu}[r|S]$ converge to the values in (A.2), respectively. Therefore, when T is sufficiently large,

$$\mathbb{E}_\nu[r] = \mathbb{E}_\mu[\mathbb{E}_{\nu|\mu}[r|S]] = \Delta t\,\mathbb{E}_\mu[kS], \tag{A.5}$$

where we used the law of total expectation for the first equality. Similarly, for sufficiently large T,

$$\mathrm{Var}_\nu[r] = \mathbb{E}_\mu\left[\mathrm{Var}_{\nu|\mu}[r|S]\right] + \mathrm{Var}_\mu\left[\mathbb{E}_{\nu|\mu}[r|S]\right] = \Delta t\,\mathbb{E}_\mu[k^2] + \Delta t^2\,\mathrm{Var}_\mu[kS], \tag{A.6}$$

where we used the law of total variance for the first equality.

Let SR denote the ex-post Sharpe ratio of the strategy over the time period $[0, t_T]$. When T is large enough, combining (A.3), (A.4), (A.5), and (A.6) yields that

$$\begin{aligned} SR &= \frac{\mathbb{E}_\mu[kS]}{\sqrt{\mathbb{E}_\mu[k^2] + \Delta t\,\mathrm{Var}_\mu[kS]}} \\ &\le \frac{\mathbb{E}_\mu[kS]}{\sqrt{\mathbb{E}_\mu[k^2]}} && \text{(A.7)} \\ &\le \sqrt{\mathbb{E}_\mu[S^2]}, && \text{(A.8)} \end{aligned}$$

where we used the positivity variance for the first inequality, and Cauchy–Schwarz Inequality for the second inequality. The result follows. ■

Proof of Theorem 2.2: In the proof of theorem 2.1, note that typically Δt is small (e.g. 1/52 for weekly re-balance); therefore, in most practical situations line (A.7) is close to an equality. In particular, when the investor takes a passive approach and allocates a constant amount of risk to the strategy at all time periods (i.e. $k = c$ for some constant c), her Sharpe ratio is slightly less than

$$SR_{passive} = \mathbb{E}_\mu[S].$$

Next, by the equality conditions of Cauchy–Schwarz Inequality, line (A.8) is an equality if and only if $k = cS$ for some constant c. In other words, the optimal risk allocation policy for the investor is to allocate $k(i) = cS(i)$ amount of risk to the strategy at time t_{i-1}. Of course, this requires the investor to have a precise estimate of the Sharpe of the strategy at time period t_{i-1}, that is the ex-ante Sharpe ratio $\hat{S}(i)$ must match the prevailing Sharpe ratio $S(i)$. And yet, even with the benefit of perfect foresight for knowing the Sharpe ratio of the strategy over each time period, the Sharpe ratio of this investor is slightly less than

$$SR_{clairvoyant} = \sqrt{\mathbb{E}_\mu[S^2]}.$$

Proof of Theorem 2.3: The proof follows a similar logic to the proof of theorem 2.1 and we utilize the notations described therein. Let $m_{3,\nu}(\cdot)$ denote the third central moment with respect to the empirical distribution ν. For T sufficiently large,

$$\begin{aligned} m_{3,\nu}(r) &= \mathbb{E}_\mu[m_{3,\nu|\mu}(r|S)] + m_{3,\mu}(\mathbb{E}_{\nu|\mu}[r|S]) + 3\operatorname{Cov}_\mu(\mathbb{E}_{\nu|\mu}(r|S), \operatorname{Var}_{\nu|\mu}(r|S)) \\ &= \mathbb{E}_\mu[0] + m_{3,\mu}(kS\Delta t) + 3\operatorname{Cov}_\mu(kS\Delta t, k^2\Delta t) \\ &= \Delta t^3 m_{3,\mu}(kS) + 3\Delta t^2 \operatorname{Cov}_\mu(kS, k^2), \end{aligned} \tag{A.9}$$

where we used the law of total cumulance (the special case for the third moment) for the first equality, and (A.2) for the second equality (because T is sufficiently large).

Let SK denote the skewness of the returns of the strategy over the time period $[0, t_T]$. When T is large enough, combining (A.6) and (A.9) yields that

$$SK = \frac{m_{3,\nu}(r)}{(\operatorname{Var}_\nu[r])^{3/2}} = \frac{3\operatorname{Cov}_\mu(kS, k^2) + \Delta t\ m_{3,\mu}(kS)}{(\mathbb{E}_\mu[k^2] + \Delta t \operatorname{Var}_\mu[kS])^{3/2}}\sqrt{\Delta t} \tag{A.10}$$

Substituting $k = c$ in (A.10) for the passive investor and simplifying yields equation (4). Substituting $k = cS$ in (A.10) for the clairvoyant investor and simplifying yields equation (5). This completes the proof. ■

Proof of Theorem 3.2: The idea of the proof is very similar to the proof of theorem 2.1; however, here instead of conditioning on the value of the Sharpe ratio, we condition on the value of the correlation matrix. Let $\vec{k}(i)$ denote the N-dimensional vector of risk allocated to factors at the beginning of time period $(t_{i-1}, t_i]$. By the assumption of the theorem, $\vec{k}(i)$ is a function of correlation matrix $\Omega(i)$. Let $r_\Pi(i)$ denote the portfolio's return over time period $(t_{i-1}, t_i]$. Given the assumption of the theorem, $r_\Pi(i)$ (conditioned on the periods correlation matrix $\Omega(i)$) is given by the normal distribution

$$r_\Pi(i)|\Omega(i) \sim N(\vec{k}(i)\mathbb{1}\Delta t, \vec{k}(i)'\Omega(i)\vec{k}(i)\Delta t)$$

With a slight abuse of the notation, write the above expression as

$$r_\Pi|\Omega \sim N(\vec{k}'\mathbb{1}\Delta t, \vec{k}'\Omega\vec{k}\Delta t), \tag{A.11}$$

where $\vec{k}$ is a function of Ω.

In particular,

$$\mathbb{E}[r_\Pi|\Omega] = \vec{k}'\mathbb{1}\Delta t \qquad \text{and} \qquad \mathrm{Var}[r_\Pi|\Omega] = \vec{k}'\Omega\vec{k}\Delta t. \tag{A.12}$$

Let ν denote the empirical distribution for $\{t_i\}_{i=1}^{i=T}$. Observe that over the time period $[0, t_T]$, the annualized average return is given by

$$\frac{1}{\Delta t}\sum_{i=1}^{T} r_\Pi(i)/T = \frac{1}{\Delta t}\,\mathbb{E}_\nu[r_\Pi], \tag{A.13}$$

and the annualized standard deviation is given by

$$\sqrt{\frac{1}{\Delta t}\sum_{i=1}^{T}(r_\Pi(i) - \bar{r}_\Pi)^2/T} = \frac{1}{\sqrt{\Delta t}}\sqrt{\mathrm{Var}_\nu[r_\Pi]}. \tag{A.14}$$

Note that by the law of large numbers, when T is sufficiently large, $\mathbb{E}_{\nu|\lambda}[r_\Pi|\Omega]$ and $\mathrm{Var}_{\nu|\lambda}[r_\Pi|\Omega]$ converge to the values in (A.12), respectively. Therefore, when T is sufficiently large,

$$\mathbb{E}_\nu[r_\Pi] = \mathbb{E}_\lambda[\mathbb{E}_{\nu|\lambda}[r_\Pi|\Omega]] = \Delta t\,\mathbb{E}_\lambda[\vec{k}'\mathbb{1}], \tag{A.15}$$

where we used the law of total expectation for the first equality. Similarly, for sufficiently large T,

$$\mathrm{Var}_\nu[r_\Pi] = \mathbb{E}_\lambda\left[\mathrm{Var}_{\nu|\lambda}[r_\Pi|\Omega]\right] + \mathrm{Var}_\lambda\left[\mathbb{E}_{\nu|\lambda}[r_\Pi|\Omega]\right] = \Delta t\,\mathbb{E}_\lambda[\vec{k}'\Omega\vec{k}] + \Delta t^2\,\mathrm{Var}_\lambda[\vec{k}'\mathbb{1}], \tag{A.16}$$

where we used the law of total variance for the first equality.

Let SR denote the ex-post Sharpe ratio of the multi-factor portfolio over time period $[0, t_T]$. When T is large enough, combining (A.13), (A.14), (A.15), and (A.16) yields that

$$SR = \frac{\mathbb{E}_\lambda[\vec{k}'\mathbb{1}]}{\sqrt{\mathbb{E}_\lambda[\vec{k}'\Omega\vec{k}] + \Delta t\,\mathrm{Var}_\lambda[\vec{k}'\mathbb{1}]}}$$

This completes the proof. ∎

References

[1] R.E. Caflisch, *Monte Carlo and Quasi-Monte Carlo Methods*, Acta Numer- ica (1998), pp. 149.

[2] S. Ghosh, S.G. Henderson, *Behavior of the norta method for correlated random vector generation as the dimension increases*, ACM Transactions on Modeling and Computer Simulation (TOMACS), 13 (3) (2003), pp. 276-294.

[3] D. Lewandowski, D. Kurowicka, H. Joe, *Generating random correlation matrices based on vines and extended onion method*, Journal of Multivariate Analysis, 100 (9) (2009), pp. 1989-2001.