Financial Instruments and Portfolio Choice
Group Project
Submission Deadline: 15/2/2024

The aim of this project is to handle (download and manipulate) big financial data and create a real-life investment strategy based on the time-series and the cross-section US stock data covering the period from 1/1973 to 12/2022. The students are expected to use a software of their choice (Python, Stata, etc.) and write codes in order to achieve this goal. The details of the project are as follows:

1) Each group will download the following data using the CRSP database (via WRDS):

   Permno, price, holding period return, number of shares outstanding, exchange code, share code.

   The data should then be filtered so that only stocks trading at prices more than $1, trading in NYSE, AMEX, or NASDAQ (i.e. with exchange codes 1, 2, or 3), and that have share codes 10 and 11 (i.e. common shares) should be included in the analysis.

   Furthermore, depending on the project allocations, each group will download either one of the four q-factors (ME, IA, ROE, or EG) as in Hou et. al. (2021), the liquidity factor (LIQ) of Pastor and Stambaugh (2003), short- and long-horizon behavioral factors (PEAD and FIN) of Daniel, Hirshleifer and Sun (2020), the regret factor (FREG) of Arisoy, Bali, and Tang (2023), and the volatility factor (ΔVIX). Furthermore, each group should download Fama and French (2015) five factors (MKT, SMB, HML, CMA, RMW) as well as the risk free-rate.

   For Fama-French factors and the risk-free rate:

   https://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html

   For Hou et al. (2021) q-factors:

   http://global-q.org/factors.html

   For Pastor and Stambaugh (2003) liquidity factor :

   https://faculty.chicagobooth.edu/faculty/lubos-pastor/data

   For Daniel, Hirshleifer and Sun (2020) behavioral factors:

   http://www.kentdaniel.net/data.php

   For Arisoy, Bali, and Tang (2023) regret factor

   https://sites.google.com/a/georgetown.edu/turan-bali/data-working-papers?authuser=0

   For VIX data:

   https://www.cboe.com/tradable_products/vix/vix_historical_data/

2) The sample period is 1/1973-12/2022.

3) Each group will be allocated to one investment strategy based on the following factors :

Group 1 – size (ME)
Group 2 – investments (IA)
Group 3 – return on equity (ROE)
Group 4 – expected growth (EG)
Group 5 – liquidity (LIQ)
Group 6 – short-horizon behavioral factor (PEAD)
Group 7 – long-horizon behavioral factor (FIN)
Group 8 – regret factor (FREG)
Group 9 – volatility factor (ΔVIX)

4) Using the nine factors defined above, each group will first calculate the exposure of each stock's excess return to their allocated factor (by using a two-factor model, i.e. MKT + their allocated factor) using 60-month rolling window estimation period, i.e., the first estimation window should be 1/1973-12/1977, the second 2/1973-1/1978, and so on, and the last estimation window should be 1/2018-12/2022.

5) Using the betas estimated in step 4, each group will then create a monthly trading strategy and form 10 portfolios (both equally- and value-weighted) each month starting in 12/1977 and ending in 11/2022.

6) Finally, next-month average raw return, CAPM alpha, Fama-French 3-factor alpha, and Fama-French 5-factor alpha of the 10 portfolios as well as the arbitrage portfolio that is long in portfolio 10 and short in portfolio 1 will be reported. The significance (t-statistics) of the return and alphas of the long-short arbitrage portfolio should also be reported. (Any graphs or supplementary tables can also be reported).

7) The project should be prepared using a Word processor or LATEX. In their report, each group should explain clearly in each step the methodology that they follow. The figures and tables should be clear and self-explanatory with proper headings, titles, and notes. The code used to prepare the project and a small extract of the merged dataset should be added to the Appendix.