

# Mandatory Assignment 2

2024-04-12

## General remarks

- You are supposed to hand in one (!) .qmd (Quarto) script with commented source code in Python or R and one (!) .pdf report in which you describe your methods and results (e.g., with figures, tables, equations, etc.)
- The report (font size 12pt; use the provided template) can be a maximum of 7 pages, including figures, tables, and other appendices. Unless otherwise stated, it should not contain any code.
- The source code has to create every figure, table, and number used for your report
- When reporting numbers in your report, please consider a meaningful rounding scheme (there is usually no need to report six digits after the comma). Make sure you only use `round()` for the final reported results and not for any intermediate steps
- Within the report, describe every significant step and make sure that tables and figures are self-explanatory by providing meaningful captions and variable names. You cannot assume that the reader of the report goes through your code in case your description may be confusing
- You increase your chance of helpful feedback as well as your final grade if you provide meaningful comments in the code that explain the purpose of each step
- **Minimum requirement:** the code must run without error or interruption. Ensure the code loads all required packages at the beginning of the script (you can assume that the user has the packages available locally, so do not include `install.packages()` commands). Attempt to solve every task. Document your struggles if you do not finish an exercise. **Hand-ins that do not fulfill this minimum requirement will not receive peer feedback and will not be considered for the final exam**
- The assignments can be written individually or by groups of a maximum of three students
- The plagiarism rules must be complied with, and please be aware of the rules for co-written assignments
- All parts must be answered in English, including comments in the code.

The deadline to hand in the file on Feedback Fruits via Absalon is **Friday, April 19th, 7 pm.**

## The momentum effect (inspired by Section 11 of Bali, Engle, and Murray (2016) and Chapter 3 of Nagel (2021))

Many studies have shown that previous stock returns can predict future stock returns in the cross-section. One of the most prominent such phenomena, documented by Jegadeesh and Titman (1993) and known as the medium-term momentum effect, is that stocks that have performed well in the medium-term past (six to 12 months) are more likely to outperform in the future.

The momentum effect is widely considered a behavioral phenomenon, as including controls for risk in statistical analyses fails to change the result. This mandatory assignment is devoted to an empirical investigation of the momentum effect. You will start by measuring momentum. Then, you will examine the relationship between momentum and future stock returns. Finally, you will consider real-world implementation issues of such a momentum strategy.

### Exercise

0. Sign your group up in Absalon by clicking on *People* and self-select into one of the available groups in group set **Groups MA 2**. Do this before anything else. To submit your final report, Feedback Fruit has to retrieve your group association, which may take a couple of hours to synchronize.
1. Read in the monthly CRSP data (`crsp_monthly`) and the monthly market excess returns (`factors_ff3_monthly`) from the `tidy_finance_*.sqlite` database.
2. For each permno-month observation in the CRSP sample, generate a new variable, `mktcap_lag_12`, which contains the firms' market capitalization 12 months before the measurement date. Compute the momentum of stock  $i$  as the relative change in market capitalization, represented in percentage terms (1.00 is a 1% return) during the period covering months  $t - 12$  until  $t - 1$ . Specifically, if  $mc_{i,t}$  denotes the market capitalization of stock  $i$  in month  $t$ , compute 12-month momentum as

$$Mom_{i,t}^{12} = 100 (mc_{i,t-1} - mc_{i,t-12}) / mc_{i,t-12}.$$

Note that `crsp_monthly` already contains a column `mktcap_lag` with  $mc_{i,t-1}$ . Do you expect differences between computing momentum as the relative change in *prices* or the relative change in *market capitalization*?

3. Next, examine the relation between momentum and future stock returns. For that purpose, perform univariate portfolio sorts in the following way:
  - Compute monthly portfolio breakpoints. Each month, use  $Mom_{i,t}^{12}$  as a sorting variable. The portfolio breakpoints are the deciles of  $Mom_{i,t}^{12}$  calculated using all stocks  $i$  in the sample during the month  $t$ .
  - To help understand the characteristics of the stocks comprising each of the  $Mom_{i,t}^{12}$  sorted portfolios, present the equal-weighted average values of  $Mom_{i,t}^{12}$  and  $mc_{i,t}$  for each of the ten portfolios.

- Compute value-weighted monthly excess returns for the decile portfolios. Report the average excess return and the CAPM alpha for the ten momentum-sorted portfolios. Finally, analyse the momentum strategy: a portfolio that goes long past winners (the highest decile) and short past losers (the lowest decile). What do you conclude about the performance of the momentum strategy? Is the strategy market neutral (i.e., does it exhibit a zero market beta)? Does the strategy deliver abnormal excess returns (i.e., a positive alpha)?
4. As a final step - and as a generalization to the momentum strategy - we consider cross-sectional stock return predictions that use only past returns as predictors. Specifically, we analyze a model where the monthly stock excess return  $r_{i,t}$  is a linear function of 60 lags of the stock's own past monthly returns and past monthly squared returns. We skip the first lag to avoid contamination with microstructure-induced biases and very short horizons. Thus, we have the regression

$$r_{i,t+1} = \sum_{k=1}^{60} b_k r_{i,t-k} + \sum_{k=1}^{60} c_k r_{i,t-k}^2.$$

Implement a machine-learning framework based on ridge regression following these steps:

- The sample comprises all stocks in the CRSP dataset. Each month, you remove small stocks with market capitalization below the 20th percentile of all stocks listed on the NYSE.
- Demean the dependent variable  $r_{i,t+1}$  month by month (therefore, the model does not include an intercept term)
- Demean all predictor variables month by month to focus purely on cross-sectional variation. Cross-sectionally standardize the predictor variables with unit standard deviation each month.
- Estimate the vector of regression coefficients  $[\hat{b}, \hat{c}]$  using ridge regression. Choose the hyperparameter with a suitable cross-validation scheme.

In the report, clearly state the estimator, the cross-validation scheme, and the resulting selected hyperparameter. Visualize all computed regression coefficients in a useful manner. Discuss if (and how) the estimates recover information related to the momentum effect or other prominent predictability patterns. How could the estimates be exploited to design a trading strategy?

## References

- Bali, Turan G., Robert F. Engle, and Scott Murray. 2016. *Empirical Asset Pricing : The Cross Section of Stock Returns*.
- Jegadeesh, Narasimhan, and Sheridan Titman. 1993. "Returns to Buying Winners and Selling Losers: Implications for Stock Market Efficiency." *The Journal of Finance* 48: 65–91.
- Nagel, Stefan. 2021. *Machine Learning in Asset Pricing*. Princeton University Press.