FRE6711 Quantitative Portfolio Management
Memo for the Final Project

Due date for the project: 05/05/2021

Prof. Papa Momar Ndiaye

April 6, 2021

1 Overview

This project aims 1) to build a factor-based model allocation namely a Long/Short Global Macro Strategy with a Target Beta and 2) to evaluate its sensitivity to variations of Beta and its sensitivity to the length of the estimators for covariance matrix and the expected returns under different market scenarios.

Students may work individually or in small teams (typically up to 3 people). Each team is required to build a investment strategy that maximizes the return of the portfolio subject to a constraint of target Beta, where Beta is the usual single factor Market risk measure. The portfolio will be reallocated (re-optimized every week for period of analysis from March 2007 to end of March 2021. For practical considerations, we will assume that our universe of investment is a set of ETFs large enough to represent the World global economy and that our factor model is the French Fama 3-factor model. The data needed for the project are freely available for download from Ken French's website for the factors historical values (http://mba.tuck.dartmouth.edu/pages/faculty/ken.french) and from Quandl (www.quandl.com) or yahoo for the ETFs.

The performance and the risk profiles of such a strategy may be very sensitive to the target Beta and the market environment. A low Beta meaning a strategy that aims to be de-correlated to the global market represented by the S&P 500, while a high Beta meaning that, having a big appetite for risk, we are aiming to ride or scale up the market risk. In addition to that, such a strategy is likely to to be quite sensitive to the estimators used for the Risk Model and the Alpha Model (for example the length of the look-back period used for estimation risk and expected returns), so it is important to understand the impact of those estimators on the Portfolio's characteristics: realized return, volatility, skewness, VaR/ CVaR and risk to performance ratios such as the Sharpe ratio.

To simplify, we will assume in this project that once the factor model built, we will use trend following estimators for the Expected returns, namely the sample mean and the sample covariance. As the quality of those sample estimators depend on the length of the look-back period, we will typically consider 3 cases: a long look-back period (≥ 120 days), a short look-back period (≤ 40 days) and a medium look-back period, and use the notational conventions Long-Term estimators (LT), Short-Term estimators (ST) and Mid-Term estimators (MT), defining therefore a Term-Structure for the Covariance and Expected Return. A similar remark on the dependance on the look-back period can be made for the estimation of the coefficients of the model as they are computed using a regression on the factors. Consequently, a central question is to assess properly the impact of the length of those regression-based estimators on the realized performance and risk indicators of the optimized portfolio.

In summary, the behavior of the optimal portfolio built from a specific combination of estimators for Covariance and Expected Return may change with the Market environment and the target Beta (a particular strategy being defined by a specific combination, for example the notation $S_{60}^{90}(0.5)$ - can be used to say that you are using 60 days for estimation of covariance, 90 days for estimation of Expected Returns and a target $\beta = 0.5$). The goal of this project is to understand, analyze and compare the behavior of strategies built using chosen combinations of return/risk estimators and Target Beta during several historical periods : before the subprime (2008) crisis, during that crisis and after the crisis. The factor model we will use, known as the French Fama 3-factor model has 3 factors, Momentum, Value and Size.

2 Investment Strategy

We will consider an portfolio optimization problem of the form:

$$\begin{cases} \max\limits_{\omega \in \mathbb{R}^n} \rho^T \omega - \lambda(\omega - \omega_p)^T \Sigma (\omega - \omega_p) \\ \sum\limits_{i=1}^{n} \beta_i^m \omega_i = \beta_T^m \\ \sum\limits_{i=1}^{n} \omega_i = 1, \ -2 \le \omega_i \le 2, \end{cases} \tag{1}$$

where

- Σ is the the covariance matrix between the securities returns (computed from the Factor Model), ω_p is the composition of a reference Portfolio (the previous Portfolio when rebalancing the portfolio and ω_p has all its components equal to $1/n$ for the first allocation) and λ is a small regularization parameter to limit the turnover;
- $\beta_i^m = \dfrac{cov(r_i, r_M)}{\sigma^2(r_M)}$ is the Beta of security S_i as defined in the CAPM Model so that $\beta_P^m = \sum\limits_{i=1}^{n} \beta_i^m \omega_i$ is the Beta of the Portfolio;
- β_T^m is the Portfolio's Target Beta, for example $\beta_T^m = -1$, $\beta_T^m = -0.5$, $\beta_T^m = 0$, $\beta_T^m = 0.5$, $\beta_T^m = 1$, $\beta_T^m = 1.5$.

The French Fama factor models are well documented in the literature but reminded here for reference. For instance, under the 3-factor model, the random return of a security is given by the formula

$$r_i = r_f + \beta_i^3(r_M - r_f) + b_i^s r_{SMB} + b_i^v r_{HML} + \alpha_i + \varepsilon_i \tag{2}$$

with $\mathbb{E}(\varepsilon_i) = 0$ in such a way that we have in terms of Expected Returns

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

In equation (2), the 3 coefficients β_i^3 , b_i^s and b_i^v and estimated by making a linear regression of the time series $y_i = \rho_i - r_f$ against the time series $\rho_M - r_f$ (Momentum Factor), r_{SMB} (Size Factor) and ρ_{HML} (Value Factor)[1]. Note that in general $\beta_i^m \neq \beta_i^3$ and needs to be estimated by a separated regression or computed directly.

3 Investment Universe and Analysis Setup

3.1 Investment Universe

We will consider the following set of ETFs that you can download from Yahoo, Google or Quandl from March 1^{st}, 2007 to June 30^{th}, 2020.

1. CurrencyShares Euro Trust (FXE)
2. iShares MSCI Japan Index (EWJ)
3. SPDR GOLD Trust (GLD)
4. Powershares NASDAQ-100 Trust (QQQ)
5. SPDR S&P 500 (SPY)
6. iShares Lehman Short Treasury Bond (SHV)
7. PowerShares DB Agriculture Fund (DBA)
8. United States Oil Fund LP (USO)
9. SPDR S&P Biotech (XBI)
10. iShares S&P Latin America 40 Index (ILF)
11. iShares MSCI Pacific ex-Japan Index Fund (EPP)
12. SPDR DJ Euro Stoxx 50 (FEZ)

[1] ρ_M for Market hence Momentum, ρ_{SMB} for Small minus Big and ρ_{HML} for High minus Low

3.2 Benchmark

The benchmark will be the Market Portfolio S&P 500 (SPY ETF)

3.3 Analysis Periods and Bactkesting

- Divide the overall analysis period into 3 sub-periods: before, during and after the subprime crisis.
- Run separate backtests for each sub-period when comparing strategies. Note that there are different angles for such comparison: (1) Impact of Beta Target on a given Term-Structure: compare $S_{40}^{200}(\beta_{T1})$ to $S_{40}^{200}(\beta_{T2})$ or (2) impact of various term structure given Beta: for example, compare $S_{40}^{200}(0.5)$ to $S_{40}^{90}(0.5)$.
- Run also a comparison over the whole period from March 1^{st}, 2007 to March 30^{th}, 2021.
- For your backtesting, rebalance your portfolios once a week.

3.4 Important remark for BackTesting

A backtesting is non anticipative. That means that if you testing a strategy at given date t, you should only use past data as inputs for your optimization.

- For example, assume that you you decide to backtest a strategy with 60 days of historical data for the covariance, 90 days for the returns. In that case you need to estimate the Betas, Covariance and expected returns excluding all information for dates greater or equal to t.
- In the case of weekly rebalancing, assume that you generate a new portfolio every week. Then you will have a run a new optimization every 5 days, say at a given sequence of dates $t_1, t_2, t_3, \ldots, t_n$. For first date t_i, use historical data to estimate all inputs, run optimization, and store the weights. For the next date, roll the historical data window, re-estimate your inputs and generate new weights. Repeat the process until you reach the date t_n.

3.5 Performance and Risk Reporting for comparing Strategies

Use a Performance Analytics Module in R, Matlab or Python as much as possible for the Risk and Performance Reporting. Below is the list of Key Indicators to report your Optimal Strategies. All daily indicators will be annualized assuming that each year has 250 business days. For example, the Reporting for a Strategy over a given period (example: from 03/01/1997 to 12/30/2008) will be provided by a summarizing Table with the following lines

- Cumulated PnL or Return
- Daily Mean Arithmetic / Geometric Return, Daily Min Return
- Max 10 days Drawdown
- Volatility
- Sharpe Ratio
- Skewness, Kurtosis
- Modified VaR, CVaR

In addition to that table:

1. Plot the evolution the graph of cumulated daily Profit and Loss (PnL) assuming that you invest \$100 at the first allocation date in Portfolio and \$100 in the S&P 500 (when you benchmark your strategies against the Market, the SPY is representing the S&P500 Index).
2. Plot and analyze the distribution of daily Returns.
3. A summarizing Table with the following lines for comparison with the underlying

   For comparison with between the strategies and the S&P, a summary table may look like:

	$S^{90}_{60}(\beta^m_T = 0.5)$	$S^{30}_{120}(\beta^m_T = 1)$	$S^{90}_{180}(\beta^m_T = 0)$	SPY
Mean Return			12	
⋮				
Max DD			8	

3.6 Tools

- Data can be downloaded from R, Matlab or Python using the APIs provided by Quandl. Alternatively, you may use native functions when available, for example in R, "get.hist.quote"[2], or "get_data_yahoo" using Pandas in Python.
- The strategies will be implemented using the Quadratic Solver in R, Matlab or Python.

[2] $fxe < -get.hist.quote(instrument = "fxe", start = "2007-01-01", end = "2018-12-01", quote = "Close")$

4 Submission of the Final Report

You have to submit the following.

1. A final report can be a Word, Latex File or PPT slides presenting your findings and conclusions about the impact of the estimators on the behavior of your strategy, and also what kind of estimators would recommend to use, when and why (before, during and after the crisis). So to repeat again, a global period of backtest from 2007 to March 2021, with 3 sub-periods (before, during and after the crisis) and 2 axes of analysis: sensitivity to the term-structure of estimators (short-term, mid-term and long-term) for covariance and expected returns and sensitivity to β.

2. The report should contain a clear description of the notations, models and strategies you have analyzed, the graphs and summarizing tables supporting your quantitative and qualitative analysis. You can include a brief description of the computational engine you have built but do not include any code or Rmarkdown output in the core of your report.

3. It is mandatory to submit also the code developed for the project (R, Matlab, Python or other) and all supporting graphs, tables and simulation results in a Zip file. The code should ready to run when unzipped and with minimal directions to the evaluators. The submitted code will be tested for comparison and it is a requirement to build your code in a modular and clearly documented manner..

Appendix

A Practical aspects

For estimation of the parameters of the factor model, you can use a cross sectional regression model by gathering all the individual securities model in a single "big" factor model. If you assume, that you have 3 factors, then the model at time t for each asset is given by

$$r_{it}^{e} = \alpha_i + \beta_i^3 (r_{Mt} - r_{ft}) + b_i^s r_{SMBt} + b_i^v r_{HMLt} + \alpha_i + \varepsilon_{it} \qquad (4)$$

with $r_{it}^{e} = r_{it} - r_{ft}$, and moreover the ε_{it} are independent of the factors and satisfy

$$cov(\varepsilon_{it}, \varepsilon_{js}) = \begin{cases} \sigma_i^2 \text{ when } i = j, \text{ and } t = s \\ 0 \text{ otherwise.} \end{cases}$$

A.1 Time Series Model for a given Security

If we consider T observations of the excess return of Security S_i stacked in a column vector $R_i = \begin{bmatrix} r^e_{i1} \\ r^e_{i2} \\ \cdots \\ r^e_{iT} \end{bmatrix}$, we have the time series regression model for Security i:

$$R_i = \mathbf{1}_T \alpha_i + F\beta_i + \varepsilon_i \text{ for } i = 1, 2, \ldots n \tag{5}$$

where

- $\beta_i = \begin{bmatrix} \beta^3_i \\ b^s_i \\ b^v_i \end{bmatrix}$ is the (3 by 1) vector of Betas
- $\mathbf{F} = \begin{bmatrix} \mathbf{f}_1' \\ \vdots \\ \mathbf{f}_T' \end{bmatrix} = \begin{bmatrix} r_{M1} - r_{f1} & r_{SMB1} & r_{HML1} \\ \vdots & \ddots & \vdots \\ r_{MT} - r_{fT} & r_{SMBT} & r_{HMLT} \end{bmatrix}$ is the is the (T x 3) matrix of observations of the factors.
- the residual term ε_i is a (T by 1) vector satisfying $\mathbb{E}(\varepsilon_i \varepsilon_i') = \sigma_i^2 \mathbb{I}_T$

The previous model is well-suited for a regression to estimate the coefficients of the model using data for the securities and the factors.

A.2 Cross Sectional Model

Alternatively, we can use a cross sectional formulation that can be useful for risk analysis including the derivation of the covariance matrix of the returns. If we define $R_t = \begin{bmatrix} r^e_{1t} \\ r^e_{2t} \\ \cdots \\ r^e_{nt} \end{bmatrix}$, the vector of all Securities excess returns at time t, then we can write

$$R_t = \alpha + \mathbf{B}\mathbf{f}_t + \varepsilon_t \text{ for } t = 1, 2, \ldots T, \tag{6}$$

where

- $B \begin{bmatrix} \beta_1' \\ \beta_2' \\ \cdots \\ \beta_n' \end{bmatrix} = \begin{bmatrix} \beta^3_1 & b^s_1 & b^v_1 \\ \vdots & \ddots & \vdots \\ \beta^3_n & b^s_n & b^v_n \end{bmatrix}$ is a N by 3 matrix,

- $\mathbf{f}_t = \begin{bmatrix} r_{Mt} - r_{f1} \\ r_{SMBt} \\ r_{HMLt} \end{bmatrix}$ is the vector of factor returns at time t.
- $\mathbb{E}(\varepsilon_t \varepsilon_t^{'} | \mathbf{f}_t) = D = diag(\sigma_1^2, \ldots, \sigma_n^2)$

The cross sectional model implies that if Ω_f is the covariance of the factors, then

$$cov(R_t) = \mathbf{B}\Omega_f\mathbf{B}^{'} + D \tag{7}$$

which implies that

$$cov(R_{it}) = \beta_i \Omega_f \beta_i + \sigma_I^2 \tag{8}$$

and

$$cov(R_{it}, R_{jt}) = \beta_i \Omega_f \beta_j \tag{9}$$