

FE630 - Final Project

Author: Sid Bhatia

Date: May 14th, 2024

Pledge: I pledge my honor that I have abided by the Stevens Honor System.

Professor: [Papa Momar Ndiaye](#)

1. Overview

1.1 Goal

The goal of this project to build and compare *two factor-based long short allocation models* with constraints on their *betas*. The first strategy considers a **target Beta** in the interval $[-0.5, 0.5]$, while the second has a target Beta in the interval $[-2, +2]$.

The first strategy operates similar to a **Value-at-Risk Utility** corresponding to **Robust Optimization**; the second strategy incorporates an **Information Ratio** term to limit the deviations from a benchmark, provided those deviations yield a 'high return.'

Once the optimization models are built, we want to *compare* the outcomes of the two models while simultaneously evaluating their sensitivity to the *length* of the estimators for the **covariance matrix** in tandem with the **expected returns** under various market regimes/scenarios.

1.2 Reallocation

The portfolios will be *reallocated* or, in other words, 'reoptimized' weekly from the beginning of **March 2007** to the end of **March 2024**. Our *investment universe* encompasses a set of exchange-traded funds (**ETFs**) which is large enough to represent the '**Global World Economy**' (as according to some).

We will utilize the [Fama–French Three-Factor Model](#) which incorporates the following factors:

- [Momentum](#)
- [Value](#)
- [Size](#).

Regarding data accessibility, these factors have historical values available for **free** from **Ken French's** [personal website](#) in tandem with Yahoo Finance.

1.3 Performance Evaluation

Naturally, the performance as well as the risk profiles of the aforementioned strategies may be (relatively) sensitive to the *target Beta* and the (current) market environment.

For example, a '**low Beta**' (essentially) means that a strategy is created with the objective or aim to be '**decorrelated**' (no linear relationship between entites) with the 'Global Market,' which, in our case, is represented by the **S&P 500** (i.e., no *systematic relationship*).

A '**high Beta**' is simply the antithesis, or opposite, of what we just discussed. In layman's terms, we have a (higher) appetite for '*risk*' (in this case, let's keep it simple and define our premise as σ or **standard deviation**) and desire to ride or 'scale up' the *market risk* (**systematic risk**).

Moreover, it's imperative that one acknowledges that such a (described) strategy is more probable to be (quite) sensitive to the *estimators* used for the **Risk Model** and the **Alpha Model** (e.g., the length of the *look-back period* utilized); therefore, it is necessary to understand and, most importantly, *comprehend* the impact of said estimators on the **Portfolio's** characteristics:

- (Realized) **Return** : μ_h
- (Historical) **Volatility** : σ_h
- **Skewness** : $(\mathbb{E}[(\frac{x-\mu}{\sigma})^3]) = \frac{\mu_3}{\sigma^3} = \frac{\kappa_3}{\kappa_2^{3/2}}$
- **VaR / Expected Shortfall**
- **Sharpe Ratio** : $S_a = \frac{\mathbb{E}[R_a - R_b]}{\sigma_a} = \frac{\mathbb{E}[R_a - R_b]}{\sqrt{\mathbb{V}(R_a - R_b)}}$

1.4 Simplification

To make it easier, we assume that once the **Factor Model** (FM) has been constructed, we will use **trend following** estimators for the **Expected Returns**. Since the quality of the estimators depend on the **look-back period**, we define three cases:

- **Long-Term Estimator (LTE)** : $LT \Rightarrow LB \in \{180 \text{ Days}\}$.
- **Mid-Term Estimator (MTE)** : $MT \Rightarrow LB \in \{90 \text{ Days}\}$.
- **Short-Term Estimator (STE)** : $ST \Rightarrow LB \in \{40 \text{ Days}, 60 \text{ Days}\}$.

Specifically, we define a **Term-Structure** for the Covariance Σ and Expected Return μ .

1.5 Synthesis

To (briefly) summarize, the behavior of a (potential) '*optimal*' portfolio built from a melting pot of *estimators* for **Covariance** and **Expected Return** may vary according to the cadence

of the '**Market**' (environment/regime) or an aforementioned strategy.

For example, the (mathematical) notation S_{40}^{90} is just fancy jargon to visually illustrate that we are using **40 days** for the covariance estimation and **90 days** for the expected returns estimations—it's not that deep.

Overall, the goal of this fun, entertaining project is to conceptualize, visualize, understand, analyze, and compare the behavior of our ideas; we want to see if we can (actually) make some \$\$\$, especially during momentous, historical (time) periods such as the **Subprime Mortgage Crisis** of 2008, the horrendous commencement of **Coronavirus SARS-CoV-2 Disease** of 2019, et cetera.

2. (Investment) Strategy

Alrighty, let's get to the fun, juicy portion; shall we?

2.1 (Mathematical) Strategic Formulation

Let's make things interesting—spicy, one may say.

Consider two strats [(clipping) of 'strategies,' as embodied in *Morphology*]:

$$(\text{Strategy I}) \quad \left\{ \begin{array}{l} \max_{\omega \in \mathbb{R}^n} \rho^T \omega - \lambda \sqrt{\omega^T \Sigma \omega} \\ -0.5 \leq \sum_{i=1}^n \beta_i^m \omega_i \leq 0.5 \\ \sum_{i=1}^n \omega_i = 1, \quad -2 \leq \omega_i \leq 2, \end{array} \right. \quad (1)$$

and

$$(\text{Strategy II}) \quad \left\{ \begin{array}{l} \max_{\omega \in \mathbb{R}^n} \frac{\rho^T \omega}{\text{TEV}(\omega)} - \lambda \sqrt{\omega^T \Sigma \omega} \\ -2 \leq \sum_{i=1}^n \beta_i^m \omega_i \leq 2 \\ \sum_{i=1}^n \omega_i = 1, \quad -2 \leq \omega_i \leq 2, \end{array} \right. \quad (2)$$

where we define the hieroglyphics used above:

- Σ is the **covariance matrix** between the securities returns (as computed from the **FF3FM**);
- $\beta_i^m = \frac{\text{Cov}(r_i, r_M)}{\sigma^2(r_M)}$ is the **Beta** (not to be confused with the **colloquial slang** usage) of some **security**) S_i as defined by the **CAPM Model** such that $\beta_P^m = \sum_{i=1}^n \beta_i^m \omega_i$ is the **Portfolio Beta**;

- $\text{TEV}(\omega) = \sigma(r_P(\omega) - r_{\text{SPY}})$ is the '**Tracking Error Volatility**', which (if you're *really nerdy*) you can derive it as such:

$$\sigma(r_P(\omega) - r_{\text{SPY}}) = \sqrt{\omega^\top \Sigma \omega - 2\omega^\top \text{Cov}(r, r_{\text{SPY}}) + \sigma_{\text{SPY}}^2} \quad (3)$$

Oh yeah, I should probably define what '**FF3FM**' means; that would (probably) be helpful.

2.2 Fama–French Three-Factor Model

So, to echo the previous sentiment, we should (almost surely) explain what is this *funky* model we kept referencing:

$$r_i = r_f + \beta_i^3(r_M - r_f) + b_i^s r_{\text{SMB}} + b_i^v r_{\text{HML}} + \alpha_i + \epsilon_i \quad (4)$$

Sorry for writing (or, to be *really technical*, **typesetting**) more hieroglyphics. We gotta keep going for a bit—stay with me!

If we assume our **white noise/error terms**, on 'average', have a (numerical) value of 0 (i.e., $\mathbb{E}[\epsilon_i] = 0$), we can derive a new goofy equation:

$$\rho_i = r_f + \beta_i^3(\rho_M - r_f) + b_i^s \rho_{\text{SMB}} + b_i^v \rho_{\text{HML}} + \alpha_i \quad (5)$$

In the new **cursive script** defined above, the 3 coefficients β_i^3 , b_i^s , and b_i^v are estimated by making a **linear regression**, or, in 'plain English', drawing a **line of best fit** of the **time series** $y_i = \rho_i - r_f$ against the other cool time series $\rho_M - r_f$ (**Momentum Factor**), r_{SMB} (**Size Factor**), and ρ_{HML} (**Value Factor**).

I feel like I'm forgetting something . . .

Oh yeah! There's an extra (nerdy) thingy we gotta verify: (generally), $\beta_i^m \neq \beta_i^3$ and needs to be estimated by a separate regression or directly computed.