FIN42120 Programming for Financial Data Science

Group Project Assignment Details

(Please read all instructions very carefully)

Due Date: Friday, 12th May 2023 at 5pm

(11th April 2023)

Submit your Python source code and write-up of your results/answers (succinctly described and as a PDF file) via the dedicated submission folder under Assignments in the Assessment section on Brightspace. Source code can be separated into multiple files and function should be appropriately named. Please be sure that all tables, figures, and code are well annotated with appropriate information to help the reader discern what you are doing. The submitted report should be attached to the appropriate submission form, which should closely adhere to the template provided in Brightspace and clearly provide all information required therein, including the assignment number, the team number, the name and student number of each team member and details on their individual contribution.[1]

1. Download monthly prices of the S&P 500 stock market index (SP500) from the website of the U.S. Federal Reserve Economic Data (https://fred.stlouisfed.org/) or Bloomberg and the Bloomberg Barclays U.S. Aggregate Bond Index (LBUSTRUU) from Blomberg for the period from December 1979 to December 2021. Download monthly data on the risk-free rate of return (among the so-called Fama and French factors) for the same period from Professor Kenneth French's data library at http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html.

   Write a Python function to compute the following statistics for U.S. stock and bond simple excess returns:

   (a) annualized mean/average
   (b) annualized volatility
   (c) annualized Sharpe ratio
   (d) skewness
   (e) kurtosis

   Report the summary statistics in a table and explain the results.

2. Divide the total sample into two periods: an in-sample from January 1980-December 1999 and an out-of-sample period from January 2000-December 2021. Using a recursive estimate approach, generate a time-series of monthly out-of-sample constant

[1] Please note that this is a requirement and, if not fulfilled, I will not accept the submission, applying any relevant late submission penalty until a proper submission is made.

expected (mean) excess return forecasts (call it the *mean benchmark forecast*) for each of the two asset classes. Write a function to do this computation.

3. Download monthly data for five variables that you believe are plausible predictors (see, for example, Rapach, Ringgenberg, and Zhou, 2016 for stocks, Lin, Wu, and Zhou, 2017, and the references therein for inspiration) of the asset class excess returns. Based on each of your five predictors and using the same recursive estimation approach, generate monthly out-of-sample excess return forecasts for each of the asset classes using the following three predictive models:

   (i) an OLS predictive regression model for each of the five predictors
   (ii) a combination forecasts of excess returns that is a simple averages of the forecasts based on the five predictors from the OLS model
   (iii) any 2 penalized linear regressions

   The above leads to a total of 8 predictive model forecasts each for stock and bonds. Compute the mean squared forecast error (MSFE) for the benchmark forecast and the ratio of MSFEs for the forecasts of the models relative to the mean benchmark forecast MSFE. Compare the performance of the predictive model forecasts relative to the mean benchmark forecasts using the Diebold and Mariano (1995) test for equal predictive ability (you should write your own function to perform this test). You should clearly state the null hypothesis being tested and also provide a discussion of your table of results. In addition, create a figure showing the time-series of the mean benchmark, combination, and two penalized linear regression excess return forecasts for each of the two asset classes.

4. Generate the out-of-sample forecasts of the (2-by-2) sample variance-covariance matrix for a portfolio of the two asset classes using the same recursive estimation window approach.

5. Using the mean benchmark excess return forecasts and sample variance-covariance matrix forecasts, construct out-of-sample optimal tangency portfolio weights for a mean-variance investor. Compute the annualized summary statistics (mean, volatility, and Sharpe ratio) for the optimal portfolio's excess return. Repeat the exercise with the 8 predictive model excess return forecasts in place of the mean benchmark excess return forecasts and compute the annualized summary statistics for the alternative optimal portfolio's excess return.

   Report the summary statistics in a table and explain the results. Does any statistical evidence of out-of-sample forecasting performance translate into economic gains/significance? In addition, create a figure showing the time-series of portfolio weights and cumulative excess returns for the optimal portfolio based on the

mean benchmark forecast and those based on the combination and penalized linear regression excess return forecasts.

6. Repeat all the above tasks using a rolling window estimation approach. Comment on your results.

References

Diebold, F. X., Mariano, R. S., 1995. Comparing predictive accuracy. Journal of Business & economic statistics 13, 253–263.

Lin, H., Wu, C., Zhou, G., 2017. Forecasting corporate bond returns with a large set of predictors: An iterated combination approach. Management Science .

Rapach, D. E., Ringgenberg, M. C., Zhou, G., 2016. Short interest and aggregate stock returns. Journal of Financial Economics 121, 46–65.