# Portfolio sorting in empirical asset pricing*

Mads Markvart Kjær**

Empirical Asset Pricing, Spring 2022

*The note provides step-by-step details on portfolio sorting and its use in empirical asset pricing. The note is prepared for use only in the Master's course "Empirical Asset Pricing". Please do not cite, circulate, or use for purposes other than this course. This note draws on the treatment of the topic found in Bali, Engle and Murray (2016), and borrows their notation for the various portfolio sorts.

**CREATES, Department of Economics and Business Economics, Aarhus University, Fuglesangs Allé 4, DK-8210 Aarhus V, Denmark. E-mail: mads.markvart@econ@econ.au.dk.

# Contents


**1 Introduction** **1**

**2 Univariate portfolio sorts** **2**

2.A Computing breakpoints . . . . . 2

2.B Portfolio formation . . . . . 3

2.C Computing portfolio returns . . . . . 4

2.D Examining the portfolio returns . . . . . 5

**3 Bivariate portfolio sorts** **6**

3.A Independent double sort . . . . . 6

3.B Dependent double sort . . . . . 8

**4 Empirical illustration** **9**

4.A Momentum portfolios based on a univariate sort . . . . . 9

4.B A momentum factor based on an independent double sort . . . . . 13

**References** **15**

# 1. Introduction

Portfolio sorts are an integral part of empirical asset pricing, and portfolio sorting is one of the most commonly applied methodologies for exploring and identifying cross-sectional relationships between asset characteristics and future returns. In a nutshell, the approach rests on the construction of portfolios based on one or more sorting variables believed to have information about predictable variation in the cross-section of future returns. As such, searching for factors that explain the variation in expected returns is equivalent to identifying variables that can predict cross-sectional variation in future returns in a universe of assets. Portfolio sorting is highly useful for this particular task as it allows us to examine cross-sectional relationships without making any *a priori* assumptions about the relationship between the sorting variable(s) and future returns. This is an important feature that allows the data to speak for us. In fact, portfolio sorts can assist in the discovery of linear and non-linear relationships alike. This flexibility notwithstanding, we typically start from an economically motived idea, theory, or observation that provides us with a set of testable implications for the relationship(s) that should emerge. A drawback of portfolio sorts, on the other hand, is that conventional methods only allow the researcher to control for a very limited set of factors when examining the cross-sectional relations of interest. Univariate sorts, for example, consider just a single sorting variable, and therefore do not condition on any other information. Portfolios sorts based on two sorting variables are referred to as bivariate (or double) sorts. Bivariate sorts can be either independent or dependent. In independent sorts, the order of the factors is inconsequential. In dependent sorts, the order of the factors matter greatly and is crucial in determining the results. Dependent double sorts are often useful when the aim is to examine the relationship between one sorting variable and future returns conditional on another sorting variable.

To build a set of portfolios, we first need to assemble a universe of assets. This universe can comprise individual stocks, stock market indices, currencies, corporate bonds, and/or commodities. Perhaps the most well known and most-studied universe of assets is the CRSP sample of individual U.S. stocks (Harvey, Liu and Zhu, 2016, Harvey and Liu, 2019). The general approach is then to select a universe of individual assets, examine one or more economically motivated variables linked to cross-sectional variation in future returns, construct breakpoints for portfolio allocations, allocate securities and compute portfolio returns, and study the returns of the portfolios. We detail the steps and some of the choices that have to be made when sorting assets into portfolios during this note.

The note ends with two empirical illustration using the CRSP monthly stock file. We first demonstrate the univariate portfolio sorting routine using the well known momentum anomaly (Jegadeesh and Titman, 1993) as our example. We will simultaneously discuss and illustrate how various choices made by the researcher can (substantially) impact the final results and conclusions. Specifically, we consider the impact of using only a subset of stocks to determine breakpoints versus using all stocks available as well as the impact of using equal- and value-weighted portfolio returns. Second, we illustrate the independent bivariate (double) sorting approach by replicating the momentum factor of Carhart (1997) from Kenneth R. French's data library using the CRSP sample.

The note proceeds as follows. Section 2 outlines the univariate sorting approach. Section 3 considers bivariate portfolio sorts, and distinguishes between independent and dependent double sorts. Last, Section 4 provides a set of empirical illustrations.

## 2. Univariate portfolio sorts

A univariate portfolio sort only considers a single sorting variable. We will denote this by $F_{i,t}$, where $i = 1,2,\ldots,N$ refers to the $i$th asset and $t = 1,2,\ldots,T$ is the time subscript. The objective is to study the cross-sectional relationship between the factor $F_{i,t}$ and the future (excess) return to the universe of assets under study. As mentioned above, a univariate portfolio sort does not allow the researcher to control for any other effects when studying the cross-sectional relation. We can distill the process of conducting and interpreting a univariate portfolio sort into four basic steps: 1) Calculate breakpoints for dividing the universe of assets into portfolios, 2) allocate assets into portfolios using the breakpoints, 3) compute portfolio returns, and 4) examine the cross-sectional variation in average (excess) returns across the portfolios. We outline the steps and choices below.

### 2.A. Computing breakpoints

The first step is to compute breakpoints for the cross-sectional distribution of the sorting variable $F_{i,t}$ for each time period $t$. The breakpoints are then used to allocate assets into portfolios. Assets with values of $F_{i,t}$ that are smaller than the first breakpoint will be allocated to the first portfolio. Assets with values of $F_{i,t}$ that are between the first and second breakpoint will be allocated to the second portfolios. And so on. Last, securities with values of $F_{i,t}$ that are larger than the largest breakpoint will be allocated to the last portfolio. Suppose that we want to form $n_p$ portfolios, then we need to determine $n_p - 1$ breakpoints every period. The number of portfolios and breakpoints are kept constant over time. The value of the $k$th breakpoint, on the other hand, is likely to vary over time. Let us denote the $k$th breakpoint for period $t$ by $\mathcal{B}_{k,t}$ for $k = 1,2,\ldots,n_p - 1$ and $t = 1,2,\ldots,T$. These breakpoints can be determined by the percentiles of the cross-sectional distribution of the sorting variable $F_{i,t}$. In particular, denote by $p_k$ the $k$th percentile of the values of $F_{i,t}$ across all available securities that determines the $k$th breakpoint for portfolio allocations and let the us define the $k$th breakpoint at time $t$ more formally as

$$\mathcal{B}_{k,t} = \text{Percentile}_{p_k}\left(F_{i,t}\right), \tag{1}$$

where $\text{Percentile}_p\left(\cdot\right)$ denotes the $p$th percentile and $F_{i,t}$ is the cross-sectional set of valid values for the sorting variable at time $t$. Importantly, the percentiles, and by extension the breakpoints, are increasing in $k$ so that $0 < p_1 < p_2 < \cdots < p_{n_p-1}$ and $\mathcal{B}_{1,t} \leq \mathcal{B}_{2,t} \leq \cdots \leq \mathcal{B}_{n_p-1}$ for all $t$. Although the percentiles are required to be strictly increasing, this is not necessarily mean that the breakpoints will be strictly increasing. As an example, suppose that many stocks share the same value of $F_{i,t}$, then two or more breakpoints may be identical.

**Choosing assets for breakpoints** While it may seem intuitive to use all available assets for the determination of breakpoints, this is not always the optimal approach to pursue. In fact, several seminal studies within empirical asset pricing determine breakpoints using only a subset of the available assets. A well known example is found within research on the cross-section of stock returns using the CRSP stock file, where it is common to employ all stocks listed on NYSE, AMEX, and NASDAQ for the study, but determine breakpoints using NYSE stocks only. That is, breakpoints $\mathcal{B}_{k,t}$ are determined using NYSE stocks at time $t$ and then these breakpoints are used to allocate stocks from NYSE, AMEX, and NASDAQ into portfolios. This approach is used, among others, in the construction of the well known Fama and French (1993) size and value factors, and we will see a similar example in our empirical illustrations below for the momentum factor. Luckily, there is nothing in the general method that prevents us from doing so — a highlight of the flexibility of the portfolio sorting methodology. In fact, this is the reason we often separate the determination of breakpoints and the allocation of assets into two distinct steps.

**Choosing the number of portfolios** An equally important issue to address is the number of portfolios $n_p$ that we wish to form. This choice is largely a trade-off between the number of assets in each portfolio and the cross-sectional variation in expected returns that can reliably be identified using the sorting variable $F_{i,t}$. If there is only a small number of assets in each portfolio, then portfolio returns are likely to be noisy and include a fair amount of idiosyncratic risk. This is mitigated by allocating a *large* number of assets to each portfolio. Yet, having a large number of assets in each portfolio reduces the number of portfolios that can be constructed and, consequently, the cross-sectional variability in portfolio returns, which can hinder the detection of cross-sectional relationships. In the end, most studies settle on a choice between three and 20 portfolios. The most common being five and ten. Overall, this is a choice that should be discussed and argued for.

**Choosing percentiles** Another flexibility in the portfolio sorting methodology is the choice of percentiles for breakpoint determination. A simple and common choice is to build portfolios whose breakpoints are determined using evenly spaced percentiles. For example, if we are interested in constructing five portfolios, then one would use the 20th, 40th, 60th, and 80th percentiles of the cross-sectional distribution of $F_{i,t}$. In this case, one can determine the percentiles for the $k = 1,2,\ldots,n_p - 1$ breakpoints as $k \times (1/n_p)$. However, while the evenly-spaced approach is arguably the most common in the literature for univariate sorts, other approaches are similarly widely used. For instance, when Fama and French (1993) build their size and value factors, they consider the 30th and the 70th percentiles of the sorting variable $F_{i,t}$ as the breakpoints. In the end, this is a choice that the researcher has to make (and argue for).

### 2.B. Portfolio formation

With the breakpoints in hand, we can turn to the second step of allocating individual securities into their respective portfolios. At each time $t$, we take all securities with values of $F_{i,t}$ that are less than or equal to the first breakpoint $\mathcal{B}_{1,t}$ and allocate them to the first portfolio (P1). Portfolio 2 (P2) contains all securities with values of $F_{i,t}$ that lie between (or equals) $\mathcal{B}_{1,t}$ and $\mathcal{B}_{2,t}$. And so on. The last portfolio contains all securities with values of $F_{i,t}$ bigger than $\mathcal{B}_{n_p-1,t}$.

More formally, suppose that we have $k = 1,2,\ldots,n_p - 1$ breakpoints defined using the chosen percentiles. Moreover, let us define $\mathcal{B}_{0,t} = -\infty$ and $\mathcal{B}_{n_p,t} = \infty$ to exhaust all possible values of $F_{i,t}$. We can then identify all assets $i$ that belong to the $k$th portfolio $P_{k,t}$ formed at time $t$ as the set of assets with values of $F_{i,t}$ that satisfy the relation

$$P_{k,t} = \left\{ i \middle| \mathcal{B}_{k-1,t} \leq F_{i,t} \leq \mathcal{B}_{k,t} \right\} \tag{2}$$

for $k = 1,2,\ldots,n_p$. Note that this approach puts all securities with the lowest values of the sorting factor $F_{i,t}$ in the first portfolio and all securities with the largest values of the sorting factor $F_{i,t}$ in the last portfolio by construction.

### 2.C. Computing portfolio returns

With individual assets allocated into their respective portfolios, we need to discuss how to construct portfolio returns. This is not necessarily an innocuous choice. We often distinguish between portfolios with equal-weighted and portfolios with value-weighted returns. Equal-weighted portfolio returns for portfolio $k$ are simply computed as

$$r_{k,t} = \frac{1}{N_{k,t}} \sum_{i=1}^{N_{k,t}} r_{i,t}, \tag{3}$$

where $N_{k,t}$ denotes the number of securities in portfolio $k$ at time $t$, and the sum is taken over all securities in the $k$th portfolio at time $t$. Value-weighted returns are computed as the weighted average of returns, where the weights are based on the relative market capitalization of asset $i$. Let $ME_{i,t}$ denote market capitalization of security $i$ at time $t$, then value-weighted portfolio returns can be defined as

$$r_{k,t} = \frac{\sum_{i=1}^{N_{k,t}} ME_{i,t-1} \times r_{i,t}}{\sum_{i=1}^{N_{k,t}} ME_{i,t-1}}. \tag{4}$$

Note that we lag the market capitalization by one period so that portfolio weights are based on information observable at the time of portfolio formation. Value-weighting is most appropriate when the individual assets are stocks (e.g., stocks from the CRSP sample). The use of value-weighted returns in empirical analyses are often viewed as more representative of the returns that an investor would have realized by implementing a given strategy. Intuitively, value-weighting places a higher weight on larger (in terms of market equity) stocks, which tend to be more liquid (and visible), and therefore cheaper and easier to trade. An often raised concern with equal-weighted portfolios is that their performances are driven by small and illiquid stocks that are hard (and expensive) to trade. That is, any observed return phenomena may then simply be driven by a liquidity premium.

Finally, in addition to computing portfolio returns, it is also custom to compute the returns to a zero-cost long-short portfolio that takes a long position in the last portfolio and a short position in the first portfolio (or vice versa if the relationship demands it). This is also often referred to as a spread portfolio. These returns will be equal- or value-weighted depending on the choice for the portfolios themselves. The return on this long-short portfolio can then be used to assess whether a cross-sectional relation can be detected. It also represent a trading strategy (i.e., factor investing). That is, if current values of the sort factor are informative about future cross-sectional differences in returns, then the long-short portfolio should deliver a positive and significant return on average. We return to this point below. We end by remarking that the average return to the traded long-short factor equals its risk price in a cross-sectional asset pricing exercise if the factor is priced.

### 2.D. Examining the portfolio returns

The last step is to examine the portfolio returns. The main objective is to determine whether there is a reliable cross-sectional relation between the sorting variable $F_{i,t}$ and future asset returns.

**Descriptive statistics** The first step is to compute descriptive statistics for the portfolios, including mean portfolio (excess) returns, standard deviations, skewness, kurtosis, and Sharpe ratios (to mention the most important). This will help determine the reliability of the cross-sectional relationship and the characteristics of the stocks in the different portfolios. Moreover, it is custom to test whether average portfolio returns are significantly different from zero, and also test whether the long-short strategy has an average return that is different from zero. An average return that is significantly different from zero is indicative of cross-sectional predictability. Usually, one regress the portfolio (long-short) returns on a constant and evaluate the significance using a standard $t$-test with HAC standard errors (e.g., Newey and West (1987)). More importantly, we look for monotonic relationships in the average returns between the first and the last portfolio. If a monotonic pattern arises, it is a strong indication that the results of due to the ability of the sorting variable $F_{i,t}$ to predict the future cross-section of returns.

**Risk-adjusted returns** Last, we want to examine whether the patterns in average portfolio returns can be attributed to cross-sectional variation in portfolio sensitivities to systematic risk factors. That is, we want to test whether the return pattern survives controlling for known risk factors identified in the asset pricing literature. Models include the Capital Asset Pricing Model, the Fama and French (1993) three-factor model, the Carhart (1997) four-factor model, the Fama and French (2015) five-factor model, or the Hou, Xue and Zhang (2015) four-factor q-theory model. We are particular interested in the values and significance of the intercept in such models as they represent risk-adjusted returns (and a mean-variance spanning test for the long-short factor). If alpha is significantly different from zero, we have evidence of returns that are not attributable to known risk factors. However, we typically inspect the loadings on the risk factors as well to identify if and how our portfolio returns are related to some of these well established risk factors.

## 3. Bivariate portfolio sorts

This section presents bivariate (or double) portfolio sorts in which the universe of assets is sorted into portfolios based on two sorting variables rather than one. As such, the bivariate sorting procedure is designed to assess the cross-sectional relations between two sorting variables, which we refer to as $F^1_{i,t}$ and $F^2_{i,t}$, and future (excess) returns. The process for implementing a bivariate portfolio sort is similar to the univariate case, but differs materially in the first steps: constructing the breakpoints and forming the portfolios. The remaining steps are identical, and we will therefore not repeat the discussion here. When moving from univariate to bivariate portfolio sorts, it becomes critically important to distinguishes between two types of sorting procedures: independent double sorts and dependent double sorts. We outline each separately below and discuss their differences.

### 3.A. Independent double sort

The independent double sort builds portfolios by sorting on two variables independently. Accordingly, we therefore construct two sets of breakpoints in each period $t$. The first set of breakpoints corresponds to values of the first sorting variable $F^1_{i,t}$. The second set of breakpoints corresponds to values of the second sorting variable $F^2_{i,t}$, and is calculated completely independently of the breakpoints for $F^1_{i,t}$ — hence the name. The independence of the sorts implies that the ordering of the sorting variables is inconsequential. It makes no difference which variable is considered the first and which is considered the second. Switching the order will not affect on the results of the analysis.

**Computing breakpoints** The procedure for an independent double sort consists of computing breakpoints for both sorting variables and allocating assets accordingly. We will follow Fama and French (1993) and refer to these allocations as "groups". Portfolios will then be formed on the basis of the intersection of the groups of assets from each of the two independent sorts. Let $n_{p_1}$ represent the number of groups that will be created based on the first sorting variable $F^1_{i,t}$ and let $n_{p_2}$ be the number of groups that will be created based on the second sorting variable $F^2_{i,t}$. The total number of portfolios that will be formed is then $n_{p_1} \times n_{p_2}$. There are therefore $n_{p_1} - 1$ and $n_{p_2} - 1$ breakpoints for the first and second sort variables, respectively.

The breakpoints for each of the two sorting variables are calculated in exactly the same way as for the univariate portfolio sort. The breakpoints used to form the groups for the first sorting variable will be denoted $\mathcal{B}^1_{k,t}$ for $k = 1,2,\ldots,n_{p_1} - 1$ and the breakpoints for the second sorting variable are $\mathcal{B}^2_{j,t}$ for $j = 1,2,\ldots,n_{p_2} - 1$. The breakpoints are determined from the percentiles of the sorting variables

$$\mathcal{B}^1_{k,t} = \text{Percentile}_{p_k}\left(F^1_{i,t}\right) \tag{5}$$

and

$$\mathcal{B}^2_{j,t} = \text{Percentile}_{p_j}\left(F^2_{i,t}\right), \tag{6}$$

where $\text{Percentile}_p\left(\cdot\right)$ denotes the $p$th percentile of the sorting variable at time $t$. It is worth noting that the availability of values for $F^1_{i,t}$ and $F^2_{i,t}$ may differ across the assets and we, as researchers, must therefore decide whether to use all available data or restrict the sample to assets with values for both sorting variables at time $t$. Moreover, the set of asset used to calculate the breakpoints for $F^1_{i,t}$ may be different from the set of assets used to calculate the $F^2_{i,t}$ breakpoints. Nevertheless, as above, neither of the two sets of assets used to determine breakpoints is necessarily the same as the set of assets that will eventually be allocated to the portfolios. For example, NYSE stocks may be used to determine the breakpoints, whereas all stocks listed on NYSE, AMEX, and NASDAQ may be allocated to the groups and the portfolios.

**Portfolio formation** We then proceed to building portfolios by allocating assets into groups and portfolios. If there are $n_{p_1}$ groups based on the first sorting variable $F^1_{i,t}$ and $n_{p_2}$ groups based on the second sorting variable $F^2_{i,t}$, then there will be $n_{p_1} \times n_{p_2}$ portfolios at each time $t$. The portfolios at time $t$ are denoted $P_{k,j,t}$, where the first subscript indicates the group of the first sort variable and the second subscript indicates that of the second sort variable. In general, the portfolios are defined as the set

$$P_{k,j,t} = \left\{ i \middle| \mathcal{B}^1_{k-1,t} \leq F^1_{i,t} \leq \mathcal{B}^1_{k,t} \right\} \cap \left\{ i \middle| \mathcal{B}^2_{j-1,t} \leq F^2_{i,t} \leq \mathcal{B}^2_{j,t} \right\} \tag{7}$$

for $k = 1,2,\ldots,n_{p_1}$, $j = 1,2,\ldots,n_{p_2}$, and where we let $\mathcal{B}^1_{0,t} = \mathcal{B}^2_{0,t} = -\infty$ and $\mathcal{B}^1_{n_{p_1},t} = \mathcal{B}^2_{n_{p_2},t} = \infty$ to exhaust all values. Last, $\cap$ is the intersection operator. Thus, for a given asset $i$ to be held in portfolio $P_{k,j,t}$, the asset must have a value of $F^1_{i,t}$ in period $t$ that is between the $k-1$st and $k$th period $t$ breakpoints for the first sorting variable, and also have a period $t$ value of $F^2_{i,t}$ between the $j-1$st and $j$th period $t$ breakpoints for the second sorting variable.

In an independent double sort, the percentage of asset held in each portfolio will not always reflect the percentiles use to determine the breakpoints. This observation originates from the sorting variables often having a non-zero correlation. As each portfolio represents the intersection of groups formed based on the independent sorts, then a positive correlation between $F^1_{i,t}$ and $F^2_{i,t}$ will lead to portfolios that contain assets with high (or low) values of both sorting variables having a disproportionately large number of assets, while those portfolios that contain assets with low values of one sort variable and high values of the other contains will hold fewer assets. The opposite is true when the sorting variables are negatively correlated. Additionally, when the breakpoint sample and the full universe of assets are not the same, then the number of assets in each portfolio will further depend on the distribution of the sorting variables in the different samples. A classic example is the 25 size and book-to-market-equity portfolios constructed in Fama and French (1992, 1993), where the number of stocks in the different portfolios varies greatly.

We end this section by noting that the remaining steps involving the computation and examination of portfolios returns are identical to the univariate case. In particular, the researcher has to choose between equal- or value-weighted portfolio returns and which asset pricing model to use for the examination of the alphas (risk-adjusted returns). The descriptive statistics reported for the independently double sorted portfolios will similarly contain the same set of statistics such as mean, standard deviations, and Sharpe ratios.

### 3.B. Dependent double sort

A dependent double sort is very similarly to its independent counterpart in that the portfolios are constructed by sorting assets based on the values of two sorting variables $F^1_{i,t}$ and $F^2_{i,t}$. The main difference is that breakpoints for the second sorting variable in the dependent sort are formed *within* each group of the first sorting variable. A dependent portfolio sort is therefore useful when the objective is to understand the relation between $F^2_{i,t}$ and future returns conditional on $F^1_{i,t}$. The relation between $F^1_{i,t}$ and returns is not explicitly of interest as such, and $F^1_{i,t}$ is therefore regarded as a control variable only. Importantly, the ordering of the sorting variables now matters greatly and switching the order may lead to very different results.

**Computing breakpoints** The calculation of breakpoints for the dependent double sort starts the same way as in the independent double sort. However, it becomes critically important to distinguish which independent variable is the control variable, $F^1_{i,t}$, and which variable is part of the relation of interest, $F^2_{i,t}$. Unlike the independent sort, the ordering of the sorting variables matters materially in a dependent double sort. We first compute breakpoints for the first sorting variable $F^1_{i,t}$ (the control variable). Let $n_{p_1}$ denote the number of groups based on $F^1_{i,t}$ and let $\mathcal{B}^1_{k,t}$ for $k = 1,2,\ldots,n_{p_1} - 1$ be the breakpoints for the first sorting variable that are calculated identically to the independent sort case. As always, breakpoints may be calculated using a different set of assets than the universe that will eventually be sorted into portfolios.

With the breakpoints for the first sorting variable $F^1_{i,t}$ in hand, we can allocate assets into $n_{p_1}$ groups based on the breakpoints $\mathcal{B}^1_{k,t}$. The following step distinguishes the dependent double sort from the independent double sort. In a dependent double sort, we perform the second sort separately for each of the $n_{p_1}$ groups of assets formed using the breakpoints for the first sorting variable. The breakpoints for the groups based on the second sorting variable $F^2_{i,t}$ will therefore be different for each of the $n_{p_1}$ groups formed by sorting on the first sorting variable. The breakpoints for the second sorting variable $F^2_{i,t}$ are therefore defined as

$$\mathcal{B}^2_{k,j,t} = \text{Percentile}_{p_j}\left(F^2_{i,t} \middle| \mathcal{B}^1_{k-1,t} \leq F^1_{i,t} \leq \mathcal{B}^1_{k,t}\right), \tag{8}$$

where $k = 1,2,\ldots,n_{p_1} - 1$, $j = 1,2,\ldots,n_{p_2} - 1$, and $p_j$ is the percentile for the $j$th breakpoint based on the second sort variable, $n_{p_2}$ is the number of groups to be formed based on the second sort variable $F^2_{i,t}$, $\mathcal{B}^1_{0,t} = -\infty$, $\mathcal{B}^1_{n_{p_1},t} = \infty$, and $\left\{F^2_{i,t} \middle| \mathcal{B}^1_{k-1,t} \leq F^1_{i,t} \leq \mathcal{B}^1_{k,t}\right\}$ is the set of values of $F^2_{i,t}$ across all entities in the sample with values of $F^1_{i,t}$ that are between $\mathcal{B}^1_{k-1,t}$ and $\mathcal{B}^1_{k,t}$. Thus, for each of the $n_{p_1}$ groups of entities formed on $F^1_{i,t}$ , there will be $n_{p_2} - 1$ breakpoints for the second sort variable $F^2_{i,t}$. This is the conditional sorting part.

At this point, a brief discussion of the choice of the number of groups to use for each of the sorting variables is warranted. As always, the objective is to find a reasonable balance between the number of assets in each portfolio and the number of portfolios. In a dependent double sort, there is one major difference in choosing the number of groups compared to an independent double sort. In a dependent double sort, because sorting based on the second sorting variable $F_{i,t}^2$ is done within each group of assets formed by the first sorting variable, correlation between the sort variables does not play a role in determining an appropriate number of breakpoints. As long as there is a sufficient number of assets in each group formed by sorting on the first sorting variable $F_{i,t}^1$ to form $n_{p_2}$ groups of assets when sorting based on the chosen percentiles of $F_{i,t}^2$, the dependent double sort should provide an accurate assessment of the relation between $F_{i,t}^1$ and future returns.

**Portfolio formation** The allocation of assets into portfolios proceeds as previously. At each time $t$, all assets in the sample are first sorted into groups based on the breakpoints determined based on the first sorting variable $F_{i,t}^1$. Assets in each of those groups are then sorted into portfolios based on the conditional breakpoints of the second sorting variable $F_{i,t}^2$. In general, we can describe the portfolio consisting of assets belonging to group $k$ of the first sorting variable $F_{i,t}^1$ and group $j$ of the second sorting variable $F_{i,t}^2$ as

$$P_{k,j,t} = \left\{i|\mathcal{B}_{k-1,t}^1 \leq F_{i,t}^1 \leq \mathcal{B}_{k,t}^1\right\} \cap \left\{i|\mathcal{B}_{k,j-1,t}^2 \leq F_{i,t}^2 \leq \mathcal{B}_{k,j,t}^2\right\} \tag{9}$$

for $k = 1,2,\ldots,n_{p_1} - 1$, $j = 1,2,\ldots,n_{p_2} - 1$. All available assets are therefore placed into one of $n_{p_1} \times n_{p_2}$ portfolios at time $t$. When the sample used to calculate the breakpoints is the same as the sample that is sorted into portfolios, then the percentage of assets in any given portfolio is easily calculated from the percentiles used to calculate the breakpoints. That is, there is a one-to-one correspondence between percentiles and the number of stocks in each portfolio. This does not hold when the breakpoints sample is not identical to the sample that is used to form the portfolios.

## 4. Empirical illustration

This section presents two empirical illustrations of the portfolio sorting methodology using the CRSP stock file as our universe of assets. We start with an example designed to highlight the univariate portfolio sorting approach and the impact of the choices made along the way. In particular, we build ten portfolios sorted on momentum (Jegadeesh and Titman, 1993). We then provide an illustration of the independent double sort by replicating the momentum factor (Carhart, 1997) following the description from Kenneth R. French's data library.

### 4.A. Momentum portfolios based on a univariate sort

To illustrate the univariate portfolio sorting approach itself, and the influence of some of the choices made by the researcher along the way, we will consider the construction of the well known momentum anomaly (Jegadeesh and Titman, 1993) using the monthly CRSP stock file. In particular, we consider common stocks (SHRCD 10 and 11) listed on the NYSE, AMEX, and NASDAQ exchanges (EXCHCD 1, 2, and 3) over the period January 1986 to December 2019. I refer to the accompanying Matlab files on Brightspace for the implementations.

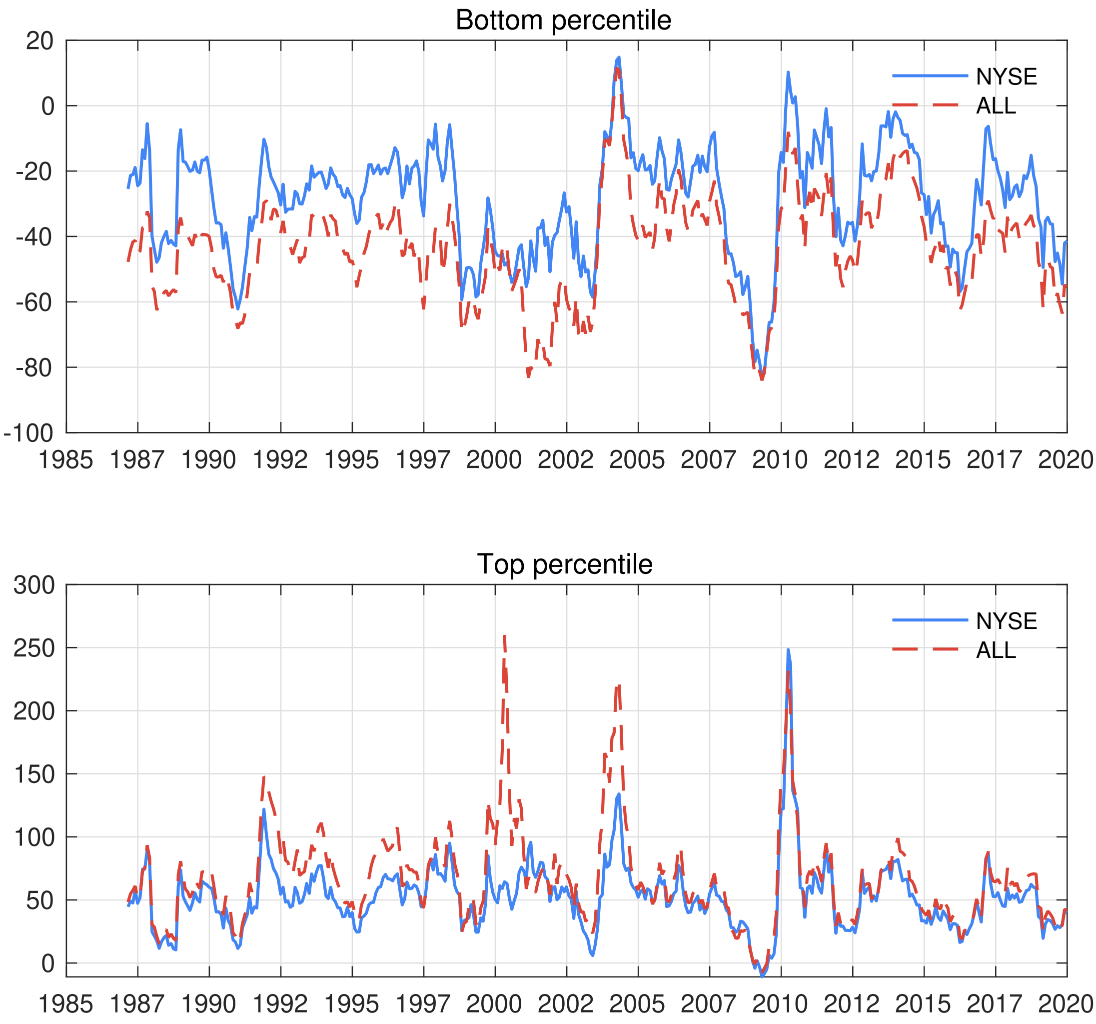
**Defining the momentum signal** Jegadeesh and Titman (1993) document strong evidence of return continuation in U.S. individual stocks for a variety of different look-back and holding periods. We will consider the definition used in, among others, Carhart (1997) and Asness, Moskowitz and Pedersen (2013), where momentum at time $t-1$ is defined as the cumulative return from $t-12$ to $t-2$ and the holding period is one month. We skip the most recent month to avoid the effect of one-month reversal in stock returns, typically attributed to liquidity and/or microstructure issues (Jegadeesh, 1990, Lo and MacKinlay, 1990). That is, the sorting variable $F_{i,t}$ is defined as

$$F_{i,t-1} = \prod_{h=0}^{10} (1 + r_{i,t-12+h}) \tag{10}$$

where $r_{i,t}$ denotes the return on stock $i$ at time $t$. This is also the definition used in Kenneth R. French's data library.

**Figure 1: NYSE versus ALL breakpoints**

This figure plots the bottom (Panel A) and top (Panel B) percentiles of the cross-sectional distribution of momentum over time. The blue lines are percentiles based on NYSE stocks only and the red lines are percentiles based on ALL stocks. The sample period is January 1986 to December 2019.

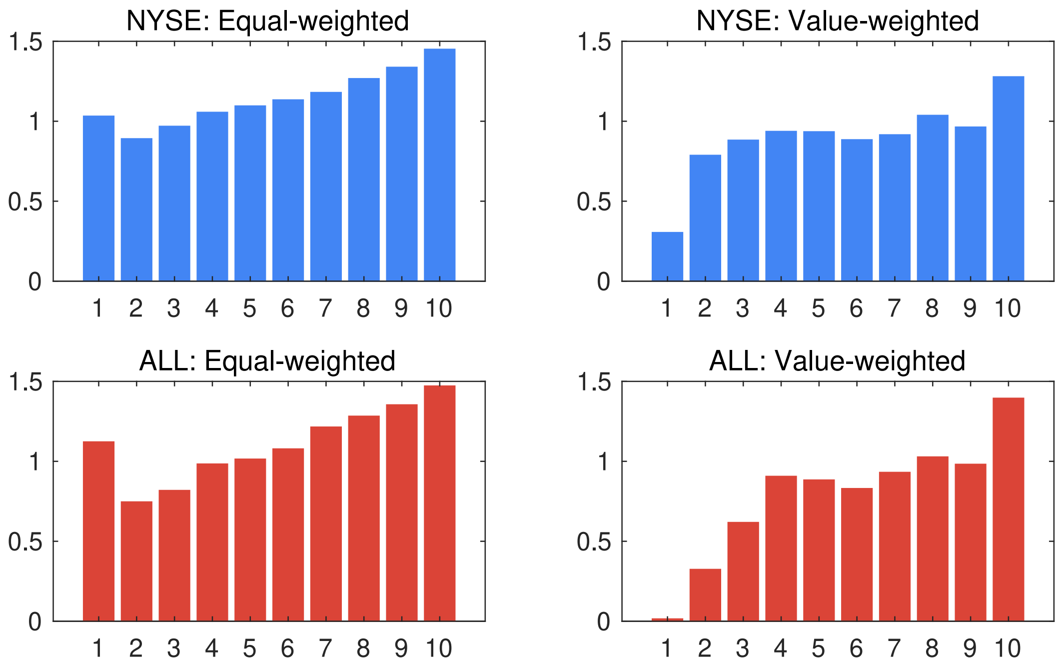


**Breakpoints and portfolio returns** With the sorting variables in hand, we can compute breakpoints, allocate stocks to their portfolios, and compute portfolio returns. We will focus on decile sorted momentum portfolios using evenly-spaced percentiles. Accordingly, we will have nine breakpoints to compute. We will consider two choices that are common in the literature: 1) using all available stocks to compute breakpoints so that all portfolios will feature an equal number of stocks every months or 2) using NYSE stocks to define breakpoints so that all portfolios feature an even number of NYSE stocks every month, but potentially an uneven number of AMEX and NASDAQ stocks. The latter may be desirable when using data over long periods as AMEX and NASDAQ stocks only became available in July 1962 and December 1972, respectively. AMEX and NASDAQ stocks also tend to be smaller than NYSE stocks. To illustrate the impact of different choices on the results, we will form breakpoints using NYSE stocks only and using ALL stocks. Similarly, for each type of breakpoint, we will compute equal- and value-weighted returns. Last, in our implementations, we follow Kenneth R. French and require the following for a stock to be included in the sample: 1) Portfolios are re-balanced every month, 2) the price at time $t-13$ is not missing, 3) the return at time $t-2$ is not missing, and 4) market equity data at time $t-1$ is not missing.

First, we compute breakpoints using (1) and plot the top and bottom percentiles of the cross-sectional distribution of momentum over time for NYSE and ALL stocks. Figure 1 illustrates that percentiles vary over time and that using only NYSE stocks can produce percentiles that are markedly different from the percentiles identified using ALL stocks. In particular, the percentiles based on ALL stocks tend to be more extreme, highlighting the influence of small and illiquid stocks on the momentum strategy. To illustrate this further, Figure 2 plots the average returns to our decile-sorted

**Figure 2: Momentum returns across choices**

This figure plots average returns to decile-sorted momentum portfolios based on either NYSE or ALL breakpoints for equal- and value-weighted portfolio returns. The sample period is January 1986 to December 2019.



momentum portfolios for our different scenarios. Several things are worth noticing. First, Equal-weighting returns produce a non-monotonic pattern, where losers (P1) have relatively high returns. Using value-weights mitigates this and produces a near-monotonic relation between momentum and future returns. Arguably, this is caused by an overweight of small and illiquid stocks in P1 under equal-weighting, whereas these stocks are assigned a smaller weight in the value-weighting scheme. These observations are the same using both NYSE and ALL stocks for the determination of breakpoints.

**Excess and risk-adjusted portfolio returns** Table 1 provides excess (Panel A) and risk-adjusted returns (Panel B) using the Fama and French (1993) factors for the value-weighted portfolios based on NYSE breakpoints. We compute excess returns by subtracting the risk-free rate from Kenneth R. French's data library. I invite you to replicate these results for the other portfolios yourself.

**Table 1: Momentum portfolios**

This table provides descriptive statistics for excess momentum returns (Panel A), including annualized mean excess returns, standard deviations, skewness, kurtosis, and Sharpe ratios. Newey and West (1987) $t$-statistics are presented in square brackets. The table also provides risk-adjusted excess returns (Panel B) along with loadings on the Fama and French (1993) factors. The sample period is January 1986 to December 2018.

| | P1 | P2 | P3 | P4 | P5 | P6 | P7 | P8 | P9 | P10 | MOM |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | Panel A: Descriptive statistics | | | | | | |
| Mean | 0.62 | 6.41 | 7.55 | 8.21 | 8.18 | 7.58 | 7.95 | 9.41 | 8.53 | 12.31 | 11.69 |
| | [0.11] | [1.55] | [2.26] | [2.82] | [3.14] | [2.96] | [3.54] | [3.91] | [3.19] | [3.40] | [2.38] |
| Std | 30.52 | 22.39 | 18.87 | 16.49 | 15.12 | 14.69 | 14.31 | 14.09 | 15.33 | 20.58 | 26.57 |
| Skew | 0.59 | 0.15 | 0.17 | -0.45 | -0.56 | -0.93 | -0.97 | -0.74 | -0.98 | -0.62 | -1.44 |
| Kurt | 6.96 | 6.77 | 7.06 | 5.44 | 6.22 | 7.12 | 7.26 | 5.55 | 7.28 | 5.47 | 10.51 |
| SR | 0.02 | 0.29 | 0.40 | 0.50 | 0.54 | 0.52 | 0.56 | 0.67 | 0.56 | 0.60 | 0.44 |
| | | | | | Panel B: Risk-adjusted returns | | | | | | |
| $\alpha$ | -12.53 | -4.31 | -1.88 | -0.42 | 0.22 | -0.27 | 0.64 | 2.26 | 1.15 | 4.51 | 17.04 |
| | [-4.11] | [-2.02] | [-1.05] | [-0.35] | [0.20] | [-0.27] | [0.56] | [2.71] | [1.01] | [2.63] | [4.26] |
| MKT | 1.56 | 1.27 | 1.10 | 1.02 | 0.96 | 0.94 | 0.90 | 0.89 | 0.94 | 1.04 | -0.52 |
| | [12.48] | [15.95] | [21.57] | [30.09] | [29.48] | [34.32] | [24.43] | [29.28] | [20.55] | [20.63] | [-3.29] |
| SMB | 0.41 | 0.11 | -0.02 | -0.09 | -0.10 | -0.09 | -0.16 | -0.09 | -0.11 | 0.38 | -0.03 |
| | [2.55] | [0.83] | [-0.15] | [-1.55] | [-1.48] | [-2.02] | [-3.48] | [-2.35] | [-1.67] | [5.69] | [-0.14] |
| HML | 0.34 | 0.37 | 0.40 | 0.33 | 0.25 | 0.27 | 0.16 | 0.13 | 0.03 | -0.30 | -0.64 |
| | [1.46] | [2.61] | [4.07] | [4.48] | [5.05] | [5.49] | [3.53] | [2.17] | [0.59] | [-3.13] | [-2.10] |
| Adj. $R^2$ | 62.92 | 70.96 | 73.56 | 81.61 | 84.96 | 86.58 | 83.17 | 83.96 | 81.08 | 76.73 | 11.06 |

First of all, excess returns are near-monotonically increasing from P1 to P10 resulting in an average long-short excess return of 11.69% per annum that is statistically significant ($t$-stat of 2.37). We also note that excess returns to losers (winners) are positively (negative) skewed and display excess kurtosis. In Panel B, we consider risk-adjusted returns, where we regress momentum returns on the Fama and French (1993) three-factor model

$$r_{k,t} - r_{f,t} = \alpha_k + b_{MKT} MKT_t + b_{SMB} SMB_t + b_{HML} HML_t + \varepsilon_{k,t}. \tag{11}$$

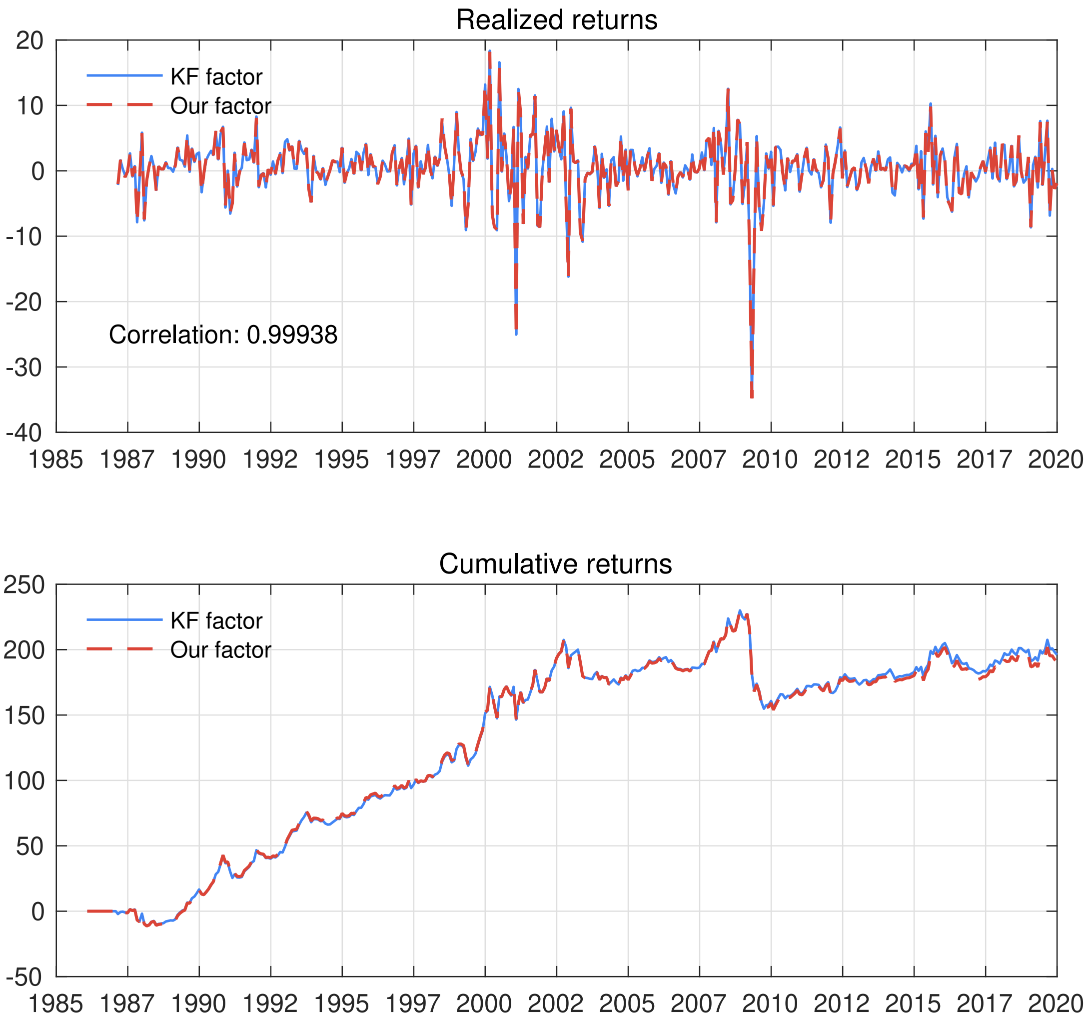
Risk-adjusted excess returns ($\alpha_k$) are monotonically increasing from P1 to P10 and significant for the momentum factor. That is, the Fama and French (1993) three-factor model cannot account for the returns to the momentum anomaly. We also note that all portfolios load positively on the market excess return (MKT), but that the momentum factor itself has a negative loading that is statistically significant. This implies that momentum cannot be explained from market risk. We also note that P1 and P10 load positively on SMB, indicating that loser and winner currencies are mostly smaller stocks. In a nutshell, winner returns are not just due to momentum picking up small stocks. The loadings on HML indicate that winners load negatively on value stocks just as the momentum strategy (Asness et al., 2013). The increase in risk-adjusted returns (17.04% relative to 11.69%) for the momentum factor can be attributed to this negative relation (together with the market).

### 4.B. A momentum factor based on an independent double sort

This section illustrates the independent double sort by replicating the momentum factor (Carhart, 1997) as available on Kenneth R. French's data library. The full universe of stocks consists of common shares listed on NYSE, AMEX, and NASDAQ. To build the long-short zero-cost factor, we first form

**Figure 3: Momentum factor and comparison to Kenneth French**

This figure plots the bottom (Panel A) and top (Panel B) percentiles of the cross-sectional distribution of momentum over time. The blue lines are percentiles based on NYSE stocks only and the red lines are percentiles based on ALL stocks. The sample period is January 1986 to December 2019.



six portfolios sorted on size (ME) and momentum (MOM). The portfolios are constructed using the independent double sort outlined above. We consider two size groups determined using the median NYSE size as the breakpoint and three momentum groups using breakpoints determined from the 30th and 70th NYSE percentiles. The intersections provide us with six portfolios: Big Losers (BL), Big Neutrals (BN), Big Winners (BW), Small Losers (SL), Small Neutrals (SN), and Small Winners (SW). For example, the SL portfolio contains all stocks in the small size group that are also in the loser momentum group. The momentum factor is then constructed from the six portfolios as

$$MOM = \frac{1}{2}\left[SW + BW\right] - \frac{1}{2}\left[SL + BL\right], \tag{12}$$

where we ignore the neutrals as is common in these types of risk factors. We note that the factor is designed to capture the risk of being a winner relative to being a loser stock while keeping the average size the same.

Figure 3 plots the return and the cumulative returns to our momentum factor. We also plot the momentum factor obtained from Kenneth R. French's data library to compare how closely we can replicate the factor. Overall, the two series are very similar and the correlation between them is 0.9994 over the full sample. The accompanying Matlab live script provides further details on the construction of the factor and the use of the independent double sort.

## References


Asness, C.S., Moskowitz, T.J., Pedersen, L.H., 2013. Value and momentum everywhere. Journal of Finance 68, 929–985.

Bali, T.G., Engle, R.F., Murray, S., 2016. Empirical Asset Pricing: The cross section of stock returns. John Wiley & Sons, Inc., Hoboken, New Jersey.

Carhart, M.M., 1997. On persistence in mutual fund performance. Journal of Finance 52, 57–82.

Fama, E.F., French, K.R., 1992. The cross-section of expected stock returns. Journal of Finance 47, 427–465.

Fama, E.F., French, K.R., 1993. Common risk factors in the returns on stocks and bonds. Journal of Financial Economics 33, 3–56.

Fama, E.F., French, K.R., 2015. A five-factor asset pricing model. Journal of Financial Economics 116, 1–22.

Harvey, C.R., Liu, Y., 2019. A census of the factor zoo. Working paper, Duke University.

Harvey, C.R., Liu, Y., Zhu, H., 2016. ... and the cross-section of expected returns. Review of Financial Studies 29, 5–68.

Hou, K., Xue, C., Zhang, L., 2015. Digesting anomalies: An investment approach. Review of Financial Studies 28, 650–705.

Jegadeesh, N., 1990. Evidence of predictable behavior of security returns. Journal of Finance 45, 881–898.

Jegadeesh, N., Titman, S., 1993. Returns to buying winners and selling losers: Implications for stock market efficiency. Journal of Finance 48, 65–91.

Lo, A.W., MacKinlay, C., 1990. When are contrarian profits due to stock market overreaction? Review of Financial Studies 3, 175–205.

Newey, W.K., West, K.D., 1987. A simple, positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix. Econometrica 55, 703–708.