Problem Set #2
Business 34902
Winter 2025

The due date for this assignment is Wednesday January 29, 8:30am (upload on canvas)

Write up the solution clearly, using tables and graphs where appropriate, preferably using LaTeX. In addition to showing the results, you must also explain and describe your results in words, as you would do if you wrote a paper. Carry out your analysis in Matlab, python, or R, and attach to your solution the code that you wrote. Structure your code carefully and insert some explanatory comments so that it is transparent to an outsider what your code does.

Data: On canvas you can find a data set: `CRSPMonthly.xlsx`. The data set contains monthly returns on the CRSP value-weighted stock market index (`rvwind`) and monthly returns on T-bills (`rf`). Use these to construct a series of excess returns. You will need these below. I use the label R_t for these excess returns below.

As a preparation first read the paper by Kelly et al., "The Virtue of Complexity in Return Prediction" at `https://doi.org/10.1111/jofi.13298`. You can skim the theoretical part. Then focus on the empirical part in section V. Now let's do a prediction exercise that is simpler than what they do in the paper, but also has a flavor of complexity.

Construct 20 predictor variables that are each 10-year moving averages of returns in overlapping windows and where these windows differ by 12 months. For predicting excess return in month t, these predictors are the average excess return in the windows $t-120$ to $t-1$, $t-132$ to $t-13$, $t-144$ to $t-25$, and so on, all the way until you have 20 predictors. Then add a vector of ones so that you will have an intercept in the regression. Now run rolling-window regressions of returns in month t on the predictors that you constructed. Let the length of the rolling window be w months. The first window is the earliest one in the data set that has w observations of all 20 predictors that you constructed. Once you have run this regression, record the results and move the window forward by one month. Then re-run the regression, record the result, and so on, until you have reached the end of the data set.

As I explain below, we will consider $w < 20$, which can mean fewer observations in each rolling window than the number of predictors. Clearly you can't estimate this with the standard OLS estimator. Instead of $(\boldsymbol{X}'\boldsymbol{X})^{-1}$ in the standard OLS estimator, use $(\boldsymbol{X}'\boldsymbol{X})^{+}$, i.e., the (Moore-Penrose) pseudoinverse. This is what Kelly et al. refer to as the "ridgeless" regression.

For each rolling regression window ending in month t construct the predicted excess return for month $t+1$, $\hat{R}_{t+1|t}$. Using the realized excess return R_{t+1}, calculate the return on a market-timing strategy based on the predicted return as follows:

$$R_{p,t+1} = R_{t+1}\omega_t, \qquad \omega_t = \hat{R}_{t+1|t} \tag{1}$$

Once you have the full time series of R_p, calculate the annualized Sharpe ratio and the information ratio (for the latter, regress $R_{p,t+1}$ on R_{t+1}, including an intercept, to obtain the timing strategy's market beta, then remove beta times R_{t+1} from $R_{p,t+1}$, then take the ratio of mean abnormal return to its standard deviation and annualize).

1. Do the above procedure for four different rolling regression window sizes: $w = 3$ (yes, only three observations, this is not a typo), $w = 12$, $w = 60$, $w = 120$. For each w, report the annualized Sharpe ratio and information ratio of the market-timing strategy. For comparison, calculate and report the Sharpe ratio of the stock market index (i.e., a strategy with weight $\omega_t = 1$ over the same sample period over which the returns of market-timing strategy with $w = 3$ are available). Are you surprised by these results? If yes, why? If not, why not? What do these results mean? Do they make economic sense?

2. For comparison, also calculate and report the Sharpe ratio and information ratio of a very simple one-month momentum strategy (over the same sample period over which the returns of the market-timing strategy with $w = 3$ are available) where the weight is simply the one-month lagged excess return

$$R_{p,t+1} = R_{t+1}\omega_t, \qquad \omega_t = R_t \tag{2}$$

Does this help understand what's going on?

3. To further dig deeper, examine what the rolling regression predictions really do. Take the case of $w = 3$. The predicted return $\hat{R}_{t+1}$ in each regression window ending in t is constructed as

$$\hat{R}_{t+1} = \boldsymbol{x}_t'(\boldsymbol{X}'\boldsymbol{X})^{+}\boldsymbol{X}'\boldsymbol{y} \tag{3}$$

where

- $\boldsymbol{x}_t$ contains the values of the predictors in t
- $\boldsymbol{X}$ predictor observations from $t-3$ to $t-1$
- $\boldsymbol{y}$ collects excess returns from $t-2$ to t

So, effectively, the construction of predicted returns forms a weighted sum of the elements of $\boldsymbol{y}$.

In the case where the windows are extremely short ($w = 3$) examine how the time series of these weights $\boldsymbol{x}_t'(\boldsymbol{X}'\boldsymbol{X})^{+}\boldsymbol{X}'$ looks like. Do you see a pattern that helps explain the performance of the market-timing strategy that you observed above? In light of this, how do you interpret the role of "complexity in return prediction" here?

4. Can you figure out what sort of properties predictors must have to deliver weights $\boldsymbol{x}_t'(\boldsymbol{X}'\boldsymbol{X})^{+}\boldsymbol{X}'$ that look like those that you observed in (3.) for $w = 3$? (Hint: look into the properties of the Moore-Penrose inverse).