Problem Set 4: Size and Value

Professor Bernard Herskovic
bernardherskovic@anderson.ucla.edu

Due on Sunday 5/24 midnight PST. This is an **individual assignment**, but you can discuss it with your classmates. If you discuss with other classmates, indicate their names in your write-up. Please submit R code as well as a separate write-up. Explain the procedure and your answers clearly in the write-up (such that someone unfamiliar with the problem could solve it). Code must be formatted as instructed in order to receive a grade. Use CCLE to submit your answers.

You should submit two files:

- .R **PS4_YourStudentID** (for example, PS4_012345678.R), with **all** code used in answering the questions written below
- .pdf **PS4_YourStudentID** (for example, PS4_012345678.pdf), with discussion on how you answered the questions written below, as well as responses to any particular questions asked

1. Prepare data for analysis. Combine necessary CRSP and Compustat datasets needed to define size and book-to-market decile portfolios as defined in Fama and French (1992b)[1], as well as the HML and SMB factors as defined in Fama and French (1993)[2]. Detail which datasets you use, how you merged them, how you calculated the portfolios, and any differences between the building of the decile portfolios and the factors. Output should be between January 1973 and December 2019.

- No function required for this question. Prepare the data you will need for the following questions of this problem set. Process should go from raw datasets to a dataset of portfolio returns. As a suggestion, the TA ended up with the following output:
  - data.table, with each row corresponding to a unique Year-Month-Decile, with columns

Variable Name	Variable type	Variable description
Year	Integer	
Month	Integer	
port	Integer	Portfolio decile (only relevant for size and book-to-market)
Size_Ret	Numeric	Return of that size portfolio
BtM_Ret	Numeric	Return of that book-to-market portfolio
HML_Ret	Numeric	Return of the high-minus-low portfolio
SMB_Ret	Numeric	Return of the small-minus-big portfolio

  - This is not a trivial exercise. Be detailed in your write-up of the decisions you made, as there are many of them that can make a significant difference in your end result.

[1] http://onlinelibrary.wiley.com/doi/10.1111/j.1540-6261.1992.tb04398.x/full
[2] http://www.sciencedirect.com/science/article/pii/0304405X93900235

- Hints
  - Basic instructions on how portfolios are constructed is available on Kenneth French's website[3]. Read Fama and French (1992, 1993) for additional **(and necessary)** information to replicate these portfolios.
  - Limit your sample to common shares.
  - Pay attention to the rebalance frequency and timing
  - To replicate value anomaly, you will need to merge CRSP and Compustat data. WRDS provides a linktable between both datasets.
  - Get stock information (prices, returns, etc.) from CRSP directly (not Compustat).
  - From Fama and French (1993): "We define book equity, BE, as Compustat value of shareholders' equity, plus balence-sheet deferred taxes and investment tax credit (if available), minus the book value of preferred stock. Depending on availability, we use redemption, liquidation, or par value (in that order) to estimate the value of prefereed stock." The timing is well described in the papers (Fama and French, 1992 and 1993), as well as in French's website. Let's be more precise about the variables:
    * Shareholders' equity (SHE): variable reported in Compustat is "Stockholders' Equity - Total" (SEQ). If not available, use "Common/Ordinary Equity - Total" (CEQ) plus "Preferred/Preference Stock (Capital) - Total" (PSTK). If not available, use "Assets - Total" (AT) minus "Liabilities - Total" (LT) minus "Minority Interest (Balance Sheet)" (MIB). If not available, use AT minus LT.
    * Deferred taxes and investment tax credit (DT): variable reported in Compustat is "Deferred Taxes and Investment Tax Credit" TXDITC. If not available, use "Investment Tax Credit (Balance Sheet)" (ITCB) plus "Deferred Taxes (Balance Sheet)" (TXDB). If not available, sum what is not missing.
    * Book value of preferred stock (PS): Use redemption value, which is variable "Preferred Stock Redemption Value" (PSTKRV). If not available, use liquidation value, which is "Preferred Stock Liquidating Value" (PSTKL). If not available, use par value, which is "Preferred/Preference Stock (Capital) - Total" (PSTK).
    * Define book equity (BE) as: $BE = SHE - PS + DT - PRBA$ (need value of SHE to compute BE, other variables included if not missing). The last variable is "Postretirement Benefit Asset" (PRBA), and you will have to get this variable from Compustat's Pension Annual data: merge to Compustat using Compustat's global variable key (GVKEY).
    * (optional) For missing value of book equity, you can use historical book equity value available n French's website—Davis, Fama and French (2000)'s data.
  - To calculate book-to-market, you will have to sum the market equity of subsidiaries. A company (permanent identifier variable: PERMCO) may have different securities (identified by PERMNO).

[3] `http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html`

2. For each size decile and the long-short portfolio, report the annualized average excess returns, annualized volatility, Sharpe Ratio, and skewness. Also report the correlation between the portfolios that you have constructed (the 10 portfolios and the long-short portfolio) and those from French's website.

3. For each book-to-market decile and the long-short portfolio, report the annualized average excess returns, annualized volatility, Sharpe Ratio, and skewness. Also report the correlation between the portfolios that you have constructed (the 10 portfolios and the long-short portfolio) and those from French's website.

4. Has the value and size anomaly worked in the past few years? Show some empirical evidence.

5. For both HML and SMB portfolios, report the annualized average excess returns, annualized volatility, Sharpe Ratio, and skewness. Report correlations between the replicated factors and the factor from French's website. Have the factors been consistent across time? Show some empirical evidence.

6. Compare and contrast using the characteristic portfolios (Fama and French 1992) and the factor portfolios (Fama and French 1993).