Mandatory Assignment 2

Asbjørn Fyhn & Emil Beckett Kolko

2024-04-19

Introduction

Jegadeesh and Titman (1993) documented that a portfolio that goes long high return stocks and short low return stocks over the past 3 to 12 months earns abnormal profits the following year. This phenomenon is known as the medium-term momentum effect. In this assignment, we first investigate this effect by estimating the returns of a long/short momentum portfolio using an ordinary least squares (OLS) regression. Next, we explore the predictions of cross-sectional stock returns using only past returns as predictors. This analysis is conducted through a machine learning framework that employs ridge regression. Instead of the positive correlation between last year's returns and future returns found in Jegadeesh and Titman (1993), we find a negative correlation. Therefore, we conclude that a portfolio that goes long low return stocks and short high return stocks over the past 12 months earns abnormal profits.

Exercise 1 and 2:

We load the "tidy_finance_python.sqlite" database and select variables from the tables "crsp_monthly" and "factors_ff3_monthly". Thereafter, we compute the momentum of stock i as described in the assignment.

We do not expect significant differences between computing momentum as the relative change in prices compared to the relative change in market capitalization. In this assignment, we use the market capitalization as opposed to the stock price. The benefit of doing so is that our computations are not affected by events that increase or decrease the number of outstanding stocks, such as stock issuances or buybacks. Such events affect stock prices without affecting returns. However, when using the market capitalization, events such as dividend payouts, decrease market capitalization without affecting returns. Therefore, there are problems with both types of calculations. Lucklily, both the stock price and market capitalization has been adjusted accordingly in the CRSP dataset. Therefore, we should not face any problems by calculating momentum with either stocks prices or market capitalization.

We present summary statistics of our dataset in Table 1. The mean of 12-Month momentum is around 20% but the distribution has long tails. Likewise, we see that market cap has a long right tail of its distribution.

Table 1: Summary statistics of the dataset

	Count	Mean	Std	Min	Median	Max
Return	3,004,264	1.25	18.08	-99.36	0.00	2,400
Excess Return	3,004,265	0.87	18.09	-100.00	-0.37	2,399
12-Month Momentum	3,004,265	21.07	163.71	-99.88	5.08	115,127
Market Cap	3,004,265	2,185.16	18,596.58	0.01	95.60	2,902,368

Exercise 3:

We will based on the described data analyze the performance of the momentum strategy. This is done by examning the performance of portfolios created based on momentum. Specifally, we compute monthly portfolio breakpoints using 12-month momentum ($mom_{i,t}^{12}$) as the sorting variable for each month t. Subsequently, we divide the companies into ten deciles based on their 12-month momentum. Specifically, the first decile encompasses companies with the lowest 12-month momentum, while the tenth decile comprises those with the highest 12-month momentum. This method ensures that each month's portfolio breakpoints accurately reflect the momentum distribution across all stocks in the sample, facilitating effective portfolio stratification and analysis.

To create portfolios, we start by equal-weighting the stocks in each decile. The average market cap and 12-month momentum for each decile portfolio is reported in Table 2. Unsurprisingly, the table shows that the average market momentum increases with the momentum decile, but also that the average market capitalization is increasing in the momentum decile until the 7th decile.

Now instead of creating equal-weighted portfolio, we create value-weighted portfolio. The weights are then determined by a stock's relative size to the other stock within the decile. The analysis continues with the value_weighted portfolios. For each decile-portfolio, we calculate the average excess return for the ten decile which is reported alongside the portfolio characteristics in Table 2.

Table 2: Average momentum and excess return

Decile	Average momentum	Average market cap	Excess Return
1	-53.91	343	0.0017
2	-30.43	1,093	0.0040
3	-17.56	1,850	0.0057
4	-7.58	2,462	0.0064
5	1.39	2,849	0.0065
6	10.51	3,118	0.0049
7	21.06	3,146	0.0059
8	35.29	3,044	0.0069
9	59.97	2,564	0.0074
10	191.96	1,388	0.0094

Using the value-weighted portfolios, we estimate alpha of each portfolio with OLS utilizing the

CAPM equation:

$$r_{i,t} - r_{f,t} = \alpha_i + \beta_i(r_{m,t} - r_{f,t}) + \epsilon_{i,t}$$

We report the alphas and their corresponding 95% confidence intervals in Figure 1.

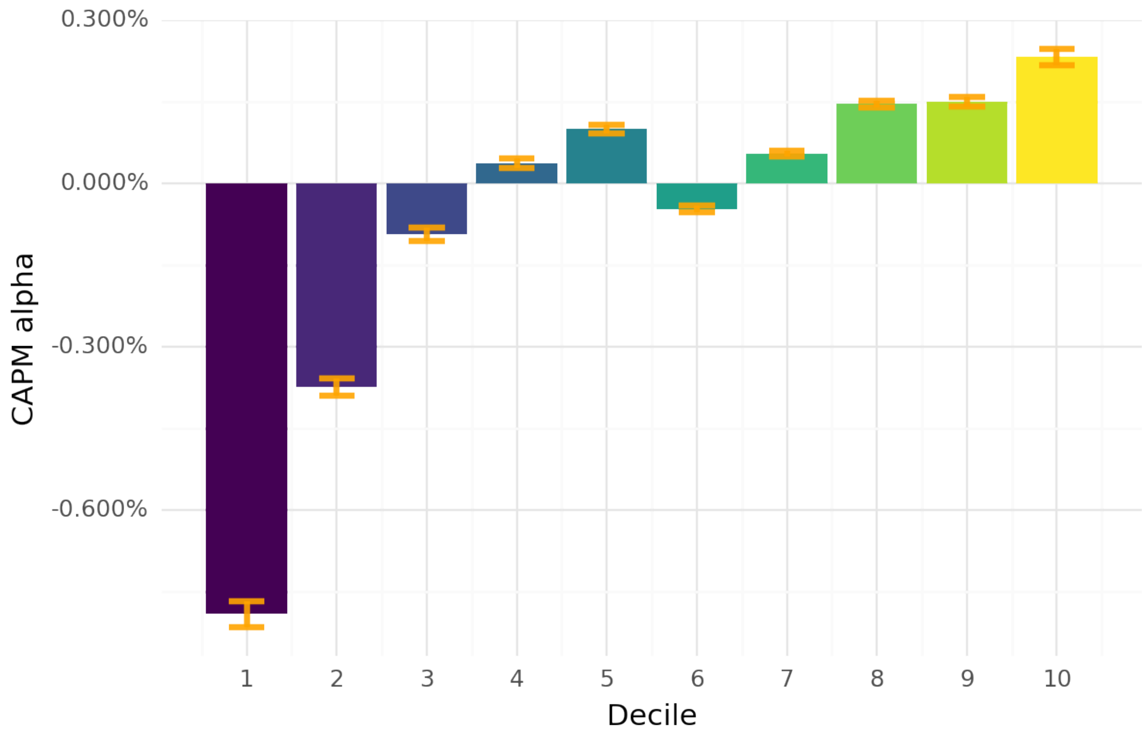

Figure 1: CAPM alphas of momentum-sorted portfolios

From Figure 1, we see that the lowest decile generates the lowest alpha and generally a higher momentum means higher alpha. This result is in accordance with Jegadeesh and Titman (1993) and represents an imperfection in the market. In the following, we test if we can exploit this imperfection with a momentum strategy that goes long past winners and short past losers. Specifically, we estimate the alpha and beta of a portfolio that shorts the 1st decile portfolio and goes long the 10th decile portfolio with OLS of the following regression model:

$$r_{p,t} = \alpha + \beta r_{m,i}$$

where $r_{p,t}$ is the return of the portfolio, and $r_{m,i}$ is the return of the market index. The alpha and beta are estimated with the Newey-West standard errors with a lag length of 12 months. The results of the estimation are presented in Table 3.

Table 3: Performance of the momentum strategy

	Estimate	Std. Error	t-Value	p-Value
alpha	0.010	0.003	3.912	0.000
beta	-0.443	0.142	-3.119	0.002

We reject the null hypothesis that the return, α, of the momentum strategy is 0. This result aligns with the result in Jegadeesh and Titman (1993) that suggest it is possible to obtain a positive α with the momentum strategy. However, our results deviate from the CAPM, which states that arbitrage strategies are not possible. Furthermore, our results show a significant negative β, which means the momentum (long/short) portfolio is negatively correlated with the market. This type of portfolio can be used to reduce the market risk of other investments with positive betas.

Next, we calculate the excess return of the (long/short) portfolio described above, and for a portfolio (high) that only goes long in the stocks of the 10th decile and a portfolio (low) that exclusively goes long in the stocks of the 1st decile. The yearly returns for the three portfolios are plotted in Figure 2

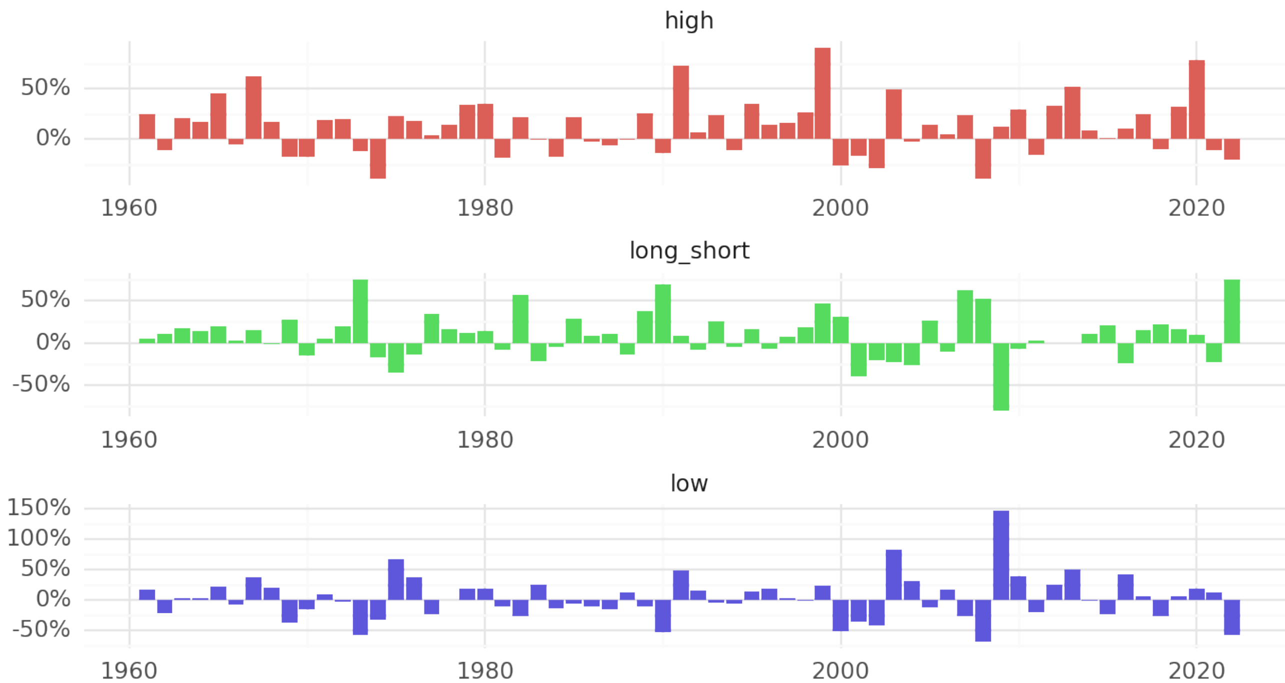

Figure 2: Annual returns of momentum portfolios

Exercise 4

We aim to estimate a model to investigate how a stock's excess return at time t is influenced by its own past returns and the magnitudes of those returns. The model is specified as

$$r_{i,t} = \sum_{k=2}^{61} b_k r_{i,t-k} + \sum_{k=2}^{61} c_k r_{i,t-k}^2$$

The coefficients, b_k represent the linear relationship between the excess return at time t and the excess return from k periods ago. As such, the a positive b_k suggests that higher returns k periods ago are associated with higher returns at time t. The squared term represents the magnitude of past returns, irrespective of their direction (positive or negative). Therefore, the coefficient c_k captures how much the volatility from k periods ago affects the current period's excess return.

We will estimate the parameters using the crsp dataset. We clean the dataset by removing the smallest stocks, specifically those with the lowest 20% market capitalization. Further, we demean the dependent variable $r_{i,t}$ by month, subtracting the mean from the series, ensuring the mean excess return in each month is zero. Additionally, we demean and standardize all lagged returns by month, dividing each by the standard deviation, making the mean equal to zero and the standard deviation across firms one.

To estimate the coefficients of this model, we will use ridge regression, a regularized version of OLS regression. Ridge regression helps reduce the variance of the estimates by adding a penalty term to the OLS regression. This penalty term is the sum of the squared coefficients multiplied by a constant, λ, a hyperparameter that must be chosen. The optimal value of λ is found through cross-validation. The regression coefficient has a closed-form solution:

$$\hat{\beta} = (X^T X + \lambda I)^{-1} X^T Y$$

Here $\hat{\beta}$ represents the estimated coefficients, X is the matrix of predictors, Y is the vector of the dependent variable, I is the identity matrix, and λ is hyperparameter. The value of λ will be chosen to minimize the estimated mean squared prediction error (MSPE):

$$\lambda^{opt} = \arg\min_{\lambda} MSPE = E\left(\frac{1}{T}\sum_{t=1}^{T}(\hat{r}_{i,t}(\lambda) - r_{i,t})^2\right)$$

where $\hat{y}_t(\lambda)$ is the predicted value of the dependent variable at time t for a given λ and y_t is the true value of the dependent variable at time t. To select the optimal hyperparameter λ, we employ a suitable cross-validation scheme. Specifically, we divide the dataset into K folds and perform cross-validation, calculating the MSPE for each fold over a range of λ values. The cross-validation procedure averages the MSPE across all K folds for each λ. The hyperparameter that results in the lowest average MSPE is then selected as the optimal λ. This approach ensures that the selected λ generalizes well to unseen data, balancing model complexity and predictive accuracy. We divide the dataset into a training set for hyperparameter selection, using cross-validation to average the MSPE across K data folds over a range of hyperparameters. The hyperparameter that yields the lowest MSPE is then selected. Choosing K involves balancing the computational burden with the desire for accurate estimates of model performance.
We tune the model over a grid of λ that is evenly spread on a logscale from 0.0 to 10000.0 with 10 datafolds.

After tuning, we set λ to the optimal value (587.802) found, and estimate the model on the remaining dataset:

In the Figure 4, we observe the coefficients for both the lagged excess returns ($r_{i,t-k}$) and their squared terms ($r^2_{i,t-k}$).
We notice a distinct pattern in the estimated coefficients:

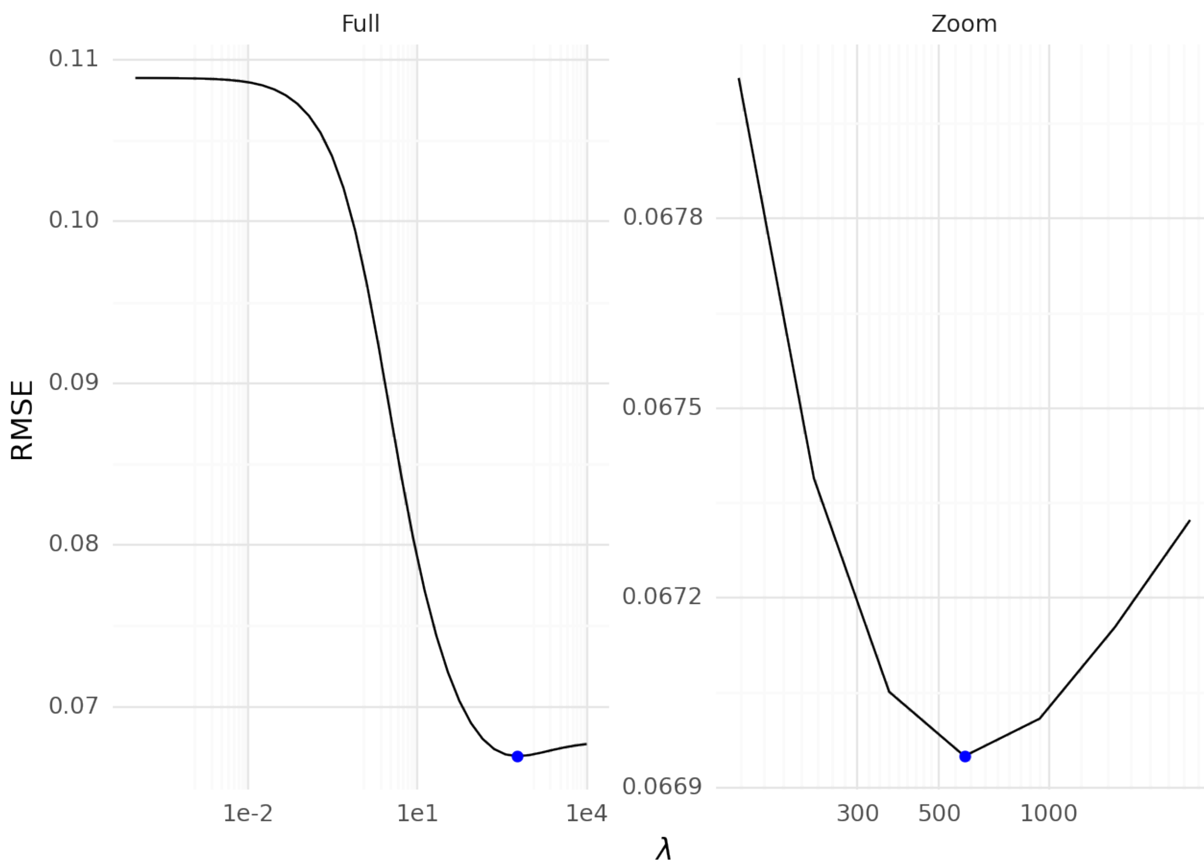

Figure 3: Root mean squared error for different penalty

- For $k < 6$, the coefficient,$\hat{b}_k$, is negative, indicating a negative relationship between past excess returns and future returns. The $\hat{c_k}$ are relatively large. Indicating a short term relationship between volatility (large move in either direction) and excess return in the current period. Any pattern in $\hat{c_k}$ is vanished.
- However, from $6 < k \leq 12$. For $k > 12$, there is a shift, with six consecutive positive coefficients, suggesting a reversal in the trend towards positive relationships.
- Beyond $12 < k$, the majority of coefficients are negative, albeit with a few exceptions. Notably, all $\hat{b}_k$ values are positive for $k \in \{12, 24, 36, 48, 60\}$, hinting at a periodicity in the data.

Interpreting these coefficient estimates, we find indications of mean reversion effects in the short run, particularly for periods under 6 months, with the effect peaking at $k = 2$ and diminishing thereafter.
In the medium run (6-12 months), there's evidence suggesting a momentum strategy. Positive excess returns experienced 6-12 months ago seem to influence the stock's performance in period t

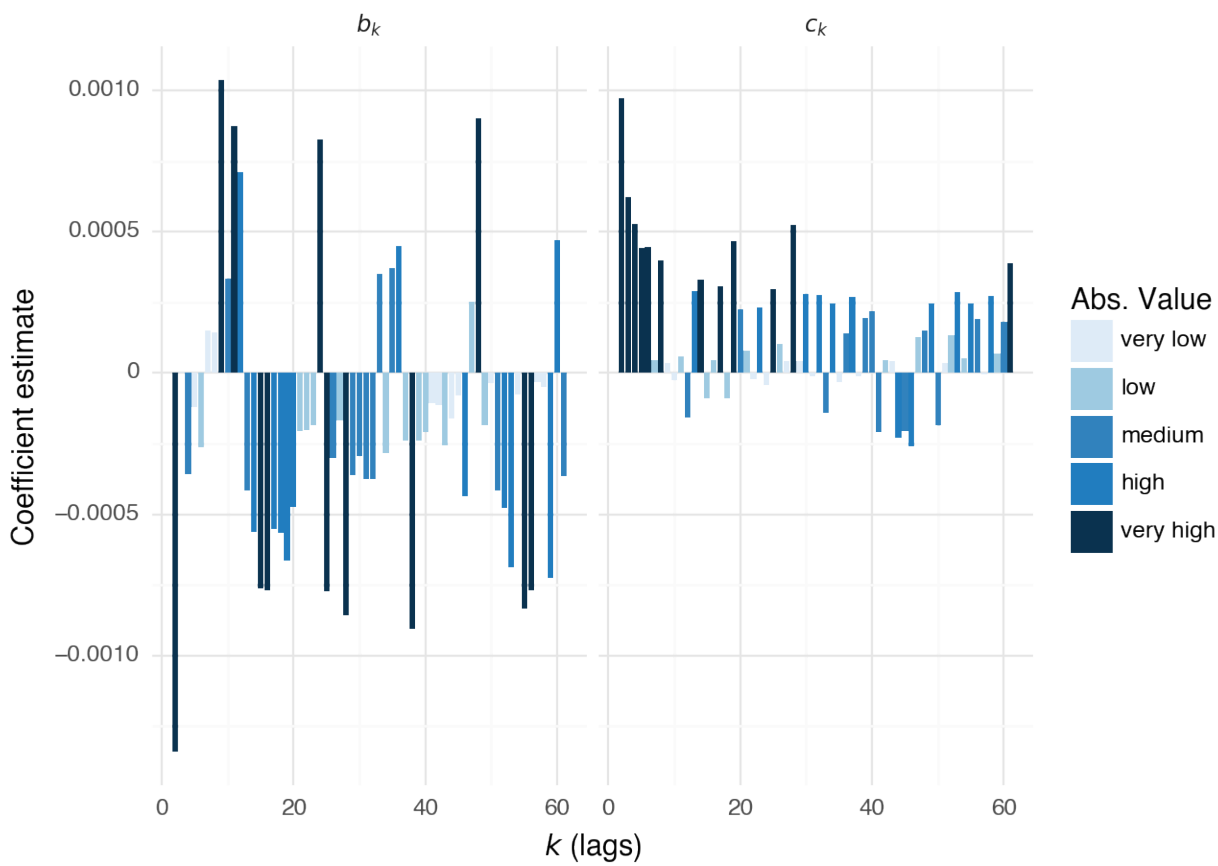

Figure 4