FBE 551 Homework #5

Due on 11/22/2025 by 11:59 pm

This is a group homework assignment. You should work in your teams, include all student names in a markdown cell at the top of the Jupyter Notebook that you submit.

Carry out the following calculations in a single Jupyter notebook. Make sure that I will be able to run it. For example, read in the files from the current directory. I.e., do not write the full path of the file in your code. Do not have your code change directories to a pathway that is probably not valid on my computer. This will allow me to run your notebook in any directory on my own PC.

Please use markdown cells to help me understand what you are doing. Clearly highlight what parts of your notebook are used to answer each question below.

Question 1

The file CRSP_Indexes.csv file contains daily adjusted closing prices for the following ETFs:

Large_Value:	large cap value stocks
Large_Growth:	large cap growth stocks
Small_Value	small cap value stocks
Small_Growth	small cap growth stocks

a) Using the FamaFrenchDaily.csv file compute the excess returns on all four ETFs. Make sure when merging to only include observations that exist in the CRSP_Indexes data. Calculate the following for each of the ETFs: annualized mean excess return, annualized standard deviation of the excess return, Sharpe ratio and t-stat for the excess return. Note: do NOT use log returns or geometric returns.

b) Now consider a strategy where you allocate 25% of your capital to each of these ETFs every day. What is the annualized excess return, annualized standard deviation of the excess return, Sharpe ratio and t-stat of the excess return of this strategy?

c) Now consider a strategy in which you allocate 100% of your portfolio to the ETF if its lagged price is above its 200-day moving average, otherwise you allocate 100% to cash and get the risk-free rate. Construct this timing strategy for each of the ETFs individually (Hint: refer to Meb Faber's S&P 500 timing strategy in Week 10). Calculate the annualized mean excess return, annualized standard deviation of the excess return, Sharpe ratio, and t-stat for the excess return for each of the four timing strategies. Which ETF seems to be the best options for implementing a timing strategy over this sample?

Question 2

Data for this question is from the Monthly tab of the PredictorData2024.xlsx file. It was downloaded from Amit Goyal's website and is an updated version of the data Goyal used in his 2011 paper with Ivo Welch on return predictability. If you open the file in excel there is a "read me" tab that describes each field. You can read in this data into Python using pd.read_excel.

The worksheet contains data that can be used to construct a number of predictors. These include:

- The D/P ratio (d/p: column N).
- The term spread (tms: column S)
- The default yield spread (dfy: column T)
- Net stock issuance (ntis: column X)

The sheet also contains market returns (ret) in column E and risk-free returns (Rfree) in column M. Use them to compute the excess return on the market.

a) Load the data from PredictorData2024.xlsx into a dataframe and drop all observations before 1950-01-01, this is our starting date. Also, keep only the variables that wer are interested in (ret, Rfree, d/p, tms, dfy, ntis) and rename "d/p" to "dp" so that the statsmodels package doesn't run into issues with the "/". Make a new variable that is the excess market return $ER_t = (ret_t - Rfree_t)$.

b) We are going to use the four predictor variables above to implement a market timing strategy. The dependent variable is ER_t and the independent variables will be the lagged predictors $dp_{t-1}, tms_{t-1}, dfy_{t-1}, ntis_{t-1}$. The framework we will use will be similar to the "walk forward" analysis we covered in Week 10: run a regression using 10 years of data to predict the next month's excess market return. Start with the first 10 years at the beginning of the sample, run a regression to predict the excess market return for the next month. Collect the prediction for that window then roll the 10-year window forward one-month and repeat. Do this until you have a predicted excess market return for every month in the sample. All regressions should use exactly 10 years of data. Since you must lag the independent variables, you will have 119 observations in each rolling window regression. Your first prediction will therefore be made in December of 1959, when you will be predicting the January 1960 return. Move the sample forward one month and repeat.

c) Make a new dataframe that just contains the excess returns ER_t and the forecast of excess returns you just estimated in part a $\widehat{ER}_t$. Make sure to drop NaNs in this dataframe since we needed 10 years of data before running the first regression. Now that we have our excess return predictions, we can use them to construct optimal portfolio weights. We will use an optimal weighting scheme similar to Hull & Qiao. Construct portfolio weights according to this equation

$$w_t = \min\left(1.5, \max\left(0.5, 100 \times \widehat{ER}_t\right)\right)$$

Note that the 100 is a lot different than the 8 used by Hull & Qiao. **I am assuming that your market excess return forecasts will not be annualized**. So, if your expected return in the next month is .01 (a 1% return in the next month or 12% annualized), then you would put 100% of your money in the

market index. Also note that the minimum weight is 0.5, rather than -0.5, as used by H&Q. Calculate the annualized average return, annualized standard deviation, Sharpe ratio and t-stat of the strategy's returns. Do the same thing for a simple "buy and hold" strategy where you simply go long the market and short the risk-free rate, earning this excess return. Would an investor benefit from trying to time the market using these predictors versus buy and hold?