Exercise 2: Construct Value Factors

The goal of this exercise is to implement and analyze different value factors.

Value Factors

Value factors buy "cheap" stocks while shorting "expensive" ones. There are many different ways to implement a value factor depending on the definition of cheap/expensive (the "characteristic"), the investment universe, the rebalancing frequency, the portfolio weighting scheme, and so on. A standard value factor is that of Fama and French (1993), which you will implement in this exercise. The methodology has changed slightly since the original paper, and we will try to replicate the version available from Kenneth French's data library. The construction of the underlying book-to-market characteristic is described here. For this exercise, you can rely on CRSP for constructing market equity, but you should supplement book equity data from Compustat, with the Moody's data used in Davis et al. (2000). Finally, you will also work with the closely related value factors from Jensen et al. (2022).

WRDS Data

Empirical asset pricing researchers heavily use data from WRDS. The most common way to access data from WRDS is via the web interface. However, this way of accessing WRDS makes it difficult for other researchers to reproduce your dataset and only gives access to a subset of the data available on WRDS.

A more powerful approach is to work directly on the WRDS server via the WRDS Cloud. Currently, the cloud supports SAS, Python, R, and Stata. For the dataset used in Jensen et al. (2022), we relied on SAS and the corresponding IDE SAS Studio for data preparation. For this exercise, you may use whatever language you prefer, but you must download all WRDS data via the cloud and not via the web interface.

JKP Data

To solve this exercise, you can get inspiration from the source code used in Jensen et al. (2022), which you can find at `https://github.com/bkelly-lab/ReplicationCrisis`. However, the construction of the book-to-market equity characteristic in Jensen et al. (2022) differs from the Fama-French method, so we recommend that you write your data extracting code from scratch. For this exercise you also need to download the US book-to-market equity factors (equal-weighted, value-weighted, capped-value-weighted) from `JKPfactors.com`.

FF Data

Download the Fama-French value factor as a comparison (HML^{FF}). You can find all their data at `https://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html`, and you should use the monthly value factor that is part of the "Fama/French 3 Factors" in their "U.S. Research Returns Data." The book equity used in Davis et al. (2000) is also available from this website.

Questions:

1. **Compute the characteristic.** Use US annual accounting data from Compustat and Moody's and monthly price data from CRSP to create the book-to-market ratio that underlies the Fama-French HML factor. Compute the book-to-market breakpoints (the 30 and 70 percentile) used for the HML factor. Describe your approach and plot how the breakpoints vary over time. *Optional: Add the Fama-French breakpoints for comparison.*

2. **Generate data set of stock returns.** Create a data set of monthly return data, describe your approach. Compute the standard deviation of returns each month, and plot the resulting series over time.

3. **Generate factor returns.** Combining the characteristic data and the return data, create a monthly return series of the HML factor in the US. Describe your approach, plot the cumulative return of your factor alongside HML^{FF}, and report the Pearson correlation between the monthly returns of the two factors.

4. **Compare factor returns.** Compare and contrast the HML factor you have created with the US Book-to-Market equity factor from JKPfactors.com (both the equal-weighted, value-weighted, capped-value-weighted). Discuss the differences.