FBE 551 Homework #3

Due on 10/14/2025 by 11:59 pm

This is a group homework assignment. You should work in a team of 2-5 students, and include all student names in a markdown cell at the top of the Jupyter Notebook that you submit.

To complete this homework assignment, you will need to use the following data sets

- CRSP_Dividends_HW3.csv: year-end betas for all stocks in the CRSP data set.
- Yearly_Betas_HW3.csv:
- FamaFrenchDaily_HW3.csv: Fama-French 3 Factors, daily returns.

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 CSV 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. (I know it may be obvious, but I want to see that you know how to use markdown.) Clearly highlight what parts of your notebook are used to answer each question below.

Start

This assignment will model a "dividend capture" strategy where you purchase stocks prior to their ex-dividend day, and sell them after receiving the rights to their dividends the following trading day. It also serves as a test of the Efficient Markets Hypothesis. If the trading strategy appears profitable, we will examine its robustness and persistence in the face of transactions costs.

Load in the 'CRSP_Dividends.csv' file to a pandas dataframe. This file includes daily data for all CRSP common stocks (SHRCD = 10 or 11) on the day which they go ex-dividend. PERMNO and PERMCO are security identifiers; DISTCD indicates the type of distribution (I have filtered for USD cash distributions; more description here: Distribution Codes | CRSP - The Center for Research in Security Prices, see page 126). PRC is the closing price; VOL is volume of shares traded that day; RET is the total return, including dividends, over the last day; SHROUT is shares outstanding; OPENPRC is the opening price for the day; RETX is the return without dividends over the last day; BID and ASK are the prevailing bid and ask quotes at the close of the day; and **vwretd** is the return on the CRSP value-weighted index (**a version of the market portfolio**).

Question 1 (10 points)

Describe your data and take reasonable steps to clean the data. For example, you should:

- Read through the remainder of this assignment and think about how you should handle observations where VOL = 0.
- Remove negative and zero dividend amounts. You should also go ahead and remove dividend amounts less than 0.01. (These are tiny dividends that may greatly impact the average ratio.)
- Negative prices indicate a stock that did not trade that day; common practice is to set those equal to the absolute value of the price.
- Remove stocks that have a market capitalization below $50M. Market cap is price times shares outstanding; SHROUT values here are in 1,000. Micro-cap stocks are often illiquid and returns coming from them can be unreliable.
- At this point, don't make any adjustments based on the OPENPRC variable.
- Be sure to document these and any other steps you take here in a markdown cell, and run a df.describe() statement to show the mean for price, dividend amount, and market cap.

Question 2 (15 points)

For each observation, calculate the closing price as of the prior day and call this 'PriorClose'. You should be able to do this using the current PRC and the return. (Think about which return to use.) Create a column of $(P_{t-1} - P_t) / D$. In a markdown cell, tell me what this ratio signifies in terms of the Efficient Market's Hypothesis. Describe the mean and range of this variable. In a markdown cell, discuss the findings – what does this variable signify? Is there a potential investment opportunity?

Suppose we were to attempt to form a profitable investment strategy by buying stocks at the close prior to the ex-dividend day, and selling them at the close the following day. Describe the annualized returns, standard deviation, and Sharpe[1] ratio for this investment strategy (use the excess return to compute the Sharpe ratio, you will need to use the daily risk-free rates from the Fama-French file here.) and display these values in percent rounded to two decimal places (eg: 0.0123 would be displayed as 1.23).

Question 3 (30 points)

As noted in Kalay's "Ex-Dividend Day Behavior" paper, looking at the close-to-close price change will incorporate a daily return impact beyond just the immediate dividend effect.

The data set "Yearly_Betas_HW3.csv" includes a year-end beta for all stocks in the CRSP data set. You should merge this into your dividend data based on **PERMNO** and **Year**. The beta for a particular stock will be from a CAPM regression using returns up until that date; think carefully about how to match up your dates in these two data sets in order to avoid look-ahead bias.

For each dividend event in your sample, compute the expected return from the CAPM model. In this sort of application, we assume that the alpha, as dictated by the CAPM, is equal to zero.[2] Compute the

[1] This is not technically a Sharpe ratio, as this is an event study with overlapping observations on single days. Do not worry about calculating portfolio returns for each day at this point – we will keep this as an event time study.
[2] In the event study example in lecture, I estimated these betas using an OLS regression and as a result, had both alpha and beta estimates. Here, I am just giving you the beta estimates from the CAPM model for all the stocks in the CRSP dataset. Don't worry about the alpha (intercept) in this case and just assume it is 0 like the CAPM model does.

abnormal return for each dividend event. Compute the mean and standard deviation of the abnormal return (annualized) along with the average beta. Compute the t-statistic and Sharpe ratio of the abnormal return. Is this abnormal return statistically significant? How does this average abnormal return compare to the annualized returns you found in Question 2? Display these values in percent rounded to two decimal places (eg: 0.0123 would be displayed as 1.23) except for the beta. Beta should be between zero and 1 and rounded to 2 decimal places.

Question 4 (15 points)

Instead of holding the stock throughout the day, suppose you were able to buy at the closing price on the prior day and sell at the open. What would be the annualized returns, standard deviation, t-stat and Sharpe ratio be for this strategy? Are these returns statistically significant? Display these values in percent rounded to two decimal places (eg: 0.0123 would be displayed as 1.23) except for the beta. Beta should be between zero and 1 and rounded to 2 decimal places.

Question 5 (10 points)

As the Kalay paper notes, transaction costs may be a reason why a strategy like this shows potential profits in event studies or backtests. Compute the transaction cost using the bid and ask prices, as we discussed in the lecture slides in Week 5. After paying for these transaction costs, what are the annualized returns to this strategy? Use the full-day close-to-close returns you calculated in Question 2. Display these values in percent rounded to two decimal places (eg: 0.0123 would be displayed as 1.23) except for the beta. Beta should be between zero and 1 and rounded to 2 decimal places.

Question 6 (20 points)

Consider the investment strategy in Question 2. Rank the data into deciles by the following metrics

- Market Capitalization
- Dividend yield (D/P) as done in the Kalay paper
- CAPM Beta

and analyze the returns, bid-ask spreads and returns net-of-costs for each decile for the investment strategy in Question 2 (the “groupby” command will be helpful here). Is there evidence that we could improve returns to this strategy by sorting the data into different deciles according to each metric? Write a markdown cell that summarizes your results and recommendation.