Portfolio Sorting

Empirical Asset Pricing

Mads Markvart Kjær

Department of Economics and Business Economics, Aarhus University, CREATES

E-mail: mads.markvart@econ.au.dk

Spring 2022

Recap

- Last lecture, we talked about how to test whether a factor is priced in the cross-section of some asset class
- Today, we will focus on how to construct factors from financial returns

Why do we care?

Portfolio sorts are central to the empirical asset pricing literature and is a commonly applied methodology

- Can some characteristic of the assets explain cross-sectional variation?

Outcome of lecture

After the lecture, you should have

- knowlegde and understading of
  - Portfolio sorting based on characteristics and covariances, the construction of zero-cost risk factors, and their implications for market efficiency
- and be able to
  - Discuss and conduct a portfolio sort using individual assets, evaluate the resulting portfolio returns, construct long-short factors, evaluate their returns, and reflect on the implications

The King of Univeriate Portfolio sort!

TURAN G. BALI

- Robert Parker Chair Professor of Finance
  McDonough School of Business
  Georgetown University
- Associate Editor
  - Journal of Financial and Quantitative Analysis
  - Management Science

A world of oppotunities!

Journal of Financial Economics 131 (2019) 619–642

Contents lists available at ScienceDirect

Journal of Financial Economics

journal homepage: www.elsevier.com/locate/jfec

Common risk factors in the cross-section of corporate bond returns☆

Jennie Bai, Turan G. Bali, Quan Wen

McDonough School of Business, Georgetown University, 3700 O St., Washington, DC, 20057, USA

Since the empirical distribution of bond returns is skewed, peaked around the mode, and has fat tails, downside risk—defined as a nonlinear function of volatility, skewness, and kurtosis—is expected to play a major role in the cross-sectional pricing of corporate bonds.

4.2. Univariate portfolio analysis

We first examine the significance of a cross-sectional relation between VaR and future corporate bond returns using portfolio-level analysis. For each month from July

A world of oppotunities!

Journal of Financial Economics 135 (2020) 725–753

Contents lists available at ScienceDirect

ELSEVIER

Journal of Financial Economics

journal homepage: www.elsevier.com/locate/jfec

Left-tail momentum: Underreaction to bad news, costly arbitrage and equity returns☆

Check for updates

Yigit Atilgan[a], Turan G. Bali[b,*], K. Ozgur Demirtas[a], A. Doruk Gunaydin[a]

[a] Sabanci University, School of Management, Orhanli Tuzla 34956, Istanbul, Turkey
[b] Georgetown University, McDonough School of Business, Washington, D.C. 20057, USA

ket betas, lower book-to-market ratios, and lower idiosyncratic volatilities. Finally, there is a highly significant, positive correlation between idiosyncratic volatility and lottery demand.

3.2. *Univariate portfolio analysis*

In this section, we perform univariate portfolio-level analysis, where deciles are formed every month by sorting stocks based on their value-at-risk metrics at the 1% level and one-month-ahead returns are calculated for each

A world of oppotunities!

The Journal of FINANCE
The Journal of THE AMERICAN FINANCE ASSOCIATION

THE JOURNAL OF FINANCE • VOL. LXIX, NO. 5 • OCTOBER 2014

The Joint Cross Section of Stocks and Options

BYEONG-JE AN, ANDREW ANG, TURAN G. BALI, and NUSRET CAKICI*

ABSTRACT

Stocks with large increases in call (put) implied volatilities over the previous month tend to have high (low) future returns. Sorting stocks ranked into decile portfolios by past call implied volatilities produces spreads in average returns of approximately 1% per month, and the return differences persist up to six months. The cross section of stock returns also predicts option implied volatilities, with stocks with high past returns tending to have call and put option contracts that exhibit increases in implied volatility over the next month, but with decreasing realized volatility. These predictability patterns are consistent with rational models of informed trading.

A world of oppotunities!

Growth Options and Related Stock Market Anomalies: Profitability, Distress, Lotteryness, and Volatility

Turan G. Bali, Luca Del Viva, Neophytos Lambertides, and Lenos Trigeorgis*

including the FISKEW factor whenever the value-minus-growth return spread appears significant in parts of the sample period. This holds using either the decile 10 minus decile 1 return spread on the book-to-market portfolios or the HML factor of Fama and French (1993). We provide a discussion and corresponding results in Section II and Figures A.1 and A.2 of the Supplementary Material.

D. Univariate Portfolio Analysis and Economic Significance

Table 5 provides further evidence concerning the economic significance of our growth-option-driven skewness measure, $E[ISKEW]_{GO}$, based on univariate portfolios. For each month, we form equal-weighted (EW) and value-weighted (VW) decile portfolios by sorting individual stocks based on their growth-option-driven expected idiosyncratic skewness, $E[ISKEW]_{GO}$, where decile 1 contains stocks with the lowest $E[ISKEW]_{GO}$ and decile 10 contains stocks with the highest $E[ISKEW]_{GO}$. Table 5 reports, by row, the average $E[ISKEW]_{GO}$, the average

A world of oppotunities!

informs
http://pubsonline.informs.org/journal/mnsc/

MANAGEMENT SCIENCE
Vol. 64, No. 9, September 2018, pp. 4137–4155
ISSN 0025-1909 (print), ISSN 1526-5501 (online)

Unusual News Flow and the Cross Section of Stock Returns

Turan G. Bali,[a] Andriy Bodnaruk,[b] Anna Scherbina,[c,d] Yi Tang[e]

[a] McDonough School of Business, Georgetown University, Washington, DC 20057; [b] College of Business Administration, University of Illinois at Chicago, Chicago, Illinois 60607; [c] Graduate School of Management, University of California at Davis, Davis, California 95616; [d] International Business School, Brandeis University, Waltham, Massachusetts 02453; [e] Gabelli School of Business, Fordham University, New York, New York 10023

Contact: turan.bali@georgetown.edu (TGB); bodnaruk@uic.edu (AB); ascherbina@ucdavis.edu (AS); ytang@fordham.edu (YT)

2.2. Volatility Shocks and Future Returns

In the remainder of the section, we establish a robust negative relation that exists between the current month's volatility shocks and the following month's returns.

2.2.1. Volatility-Shock-Sorted Portfolios. To establish the negative predictive ability of volatility shocks on future returns, every month we sort stocks into decile portfolios based on $IVOL^{shock}$. We then calculate next month's equal- and value-weighted portfolio returns and the return differentials between

A world of oppotunities!

informs.
http://pubsonline.informs.org/journal/mnsc/

MANAGEMENT SCIENCE
Vol. 63, No. 11, November 2017, pp. 3760–3779
ISSN 0025-1909 (print), ISSN 1526-5501 (online)

Dynamic Conditional Beta Is Alive and Well in the Cross Section of Daily Stock Returns

Turan G. Bali,[a] Robert F. Engle,[b] Yi Tang[c]

[a] McDonough School of Business, Georgetown University, Washington, DC 20057; [b] Stern School of Business, New York University, New York, New York 10012; [c] School of Business, Fordham University, New York, New York 10023

Contact: turan.bali@georgetown.edu (TGB); rengle@stern.nyu.edu (RFE); ytang@fordham.edu (YT)

3. Market Beta and the Cross Section of Daily Returns

This section examines the significance of a cross-sectional relation between the unconditional beta, the dynamic conditional beta, and daily stock returns based on the long-short equity portfolios. First, we perform univariate portfolio-level analysis for the unconditional measures of market beta. Second, we test the predictive power of the DCC beta based on univariate portfolio-level analysis. Finally, we provide average portfolio characteristics of the DCC beta–sorted port-

A world of oppotunities!

Liquidity Shocks and Stock Market Reactions

Turan G. Bali
McDonough School of Business, Georgetown University

Lin Peng
Zicklin School of Business, Baruch College, City University of New York

Yannan Shen
McCallum Graduate School of Business, Bentley University

Yi Tang
Schools of Business, Fordham University

2. Cross-Sectional Relation Between Liquidity Shocks and Stock Returns

The significantly positive correlation between liquidity shocks and one-month-ahead stock returns suggests that negative liquidity shocks (reductions in liquidity) are related to lower future stock returns, and vice versa. In this section, we perform formal analysis, and show that the pricing effect documented in this paper cannot be explained by other risk factors and stock characteristics that are known to predict future stock returns in the cross-section.

2.1 Univariate portfolio-level analysis

We begin our empirical analysis with univariate portfolio sorts. For each month, we sort common stocks trading on NYSE/AMEX/NASDAQ into decile

A world of oppotunities!

Hybrid Tail Risk and Expected Stock Returns: When Does the Tail Wag the Dog?

Turan G. Bali
Georgetown University

Nusret Cakici
Fordham University

Robert F. Whitelaw
New York University and NBER

3. Preliminary Evidence

Given the number of potential control variables, that is, other stock characteristics that may influence returns, the Fama-MacBeth cross-sectional regression approach may be the natural way to examine the predictive power of measures of tail risk. We turn to these regressions in Section 4; however, to get an initial feel for the data, we first look at univariate sorts on the basis of our three tail risk measures and the associated characteristics of the portfolios.

3.1 Average returns for univariate portfolio sorts

Table 1 presents the average monthly returns for the equal-weighted and value-weighted decile portfolios that are formed by sorting the NYSE, AMEX, and NASDAQ stocks based on our three tail risk measures—

A world of oppotunities!

Journal of Financial Economics 126 (2017) 471–489

Contents lists available at ScienceDirect

Journal of Financial Economics

journal homepage: www.elsevier.com/locate/jfec

ELSEVIER

Is economic uncertainty priced in the cross-section of stock returns?☆

Turan G. Bali[a,*], Stephen J. Brown[b,c], Yi Tang[d]

[a] McDonough School of Business, Georgetown University, 37th and O Streets, N.W., Washington, D.C. 20057, USA
[b] Monash Business School, Monash University, 900 Dandenong Road, Caulfield East, Victoria 3145, Australia
[c] Stern School of Business, New York University, 44 West fourth Street, New York, NY 10012, USA
[d] Gabelli School of Business, Fordham University, 45 Columbus Avenue, New York, NY 10023, USA

4. Empirical results

In this section, we conduct parametric and nonparametric tests to assess the predictive power of the uncertainty beta over future stock returns. First, we start with univariate portfolio-level analyses. Second, we discuss average stock characteristics to obtain a clear picture of the composition of the uncertainty beta portfolios. Third, we conduct bivariate portfolio-level analyses to examine the predictive power of the uncertainty beta after controlling for

A world of oppotunities!

JOURNAL OF FINANCIAL AND QUANTITATIVE ANALYSIS Vol. 56, No. 5, Aug. 2021, pp. 1653–1678

doi:10.1017/S0022109020000538

The Macroeconomic Uncertainty Premium in the Corporate Bond Market

Turan G. Bali
Georgetown University McDonough School of Business
Turan.Bali@georgetown.edu (corresponding author)

Avanidhar Subrahmanyam
University of California at Los Angeles Anderson Graduate School of Management
subra@anderson.ucla.edu

Quan Wen
Georgetown University McDonough School of Business
Quan.Wen@georgetown.edu

III. Empirical Results

In this section, we conduct parametric and nonparametric tests to assess the predictive power of the uncertainty beta over future corporate bond returns. We start with univariate portfolio-level analyses, presenting the average returns, alphas, and average bond characteristics of β^{UNC}-sorted portfolios. Second, we conduct bivariate portfolio-level analyses to examine the predictive power of the uncertainty beta after controlling for well-known measures of systematic risk, liquidity, and bond characteristics. Third, we provide an alternative risk-based explanation of the

A world of oppotunities!

Journal of Financial Economics 99 (2011) 427–446

Contents lists available at ScienceDirect

Journal of Financial Economics

journal homepage: www.elsevier.com/locate/jfec

Maxing out: Stocks as lotteries and the cross-section of expected returns ☆

Turan G. Bali [a,1], Nusret Cakici [b,2], Robert F. Whitelaw [c,d,*]

[a] Department of Economics and Finance, Zicklin School of Business, Baruch College, One Bernard Baruch Way, Box 10-225, New York, NY 10010, United States
[b] Department of Finance, Fordham University, Fordham University, 113 West 60th Street, New York, NY 10023, United States
[c] Stern School of Business, New York University, 44 W. 4th Street, Suite 9-190, New York, NY 10012, United States
[d] NBER, United States

2.2. Univariate portfolio-level analysis

Table 1 presents the value-weighted and equal-weighted average monthly returns of decile portfolios that are formed by sorting the NYSE/Amex/Nasdaq stocks based on the maximum daily return within the previous month (MAX). The results are reported for the sample period July 1962–December 2005.

Portfolio 1 (low MAX) is the portfolio of stocks with the lowest maximum daily returns during the past month,

Portfolio sorts

- Portfolio sorts are useful in identifying and assessing variables that can predict cross-sectional variation in future returns
- Typical sorting variables are *characteristic* such as size, value, CAPM-β, idiosyncratic volatility, or past returns of the assets
- But the sorting variable can also be exposure to economically motivated risk factors, e.g., market-wide volatility or macroeconomic risks

Portfolio sorts

- Once sorted, we are interested in;
  - Assessing the cross-sectional relation between the sorting variable(s) and average returns
  - The returns to a zero-cost long-short (spread) portfolio

Advantages

- Portfolio sorts offer several advantages as a methodology

  1. Portfolio sorts does not require any a priori assumptions about the cross-sectional relationship between the sorting variable and expected returns
  2. Sorting stocks into portfolios can assist in the discovery of non-linear and linear cross-sectional relations alike
  3. It is highly flexibly and lets the researcher control much of the setup

Disadvantages

- On the other side, there are also a few drawbacks with the approach

    1. It is only possible to control for a very limited set of factors when examining the cross-sectional relations of interest
    2. The univariate (bivariate) approach only considers one (two) sorting variable(s) without controlling for other potentially important factors
    3. Many choices are left to the researcher without any clear guidance on *correct* implementations (e.g., data-snooping concerns)

Univariate Portfolio Sort

Univariate portfolio sort

- A univariate portfolio sort considers a single sorting variable $F_{i,t}$ for each security $i = 1,2,\ldots,N$ at time $t = 1,2,\ldots,T$
- The objective is to study the cross-sectional relationship between the factor $F_{i,t}$ and the future (excess) return to individual assets

Steps in univariate portfolio sorting

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 in a meaningful way
4. Examine the cross-sectional variation in average portfolio (excess) returns

Step 1: Computing breakpoints

- The first step is to compute breakpoints for the cross-sectional distribution of the sorting factor $F_{i,t}$ for each time period t
- Suppose that we wish to form n_p portfolios, then we need $n_p - 1$ breakpoints for portfolio formation
- Let p_k denote the kth percentile of the values of $F_{i,t}$ across all available assets and denote by $\mathcal{B}_{k,t}$ the kth breakpoint at time t, then

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

- 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

Choices in breakpoint determination

Choices in breakpoint determination

There are three key choices left to the researcher in determining the breakpoints

1. **Choosing the assets for the breakpoints:** We can determine the breakpoints using all assets or a subset of the assets. As an example, one can use all available stocks on the CRSP tape or only NYSE stocks
2. **Choosing the number of portfolios:** 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 factor $F_{i,t}$
3. **Choosing percentiles:** The percentiles can be evenly spaced or unevenly spaced. Fama and French (1993), as en example, use the 30th and 70th percentiles

- In the end, we need to make economically motivated and well-argued choices for these parameters that suit the study at hand

Step 2: Portfolio formation

- Suppose that we have $k = 1, 2, \ldots, n_p - 1$ breakpoints defined using the chosen percentiles, and define $\mathcal{B}_{0,t} = -\infty$ and $\mathcal{B}_{n_p,t} = \infty$ to exhaust all possible values of the sorting variable $F_{i,t}$
- We can then identify all securities i that belong to the kth portfolio formed at time t as the set of securities with values of $F_{i,t}$ that satisfy the relation

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

- 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

Step 3: Computing portfolio returns

- Let $N_{k,t}$ denote the number of securities in portfolio k at time t, then equal-weighted portfolio returns for portfolio k are computed as

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

  where the sum is taken over all securities in the kth portfolio at time t

- Let $ME_{i,t}$ denote market value of security i at time t, then value-weighted returns are defined as (we measure $ME_{i,t}$ at portfolio formation)

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

- Value-weighting is most appropriate when the securities are stocks, e.g., U.S. individual stocks from the CRSP sample
- Value-weighting alleviates issues with assigning *too large* weights to small and illiuid securities that are hard and expensive to trade

Step 4: Examining portfolio returns

- The main objective here is to determine whether there is a reliable cross-sectional relation between the sorting variable $F_{i,t}$ and future asset returns in the cross section

  1. The first step is to compute descriptive statistics for the portfolio (excess) returns and the long-short portfolio

Step 4: Examining portfolio returns, cont

2. We then look for monotonic relationships in the average returns between the first and the last portfolios

   $\rightarrow$ Patton and Timmermann (2010) provide a test to test for monotonicity in portfolio returns/factor exposure!

The Patton Timmermann test for monotonicity

- Let $\Delta_i = \mu_i - mu_{i-j}$ denote the difference between the average return of portfolio i and $i-1$
- We are then interested in testing whether all Δ_i's is greater than 0 or not. Meaning that

$$H_0 : \Delta \leq 0 \tag{5}$$
$$H_1 : \Delta > 0 \tag{6}$$

- The test statistic is simply $J_T = \min \Delta$
- The p-value is estimated using a block stationary bootstrap:
  1. Generate B random series of length T with average block length
  2. Calculate under the null $J_T^b = \min(\hat{\Delta} - \min \Delta)$ for each bootstrap ()
  3. P-value is given as $\hat{p} = \frac{1}{B} \sum_{b=1}^{B} 1_{J_T^b > J_T}$

Step 4: Examining portfolio returns, cont

3. Next, we test whether the return pattern survives controlling for known risk factors identified in the asset pricing literature

- This amounts to testing intercepts (α) in time series regression using (some) asset pricing models

  $\rightarrow$ In the end the evaluation depends on the choice of asset pricing model! (FF3, FF5, Carhart 4-factor, etc.)

- We will illustrate the method and the impact of the choices using the momentum anomaly later in these slides

Bivariate Portfolio Sort

Bivariate portfolio sorts

- We now turn to bivariate (or double) portfolio sorts in which the universe of assets is sorted into portfolios based on two sorting variables rather than one
- Bivariate portfolio sorts are useful when we want to condition on (or control for) more than one sorting variable
- Bivariate sorts differ mainly in the construction of breakpoints and portfolio formation. The remaining steps are identical to the univariate case

Type of bivariate sorts

Types of bivariate sorts

In general, when considering bivariate portfolio sorts, we need to distinguish between independent and dependent sorts

- **Independent double sorts:** The ordering of the sorting variables does not matter
- **Dependent double sorts:** The ordering of the sorting variables is critically important

Independent double sort

- The independent double sort builds portfolios by sorting on two variables $F_{i,t}^1$ and $F_{i,t}^2$ independently

- We create n_{p_1} groups based on $F_{i,t}^1$ and n_{p_2} groups based on $F_{i,t}^2$ for a total of $n_{p_1} \times n_{p_2}$ portfolios

- The breakpoints for the two sorting variables are then defined as

$$\mathcal{B}_{k,t}^1 = \text{Percentile}_{p_k}\left(F_{i,t}^1\right) \tag{7}$$

$$\mathcal{B}_{j,t}^2 = \text{Percentile}_{p_j}\left(F_{i,t}^2\right), \tag{8}$$

- We are still facing the same choices for breakpoint determination as above:
  - Which assets should we use?
  - How many groups should we employ?
  - And what percentiles should be considered?

Building portfolios in the independent sort

- We create a total of $n_{p_1} \times n_{p_2}$ portfolios based on the groups identified using the sorting variables independently

- The portfolios are defined as the intersection of the groups based on the two sorting variables

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

  where $\cap$ is the intersection operator and $k = 1,2,\ldots,n_{p_1}$ and $j = 1,2,\ldots,n_{p_2}$ refers to the groups

- Portfolios are formed on the basis of the intersection of the groups of assets from each of the two independent sorts

- The remaining steps involving the calculation and examination of portfolio returns are equivalent to the univariate case. In a nutshell, we wish to establish if a reliable cross-sectional relationship exists

Dependent double sort

- The dependent double sort similarly builds portfolios by sorting on two variables $F^1_{i,t}$ and $F^2_{i,t}$. However, $F^1_{i,t}$ is now a control variable
- 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
- The n_{p_1} groups and breakpoints $\mathcal{B}^1_{k,t}$ for $k = 1,2,\ldots,n_{p_1} - 1$ for the first sorting variable is constructed identically to the independent sort case
- The breakpoints for the second sorting variable $F^2_{i,t}$ are now different and instead defined as

$$\mathcal{B}^2_{k,j,t} = \text{Percentile}_{p_j}\left(F^2_{i,t} \middle| \mathcal{B}^1_{k-1,t} \le F^1_{i,t} \le \mathcal{B}^1_{k,t}\right), \tag{10}$$

- Note that the order of the sorting variables is now critically important and a switch in ordering can lead to vastly different results

Building portfolios in the dependent sort

- 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 $F_{i,t}^2$
- The portfolios are defined as the intersection of the conditional groups based on the two sorting variables

$$P_{k,j,t} = \left\{ i | \mathcal{B}_{k-1,t}^1 \leq F_{i,t}^1 \leq \mathcal{B}_{k,t}^1 \right\} \cap \left\{ i | \mathcal{B}_{k,j-1,t}^2 \leq F_{i,t}^2 \leq \mathcal{B}_{k,j,t}^2 \right\} \tag{11}$$

  where $\cap$ is the intersection operator and $k = 1,2,\ldots,n_{p_1}$ and $j = 1,2,\ldots,n_{p_2}$ refers to the groups
- The remaining steps involving the calculation and examination of portfolio returns are equivalent to the univariate case

The CRSP universe

The CRSP stock file

- We want to study cross-sectional patterns in expected returns using an appropriate universe of financial assets
- The main source for data on U.S. individual stocks is the Center for Research in Security Prices (CRSP) stock file
  - CRSP is maintained by the University of Chicago's Booth School of Business
  - The CRSP tape provides data from December 31, 1925 up to today
  - The CRSP tape is hosted through Wharton Research Data Services (WRDS)
  - Also contains data on market indices, stock market factors, and bonds
  - The access link is here: https://wrds-www.wharton.upenn.edu
- For more information and details about log-in (signing up) and how to use the CRSP web-based access at WRDS, see crspNotes.pdf on Brightspace

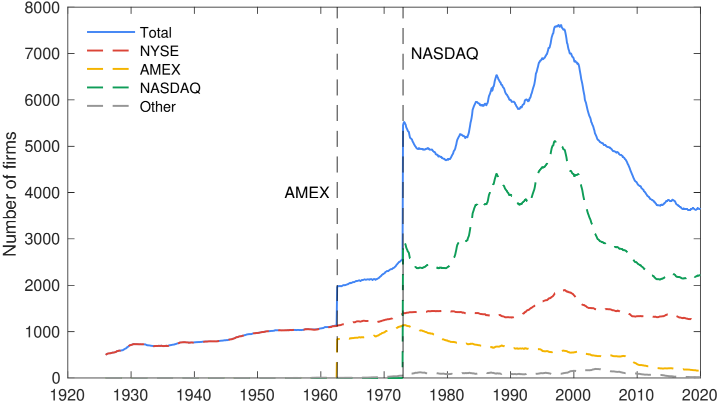
Overview of data

- CRSP contains monthly and daily data on U.S. individual stocks. The most important stock variables for our purpose here are

    1. **PERMNO:** Every stock issue is assigned a unique PERMNO that does not change over time, even if the company name, ticker, or exchange do. PERMNO is the principal identifier of a stock in CRSP
    2. **SHRCD:** Every stock is issued a share code and we can use it to identify U.S. common stocks (SHRCD 10 and 11)
    3. **EXCHCD:** A stock's exchange code indicates the exchange on which the security is listed, e.g., NYSE, AMEX, or NASDAQ (EXCHCD 1, 2, and 3, respectively)
    4. **PRC:** The price of the security is the closing price or the negative bid/ask average for a trading day (so always use the absolute price)
    5. **RET:** The holding-period return of a security including dividend payments and share repurchases and takes splits into account
    6. **SHROUT:** The number of shares outstanding is the number of publicly held shares (recorded in thousands) and is useful for computing market equity and value-weighted portfolio returns

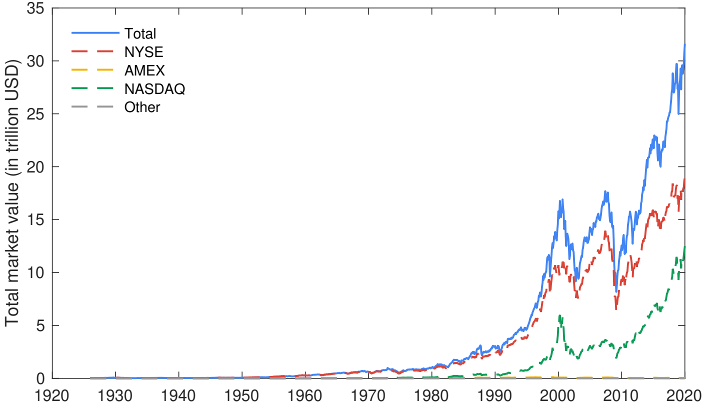
Number of firms in CRSP by exchange

- The number of firms available in the CRSP sample varies considerably over time and across exchanges
- AMEX is included July 1962 and NASDAQ in December 1972



Market equity of firms in CRSP by exchange

- The total market value of stocks listed across the different exchanges differ greatly. Most originates from stocks on NYSE and NASDAQ, whereas very little originates from AMEX and other stocks



Empirical illustrations

Momentum portfolios

- To illustrate univariate portfolio sorting, we consider the construction of the momentum anomaly (Jegadeesh and Titman, 1993) using the CRSP sample
- We consider common stocks (SHRCD 10 and 11) listed on the NYSE, AMEX, and NASDAQ exchanges (EXCHCD 1, 2, and 3) from January 1986 to December 2019

Momentum signal

We define the momentum signal as in Carhart (1997) and Asness et al. (2013), where momentum at time $t-1$ is defined as the cumulative return from $t-12$ to $t-2$

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

where $r_{i,t}$ is the return on stock i at time t, and we skip the most recent month to avoid short-term reversal effects (Jegadeesh, 1990, Lo and MacKinlay, 1990)

Choices and requirements for stocks

- We illustrate the impact of the choices open to the researcher by building momentum portfolios using

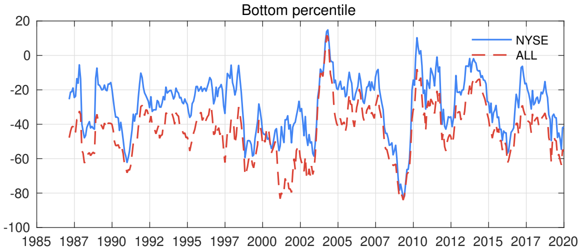
  1. Breakpoints based on NYSE and ALL stocks
  2. Equal- and value-weighted portfolio returns

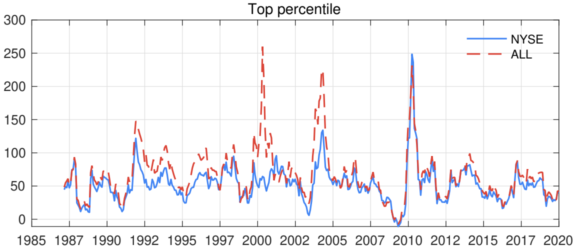
- In our implementations, we follow Kenneth R. French and require the following for a stock to be included in the sample

  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
  4. Market equity data at time $t-1$ is not missing

Breakpoints

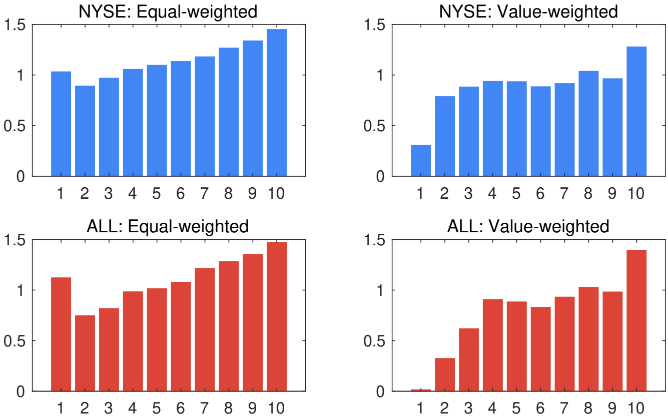
- First, consider the differences in top and bottom percentiles when using NYSE and ALL stocks, respectively





Momentum returns

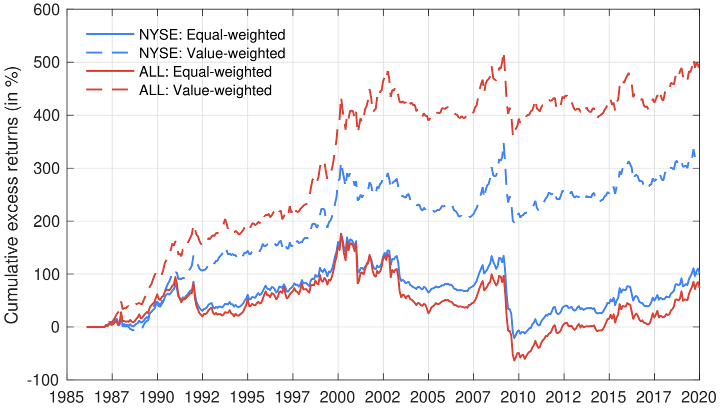
- Consider next the differences in return patterns originating from breakpoint choices and using equal- and value-weighted portfolio returns



- Applying the Patton and Timmermann (2010) test yields rejection for all sorts excect ALL: Equal-weighted which delivers a p-value of roughly 0.13

Cumulative excess momentum returns

- Consider also the cumulative excess returns to a momentum strategy in stocks for the same breakpoint choices and equal- and value-weighted portfolio excess returns



Momentum returns

- Last, we consider descriptive statistics for the NYSE-based, value-weighted momentum portfolio excess returns and risk-adjusted excess returns using the Fama and French (1993) three-factor model

$$r_{k,t} - r_{f,t} = \alpha_k + b_{MKT} MKT_t + b_{SMB} SMB_t + b_{HML} HML_t + \varepsilon_{k,t} \qquad (13)$$

	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											
α	**-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

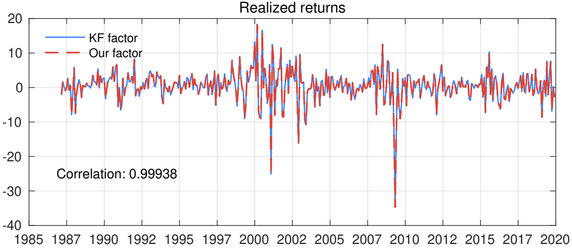
A momentum factor

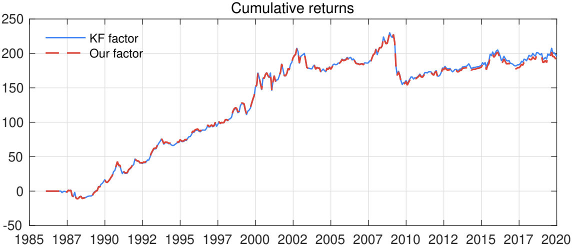
- We illustrate an independent double sort by replicating the momentum factor (Carhart, 1997) available on Kenneth French's data library
- The full universe of stocks consists of common shares (SHRCD 10 and 11) listed on NYSE, AMEX, and NASDAQ
- We independently sort stocks based on size (ME) and momentum (MOM), where 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)
- The momentum factor is then constructed from the six portfolios as follows

$$MOM = \frac{1}{2}\left[SW + BW\right] - \frac{1}{2}\left[SL + BL\right], \tag{14}$$

Comparison with Kenneth French

- The momentum factor constructed here as a correlation of 0.9993 with the factor obtained from Kenneth French, and the series are very similar





Potential projects

Potential projects

- Portfolio sorting is not unique to stocks, but the method is equally applicable to currencies, bonds, commodities, and pretty much any asset with a return
- Sort a universe of assets into portfolios based on a *new* sorting variable and evaluate whether the inherent risk is priced in the cross-section
- Re-evaluate an existing sorting factor using updated data, more subsamples, and/or different choices for the data and construction of portfolios
- Construct an existing set of anomaly portfolios and investigate a well-motivated risk-based (or behavioral) explanation
- Consider new double sorts where you take one well known variable together with a new variable (e.g., attention, sentiment, etc,) and explore their joint dynamics (see, e.g., Medhat and Schmeling (2020) for a recent example)
- Build a new cross-section of test assets to evaluate existing risk factors and asset pricing models

References

Asness, C. S., T. J. Moskowitz, and L. H. Pedersen (2013). Value and momentum everywhere. *Journal of Finance 68*(3), 929–985.

Carhart, M. M. (1997). On persistence in mutual fund performance. *Journal of Finance 52*(1), 57–82.

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

Jegadeesh, N. (1990). Evidence of predictable behavior of security returns. *Journal of Finance 45*(3), 881–898.

Jegadeesh, N. and S. Titman (1993). Returns to buying winners and selling losers: Implications for stock market efficiency. *Journal of Finance 48*(1), 65–91.

Lo, A. W. and C. MacKinlay (1990). When are contrarian profits due to stock market overreaction? *Review of Financial Studies 3*(2), 175–205.

Medhat, M. and M. Schmeling (2020). Short-term momentum. *Review of Financial Studies*, Forthcoming.

Patton, A. J. and A. Timmermann (2010). Monotonicity in asset returns: New tests with applications to the term structure, the capm, and portfolio sorts. *Journal of Financial Economics 98*(3), 605–625.