

# Independent bivariate sorts of portfolio (Size-MOM and Size-STR)

Václav Jisl, Václav Šmíro

Institute of Economic Studies, Charles University

March 31, 2024

## 1. Introduction

In this paper we investigate the role of key firm characteristics in explaining the cross-sectional variation of stock returns in the U.S. equity market. Specifically, we sort a subset of S&P 500 firms into portfolios based on two independent bivariate sorts. Size-MOM (momentum), and Size-STR (short term reversal). We evaluate the time series properties of these portfolios and their performance using the Fama French model. Furthermore, we estimate Fama MacBeth cross sectional regressions to assess the joint explanatory power of the sorting variables.

Our objective is to empirically test theoretical insights from asset pricing literature, which suggest that these characteristics are fundamental in determining expected returns. Specifically, we compute 12-month momentum and 1-month short term reversal signals. We sort stocks into 3x3 portfolios by size and either momentum or short term reversal and calculate both equal-weighted and value-weighted portfolio returns over time. We evaluate performance using standard asset pricing models, including CAPM and Fama-French 5-factor extensions with momentum and liquidity, proxied via short term reversal.

## 2. Data description

In the study, we utilize a dataset covering a subset of the S&P 500 index. Different data frames available for our analysis cover different periods, but in general span over majority of the period between 2000 and 2025. The main entry data source is obtained from the class provided material and includes objects covering firm level information about some of their basic characteristics. Specifically, the source file contains the following objects:

- **Symbols:** A list of tickers representing the S&P 500 constituents. This serves as a reference index for all firm specific data. The set includes all tickers included in the index at the time of creation of the dataset. We, however, only analyzed a subset of this data, hence we had to filter accordingly and only use part of the data provided.
- **OHLCV\_sap500:** This object provides daily historical market data for each firm in the panel. For every stock, we observe open, high, low, close, and adjusted close prices, along with daily trading volume.
- **MktCap\_sap500:** This dataset contains the daily market capitalization of each firm over the sample period. It allows for the construction of value-weighted portfolios and the sorting of firms based on size.

- **book\_value\_sap500:** For each firm, we observe periodic accounting-based indicators including stock price, book value per share, and the price to book ratio. These values are reported quarterly.

Each of these objects is structured as a list of time series, indexed by firm tickers. This structure facilitates firm level analysis over time and enables the construction of cross-sectional asset pricing tests.

On top of this we also work with several factor-specific data sources. These include long-term daily data spanning over some hundred years. These materials are formatted in rather straight forward manner only including date stamp and respective values.

### 3. Data manipulation

In order to be able to perform our analysis, we have to prepare and transform raw available financial datasets into analyzable formats suitable for portfolio construction. The workflow combines data filtering, return calculation, anomaly signal extraction, and merging with factor models. The goal is to create clean, aligned, and interpretable datasets that support empirical testing of cross-sectional return patterns.

To ensure consistency, the first step is filtering available data to only include stocks listed in our scope. Stocks with missing or malformed adjusted close columns were also filtered out. An example of a problematic ticker we had to omit was TSCO. We then computed log returns using the adjusted close prices for each stock. A safe wrapper function had to be employed to handle missing values and misaligned data. Factor data (RF, MOM, STR) were then read from provided CSV files and also cleaned to remove missing values, duplicate dates, and malformed entries.

When calculating momentum and short term reversal we had to be careful to properly distinguish effects of both these variables. For momentum a 12-month cumulative return was calculated, excluding the most recent month. This was implemented using a 252-day rolling window and skipping the last 21 days. Similarly, a 1-month reversal signal was computed by excluding the last day of return in a 21-day rolling window. Only then, for both momentum and short-term reversal strategies, stocks could be sorted into terciles based on size and the respective signal.

### 4. Empirical methodology

As part of the empirical methodology, portfolios are constructed based on two firm-level characteristics. One set of portfolios is sorted by firm size (Size) and momentum (MOM), and the other by firm size (Size) and short-term reversal (STR).

Bivariate portfolios are independently sorted based on these characteristics. Stocks are first ranked according to market capitalization (Size) and either momentum or short-term reversal signals. The sorting is based on cross-sectional percentiles: firms below the 30th percentile are classified as “Low,” those between the 30th and 70th percentile as “Medium,” and those above the 70th percentile as “High.” This sorting procedure results in 3×3 portfolios for each characteristic pair (Size–MOM and Size–STR).

The momentum signal is calculated as the cumulative return over the previous 12 months, excluding the most recent month, in order to avoid contamination from short-term reversal

effects. The short-term reversal signal is measured as the return over the most recent month, excluding the final trading day. Firm size is measured using market capitalization at the end of the month.

After sorting, the portfolios are formed and both equally-weighted and value-weighted returns are calculated. Then we evaluate the average monthly returns across all portfolios and the spreads between high and low groups, to assess the economic significance of the sorting variables.

We then proceed with factor return analysis. To determine whether the portfolio returns are statistically significant, it is necessary to adjust them for systematic risk using the CAPM model, the Fama–MacBeth regression, and the Fama–French three-factor model, extended to include liquidity and momentum factors.

The difference between the CAPM model and the two-step Fama–MacBeth methodology lies in their approach to quantifying the relationship between asset returns and explanatory variables.

### **CAPM model**

Capital Asset Pricing Model is a single-factor time-series regression model that explains the expected return of a portfolio (or asset) as a linear function of its exposure to the market risk premium. The model has the following form:

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

Where  $R_{i,t}$  is the return of the asset,  $R_{f,t}$  is the risk-free rate, and  $R_{m,t}$  is the market return. The coefficient  $\beta_i$  captures the level of systematic risk, while  $\alpha_i$  represents any excess return not explained by the market.

### **Fama–MacBeth regression**

In contrast, the Fama–MacBeth regression method is based on a cross-sectional approach and is used to examine whether specific firm-level characteristics explain differences in returns across stocks. In each period, a cross-sectional regression of future returns on characteristics such as size (Size), momentum signal (MOM), or short-term reversal (STR) is performed.

$$R_{i,t+1} = \gamma_{0,t} + \gamma_{1,t} \cdot Size_{i,t} + \gamma_{2,t} \cdot MOM_{i,t} + \gamma_{3,t} \cdot STR_{i,t} + \varepsilon_{i,t+1}$$

The resulting regression coefficients  $\gamma_{j,t}$  are then averaged over time and serve as an estimate of the long-term predictive power of the individual variables, which is represented by the equation below:

$$\bar{\gamma}_j = \frac{1}{T} \sum_{t=1}^T \gamma_{j,t}$$

The average intercept value can be interpreted as an  $\alpha$  estimator. This method thus allows testing whether the selected characteristics provide statistically and economically significant information about future returns, with the associated t-statistic indicating whether a given factor predicts returns across stocks.

### **Fama–French factor model**

In contrast to the two methods described above, the Fama–French factor model is constructed as a multifactor time-series regression, where portfolio returns are explained through exogenously defined factors that represent systematic risks. In addition to the market risk premium (MKT) included in the CAPM model, it also incorporates size (SMB) and value versus growth (HML) factors, and is further extended in our empirical analysis to include momentum (UMD) and liquidity (LIQ) factors.

$$R_{p,t} - R_{f,t} = \alpha_p + \beta_1 \cdot MKT_t + \beta_2 \cdot SMB_t + \beta_3 \cdot HML_t + \beta_4 \cdot UMD_t + \beta_5 \cdot LIQ_t + \varepsilon_{p,t}$$

For better clarity, the factors are listed below:

- **MKT**: market risk premium
- **SMB**: size factor (small minus big)
- **HML**: value factor (high minus low)
- **UMD**: momentum factor (up minus down)
- **LIQ**: liquidity factor (Pastor–Stambaugh)

The Fama–French model is primarily used for analyzing portfolio returns and for testing whether a statistically significant excess return ( $\alpha$ ) persists after controlling for commonly accepted risk factors. While the Fama–MacBeth regression tests the predictive power of firm-level characteristics at the individual stock level, the Fama–French model assesses whether portfolio returns correspond to exposures to widely accepted market factors.

It is important to note that the Fama–French model is estimated using Newey–West adjusted standard errors, which correct the standard errors of regression coefficients to make them robust to heteroskedasticity and autocorrelation.

Moreover, the importance of specific factors varies across different portfolio constructions. For Size–MOM portfolios, the UMD factor is essential, as it captures systematic momentum risk. The SMB and HML factors help control for size and value exposures, respectively.

In contrast, for Size–STR portfolios, the momentum factor is not informative, as STR strategies are negatively correlated with momentum. Instead, the liquidity factor (LIQ) becomes particularly relevant, since STR strategies often involve illiquid and temporarily overreacting stocks.

To evaluate the performance of such portfolios, it is essential to estimate the alpha coefficient ( $\alpha$ ) which reflects the average return of the portfolio after adjusting for systematic risk represented by factor exposures. The interpretation of alpha is as follows:

- If  $\alpha > 0$ : the portfolio generates a positive excess return

- If  $\alpha = 0$ : the return is fully explained by known risk factors
- If  $\alpha < 0$ : the portfolio is underperforming relative to its risk profile

Therefore, alpha estimation serves as a key metric in determining whether a portfolio's return is purely a result of factor exposure or if an unexplained return component remains.

## 5. Results

### a) Bivariate sorting results:

#### i) Size – Momentum sorting

We construct 3×3 bivariate portfolios sorted independently on firm Size and Momentum (MOM). Table 1 shows the distribution of 1 349 632 observations across the 9 portfolios. While the portfolios are reasonably balanced, some warnings indicated non-unique breaks, potentially due to low cross-sectional dispersion in firm sizes. Nevertheless, portfolios are well-populated, allowing for reliable cross-sectional analysis.

Table 1: Number of Observations in Size–Momentum Portfolios

	<b>Momentum Q1 (Low)</b>	<b>Momentum Q2 (Medium)</b>	<b>Momentum Q3 (High)</b>
<b>Size Q1 (Small)</b>	182,710	176,297	136,766
<b>Size Q2 (Medium)</b>	127,604	192,825	133,656
<b>Size Q3 (Big)</b>	94,547	170,792	134,435

We compute equal-weighted and value-weighted daily returns for each of the 3×3 Size–Momentum sorted portfolios. For each trading day, the return is calculated as the arithmetic mean of all stocks within the respective portfolio.

Table 2: Samples of Daily Returns (Equal-Weighted and Value-Weighted)

<b>Date</b>	<b>Size Group</b>	<b>Momentum Group</b>	<b>EW Return</b>	<b>VW Return</b>
2001-01-02	1	1	-0.03504712	-0.035117398
2001-01-02	1	2	-0.02323731	-0.003843025
2001-01-02	1	3	-0.03360554	-0.023203186
2001-01-02	2	1	-0.04117766	-0.043542768
2001-01-02	2	3	0.01064717	0.010647172
2001-01-03	1	1	0.06984971	0.081398499
...	...	...	...	...
2024-02-27	2	1	-0.00088042	-0.00240182
2024-02-27	2	2	0.00457777	0.00677125
2024-02-27	2	3	0.00070941	0.00308215
2024-02-27	3	1	0.00712620	0.00528094
2024-02-27	3	2	0.00377927	0.00440145

Date	Size Group	Momentum Group	EW Return	VW Return
2001-01-02	1	1	-0.03504712	-0.035117398
2001-01-02	1	2	-0.02323731	-0.003843025
2001-01-02	1	3	-0.03360554	-0.023203186
2001-01-02	2	1	-0.04117766	-0.043542768
2001-01-02	2	3	0.01064717	0.010647172
2001-01-03	1	1	0.06984971	0.081398499
...	...	...	...	...
2024-02-27	3	3	-0.00123237	-0.00102509

Table 2 presents a sample of daily returns at the beginning and end of the sample period. On January 2, 2001, returns were predominantly negative, with the (Size = 2, Momentum = 1) portfolio showing  $-4.12\%$  (EW) and  $-4.35\%$  (VW). In contrast, portfolios with high momentum, such as (Size = 2, Momentum = 3), delivered positive returns of  $1.06\%$  across both weighting schemes.

By February 27, 2024, daily returns were more moderate across all portfolios. For example, the (Size = 3, Momentum = 1) portfolio returned  $0.71\%$  (EW) and  $0.53\%$  (VW), while (Size = 3, Momentum = 3) produced a slight negative return.

## ii) Size – Short-term reversal sorting (STR):

To further explore cross-sectional return patterns, we construct  $3 \times 3$  portfolios based on independent sorts of firm Size and short-term reversal (STR). Table 2 summarizes the number of observations allocated to each portfolio over the sample period, totaling 1,405,534 observations.

Although the groupings are generally well distributed, we encountered warnings indicating non-unique breakpoints in the Size variable. This likely reflects limited variation in firm sizes at specific points in time.

Table 3: Number of Observations in Size–STR Portfolios

	STR Q1 (Low)	STR Q2 (Medium)	STR Q3 (High)
Size Q1 (Small)	188,261	194,708	163,229
Size Q2 (Medium)	123,899	189,705	129,356
Size Q3 (Big)	109,612	177,582	129,182

We compute equal-weighted and value-weighted daily returns for each of the  $3 \times 3$  Size–STR sorted portfolios. For each trading day, the equal-weighted return is calculated as the average return of all stocks within the given portfolio, while the value-weighted return is based on market capitalization weights.

Table 4: Samples of Daily Returns (Equal-Weighted and Value-Weighted)

Date	Size Group	STR Group	EW Return	VW Return
2024-02-27	2	1	0.00411700	0.00375152
2024-02-27	2	2	0.00243828	0.00095213
2024-02-27	2	3	0.00235196	0.00671175
2024-02-27	3	1	0.00179421	0.00513443
2024-02-27	3	2	0.00546365	0.00249500
2024-02-27	3	3	0.00053883	-0.00275826

Table 4 represents the final day in the sample. Unlike the Size–Momentum portfolios, these returns are generally modest and positive, with the exception of a slight negative value-weighted return in the (Size = 3, STR = 3) portfolio. The (Size = 3, STR = 2) portfolio exhibited the highest equal-weighted return of 0.55%, suggesting short-term reversal effects among large firms. Differences between EW and VW returns indicate the influence of firm size and weighting method on return dynamics.

## b) Fama–French model

To explain cross-sectional variation in portfolio returns, we employ the Fama–French model augmented with momentum (MOM) and liquidity (LIQ) factors. Before running the regressions, we cleaned and aligned the daily factor dataset with our portfolio returns.

The final dataset spans the period from November 1926 to December 2024, comprising 25,800 daily observations. Table 5 illustrates sample values from the beginning and end of the dataset. Factor returns are reported in daily percentages and are generally small, typically ranging from  $-1\%$  to  $+1\%$ . Notably, factors such as MOM and LIQ exhibit greater variability, reflecting their higher short-term volatility.

Table 5: Sample of Daily Fama–French Factor Returns

Date	Mkt.RF	SMB	HML	RF	MOM	LIQ
1926-11-03	0.20	−0.20	−0.33	0.013	0.54	0.81
1926-11-04	0.59	−0.12	0.65	0.013	−0.51	0.37
...	...	...	...	...	...	...
2024-12-30	−1.09	0.12	0.74	0.017	0.09	0.20
2024-12-31	−0.46	0.00	0.71	0.017	−1.06	0.86

In the next step, we estimate risk-adjusted alphas for each of the Size–Momentum and Size–STR portfolios by running time-series regressions of portfolio excess returns on these factors. We apply Newey–West adjusted standard errors to account for autocorrelation and heteroskedasticity in the residuals.

The table 6 below reports the estimated alphas and corresponding  $t$ -statistics for four portfolio strategies: equal- and value-weighted variants of the Size–MOM and Size–STR portfolios. All alphas are calculated using daily returns.

Table 6: Estimated Alphas and t-statistics for Fama–French model

Portfolio	Alpha	t-stat
MOM Equal-Weighted	−0.00532	−52.81
MOM Value-Weighted	−0.00524	−52.32
STR Equal-Weighted	−0.00572	−53.90
STR Value-Weighted	−0.00567	−53.78

The results show that all four portfolios exhibit strongly negative and highly statistically significant alphas. This implies that, even after controlling for well-known risk factors, the portfolios significantly underperform. Such a result would normally suggest poor risk-adjusted performance. However, the magnitude of the alphas is large for daily return data and is thus unrealistic.

This discrepancy raises concerns about the specification or scaling of returns. Possible explanations include incorrect units (e.g., percent vs. decimal), overfitting due to high-frequency noise, or structural breaks in the factor-return relationship.

To isolate the performance of extreme momentum strategies, we compute **3–1 spreads** by subtracting the returns of low-momentum portfolios (Q1) from high-momentum portfolios (Q3), within each size group. We construct both equal-weighted (EW) and value-weighted (VW) spreads, and merge these returns with Fama–French daily factors. The excess returns are then regressed on the factor set using Newey–West adjusted standard errors.

Table 7: Estimated Alphas and t-statistics for 3–1 Momentum Spreads

Portfolio	Alpha	t-stat
3–1 Equal-Weighted	−0.00525	−26.28
3–1 Value-Weighted	−0.00531	−26.04

Once again, the estimated alphas are statistically significant but economically puzzling. Both momentum spread portfolios yield strong negative alphas, contrary to the typical expectation of positive excess performance for long-short momentum strategies. The large magnitude of the alpha estimates (around −0.5% per day) reinforces the earlier concern that the regression may suffer from data scaling issues, portfolio misalignment, or omitted variable bias.

Despite multiple adjustments and validation attempts, we were unable to reconstruct results fully consistent with benchmark references, such as those demonstrated in lecture materials. This highlights the sensitivity of alpha estimation to data construction and portfolio formation details.

### c) Fama–MacBeth Regression

In the final part of our analysis, we shift to cross-sectional estimation of risk premia using the two-step Fama–MacBeth regression framework. First, we compute firm-level characteristics and assemble them into daily panels. Each panel includes returns and relevant covariates (e.g., size, momentum, short-term reversal). We merge all stock-level panels and clean the dataset by removing missing values.

In the second step, we perform daily cross-sectional regressions of returns on firm characteristics. The resulting daily slope coefficients are then averaged across time, and standard errors are computed to assess significance.

The table 8 shows a representative cross-section of portfolio returns and factor exposures used in our Fama–MacBeth regressions. While the excess returns exhibit clear variation across size and momentum groups, their relationship with the factor values appears unstable. Some high-momentum portfolios yield negative excess returns despite favorable factor conditions, suggesting weak alignment between return patterns and pricing factors.

Table 8: Daily Cross-Sectional Results for Fama–MacBeth regression

Date	Size	MOM	EW Return	VW Return	Mkt.RF	SMB	HML	MOM	LIQ	EW	VW
24-02-27	2	1	−0.00088	−0.00240	0.27	1.19	−0.44	−0.92	−0.44	−0.02188	−0.02340
24-02-27	2	2	0.00458	0.00677	0.27	1.19	−0.44	−0.92	−0.44	−0.01642	−0.01423
24-02-27	2	3	0.00071	0.00308	0.27	1.19	−0.44	−0.92	−0.44	−0.02029	−0.01792
24-02-27	3	1	0.00713	0.00528	0.27	1.19	−0.44	−0.92	−0.44	−0.01387	−0.01572
24-02-27	3	2	0.00378	0.00440	0.27	1.19	−0.44	−0.92	−0.44	−0.01722	−0.01660
24-02-27	3	3	−0.00123	−0.00103	0.27	1.19	−0.44	−0.92	−0.44	−0.02223	−0.02203

Across the selected day, all portfolios report **negative excess returns**, and differences between equal-weighted and value-weighted outcomes are small. Despite reasonable structure in the data, the explanatory power of the regressors seemed limited, and the daily signal-to-noise ratio remained low throughout the sample. In the end, we were not able to successfully finish the regression. The semi-results were disappointing, and we just couldn't make the last step work.

## 6. Conclusion

In summary all portfolios yielded highly statistically significant **negative alphas**, suggesting substantial underperformance even after adjusting for common risk factors. However, the absolute magnitude of daily alphas (around −0.5%) was implausibly large, raising questions about data scaling, portfolio formation logic, or unaddressed biases.

To further assess pricing relationships, we attempted to implement a Fama–MacBeth cross-sectional regression. We successfully constructed daily panels of returns and factor exposures and verified the structure of the input data. However, despite several iterations, we were not able to complete the final step of the estimation procedure. The semi-results were inconclusive and ultimately fell short of producing reliable and interpretable risk premia.