Factoring Characteristics into Returns:

A Clinical Study on the SMB and HML Portfolio Construction Methods

Marie Lambert, Boris Fays, and Georges Hübner*

This version: **January 2020**

Abstract

Factor performance is highly sensitive to the number of stocks composing its long and short basis portfolios. We examine three methodological choices that have an impact on portfolio diversification: the (in)dependence and the (a)symmetry of the stock sorting procedure and the sorting breakpoints. We show that these methodological choices have to be considered jointly and that a dependent (D) sort that starts with the control variables with whole sample or "name" (N) breakpoints and that performs a symmetric (S) sort on characteristics minimizes the biases from unpriced risks. This paper also demonstrates that the biases introduced by currently popular sorting methodologies can become very severe under specific market conditions and are not driven by small capitalizations. This alternative framework generates much stronger "turn-of-the-year" size and "through-the-year" book-to-market effects than what is conventionally documented.

JEL classification: G10; G11; G12

Keywords: Portfolio sorting; Factor performance; Factor construction methods

---

* Marie Lambert holds the Deloitte Chair of Financial Management and Corporate Valuation at the University of Liège, HEC Liège, Belgium. She is Research Associate at the EDHEC Risk Institute and QMI research fellow. Boris Fays is a PhD candidate at the University of Liège, HEC Liège, Belgium. Georges Hübner is Professor of Finance at the University of Liège, HEC Liège, Belgium.

Corresponding author: Marie Lambert, Professor, Tel: (+32) 4 2327432. Fax: (+32) 4 2327240. E-mail: marie.lambert@uliege.be. Mailing address: HEC Liège, 14 rue Louvrex, B-4000 Liège, Belgium.

Earlier versions have benefitted from comments by Eser Arisoy, Marius Zoican, Kewei Hou, Eric Zitzewitz, Dmitry Makarov, Jeroen Derwall, Dan Galai, Pierre Armand Michel, and Christian Wolff as well as the participants in the December meeting of the French Finance Association 2014 (Paris), the 28th Australian Finance and Banking Conference 2015 (Sydney), the Financial Management Association 2017 (Boston) and the participants in research seminars at the University of Paris Dauphine, at HEC Montréal, at the University of Hohenheim, at the University of Bologna, at the University of Strasbourg and at Ghent University. Marie Lambert and Georges Hübner acknowledge the financial support of Deloitte (Belgium and Luxembourg). All remaining errors are ours.

Electronic copy available at: https://ssrn.com/abstract=3179655

The literature contains a variety of multifactor models that attempt to explain security returns using stock "attributes," such as market capitalization, book-to-market, and investments (see Barillas and Shanken (2018), Fama and French (2015, 2018), Hou et al. (2015a, 2018, 2019), Stambaugh and Yuan (2017)). More than 300 characteristic-based anomalies are documented in Harvey et al. (2016). The empirical implementation of these fundamental multifactor models faces the common challenge of constructing mimicking or hedge portfolios that capture the marginal returns associated with a unit of exposure to each attribute. However, "one has little guidance on how to construct better factors" (February 5th, 2019, Editor Amit Goyal, Review of Finance).[1] The recent literature has mainly focused on finding and creating factors and factor models that best explain and represent systematic risk, which is a central issue in asset pricing. The proposed models differ not only in terms of the stock attributes used as a proxy for risk but also in how the factors are constructed. Although these papers show that these construction rules are important, there are still ad-hoc choices with little guidance on the choice of the cutoff rule as well as the type of sort to be performed across characteristics. To address this issue, our paper works in a delimited framework (size and book-to-market) and investigates the impact of the different portfolio sorting techniques for constructing characteristic-based portfolios. With Fama and French (1993, hereafter FF) as the established benchmark, we study the various options for constructing risk factors using this framework over a very long period (51 years). On the one hand, we show that the performance of risk factors depends on the diversification of the portfolio constituents. Our findings provide insights on Lehmann and Modest (2005), who observe an outperformance of basis portfolios in terms of mean and variance when formed on a larger but finite sample. On the other hand, the

[1] The article can be found at the following address: http://revfin.org/which-factors/.

Electronic copy available at: https://ssrn.com/abstract=3179655

findings show that the number of stocks in certain portfolios vary over time according to the methodology used and that these variations could be systematically related to economic conditions.

Recent papers have already shown the importance of sequential sorting (Chan et al. (2009)) and the consequences of moving from NYSE to NYSE-AMEX-NASDAQ (whole-sample, name[2]) breakpoints (Hou et al. (2018)). Kogan and Tian (2015) and Hou et al. (2019) show that the performance of multifactor models is sensitive to these factor methodologies. However, the questions of the (a)symmetry of the sort and the use of pre-versus post-sorting within a sequential sorting have not received much attention in the literature. In addition, the remaining gap in this literature is the question of the joint effects of all portfolio sorting methodological choices, which is the focus of our paper. We establish that the empirical performance of the factors is highly sensitive to the number of stocks in portfolios. This number itself depends on the economic context. Our theoretical framework assumes that sorting stocks on characteristics consists of a sort on expected returns. Cochrane (2011) formally shows the effect of portfolio sorts on the Sharpe ratio of spread portfolios and the distance to the Sharpe of the true underlying risk factor. The distance inversely depends on the number of stocks sorted into the short and long legs of the spread: a lower number of stocks will reduce the spread portfolio's Sharpe ratio and *t*-statistic by introducing idiosyncratic or unpriced risk. The underperformance is more severe when the level of diversification varies among the portfolios constituting the risk factors, i.e., in case of an unequal allocation of stocks into portfolios. We demonstrate the biases introduced by such an unequal number of stocks into portfolios when building risk factors. We also show that these biases impact factor performance.

[2] While Hou et al. (2019) refer to the breakpoints considered for all the stocks in the three main U.S. equity markets as "NYSE-AMEX-NASDAQ", whereas this paper refers to the same breakpoints as “name” breakpoints.

Electronic copy available at: https://ssrn.com/abstract=3179655

On the methodological side, we use the same criterion as Barillas and Shanken (2018) and Fama and French (2018), namely the maximum squared Sharpe ratio, to assess the performance of our spread portfolios. However, the stance of our study is different and upstream from their papers. Our goal is not to compare model specifications or make the selection among anomaly based factors.[3] Rather, we examine the impact of the factor construction methods, especially of the portfolio sort, on the mean-variance risk-return properties of the factors, which is why we decompose each of the factor construction choices using a clinical study on a single set of factors rather than conducting a comparison among different anomaly based factors. The horse race that we carry out is not one *between* different models but one between different ways of creating risk premiums *within* the same model.

The findings of our research are, in our opinion, economically significant. Our alternative framework generates much stronger "turn-of-the-year" size and "through-the-year" book-to-market effects than what is conventionally documented.[4] The better method of identifying the size and value factors that we uncover for the US market has implications for factor investing and asset pricing. Moreover, this question contributes to the recent debate on the difficulty of inferring independent information about average returns (Green et al. (2017)) and the difficulty of producing

[3] We do not aim to test whether our factors could subsume other existing factors among the more than 300 currently listed factors. This question is beyond the scope of this paper due to the issues of multiple testing and limited sample data.

[4] See for instance van Dijk (2011) for a review of literature on the effect related to the size premium as well as Moller (2008) for an empirical investigation on the "turn-of-the-year" size effect.

Electronic copy available at: https://ssrn.com/abstract=3179655

risk factors exempted from unpriced risks (Daniel et al. (2018a)).[5] These issues are important because better constructed factors can improve asset pricing tests and might lead to the acceptance of a model that would otherwise be rejected because of poor measurement. We also contribute to another important recent finding in this field made by Grinblatt and Saxena (2019), who show that mimicking risk factors can be improved by using optimal weights of basis portfolios rather than traditional equal weighting of the long and short positions. The authors do not compare ways of constructing these underlying "basis portfolios" or challenge their configuration. For this reason, we think that the two methods (Dependent-Symmetric-Name (DSN) and the Grinblatt and Saxena method) are not directly comparable or in competition. However, testing the joint combination of the approaches of Grinblatt and Saxena (2019) and ours could potentially lead to empirical factors closer to the true mean-variance efficient factors. This research would indeed combine the merits of the two different approaches. We consider this test to be nonetheless out of the scope of this research as we do not claim that our DSN approach produces mean-variance efficient factors. On the contrary, we show that the DSN approach limits ad hoc sorting choices and biases and brings significant empirical outperformance to current sorting methods.

The rest of the paper is organized as follows. Section 1 reviews the literature on portfolio sorting. Section 2 presents our theoretical framework and potential biases induced by factor construction choices. Section 3 applies our theoretical framework for size and value factors and describes the dataset of US stocks used to perform this clinical exercise. Section 4 investigates the particular effect of breakpoints. Section 5 performs mean-variance spanning tests on a series of

[5] For instance, on the issue of identifying a parsimonious set of risk factors that carry independent information about stock returns, Daniel et al. (2018a, p. 4) write that "*any cross-sectional correlation between firm-characteristics and firm exposures to unpriced factors will result in the factor-portfolio being inefficient*".

Electronic copy available at: https://ssrn.com/abstract=3179655

basis portfolios and against competing models. Section 6 conducts robustness tests. Section 7 concludes the paper.

1. Literature Review

The Fama and French (1993) three-factor model and its extension to momentum by Carhart (1997) have become the benchmarks of empirical asset pricing. Using datasets from the merger of the Center for Research in Security Prices (CRSP) and Compustat, FF consider two independent methods for scaling US stocks, including an annual two-way sort on market equity and an annual three-way sort on book-to-market according to NYSE breakpoints (quantiles). Next, they construct six value-weighted (two-dimensional) portfolios at the intersections of the annual rankings (performed each June of year *y* based on the fundamentals displayed in December of year *y*-1). The size or SMB factor measures the return differential between the average small-cap and the average large-cap portfolios, whereas the book-to-market or High-minus-Low (HML) factor measures the return differential between the average value and the average growth portfolios. The resulting so-called "Fama-French three-factor model" has become a central version of empirical asset pricing models.

These portfolio sorting choices have already been questioned by the literature. Kogan and Tian (2015) show that the performance of a multifactor model in sample is sensitive to the factor construction method and that the performance is not persistently out-of-sample. Hou et al. (2018) document that the factors' performance is sensitive to the choice of breakpoints, while Chan et al. (2009) plead for the construction of attribute-matched portfolios formed under dependent sorting. However, their performance attribution approach requires knowledge of the fund holdings to perform the matching.

Electronic copy available at: https://ssrn.com/abstract=3179655

This section reviews the main findings in terms of bias in stock attributes and sorting procedures. The rest of the paper will contribute to the literature by considering the joint implications of sorting features, i.e., adding the questions of a symmetric sort and of a pre-versus post-sorting within a dependent sorting, on the mean-variance performance of the factors.

Correlation bias in stock attributes. To study the performance of US equity funds, Daniel et al. (1997), Daniel and Titman (1998), and Wermers (2004) favor a characteristic-based approach over traditional regression-based analyses. Chan et al. (2009) emphasize that regression-based analyses fail to value the performance of passive portfolios because the correlation bias in stock attributes drives stock allocation into portfolios. In other words, if the intrinsic correlation between size and value characteristics is not controlled by factor construction, then it cannot be properly controlled for in a regression. For instance, the abnormal returns of a small-value portfolio might be underestimated when small growth stocks outperform: small-value portfolios would load heavily on the size factor inflated by the outperforming small growth stocks (without appropriate controls for building the size factor). Hence, the regression-based analysis does not completely resolve the intrinsic correlation between factors through orthogonalization. Daniel et al. (2018a) resurrect the multifactor approach and show that it can be equivalent to a characteristic-attribution model if the factors are hedged for unpriced risks. To do so, they use industry-level breakpoints to allocate stocks into portfolios and to ensure that the long and short legs of the factors are characteristic-balanced/neutral portfolios, and they show that the Sharpe ratio of the factors might be improved significantly.

Dependent sort. The comprehensive study of Chan et al. (2009) unambiguously outlines the advantages of attribute-matched portfolios over regression-based analyses for assessing portfolio performance. They also justify the use of a sequential over an independent sort by

Electronic copy available at: https://ssrn.com/abstract=3179655

identifying adequate size and value benchmarks. The authors prove the practical relevance of a sequential sort based on its similarity to the construction mechanism of renowned benchmark indices widely used by institutional investors, such as Russell, Standard and Poor's and Wilshire. However, although their study focuses on the quality of benchmarks produced with alternative stock classification procedures, it does not investigate the properties of the methodological choices underlying the sorting procedure.

Several other examples of the use of a conditional sorting procedure exist, especially when data are scarce or in international studies (Daniel et al. (1997), Daniel and Titman (1998), Ang et al. (2006), and Novy-Marx (2013)). Agarwal et al. (2009) sort hedge funds into portfolios using the same sequential approach with the objective of estimating higher-moment risk factors. In international asset pricing, Liew and Vassalou (2000) adapt the approach with a triple conditional sort to compute size, value and momentum factors for various countries. Performing a dependent sort, however, poses several challenges that have not been investigated, such as the ordering of the sort or the joint effect of alternative portfolio construction choices.

Breakpoints. Little has been written about the choice of breakpoints and cutoffs. A few studies use 20/80 cutoffs instead of the traditional 30/70 FF original split (Stambaugh and Yuan (2017), Daniel et al. (2018b)) and whole-sample breakpoints (see, Stambaugh and Yuan (2017)). Hou et al. (2018) advocate for the use of NYSE breakpoints and 30/70 to avoid the effect of micro-caps. Thus far, its joint effects with other sorting methodologies and the direct consequences of alternative cutoffs have not been investigated in full detail, which is very important given that portfolio sorts are performed on characteristics that are assumed to be correlated with returns but for which we do not know the distribution.

Electronic copy available at: https://ssrn.com/abstract=3179655

Symmetry. The asymmetric sorting using a double sort on size and a triple sort on other criteria, such as book-to-market, has not been questioned thus far by the literature. Originally, the methodological choice of a 2x3 sort using NYSE breakpoints was performed to ensure the same market capitalization in the different portfolios (Fama and French (1993)). The lower dimensionality in the size sort compensates for the high correlation between the sorted characteristics when stocks are sorted into portfolios using an independent sort. An independent sort on highly correlated characteristics could not be extended from a 2x3 sort to a 3x3 or 3x3x3 sort because the correlation between characteristics would potentially produce empty portfolios, which is supported by the simulation results presented in Tables 1 and 2.[6]

Other sorting features. We also check whether our framework can accommodate a third attribute. As a practical example, we investigate momentum effects when pricing size and value factors. Recent papers show the effect of news events on stock returns (Li et al. (2008), Savor and Wilson (2016)) and a return clustering effect for market anomalies around news events (Bowles et al. (2017), and Engelberg et al. (2018)). Sorting on momentum might constitute a control for the release of news. Our results are robust to the inclusion of this additional control.

2. Theoretical Framework

We build on and extend the framework of Cochrane (2011, Appendix B Asset Pricing as a Function of Characteristics, p. 1097) to show how the stock imbalance between the long and short legs of the spread portfolio as well as the lack of diversification of these portfolios will lead the Sharpe ratio of the spread portfolio to deviate from the Sharpe ratio of the true common factor.

[6] We acknowledge that there is a natural limit to this extension as the more refined the sorting, the more idiosyncratic risk we have in the smaller buckets. We thank an anonymous referee for raising this point.

Electronic copy available at: https://ssrn.com/abstract=3179655

2.1. Portfolio Diversification and Effects on Factor Construction Biases

Cochrane (2011) posits the following relationship between characteristics and expected returns:

$$E\left(R^i - R^j\right) = b\left(C_i - C_j\right) \tag{1}$$

where *C* stands for the characteristics (i.e., size or book-to-market) of portfolio *i* or *j*.

Assuming that the sort on characteristics corresponds to an expected return sort with an underlying common risk factor *f*, Cochrane (2011) defines the variance in the spread portfolio as

$$\sigma^2\left(R^i - R^j\right) = (\beta^i - \beta^j)^2\sigma(f)^2 + 2\frac{\sigma_\varepsilon^2}{N} = \frac{b^2}{E(f)^2}(C^i - C^j)^2\sigma(f)^2 + 2\frac{\sigma_\varepsilon^2}{N} \tag{2}$$

where β stands for the exposure of portfolio *i* or *j* to the common factor *f*, *N* is the number of stocks within the spread portfolios, and σ_ε^2 is the idiosyncratic variance in the individual stocks composing the spread portfolios.

We extend this equation with subscripts *i* and *j* for the number of stocks in the long and short leg of the spread portfolios, respectively, because they might differ.

$$\begin{aligned}\sigma^2\left(R^i - R^j\right) &= (\beta^i - \beta^j)^2\sigma(f)^2 + \frac{\sigma_\varepsilon^2}{N^i} + \frac{\sigma_\varepsilon^2}{N^j} \\ &= \frac{b^2}{E(f)^2}(C^i - C^j)^2\sigma(f)^2 + \frac{\sigma_\varepsilon^2}{N^i} + \frac{\sigma_\varepsilon^2}{N^j}\end{aligned} \tag{3}$$

The Sharpe ratio of the spread portfolio return can now be defined as

$$\frac{E\left(R^i - R^j\right)}{\sigma\,(R^i - R^j)} = \frac{E(f)}{\sigma(f)}\frac{b\left(C_i - C_j\right)}{\sqrt{b^2\left(C_i - C_j\right)^2 + \frac{\sigma_\varepsilon^2}{N^i}\frac{E(f)^2}{\sigma(f)^2} + \frac{\sigma_\varepsilon^2}{N^j}\frac{E(f)^2}{\sigma(f)^2}}} \tag{4}$$

Defining a factor F as the spread on one risk characteristics C*, controlling for three levels of control (low, medium and high), and assuming for the sake of simplicity that the portfolio spreads

Electronic copy available at: https://ssrn.com/abstract=3179655

for different levels of the controls are perfectly correlated (the risk premium related to C* should be the same across the control variables), and we can generalize the formula as

$$\frac{E(F)}{\sigma(F)} = \frac{E(f)}{\sigma(f)} \frac{\sum_{i=L,M,H} b(C_{Li} - C_{Si})}{\sum_{i=L,M,H} \sqrt{b^2(C_{Li} - C_{Si})^2 + \frac{\sigma_\varepsilon^2}{N_{Li}} \frac{E(f)^2}{\sigma(f)^2} + \frac{\sigma_\varepsilon^2}{N_{Si}} \frac{E(f)^2}{\sigma(f)^2}}} \tag{5}$$

where S and L denote the "short" leg and "long" leg, respectively, within the three levels of controls (L, M, H), and *f* is the true underlying common risk factor.

From equation (5), it follows that the Sharpe ratio and *t*-statistics of the spread portfolio will be closer to those of the true common factor for a large characteristics spread or for better diversified portfolios, all else being equal. The problem consists of ensuring the proper diversification of the portfolios because a finer sort might diminish the number of stocks in portfolios.

2.2. Impact for Risk-Return Properties

This subsection goes further by showing the deviation between the Sharpe of the spread portfolio and the Sharpe of the true common factor due to an imbalance in the number of stocks between the short (denoted by S) and long (denoted by L) leg portfolios forming the return spread (both short and long portfolios are considered here to be diversified across the different levels of controls for the sake of simplicity).

The Sharpe ratio of the spread portfolios can be written as

$$SR(spread) = SR(f) \frac{b(C_L - C_S)}{\left[b^2(C_L - C_S)^2 + \frac{\sigma_\varepsilon^2}{N_L} SR(f)^2 + \frac{\sigma_\varepsilon^2}{N_S} SR(f)^2\right]^{1/2}} \tag{6}$$

Collecting the term $SR(f)$ of the denominator and squaring the equation, we obtain the squared Sharpe ratio (SSR), which is equal to

Electronic copy available at: https://ssrn.com/abstract=3179655

$$SR(spread)^2 = \frac{b^2(C_L - C_S)^2}{\frac{b^2(C_L - C_S)^2}{SR(f)^2} + \frac{\sigma_\varepsilon^2}{N_L} + \frac{\sigma_\varepsilon^2}{N_S}} \qquad (7)$$

Rearranging the terms, setting the common denominator and collecting the like terms N_S/N_L gives

$$SR(spread)^2 = \frac{N_S}{N_L}\left(\frac{N_L\, b^2(C_L - C_S)^2 SR(f)^2}{N_L\, \frac{N_S}{N_L} b^2(C_L - C_S)^2 + \sigma_\varepsilon^2\, SR(f)^2 + \frac{N_S}{N_L}\, \sigma_\varepsilon^2\, SR(f)^2}\right) \qquad (8)$$

For illustrative purposes, setting the unknown Sharpe ratio to 1, that is, $SR(f) = 1$, the equation becomes

$$SR(spread)^2 = \frac{N_S}{N_L}\left(\frac{N_L\, b^2(C_L - C_S)^2}{N_L\, \frac{N_S}{N_L} b^2(C_L - C_S)^2 + \sigma_\varepsilon^2 + \frac{N_S}{N_L}\, \sigma_\varepsilon^2}\right) \qquad (9)$$

The SSR measure has been used since the research of Treynor and Black (1973) and has been used recently in Daniel et al. (2018a), Fama and French (2018) and Barillas and Shanken (2018) to compare the performance of asset pricing models and risk factors.

Figure 1a presents the SSR as a function of the ratio of stocks found in the short leg over the long leg of the spread (N_S/N_L) setting the unobservable variables $SR(f), b(C_L - C_S)$ and σ_ε to 1, 10%, and 30%, respectively. The results show that the SSR converges to 1 for N_S/N_L equal to 1 (or log(N_S/N_L) equal to zero) and a large sample of stocks (i.e., high N_S+N_L). For a N_S/N_L value inferior or superior to 1, the number of stocks being constant, the SSR rapidly drops. Figure 1b presents the same relationship when $b(C_L - C_S)$ is set equal to 5% and shows that the effect is greater in the case of a weaker relationship between the characteristics and the return. This stylized fact elaborates on Lehmann and Modest (2005), who show the impact of sample size in the performance of basis portfolios sorted along sample characteristics. We extend that claim and

Electronic copy available at: https://ssrn.com/abstract=3179655

show that the balanced number of stocks between the long and short legs of hedge portfolios (N_S/N_L) also contribute to greater mean-variance performance of risk factors.

Figure 1

Squared Sharpe Ratio as a Function of the Number of Stocks in the Spread Portfolios N_S and N_L

The figures schematize the relationship between the Squared Sharpe ratio (SSR) formed from a spread of portfolios and the number of stocks featured in these portfolios. The number of stocks present in the long and short leg portfolios are denoted N_L and N_S, respectively. We present the SSR as a function of the ratio of stocks found in the short leg over those found in the long leg of the spread ($\frac{N_S}{N_L}$), and the unobservable variables $SR(f)$ and σ_ε are set equal to 1 and 30%, respectively. Figure 1a shows the SSR when $b(C_L - C_S)$ is set equal to 10%, while Figure 1b presents the results when $b(C_L - C_S)$ is set equal to 5%. The left plots present the squared Sharpe ratio when the *x*-axis has a linear scale, while the right plots use a log-scale for the *x*-axis.

Figure 1a: $b(C_L - C_S)$ is set equal to 10%

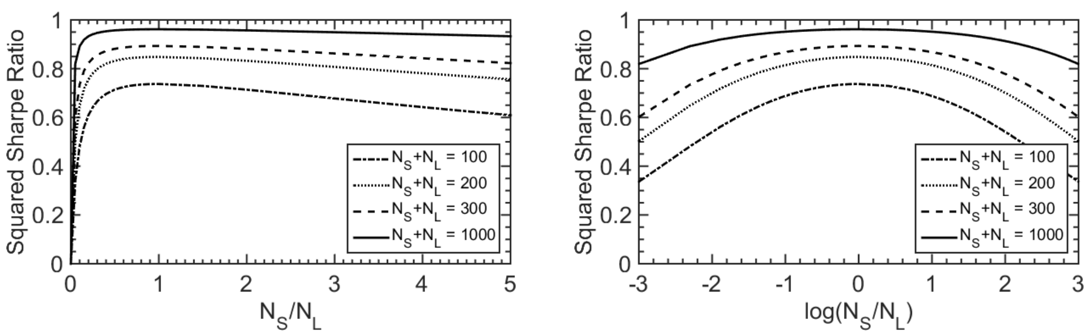
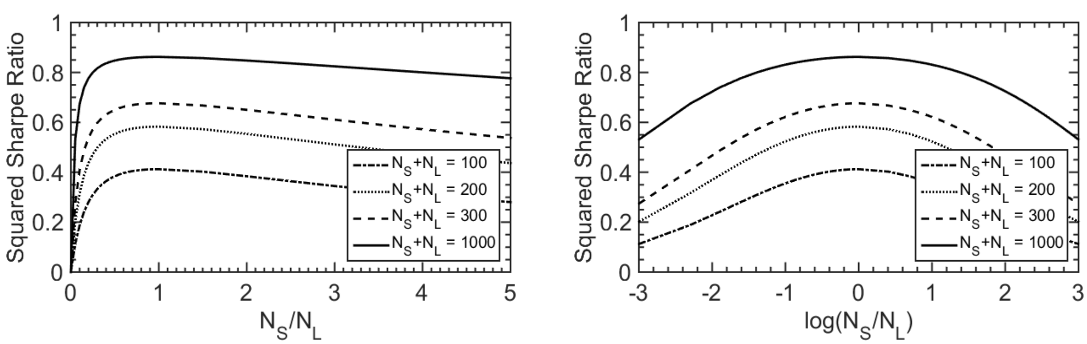


Figure 1b: $b(C_L - C_S)$ is set equal to 5%



Electronic copy available at: https://ssrn.com/abstract=3179655

2.3. Independent, Dependent Sorted Portfolios and Stock Imbalance

An independent 2x3 sort is compared to a dependent sort on the same characteristics in Figure 2. An independent sort slices the stock universe according to two independent scales on characteristics X and Y (Panel A), while a dependent sort proceeds in two steps to form X/Y portfolios. It can start with either characteristic Y (Panel B: first line) or characteristic X (Panel B: second line). Compared with the independent sort, the dependent sort adjusts the breakpoints in the second sorting step considering the correlation between the characteristics.

Figure 2 compares the stock allocation into portfolios under a dependent and an independent sort. Panels A and C (resp. B and D) illustrate the case of an independent (resp. dependent) sort on negatively correlated characteristics (such as book-to-market and market capitalization). Panels A and B depict a situation in which the two fundamentals are correlated at -30%, whereas Panels C and D consider a perfect negative correlation (-100%) between the two characteristics on which the sort is performed. The figure shows that the high level of correlation produces imbalanced portfolios under an independent framework. The figures also illustrate how the adjustments of the breakpoints under the sequential sort allow for the even split of stocks into portfolios. When the characteristics are perfectly correlated, an independent sort would even produce empty portfolios as shown in Panel C.

Our main working hypothesis is that the original independent 2x3 sort used to construct the size and value factors will lead to underdiversification of the spread portfolios and therefore the underperformance of the factor. We split this hypothesis into two sub-hypotheses:

(H1) An independent 2x3 sort will induce the underdiversification of the spread portfolios.

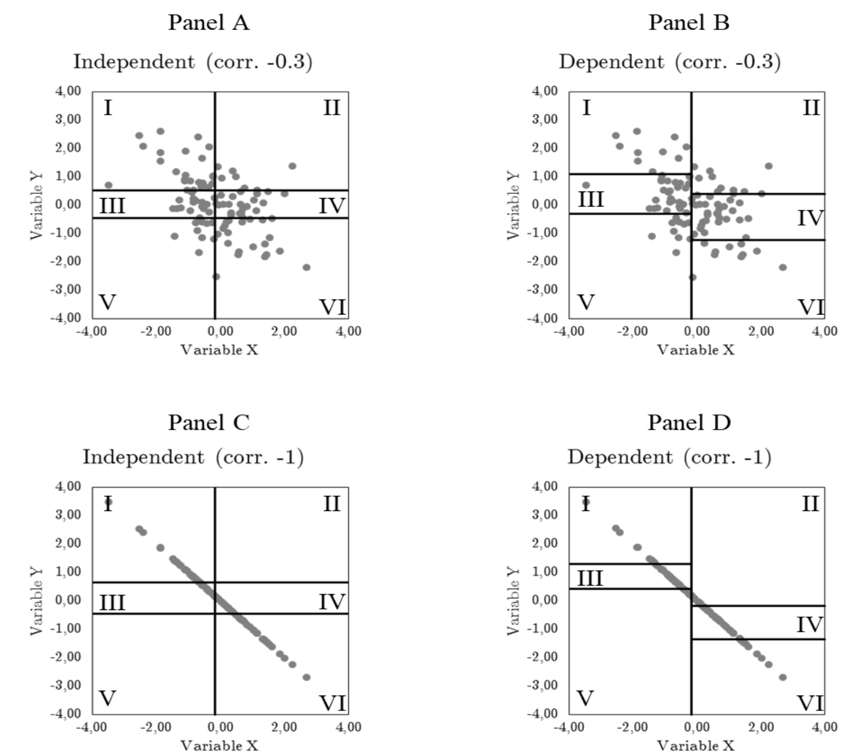
(H2) Diversification of the spread portfolios will improve factor performance.

Electronic copy available at: https://ssrn.com/abstract=3179655

Figure 2

Independent *vs.* Sequential Sorting: Allocation into Portfolios[7]

The figures illustrate the allocation of 100 stocks sorted on variables X and Y across six portfolios. Panel A (Panel B) shows the allocation according to an independent (dependent) sort when the correlation between characteristics *x* and *y* is -30%. Panel C (Panel D) shows the allocation according to an independent (dependent) sort when the correlation between characteristics X and Y is perfectly negative (-100%).



2.4. Empirical Biases Induced by Portfolio Sorting Choices

From equation (5), it follows that the Sharpe and *t*-statistics of the spread portfolio will be closer to those of the true common factor for a large characteristics spread, all else being equal.

[7] We thank Nick Baltas for suggesting this analysis.

Electronic copy available at: https://ssrn.com/abstract=3179655

The problem consists of ensuring the portfolios are properly diverse because a finer sort might diminish the number of stocks in portfolios.

We illustrate this theoretical framework in the case of a two-dimensional sort with perfectly negatively correlated variables. We consider a first sort on market capitalization (ME) (each year in June) and a second sort on the inverse of market capitalization, i.e., –ME. The aim of the exercise is to show that sorting on two characteristics with a perfect negative correlation should deliver the same pricing effect for a Small Minus Big factor and a High Minus Low (or in this case, a Big Minus Small) factor unless the sorting methodology produces undesirable effects. We construct a SMB factor with market capitalization (ME) and a HML factor with the inverse of market capitalization (–ME).

We first display in Table 1 the number of stocks that fall into the portfolios for the different sorting configurations, i.e., the choice of the sort (independent or dependent), the sort scaling (2x3 or 3x3) and the definition used for the breakpoints (NYSE or all names).

[Table 1 near here]

The table shows that an independent 2x3 sort produces a strong imbalance in the number of stocks in each portfolio, i.e., from 0 to 2226 stocks using NYSE breakpoints and from 0 to 1021 stocks using name breakpoints. The problem worsens under a symmetric 3x3 sort. A dependent sort with name breakpoints produces a much more diversified portfolio across the different configurations: the best diversification is achieved under a symmetric sort.

These results support hypothesis H1, which states that an independent 2x3 sort will result in the under-diversification of the spread portfolios.

Under Cochrane's (2011) framework, this problem in stock allocation into portfolios could affect the risk premium's performance. To measure the bias it induces, we construct different size

Electronic copy available at: https://ssrn.com/abstract=3179655

premia using either the first or the second sort and report their descriptive statistics and correlations in Table 2. By construction, they should display similar descriptive statistics and should be perfectly correlated.

[Table 2 near here]

Although correlations between the two size premia produced under the 2x3 independent NYSE framework appear to be very high (98%), the premia's statistics (i.e., mean return and volatility) strongly differ (Panel A, Table 2). However, the standard method fails when we try to extend it to name breakpoints or symmetric double sorting, thus inducing a finer stratification of US stocks into portfolios. Per construction, the pairs of factors should similarly price the size effect; however, this similarity is not achieved as shown by the correlation coefficients for these extensions of the original framework. The 3x3 independent sort (Panel C, Table 2) induces an imbalanced stratification of stocks into portfolios (cfr. Panel C, Table 1). The independent framework fails to price the size effects when extended to a symmetric sort, which illustrates the biases introduced by independent sorting on correlated characteristics. These findings support hypothesis H2 stating that under-diversification of the spread portfolios will cause the underperformance of the factor.

Under a dependent symmetric (3x3) framework, the twin premia's correlation is close to 100%, with very similar descriptive statistics as per the definition. This finding demonstrates that the dependent symmetric framework can be applied to highly correlated characteristics without introducing measurement biases. In addition, the *t*-statistics are the highest for the dependent 3x3 name breakpoint configuration (Panel H, Table 2). This evidence should be linked with the results of Table 1, Panel H, which shows that this sorting methodology maximizes diversification across the constituting portfolios. Under the framework of Cochrane (2011) described by equation (5), a

Electronic copy available at: https://ssrn.com/abstract=3179655

high level of diversification within portfolios, i.e., high values for N_{SI} and N_{BI}, indeed decreases the distance between the Sharpe or the *t*-statistics of the benchmark portfolio (or spread portfolio) and those of the underlying risk factor (Panel H, Tables 1 and 2), which also supports H2: Diversification of the spread portfolios within the symmetric dependent sort improves the factors' performance.

3. Empirical Economic Significance of the Biases: The Case of Size and Value Factors

We examine three methodological sorting choices that have an impact on portfolio diversification: (i) independence, (ii) asymmetry, and (iii) breakpoints. We construct the size and value factors by *preconditioning* the sorting procedure on the control variable and ending the sort with the variable to be priced, such as in Lambert and Hübner (2013).[8] Beyond the dependent sorting, we consider the choice of the breakpoints. The breakpoints used as a scale to allocate stocks into level portfolios can be defined either using the whole sample (i.e., using all firms and all names) or using only the firms from the NYSE; for the sake of simplicity, we refer to these breakpoints as "name" and NYSE, respectively. We also consider the impact of the number of portfolio buckets and consider *NxN* sorts. Factor performance proves to be highly sensitive to the distribution of stocks in portfolios. We call this method DSN which refers to a (D)ependent sort that starts with the control variables and (S)ymmetric splits on all-(N)ame sample breakpoints.

[8] The sequential sort can be performed by preconditioning either on the control variables or on the characteristics to be priced (i.e., postconditioning on the control variables). The two procedures do not capture the same pricing effects. Unlike the postconditioning approach, the preconditioning approach ensures that the risk factor is an equally weighted average of the spreads for each level of control. Results are available upon request.

Electronic copy available at: https://ssrn.com/abstract=3179655

3.1. Data: Reproducing the Fama and French (1993) Standard Method

Since the purpose of this paper is to build a framework that allows for a robust comparison considering the original FF approach as a standard, we strictly follow their stock selection methodology to construct our risk factors. The period ranges from July 1963 to December 2014 and includes all NYSE, AMEX, and NASDAQ stocks collected from the merger between the CRSP and Compustat databases. The analysis covers 618 monthly observations. The market risk premium corresponds to the value-weighted return on all US stocks minus the one-month T-bill rate from Ibbotson Associates (from Kenneth French's website). We consider stocks that fully match the following lists of filtering criteria: a CRSP share code (SHRCD) of 10 or 11 at the beginning of month *t*; an exchange code (EXCHCD) of 1, 2 or 3, available shares (SHROUT) and price (PRC) data at the beginning of month *t*; available return (RET) data for month *t*;[9] at least two years of listing on Compustat to avoid survival bias (Fama and French (1993)) and a positive book-equity value at the end of December of year *y*-1. Thus, our sample varies over time. For instance, from 5,612 stocks available as of December 2014, our conditions restrict the usable sample to 3,335 stocks (for 2014).

As in Fama and French (1993), we define the book value of equity as stockholders' equity reported by Compustat (SEQ) plus balance sheet deferred taxes and the investment tax credit (TXDITC). If available, we decrease this amount by the book value of preferred stock (PSTK). If the book value of stockholders' equity (SEQ) plus the balance sheet deferred taxes and investment

---

[9] When available, we include for each stock the delisting return recorded in the CRSP Event files. Using this approach allows us to replicate closely the traditional Fama-French Size and Value factors, which constitute our baseline and benchmark model.

Electronic copy available at: https://ssrn.com/abstract=3179655

tax credit (TXDITC) is not available, we use the firm's total assets (AT) minus total liabilities (LT).

Book-to-market is the ratio between book common equity for the fiscal year ending in calendar year *t*-1 and the market equity as of December *t*-1. Market equity is defined as the price (PRC) of the stock times the number of shares outstanding (SHROUT) at the end of June in year *y* to construct the size factor and at the end of December of year *y*-1 to construct the value factor.

Carhart (1997) completes the FF three-factor model by computing a momentum (i.e., a *t*-2 until *t*-12 cumulative prior-return) or UMD (up minus down) factor that reflects the return differential between the highest and the lowest prior-return portfolios.

3.2. Triple Sort

This subsection extends the method based on the conditional sorting procedure using a triple sort on name breakpoints, which can be viewed as an extension of the approach with two control variables and one pricing factor. We consider three risk dimensions (size, value and momentum) with preconditioning on momentum to control for the business cycle, earnings surprises and profitability shocks. The DSN factor construction proceeds as follows for the value (resp. size) factor. It starts by breaking up the NYSE, AMEX, and NASDAQ stock universe into three groups according to an initial control on momentum. It then successively decomposes each of the three momentum-portfolios into three sub-portfolios according to a second control criterion, i.e., market capitalization (resp. book-to-market). The final split will create for each of the nine portfolios three new sub-portfolios according to a third criterion to be priced, i.e., book-to-market (resp. market capitalization). The US stock universe would thus be composed of 27 portfolios used to reconstruct a single value (resp. size) factor. Similar to Fama and French (1993), all portfolios are value-weighted. The rebalancing is performed on an annual basis at the end of June of year *y*.

Electronic copy available at: https://ssrn.com/abstract=3179655

Under this framework (i.e., 3x3x3 and name breakpoints), the size factor is based on the outperformance of stocks with low equity compared to those with high market capitalization within the control subportfolios.[10] This practice ensures that the stocks with high book-to-market due to very low market capitalization do not drive up the HML premium. As in Fama and French (1993, p. 12), we refer to these stocks as "fallen angels", which refer to “big firms with low stock prices”. Moreover, a stock whose characteristics remain unchanged may move to another book-to-market classification even if the full book-to-market cross-section does not change in a year. This movement could occur if the stock returns follow an upward trend that would inflate its market value and inaccurately affect its book-to-market ratio. Independent sorting would miss this information and incorrectly determine a low book-to-market ratio. Such flexibility in stock migration is certainly a core element of the sequential procedure since it aims to ensure that the classification of one of the priced variables (e.g., book-to-market) is not affected by the controls (e.g., market equity).

3.3. Bias 1: Underestimation of the Value Effect

This subsection compares the performance of a dependent and symmetric sort performed on name breakpoints (DSN) to the original FF framework. We construct a type of information ratio of the DSN factors on the original FF factor. To compute the information ratio, we perform the following regression on the 252 daily returns observed after the formation of the portfolio (from the 1st of July of year *y* to the 30th of June *y*+1),

$$HML_{t,y}^{DSN} = \alpha_y + \beta_y HML_{t,y}^{FF} + \varepsilon_{t,y} \tag{10}$$

---

[10] A book-to-market ratio (second sort) of 0.5 may put a stock in the high book-to-market portfolio in one momentum-size portfolio (first sort), in the medium book-to-market in another, and in the low book-to-market in a third depending on the cross-sectional variation in the subportfolios.

Electronic copy available at: https://ssrn.com/abstract=3179655

with $HML_{t,y}^{DSN}$ representing the value factor from a dependent 3x3x3 sort on name breakpoints and $HML_{t,y}^{FF}$ representing the value factor from an independent sort. The regression thus provides 52 yearly observations (from July 1963 to December 2014) of intercepts α_y and SDs of $(\varepsilon_{t,y})$. The yearly information ratio (IR) is given by

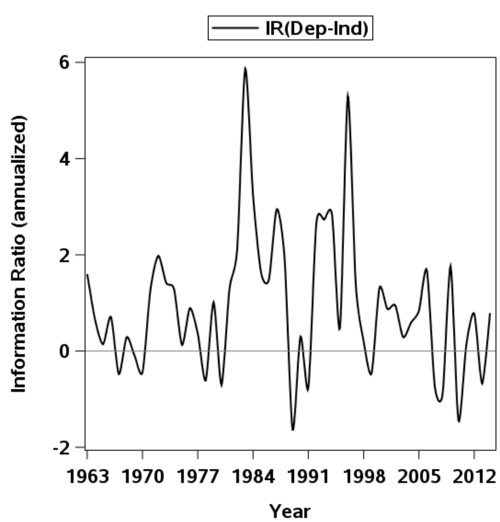
$$IR_y = \frac{\alpha_y}{\sigma(\varepsilon_{t,y})} \tag{11}$$

Figure 3 displays the evolution of these ratios for HML over the period. The IR for the value factor (HML) has an average value of 0.94, a SD of 1.46, a minimum value of -1.64 and a maximum value of 6.55; more interestingly, however, it is positive for 77% of the 52 yearly observations.

Figure 3

Information Ratio of the Value and Size Factors

The figure shows the evolution of the information ratio between the dependent and independent value factors, and it shows the information ratio from the spread of the dependent and independent value factors. The sample period ranges from July 1963 to December 2014.



Electronic copy available at: https://ssrn.com/abstract=3179655

The following analysis further investigates these results. We examine whether the outperformance of the DSN value factor is related to the allocation of stocks into portfolios and the level of diversification of constituent portfolios as suggested by our theoretical framework developed in Section 2. Finally, we investigate whether stock allocation into portfolios and factors' performance is sensitive to economic conditions.

Table 3 displays the mean return and the volatility of value factors formed from seven different sorting models. Models 1 to 4 present incremental changes to the configuration of the independent sort using NYSE or name-breakpoints and asymmetric (2x3) to symmetric (3x3) size and value splits. Models 5 to 7 present incremental changes to the configuration of the dependent sort using NYSE or name-breakpoints and asymmetric (2x3) to symmetric (3x3) size and value. An extended 3x3x3 DSN sort is also displayed. Over the full sample period, the SSR is maximized when the name breakpoint is used together with a symmetric sort under both a dependent and independent framework. The SSR increases from 0.29 to 0.88 (independent sort) and from 0.52 to 0.61 (dependent sort) when moving from an asymmetric to a symmetric sort using name breakpoints. The table also shows the important underperformance of using an asymmetric sort with name breakpoints.

As reported in Table 3 for the value factor, in the post-1973/NASDAQ period, name-breakpoints lead to a more balanced N_S/N_L and exhibit *lower* volatility compared to NYSE breakpoints only when the sort is *dependent* (Model 5). These findings explain the outperformance of the factor, which is related to lower volatility. However, the higher Sharpe ratios achieved by using name-breakpoints and a symmetric independent sort (Model 4) predominantly come from

Electronic copy available at: https://ssrn.com/abstract=3179655

higher excess returns.[11] This analysis verifies that the source of the outperformance is indeed related to a decrease of volatility due to better diversified portfolios under a DSN approach.

[Table 3 near here]

Next, we examine whether the determinants of outperformance are related to the balance in the number of stocks in the constituent portfolios and the balance in the long and short legs of the factor as suggested by our theoretical framework developed in Section 2. We also consider different macroeconomic variables to test our previous hypotheses regarding the link between sorting methodologies and market conditions. This would also inform us whether the biases are more economically significant in some specific conditions.

Figure 4 investigates the link between the imbalance in the number of stocks in the short and long legs of the value (HML) factors and their relative performance level as well as their dependency on economic conditions, and it also shows the outperformance of the dependent 3x3x3 name framework over the independent FF framework and indicates that a link occurs with the imbalance of the numbers of stocks into the FF legs of the value factor. In this section, N_S corresponds to the short leg of the factor, i.e., the growth portfolios, while N_L corresponds to the long leg of the factor, i.e., the value portfolios. The outperformance between the two factors is closely related to the imbalance between the long and short legs of the FF value factor: when N_S/N_L is superior to 1, the higher the ratio is, the higher the performance of the dependent sorting method;

[11] In the post-1973/NASDAQ period, for Model 4, the annualized mean return of the spread of 8.70% predominantly comes from shorting growth stocks (-0.08%) rather than buying value stocks (8.61%), while in Model 5, the short position in growth stocks yields an average excess return of 1.70% while buying value stocks delivers an average excess return of 8.20%. The results are available upon request.

Electronic copy available at: https://ssrn.com/abstract=3179655

when N_S/N_L is inferior to 1, we observe the expected inverse relationship.[12] This finding supports our previous evidence that underperformance occurs when the ratio deviates from 1. The ratio of N_S/N_L with a dependent and symmetrical sort is always close to 1, as expected.

Figure 4

Relationship Between the Information Ratio of the Value Factors and the Number of Stocks

The figures present the relationship between the information ratio of the value (HML) factors and the ratio of the number of stocks present in the short and long legs of the factors. Figure 4a shows the relationship for the ratio of the number of stocks present in the short and long legs of the independent value factor, while Figure 4b shows the results for the dependent value factor. The sample period ranges from July 1963 to December 2014.

Figure 4a: Independent Sort

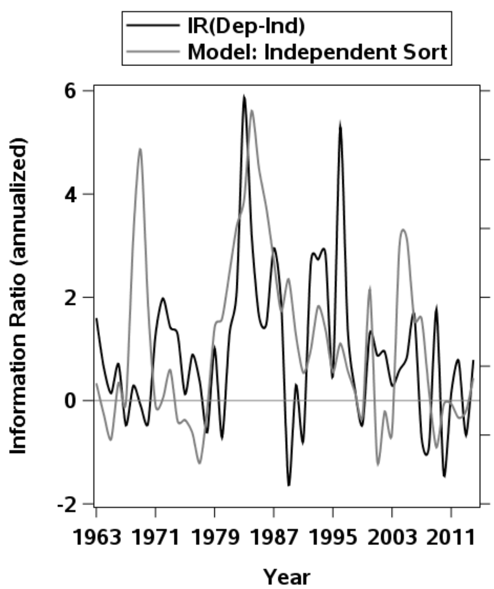
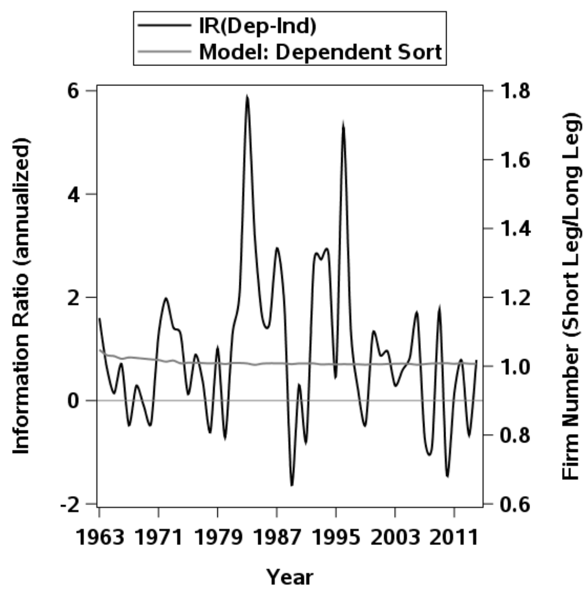


Figure 4b: Dependent Sort



---

[12] As one of the referees correctly points this is not true at the beginning of the period. We acknowledge that in 1969, there is almost a nil correlation between the size and value rankings due to an important decline in firm market capitalization overall (bear market). The important value of N_S/N_L in the independent framework is therefore not the result of correlation issues in rankings but only a consequence of the sharp rise of in the number of small cap firms. In this very particular context, the performances of the two sorting frameworks are very similar.

Electronic copy available at: https://ssrn.com/abstract=3179655

We then perform the following regression analysis to understand the source of out- and sub-performance,

$$IR_y = \alpha + \sum_{i=1}^{k} \beta^i \; C_y^i + \varepsilon_y \tag{12}$$

where C_y^i represents the control variables. First, we control for macroeconomic information with the cyclically adjusted price earnings ratio (CAPE), the consumer price index (CPI), and the long-term interest rate (Rate GS10); these three macro variables are retrieved from Robert Shiller's website. Second, we construct variables to control for market concentration with the diversity measure (Diversity), the total market size (TME) and the total level of book-to-market ratio (TBM), all of which are based on the NYSE-AMEX-NASDAQ market. These three variables are computed as of the portfolio formation date (beginning of July of each year *y*), and the diversity measure is defined as

$$Diversity(w) = \left[\sum_{i=i}^{n} w_i^p\right]^{\frac{1}{p}} \tag{13}$$

with w_i representing the market capitalization weight for the i^{th} stock and *p* set equal to ½.[13]

Third, we control for the time-series innovations for these variables as explanatory variables and use the notation Δ to differentiate the variables from their innovations.

Fourth, we control for the diversification of the constituents of the portfolio spread by using the variation in the diversity measure and total number of firms in the portfolios formed in the independent (2x3) and dependent (3x3x3) sorts, which are denoted as $\sigma_{diversiy}^{Ind}$, $\sigma_{diversity}^{Dep}$, σ_{Firm}^{Ind}

[13] The diversity measure was originally used in the study of Fernholz et al. (1998). The measure is also closely related to the reciprocal of the Herfindahl Index.

Electronic copy available at: https://ssrn.com/abstract=3179655

and σ_{Firm}^{Dep}. Finally, we look at the ratio N_S/N_L for the FF portfolios, i.e., the quotient between the average number of firms in the short leg portfolios and that in the long leg portfolios.

Table 4 presents the results of the regression analysis from equation (12) for the IR computed on the value (HML) factors. For clarity, we do not display the variables CPI, Rate GS10, Diversity, ΔDiversity, TME, ΔTME, TBM, ΔTBM, σ_{Firm}^{Ind} and σ_{Firm}^{Dep}, which have been tested but do not impact the temporal evolution of the IR for the dependent and independent value factors.

[Table 4 near here]

Model 1 of Table 4 shows the significant outperformance of the dependent compared to the independent sorting procedures for the value premium. Model 2 demonstrates that there is a significant relationship between the absolute value of the ratio (N_S/N_L-1) and the outperformance of the dependent sorting procedure, which establishes a link between the imbalance in terms of number of stocks in the long and short legs of the spread and the performance of the factor (as also depicted in Figure 4a). Model 3 shows that an equal stratification of stocks across portfolios matters for the performance of the factors, with a higher inequality of the FF allocation corresponding to a higher performance of the dependent sorting. Finally, Model 4 shows the importance of the macroeconomic variables for the outperformance of the dependent framework. As expected, the level of equity market multiple (CAPE) is a negative determinant of the outperformance of the dependent framework, which is inconsistent with the independent framework, whose breakpoints are driven by a momentum effect.

Table 5 examines the determinants of the ratio N_S/N_L. The imbalance in the number of stocks between the short and long legs for the value stock defined under an independent sorting on NYSE breakpoints will increase in periods characterized by a high multiple of market equity.

[Table 5 near here]

Electronic copy available at: https://ssrn.com/abstract=3179655

Our results show that the stock imbalance between the long and short legs depends upon the levels of the equity multiple, diversity measure and interest rates.

In the FF framework, an increase in the stock market valuation will lead to an incorrect allocation of value stocks into growth stocks and to an overweighting of the number of stocks in the growth portfolios. This finding is consistent with Chan et al. (2009) and further investigated in Figure 5 from Section 4. This evidence explains the positive relationship between N_S/N_L and the stock market valuation when N_S/N_L is greater than 1 (Model 2). However, an imbalance towards value stocks is expected to be found under a stock market with low valuation, which might drive up the book-to-market ratios of firms. Table 5 also confirms the sensitivity of the number of stocks in portfolios to market conditions.

3.4. Bias 2: Overestimation of the Size Effect

Table 6 shows that any change brought to the original size factor construction methodology leads to negative factor returns. The biggest effect is found when jointly considering name breakpoints, symmetric sorting and dependent sorting.

[Table 6 near here]

Using a similar procedure as for the value factors (see equations (10) and (11)), we compute the time series of information ratios for the size factors. The average IR of the size factor over the period is -0.83, with a standard deviation of 1.31, a minimum value of -4.01 and a maximum value of 2.94. Only 21% of the 52 yearly observations are positive. The DSN size factor delivers significantly lower risk-adjusted returns compared to the original dependent framework.

We investigate further the implications of each configuration in Table 7, which reports the *t*-statistics of strategies that are only invested in the size premium during one particular month of the year.

Electronic copy available at: https://ssrn.com/abstract=3179655

[Table 7 near here]

A change of the factor construction to name breakpoints and symmetric splits evidence relates the size premium to a beginning of the year calendar effect. This finding is consistent with the studies of Reinganum (1981), Roll (1981), and Keim (1983), who already claimed that the size premium was primarily driven by a "turn-of-the-year effect."

4. Understanding the Effects of Sorting Breakpoints

In this section, we present the incremental effect of the breakpoints on the performance of the factor. Moreover, we provide an empirical proof that the outperformance described in previous sections is not driven by small/micro stocks.

4.1. Name Breakpoints vs. NYSE Breakpoints

The traditional 2x3 independent sort of Fama and French (1993) is performed using NYSE breakpoints. Figure 5 shows that the breakpoints used for book-to-market characteristics are almost unchanged across the sample period (1963, 1994, 2001 and 2014). However, the breakpoints for market capitalization vary widely under changing market conditions. The NYSE size breakpoints increase in favorable market conditions, which induces a market effect in the Fama and French (1993) size premium and a consequent reversal in the HML effect. Sorting stocks according to the breakpoints defined on the entire sample introduces relatively resilient allocation keys into portfolios. Note that under this construction, NASDAQ stocks are largely represented in the small-cap portfolios and represent the main risk dynamics of this subportfolio.

Panels A-C-E-G (resp. Panels B-D-F-H) display the yearly values of the NYSE (resp. name) breakpoints for market capitalization and book-to-market under a 2x3 independent sorting of stocks. Panels G and H illustrate the momentum/market effect induced in the portfolios sorted using the NYSE breakpoints. To be included in a large-cap portfolio, a given stock needs to be

Electronic copy available at: https://ssrn.com/abstract=3179655

above the threshold defined by the current market conditions. The definition of large caps is much more stable across time using whole sample breakpoints. The performance of the size factor under the FF framework will be affected by an increase in the stock market valuation. In these specific periods, the long leg of the size factor will mix small and medium-to-large capitalizations and will be driven by a momentum effect from the medium-to-large cap. Our interpretation is that the DSN method, which includes Dependent, Name or Symmetric breakpoints, can control for these effects and disentangle the size effect from the momentum (and business cycles) effects.

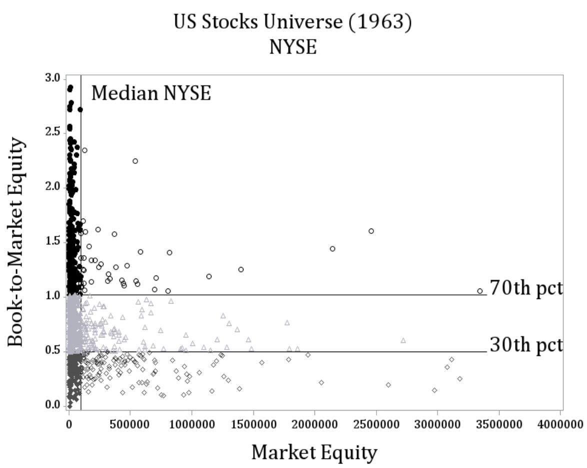
Electronic copy available at: https://ssrn.com/abstract=3179655

Figure 5

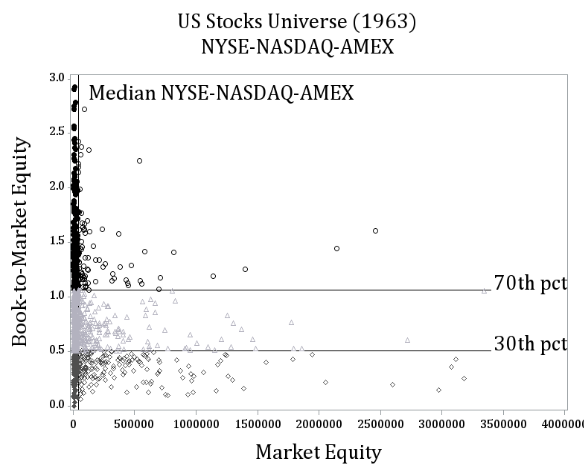
NYSE *vs.* Name Breakpoints

Figures A to H report the stratification of the US stock universe in size (small and big) and book-to-market (low, medium and high) buckets in 1963, 1994, 2001, and 2014. The *x*-axis refers to the market equity and the *y*-axis refers to the book-to-market equity. The panels on the left use the NYSE breakpoints, whereas the panels on the right use the whole sample to estimate the breakpoints. For better clarity of the breakpoints, outliers are not reported, the *x*-axis is capped between 0 and $4,000 billion, and the *y*-axis is truncated between 0 and 3[14].

Panel A



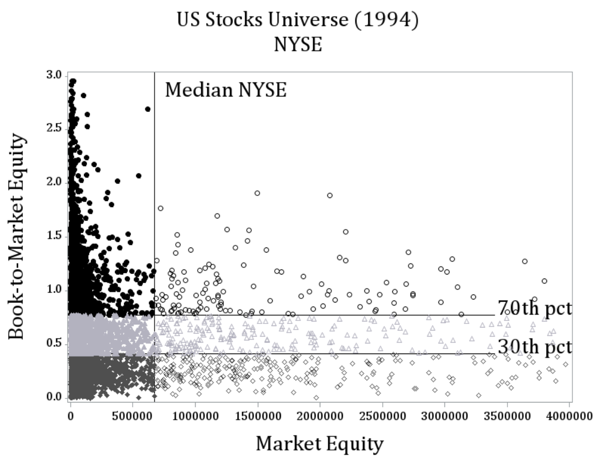
Panel B



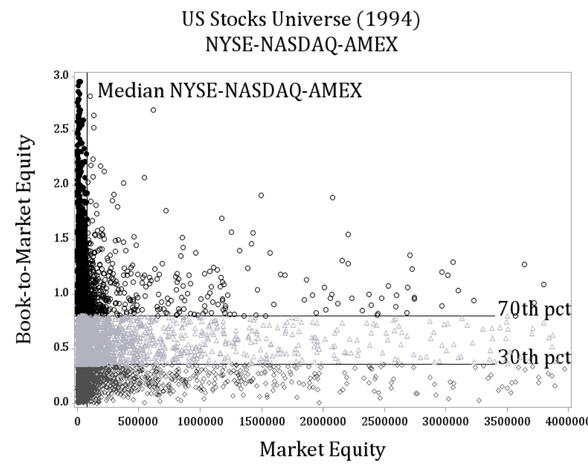
[14] The exercise could also be performed without truncating axes on a log-scale, which leads to equivalent interpretations.

Electronic copy available at: https://ssrn.com/abstract=3179655

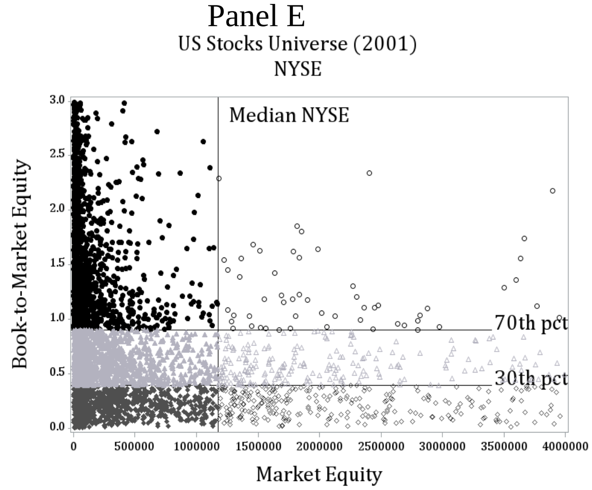
Panel C



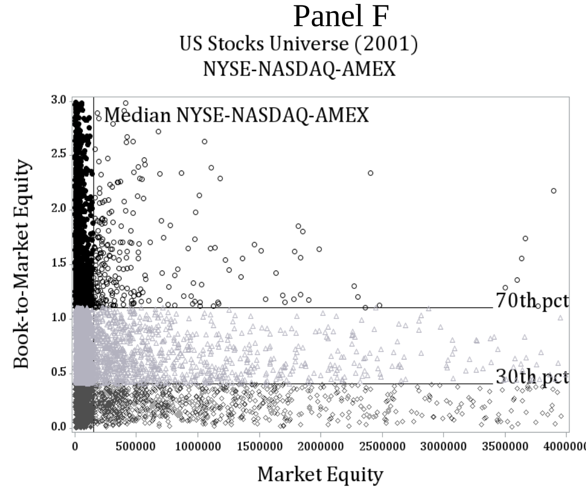
Panel D



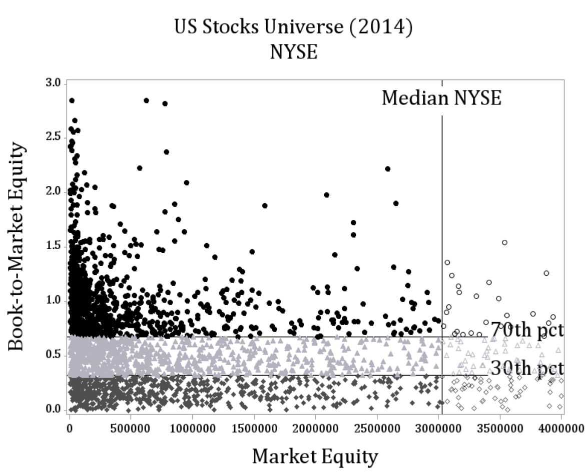
Panel E



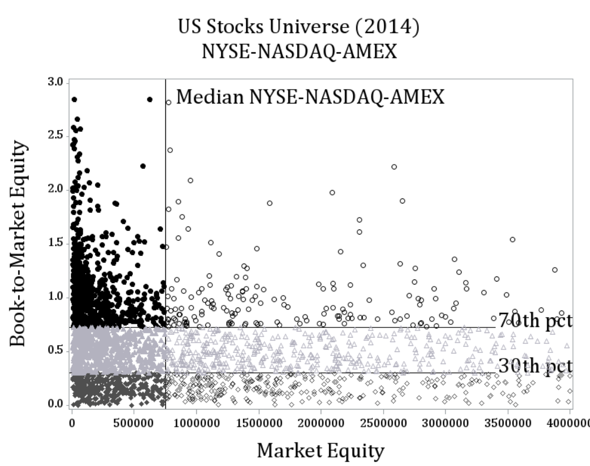
Panel F



Panel G



Panel H



Electronic copy available at: https://ssrn.com/abstract=3179655

4.2. Name Breakpoints vs. NYSE Breakpoints under a Sequential Framework

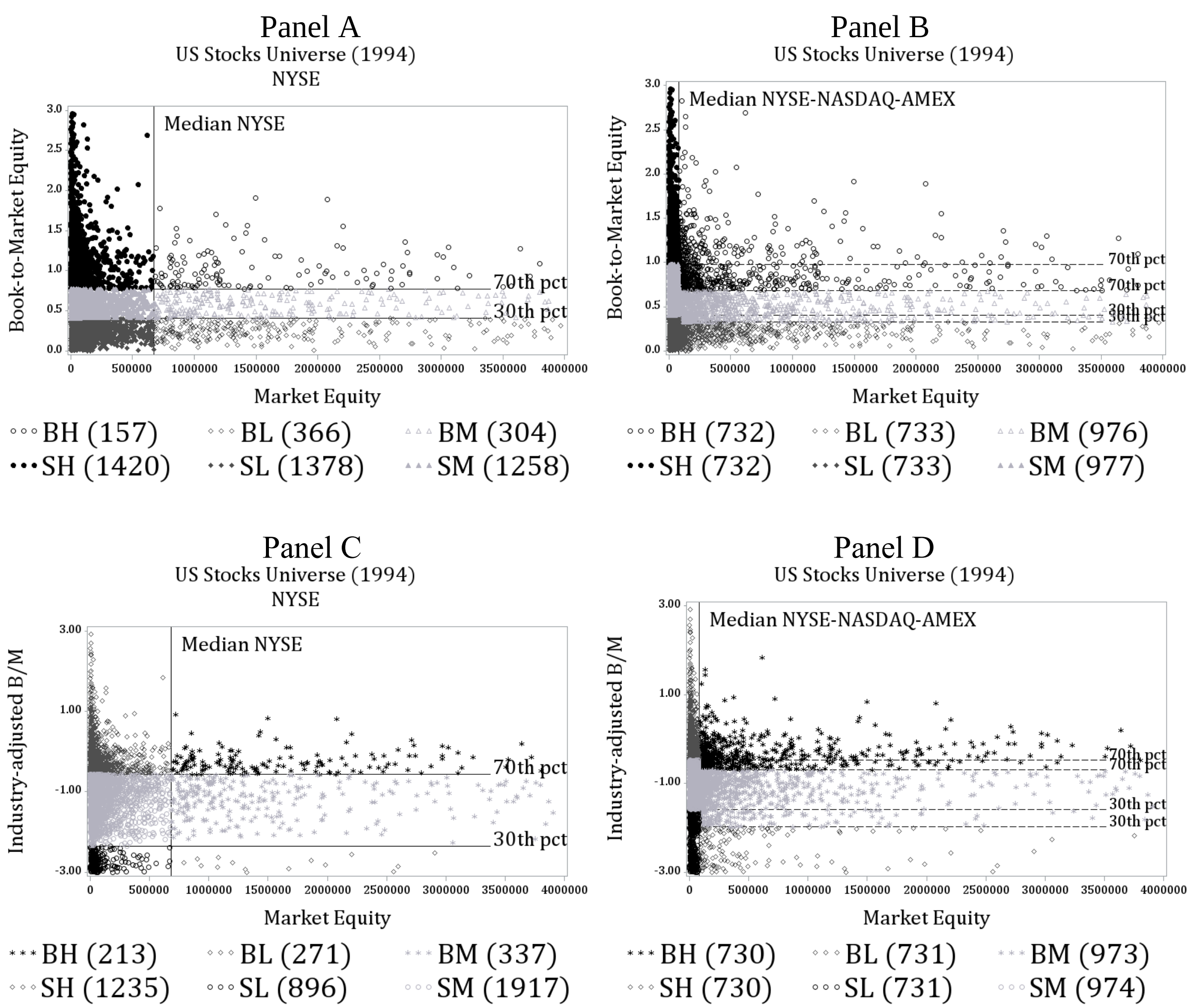
One frequently cited reason for using NYSE breakpoints is that this approach places more stocks in the low size portfolios to capture a higher percentage of the small capitalization universe in that portfolio. A whole sample approach takes another perspective by having an exogenous definition of a small stock and a classification independent of current market conditions, which might induce various levels of capitalization across portfolios. More specifically, a NYSE framework seeks a balance between the different portfolios (small and large portfolios) based on the total market capitalization included in each portfolio. However, combining the use of whole sample breakpoints with a sequential framework seeks to create a balance between portfolios based on the number of stocks. Consequently, a 2x3 independent sort will induce an imbalance in the number of stocks in portfolios to counter the capitalization effect while the use of whole sample breakpoints under a sequential sort would create an imbalance in market capitalization but the same allocation in terms of number of stocks. By applying the method based on size and book-to-market dimensions, many stocks fall into the small-value corner (as noted by Cremers et al. (2012)) under an independent 2x3 sort. This classification bias might have unpleasant practical consequences. For instance, Chan et al. (2009, p. 4579) indicate that "*many of the stocks that a large-value manager would hold in practice are classified as large-growth stocks under an independent sort procedure*". They also note that this effect is more pronounced at the end of the 1990s. Figure 6 illustrates the cross-effects of name versus NYSE breakpoints with the dependency of the sort on the allocation of stocks into portfolios, and it also considers the effect of industry-adjusted breakpoints.

Electronic copy available at: https://ssrn.com/abstract=3179655

Figure 6

Relative Stock Distribution Under Independent *vs* Dependent Sorts and Industry-Adjusted Breakpoints

Figures A to D report the stratification of the US stock universe in size (small and big) and book-to-market (low, medium and high) buckets in 1994. The *x*-axis refers to the market equity and the *y*-axis to the book-to-market equity. Panel A uses NYSE breakpoints and independent sorting, whereas Panel B uses name breakpoints together with a dependent sorting. Panel C (resp. D) replicates Panel A (resp. B) using industry-adjusted breakpoints as defined in Daniel et al. (2018a). We report in parentheses the number of stocks falling into the portfolios. To clarify the breakpoints, outliers are not reported, the *x*-axis is capped between 0 and $4,000 billion, and the *y*-axis is truncated between 0 and 3 (-3 and 3) for Panels A and B (C and D).



Electronic copy available at: https://ssrn.com/abstract=3179655

In Panel A, which presents the independent sorting with NYSE breakpoints, the number of stocks is 157 in the large-value portfolio and 366 in the large-growth portfolio, thus supporting the observations of Chan et al. (2009). However, we do not observe such a discrepancy under a dependent framework, which allocates 732 stocks into the large-value portfolio and 733 into the large-growth portfolio (see Panel B). Following Daniel et al. (2018a), we replicate the analysis using industry-adjusted breakpoints. Comparing the number of stocks into portfolio of Panels A and C, we observe a strong impact of industry-adjusted breakpoints for the original framework made of independent, asymmetric and NYSE breakpoints. We do not observe such big changes in the DSN framework. For instance, the FF method classifies 1378 stocks into the small growth portfolios using NYSE breakpoints. This number is reduced to 896 using industry-adjusted NYSE breakpoints. Under the DSN framework, the changes are marginal at 733 versus 731 stocks.

Finally, we investigate the distribution of value stocks among the 48 industries of the Standard Classification Industry (SIC) using a Fama and French (1993) approach or a DSN approach (3x3x3). The stratification in industries of value stocks over the full sample period almost does not differ between the two approaches for 47 out of 48 industries.[15] However, we observe a big difference for the utility sector. The weight of utility stocks in value portfolios is approximately 14.2% in the original Fama and French (1993) framework compared to 8.6% in the DSN framework. This finding suggests that the DSN framework controls for industry exposure and leads to a reduced concentration compared with that in the original framework. The results are non-tabulated but available upon request.

[15] The weights differ by less than 1% for 40 industries and between 1% and 2% for 7 other industries. The results are non-tabulated but available upon request.

Electronic copy available at: https://ssrn.com/abstract=3179655

5. Mean-variance Spanning Tests

We perform spanning tests of the size and value factors when extended to 3x3 or 3x3x3 and/or name breakpoints as well as the original configuration: (i) on a set of various basis portfolios and (ii) with regard to competing multifactor models.

5.1. Spanning Test on Basis Portfolios

We first test whether the DSN factors can span a variety of basis portfolios. We consider four sets of test assets: (1) the FF 25 portfolios formed on size and book-to-market (5x5), (2) the DSN version of the 25 portfolios formed on size and then book-to-market (5x5), (3) the DSN version of the 25 portfolios formed on book-to-market and then size (5x5), and (4) the 12 industry portfolios.[16]

Our spanning test follows the "sequential test" of Harvey and Liu (2019)[17], which evaluates the incremental contribution of a set of candidate factors towards a baseline model for explaining the cross-section of expected returns of test assets. The baseline model contains the value-weighted market return in excess of the risk-free rate (Rm-Rf) and the momentum factor (WML)[18], which are both obtained from Kenneth French's website. The candidate factors are size (SMB) and value (HML) factors constructed according to different sorting methods.

Harvey and Liu (2019) define a scaled intercept (SI) as

[16] The portfolios from specification (1) and (4) are obtained from Kenneth French's website.

[17] The sequential test offers several advantages with regard to a more classical GRS test. We refer the reader to Harvey and Liu (2019) for further information.

[18] The four-factor Carhart (1997) model is still largely used in the literature for modeling stock returns. It has also been intensively used as a basis for a substantial number of extensions (e.g., Fama and French (2018)).

Electronic copy available at: https://ssrn.com/abstract=3179655

$$SI = \frac{median(\{|a_i^g|/s_i^b\}_{i=1}^J - \{|a_i^b|/s_i^b\}_{i=1}^J)}{median(\{|a_i^b|/s_i^b\}_{i=1}^J)} \tag{14}$$

where $median(.)$ is the median value of the ratio $|a_i^g|/s_i^b$ or $|a_i^b|/s_i^b$, the superscript b is for the baseline model, the superscript g is for the augmented model, the subscript i refers to the *i-th* portfolio among the J test assets, and s denotes the standard errors for the regression intercept a.

A significant reduction of the absolute intercept when adding a candidate factor to a baseline will lead to a negative value of SI, which can be interpreted as the candidate factor adding incremental explanatory power to the baseline model. The statistical significance of the measure is defined with bootstrap simulations that define a level of confidence for the measure and control for multiple testing.

[Table 8 near here]

Table 8 presents the first sequence[19] of the test of Harvey and Liu (2019), and it identifies the most appropriate sorting configuration to construct a set of size and value factors that spans the test assets. The results show that the DSN (3x3x3) size and value factors bring the highest explanatory power to the baseline model when pricing the FF 25 portfolios formed on size and book-to-market (5x5). When considering the DSN test assets, a dependent sort leads to the highest incremental contribution. For the industry portfolios, however, none of the settings improve the explanatory power of the baseline model as evidenced by the multiple test p-value of 58.1%.

[19] In our test, the multiple test p-value of the second sequence is always greater than 10 percent and consequently do not report these results.

Electronic copy available at: https://ssrn.com/abstract=3179655

5.2. Spanning Test on Other Multifactor Models

We test whether the modified factors provide alpha with regard to the Carhart four factors as well as with regard to the *q*-factor model of Hou et al. (2019) or the Stambaugh-Yuan (2017) model. For that, we perform the F-test (F_1) from Kan and Zhou (2012), which consists of a test of whether $\alpha = 0$ in the following regression

$$R_2^t = \alpha + \beta R_1^t + \varepsilon^t \quad (15)$$

where R_1^t is the K-factor model against which R_2^t is tested. Here, R_1^t is Carhart four factors while R_2^t corresponds to one candidate among the set of size and value factors.

The F-test is written as

$$F_1 = \left(\frac{T - K - N}{N}\right)\left(\frac{\hat{a} - \hat{a}_1}{1 + \hat{a}_1}\right) \quad (16)$$

where *T* is the number of observations, *K* is the number of benchmark assets, *N* is the number of test assets, and $\hat{a}_1 = \hat{\mu}_1{}' \hat{V}_{11}^{-1} \hat{\mu}_1$ is the SSR of the benchmark assets, with $\hat{V}_{11}$ denoting the variance-covariance matrix and $\hat{\mu}_1$ representing the vector of mean return of the benchmark assets. In addition, $\hat{a}$ takes the same notation as $\hat{a}_1$ but refers to the benchmark assets (R_1^t) plus the new test asset (R_2^t).

The mean-variance spanning test is first performed with regard to the Carhart four factor model. Table 9 presents the results.

[Table 9 near here]

The results from Panel A of Table 9 show that when extended to 3x3 or 3x3x3, the modified size and value factors constructed using an independent sort on NYSE breakpoints do not show significant additional return relative to the existing Carhart four factors. However, any extension of the value factor (2x3, 3x3 or 3x3x3) to name breakpoints add monthly return of approximately 0.23%-0.29% to the original Carhart four factors.

Electronic copy available at: https://ssrn.com/abstract=3179655

Table 10 replicates the spanning test with regard to the Hou et al. (2015a) *q*-Factor (as the baseline model). The factors, i.e., SMB, IA, and ROE, are based on triple independent sort, and HML is not considered. These factors have been shown in the literature to subsume the Fama and French (1993) three-factor model.

[Table 10 near here]

The table reports the significance of the SMB factor notwithstanding the configuration. As already shown by Hou et al. (2015b), the value factor of Fama and French (1993) is subsumed by their model. However, considering the following changes in the sorting procedure, either a combination of name and symmetric sorting or a full DSN sorting, the value factor gain significance and provides significant alphas on the *q*-factor model.

Finally, we consider in Table 11 how our factors compete with the management (MNGMT) and performance factors of Stambaugh-Yuan (2017). Both relate to profitability and investment measures, select different quantiles (20/80) to form the usual 30/70 and use "Name" breakpoints (NYSE-AMEX-NASDAQ, instead of NYSE only).

[Table 11 near here]

Table 11 shows that any configuration of the size factor is redundant to the Stambaugh and Yuan (2017) model. However, the value factor designed either by combining symmetric sort with name breakpoints or using a D(S)N procedure is not subsumed by the model.

6. Robustness Tests

We check the robustness of our results with regard to (i) the use of equally weighted portfolios rather than value-weighted portfolios, (ii) the influence of micro-caps, (iii) the use of different breakpoints, and (iv) the potential extension of the DSN method to other factors.

Electronic copy available at: https://ssrn.com/abstract=3179655

6.1. Equally Weighted Basis Portfolios

To test the resilience of the DSN method, we re-construct the size and value factors using equally weighted portfolios for different sorting configurations. We replicate a mean-variance spanning test in Table 9. This test will be informative about the performance of the size and value factors in the presence of micro-caps because they receive higher weight in this setting. The results are displayed at Table 12.

[Table 12 near here]

Our previous results hold when all factors are defined using equally weighted portfolios. Indeed, regarding the value factors, configurations using name breakpoints (jointly or not with dependent sorting) provide positive risk-adjusted return (alpha) to the original equally weighted configuration of the factors (i.e., independent, 2x3 and NYSE breakpoints). The alpha is the highest when combining a dependent sorting with name breakpoints. Please note that the use of NYSE breakpoints with an independent sort in the presence of micro-caps would result in inefficient factors. Name breakpoints (jointly or not with a dependent sorting) provide abnormal returns to the original 4 factor Carhart model in an equally weighted setting. The highest improvement is found for the DSN setting.

In Table 13, we replicate Table 3 for the HML factor using equally weighted basis portfolios.

[Table 13 near here]

Comparing Table 13 to Table 3, we observe that all *t*-stats are inflated in an equally weighted setting with regard to a value-weighted setting. However, the increase is quite moderate in the dependent symmetric sort (*t*-stats change from 5.51 to 7.99 in Model 6 and from 5.63 to 8.26

Electronic copy available at: https://ssrn.com/abstract=3179655

in Model 7) compared to the independent framework (*t*-stats changes from 3.85 to 7.49 in Model 1, i.e., almost double, as also shown in Hou et al. (2019)).

6.2. Micro-caps

To test whether our results are the result of a dependence on micro-cap stocks, we replicate the mean-variance spanning test of Table 9 but excluding stocks with market capitalization below the 20th percentile based on the NYSE size breakpoints.[20] Table 14 presents the results.

[Table 14 near here]

Although the incremental return of the value factor using alternative sorting rules is slightly reduced, the results are still positive and significant at the 5% confidence level. Especially, the DSN 3x3 or 3x3x3 still provides incremental returns of 21.5 bps and 17 bps, respectively (with regard to 0.26 and 0.24 bps at Table 9). As in Table 9, the size factors display no incremental power on the Carhart 4-factor original model. The alphas are all negative but not significant except for one configuration. Please note that micro-caps were not excluded from the dependent test assets, which might constitute an unfair testing framework and explain the negative (yet insignificant) coefficients.

6.3. Breakpoints

We also test the robustness of our DSN methodology for various breakpoints. We use as a baseline our 3x3x3 approach in a pre-conditioning setting and perform a mean-variance spanning test similar to Table 9. Table 15 reports the results.

[Table 15 near here]

[20] This threshold is quite conservative as it eliminates much more stocks than if one had used name breakpoints.

Electronic copy available at: https://ssrn.com/abstract=3179655

Our results are robust to various thresholds, namely 30-66, 30-70, and 20-80. However, the 10-90 split for value factors under a DSN framework is not delivering the same outperformance. This sort gives such a large weight on micro-caps that the only incremental effect that stays significant is coming from the DSN size factor. The outperformance might therefore be inflated by a micro-cap factor.

6.4. Applying the DSN Method to Investment and Profitability Factors

We construct the investment (CMA) and profitability (RMWo) factors of Fama and French (2015) as well as the profitability (RMWgp) factor of Novy-Marx (2013) in an independent and a DSN (3x3x3) framework. We use the size and value characteristics as control variables in that specific order for the DSN sorting. Table 16 presents summary statistics for each pair of factors.

[Table 16 near here]

The investment factor displays a significantly higher *t*-stat (6.68) under a DSN sorting compared to the original framework (3.95). Similar improvements are found for the profitability factor of Novy-Marx (2013) (*t*-stat improved from 2.01 to 5.98 under a DSN framework). Note that the profitability factor of Fama and French (2015) displays *t*-statistics of the same magnitude under the DSN or the original settings. The factor is robust in a DSN setting.

7. Conclusions

The correct choice of risk factors in an asset pricing model is necessary but not sufficient to obtain meaningful empirical evidence. The quality of the factor construction methodology can substantially alter asset pricing tests. If the model adequately reflects the underlying drivers of systematic risk priced on the market, then sorting stocks on specific characteristics (such as size and value) is equivalent to sorting them on the basis of their expected returns. We claim that

Electronic copy available at: https://ssrn.com/abstract=3179655

performing a naive portfolio sort to form a portfolio spread, such as the SMB or HML premia, can lead to the definition of biased risk factors regardless of the characteristics. We measure the magnitude of the bias in the case of highly correlated variables and show that under an independent sort, the error increases with the symmetry of the multiple sort and the use of whole sample breakpoints. The bias becomes nonsignificant when performing a dependent and symmetric sort on characteristics. We show that the consequences of the sorting framework for the mean-variance performance of the size and book-to-market risk factors are statistically and economically very significant. By performing this in-depth analysis, our paper aims to fill a gap in the literature by providing both an empirical and a theoretical framework for factoring characteristics into returns. This finding is particularly important because it can lead to the acceptance of models that would otherwise be rejected because of poor measurement.

Our main result is that the sorting options, including the symmetry of the sort, its (in)dependence and the definition of breakpoints, affect the empirical performance risk factors. Factors formed on well-diversified and well-balanced (in terms of risk and number of stocks) constituent portfolios deliver better mean-variance properties. These properties, moreover, depend on economic conditions. Our results are grounded in asset pricing theory, which posits that the distance between the constructed benchmark and its true underlying factor increases with the poor allocation of stocks into the constituent portfolio (leading to poor diversification across the portfolio constituents and imbalance between the long and short legs) but decreases with the strength of the linear relationship between characteristics and returns.

The allocation into portfolios under an independent asymmetric sorting on NYSE breakpoints as in the FF framework (which has been the source of a multitude of applications over the last 25 years) renders the allocation of stocks into value or growth portfolios highly sensitive

Electronic copy available at: https://ssrn.com/abstract=3179655

to equity market diversity, equity multiples levels, market capitalization, interest rate and CPI. A dependent and symmetric sort, which is more resilient to market conditions, will as a consequence significantly outperform the original factor in times of low market capitalization and equity/consumer price. The level of valuation of the equity market (equity multiple) is also relevant, especially to the size factors that are inflated by an independent NYSE 2x3 sort. The DSN sorting method reveals a clear seasonal effect for the size factor and significantly reduce its significance. The sensible tie between our empirical results and theory is due to the number of stocks allocated into portfolios, in particular through the difference between the long and short legs of the spread.

Naturally, the evidence presented here is mainly limited and restricted to the FF-Carhart set of original factors, although those factors are very influential in the empirical asset pricing literature and often considered as benchmarks to evaluate sorting procedures (e.g., De Nard, Ledoit and Wolf (2019)). Beyond the original size-value-momentum four-factor model, our article paves the way for the systematic use of a dependent approach (preferably preconditioning on the control variable(s)) for the construction of spread portfolios that mimic multidimensional risk factors. The preliminary results show that a DSN method makes investment and profitability factors even stronger. It also shows that the newly constructed size and value factors are not subsumed by factors recently discovered in the literature. The empirical asset pricing literature has indeed witnessed a multiplication of *K*-factor models rooted in the FF tradition, such as the extended 5-factor model (Fama and French (2016)), the *q*-factor model (Hou et al. 2015a), and the recent mispricing factors (Stambaugh and Yuan (2017)). Additional research is needed to revisit the significance of the large set of factors using an alternative sorting procedure. Another important question is whether more accurate portfolio construction processes could lead to greater parsimony

Electronic copy available at: https://ssrn.com/abstract=3179655

in the design of factor models. Our methodological discussion could contribute to answering these important questions. These research directions occupy a prominent position in our future research agenda.

Electronic copy available at: https://ssrn.com/abstract=3179655

References

Agarwal, V., G. S. Bakshi, and J. Huij. 2009. Do higher-moment equity risks explain hedge fund returns? *Robert H. Smith School Research Paper No. RHS 06-153*

Ang, A., R. J. Hodrick, Y. Xing, and X. Zhang. 2006. The cross-section of volatility and expected returns. *The Journal of Finance* 61:259-99.

Barillas, F., and J. Shanken. 2018. Comparing Asset Pricing Models. *The Journal of Finance* 73 (2): 715–54.

Bowles, B., A. V. Reed, M. C. Ringgenberg, and J. R. Thornock. 2017. *Anomaly time*. Working Paper. *SSRN Electronic Journal*.

Carhart, M. M. 1997. On persistence in mutual fund performance. *The Journal of Finance* 52:57-82.

Chan, L. K. C., S. G. Dimmock, and J. Lakonishok. 2009. Performance: issues and evidence. *Review of Financial Studies* 22:4554-99.

Cochrane, J. H. 2011. Presidential address: discount rates. *The Journal of Finance* 66:1047-108.

Cremers, M., A. Petajisto, and E. Zitzewitz. 2012. Should benchmark indices have alpha? Revisiting performance evaluation. *Critical Finance Review* 2:1-48.

Daniel, K., M. Grinblatt, S. Titman, and R. Wermers. 1997. Measuring mutual fund performance with characteristic-based benchmarks. *The Journal of Finance* 52:1035-58.

Daniel, K., and S. Titman. 1998. Characteristics or covariances? *Journal of Portfolio Management* 24:24-33.

Daniel, K., L. Mota, S. Rottke, and T. Santos. 2018a. The cross-section of risk and return. Working Paper. *SSRN Electronic Journal*.

Electronic copy available at: https://ssrn.com/abstract=3179655

Daniel, K., D. Hirshleifer, and L. Sun. 2018b. Short- and Long-Horizon Behavioral Factors. Working Paper. *SSRN Electronic Journal*.

De Nard, G., O. Ledoit, and M. Wolf. 2019. Factor models for Portfolio selection in large dimensions: The good, the better and the ugly. *Journal of Financial Econometrics*, forthcoming.

Engelberg, J., R. D. Mclean, and J. Pontiff. 2018. Anomalies and news. *SSRN Electronic Journal*.

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.

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

Fama, E. F., and K. R. French. 2016. Dissecting anomalies with a five-factor model. *Review of Financial Studies* 29:69-103.

Fama, E. F., and K. R. French. 2018. Choosing Factors. *Journal of Financial Economics* 128 (2): 234–52.

Fernholz, R., R. Garvy, and J. Hannon. 1998. Diversity-weighted indexing. *Journal of Portfolio Management* 24:74–82.

Green, J., J. R. M. Hand, and X. F. Zhang. 2017. The characteristics that provide independent information about average U. S. Monthly stock returns. *Review of Financial Studies* 30: 4389-436.

Grinblatt, M. and K. Saxena. 2019. When Factors Don't Span Their Basis Portfolios. *Journal of Financial and Quantitative Analysis* 53 (6):2335-2354.

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

Electronic copy available at: https://ssrn.com/abstract=3179655

Harvey, C. R., and Y. Liu, 2019. Lucky Factors. Working Paper.

Hou, K., C. Xue, and L. Zhang. 2015a. Digesting anomalies: an investment approach. *Review of Financial Studies* 28:650-705.

Hou, K., C. Xue, and L. Zhang. 2015b. A comparison of new factor models. NBER Working Paper No. 20682. *SSRN Electronic Journal*.

Hou, K., C. Xue, and L. Zhang. 2018. Replicating anomalies. *Review of Financial Studies* hhy131:1-115.

Hou, K., H. Mo, C. Xue, and L. Zhang. 2019. Which factors? *Review of Finance* 23:1-35.

Kan, R., and G. Zhou. 2012. Tests of Mean-Variance Spanning. *Annals of Economics & Finance* 13:139–187.

Keim, D. B., 1983. Size-related anomalies and stock return seasonality: further empirical evidence, *Journal of Financial Economics* 12:13–32.

Kogan, L., and M. Tian. 2015. Firm characteristics and empirical factor models: a model-mining experiment. Working Paper.

Lambert, M., and G. Hübner. 2013. Comoment risk and stock returns. *Journal of Empirical Finance* 23: 191–205.

Lehmann, B.N., and D.M. Modest. 2005. Diversification and the Optimal Construction of Basis Portfolios. *Management Science* 51:581–598.

Li, X., J. Miffre, C. Brooks, and N. O'Sullivan. 2008. Momentum profits and time-varying unsystematic risk. *Journal of Banking and Finance* 32:541–558.

Liew, J., and M. Vassalou. 2000. Can book-to-market, size and momentum be risk factors that predict economic growth? *Journal of Financial Economics* 57:221-45.

Electronic copy available at: https://ssrn.com/abstract=3179655

Moller, N., and S. Zilca. 2008. The evolution of the January effect. *Journal of Banking and Finance* 32:447–457.

Novy-Marx, R. 2013. The other side of value: the gross profitability premium. *Journal of Financial Economics* 108:1-28.

Reinganum, M. R., 1981. Misspecification of asset pricing: empirical anomalies based on earnings' yields and market values, *Journal of Financial Economics* 9:19–46.

Roll, R., 1981. A possible explanation of the small firm effect, *Journal of Finance* 36:879–888.

Savor, P., and M. Wilson. 2016. Earnings announcements and systematic risk. *The Journal of Finance* 71:83–138.

Stambaugh, R. F., and Y. Yuan. 2017. Mispricing factors. *Review of Financial Studies* 30:1270-315.

Treynor, J. L., F. Black. 1973. How to use security analysis to improve portfolio selection. *Journal of Business* 46:66-86.

van Dijk, M. A., 2011. Is size dead? A review of the size effect in equity returns. *Journal of Banking and Finance* 35:3263–3274.

Wermers, R. 2004. Is money really smart? New evidence on the relation between mutual fund flows, manager behavior, and performance persistence. Working Paper. *SSRN Electronic Journal*.

Electronic copy available at: https://ssrn.com/abstract=3179655

Table 1

Stock Distribution among Portfolios sorted on Correlated Characteristics

Panels A to H report the stock distribution for portfolios sorted on market capitalization (ME) and (–ME) using a double sorting. The scale on the first sort on ME ranges from S to B (small (S), medium (M), and big (B)), and the scale of the second sort on –ME from L to H (low (L), medium (M), and high (H)). We report the mean, standard deviation (SD), minimum and maximum of the stocks in each portfolio. The sample period is July 1963 to December 2014.

Panel A: Independent - 2x3 - NYSE									
	SL	**SM**	**SH**	**BL**	**BM**	**BH**			
Mean	0	416	2226	394	309	0			
SD	0	131	907	98	85	0			
Min	0	133	439	182	124	0			
Max	0	696	4085	638	525	0			
Panel B: Independent - 2x3 - Name									
	SL	**SM**	**SH**	**BL**	**BM**	**BH**			
Mean	0	671	977	1021	674	0			
SD	0	241	349	367	243	0			
Min	0	177	255	269	177	0			
Max	0	1152	1726	1727	1151	0			
Panel C: Independent - 3x3 - NYSE									
	SL	**SM**	**SH**	**BL**	**BM**	**BH**	**ML**	**MM**	**MH**
Mean	0	1	2226	393	0	0	1	724	0
SD	0	0	907	98	0	0	0	214	0
Min	0	1	439	181	0	0	1	256	0
Max	0	2	4085	637	0	0	1	1220	0
Panel D: Independent - 3x3 - Name									
	SL	**SM**	**SH**	**BL**	**BM**	**BH**	**ML**	**MM**	**MH**
Mean	0	1	977	1021	0	0	1	1345	0
SD	0	0	349	367	0	0	0	483	0
Min	0	1	255	268	0	0	1	353	0
Max	0	1	1726	1726	0	0	1	2302	0
Panel E: Dependent - 2x3 - NYSE									
Portfolios (first sort on the variable used for the HML factor, i.e., -ME)									
	LS	**MS**	**HS**	**LB**	**MB**	**HB**			
Mean	196	362	1097	197	363	1129			
SD	49	107	447	49	107	461			
Min	92	129	217	90	128	222			
Max	319	611	2043	319	610	2042			
Portfolios (first sort on the variable used for the SMB factor, i.e., ME)									
	SL	**BL**	**SM**	**BM**	**SH**	**BH**			
Mean	806	213	1066	281	770	209			
SD	315	55	417	72	301	54			
Min	177	92	230	123	165	91			
Max	1414	349	1885	466	1413	348			

Electronic copy available at: https://ssrn.com/abstract=3179655

Panel F: Dependent - 2x3 - Name						
Portfolios (first sort on the variable used for the HML factor, i.e., -ME)						
	LS	**MS**	**HL**	**LB**	**MB**	**HB**
Mean	509	671	477	513	675	500
SD	182	241	170	184	243	180
Min	135	176	127	134	178	128
Max	864	1152	863	863	1151	863
Portfolios (first sort on the variable used for the SMB factor, i.e., ME)						
	SL	**BL**	**SM**	**BM**	**SH**	**BH**
Mean	504	513	668	677	477	506
SD	180	184	240	244	170	182
Min	133	134	173	178	126	134
Max	864	864	1151	1151	863	863

Panel G: Dependent - 3x3 - NYSE									
Portfolios (first sort on the variable used for the HML factor, i.e., -ME)									
	LS	**MS**	**HS**	**LB**	**MB**	**HB**	**LM**	**MM**	**HM**
Mean	127	266	1489	111	178	245	156	280	492
SD	33	88	612	26	48	95	40	82	221
Min	57	81	262	53	76	66	72	100	111
Max	201	445	2821	183	313	440	254	463	927
Portfolios (first sort on the variable used for the SMB factor, i.e., ME)									
	SL	**ML**	**BL**	**SM**	**MM**	**BM**	**SH**	**MH**	**BH**
Mean	245	179	112	492	280	156	1489	266	126
SD	95	48	26	220	82	40	613	88	33
Min	66	77	54	113	100	71	261	80	56
Max	441	314	183	920	463	254	2825	444	200

Panel H: Dependent - 3x3 - Name									
Portfolios (first sort on the variable used for the HML factor, i.e., -ME)									
	LS	**MS**	**HS**	**LM**	**MM**	**HM**	**LB**	**MB**	**HB**
Mean	305	402	280	408	538	396	309	405	300
SD	109	145	100	146	193	143	111	146	108
Min	81	105	76	108	143	102	80	106	77
Max	519	691	518	690	922	691	518	690	517
Portfolios (first sort on the variable used for the SMB factor, i.e., ME)									
	SL	**ML**	**BL**	**SH**	**MH**	**BH**	**SM**	**MM**	**BM**
Mean	302	406	309	280	402	304	397	538	407
SD	108	146	111	100	145	109	143	193	146
Min	78	107	80	76	104	80	102	143	108
Max	519	691	518	518	690	517	690	922	691

Electronic copy available at: https://ssrn.com/abstract=3179655

Table 2

Measuring Bias in the Factor Construction Method

Panels A through H display summary statistics for the size premia constructed using different configurations. We construct a small minus big (SMB) factor based on a sorting on market capitalization (ME) and a high minus low (HML) factor based on a sorting on "minus" market capitalization (-ME). Portfolios are value-weighted. We report the mean, standard deviation (SD), Sharpe ratio (SR), sum, minimum, maximum, and *t*-statistics of the factors' returns as well as the correlations between the SMB and HML factors. *, **, and *** indicate statistical significance at the 0.1, 0.05 and 0.01 levels, respectively. The sample period is July 1963 to December 2014.

Factor construction	# Obs	Mean	SD	SR	Sum	Min	Max	*t*-stat	Correlation Matrix	
Panel A: Independent - 2x3 - NYSE										
SMB	618	0.136	1.651	0.08	83.954	-10.004	9.998	2.045	1	
HML	618	0.164	1.962	0.08	101.336	-9.331	12.65	2.078	0.98041***	1
Panel B: Independent - 2x3 - Name										
SMB	618	0.151	2.102	0.07	93.178	-8.786	14.983	1.783	1	
HML	618	0.204	2.411	0.08	126.171	-8.586	14.941	2.105	0.97861***	1
Panel C: Independent - 3x3 - NYSE										
SMB	618	0.400	4.105	0.10	247.04	-16.863	19.502	2.421	1	
HML	618	-0.244	3.248	-0.08	-150.934	-18.457	16.306	-1.869	-0.0184	1
Panel D: Independent - 3x3 - Name										
SMB	618	0.705	7.601	0.09	435.411	-21.778	77.421	2.304	1	
HML	618	-0.097	4.112	-0.02	-59.800	-26.556	16.624	-0.585	0.0579	1
Panel E: Dependent - 2x3 - NYSE										
SMB	618	0.138	1.264	0.11	85.005	-4.122	7.008	2.706	1	
HML	618	0.193	2.351	0.08	119.105	-7.380	11.086	2.038	0.88938***	1
Panel F: Dependent - 2x3 - Name										
SMB	618	0.117	1.367	0.09	72.300	-6.708	8.798	2.128	1	
HML	618	0.231	2.36	0.10	142.796	-6.815	11.168	2.434	0.8237***	1
Panel G: Dependent - 3x3 - NYSE										
SMB	618	0.104	1.476	0.07	63.960	-6.986	9.285	1.743	1	
HML	618	0.104	1.478	0.07	64.238	-7.090	9.274	1.748	0.9993***	1
Panel H: Dependent - 3x3 - Name										
SMB	618	0.19	1.765	0.11	117.177	-7.679	11.161	2.670	1	
HML	618	0.193	1.763	0.11	119.366	-7.685	11.091	2.724	0.99953***	1

Electronic copy available at: https://ssrn.com/abstract=3179655

Table 3

Value Factors: Summary Statistics

The table shows the annualized arithmetic mean, volatility (SD) (both in %) as well as the Sharpe Ratio (SR), Squared Sharpe Ratio (SSR) and *t*-stats of the value factors constructed under seven sorting configurations (Models 1 to 7) using value-weighted basis portfolios. We consider a full sample period of daily returns from the 1st of July 1963 to the 31st December 2014 and two sub-periods corresponding to pre- and post-NASDAQ introduction in 1973.

	Model 1: Independent, Asymmetric (2x3), NYSE-BPs	Model 2: Independent, Asymmetric (2x3), Name-BPs	Model 3: Independent, Symmetric (3x3), NYSE-BPs	Model 4: Independent, Symmetric (3x3), Name-BPs	Model 5: Dependent, Asymmetric (2x3), Name-BPs	Model 6: Dependent, Symmetric (3x3), Name-BPs	Model 7: Dependent, Symmetric (3x3x3), Name-BPs
Full Sample							
Mean	4.55	7.69	4.53	8.45	6.12	6.86	6.31
SD	8.40	8.35	8.45	9.03	8.50	8.76	7.91
SR	0.54	0.92	0.54	0.94	0.72	0.78	0.80
SSR	0.29	0.85	0.29	0.88	0.52	0.61	0.64
t-stat	3.85	6.46	3.81	6.54	5.09	5.51	5.63
Pre-1973							
Mean	4.74	5.42	4.63	6.92	4.28	5.06	5.23
SD	5.82	6.82	5.73	6.99	6.79	6.76	6.23
SR	0.81	0.79	0.81	0.99	0.63	0.75	0.84
SSR	0.66	0.63	0.65	0.98	0.40	0.56	0.70
t-stat	2.51	2.43	2.48	3.01	1.94	2.30	2.58
Post-1973							
Mean	4.28	8.12	4.22	8.70	6.41	7.20	6.40
SD	8.93	8.69	9.00	9.49	8.87	9.20	8.30
SR	0.48	0.93	0.47	0.92	0.72	0.78	0.77
SSR	0.23	0.87	0.22	0.84	0.52	0.61	0.59
t-stat	3.05	5.82	2.98	5.71	4.54	4.90	4.85

Electronic copy available at: https://ssrn.com/abstract=3179655

Table 4

Determinants of the Information Ratio on the Value Factor

The table presents the coefficients of an OLS regression (*t*-statistics of the coefficients are reported in parentheses). The dependent variable is the information ratio of the DSN value (HML) factor over its FF equivalent. The independent variables are *macroeconomic variables* (namely (i) the cyclically adjusted price earnings ratio (CAPE), (ii) the consumer price index (CPI), and (iii) the long-term interest rate (Rate GS10)[21]) as well as *specific portfolio variables* (namely, (i) the dispersion in the diversity measure of the portfolio constituents in respectively the independent (2x3) and dependent (3x3x3) sorts, $\sigma_{diversity}^{Ind}$, and $\sigma_{diversity}^{Dep}$, where diversity is measured at the portfolio formation date (beginning of July of each year *y*) according to equation (13); (ii) the ratio N_S/N_L among the FF portfolios, i.e., the quotient between the average number of firms in the short (S) and long (L) leg portfolios; and (iii) the absolute value of the ratio (N_S/N_L -1)). *, **, and *** indicate statistical significance at the 0.1, 0.05 and 0.01 levels, respectively. The sample period is July 1963 - December 2014.

Model	(1)	(2)	(3)	(4)
Intercept	0.932***	-1.271	-1.779*	0.247
	(4.501)	(-1.457)	(-1.815)	(0.198)
CAPE				-0.059*
				(-1.832)
ΔCPI				-21.885***
				(-2.716)
ΔRate GS10				0.183
				(0.166)
$\sigma_{diversity}^{Ind}$			0.019**	0.021**
			(2.389)	(2.593)
$\sigma_{diversity}^{Dep}$			-0.138**	-0.147**
			(-2.085)	(-2.086)
$\left\|\frac{N_S}{N_L} - 1\right\|^{FF}$		2.065**	1.940**	1.714**
		(2.591)	(2.550)	(2.236)
R²	0.000	0.120	0.237	0.359
R²-Adj	0.000	0.103	0.188	0.272

[21] All variables are obtained from Robert Shiller's website.

Electronic copy available at: https://ssrn.com/abstract=3179655

Table 5

Determinants of the Ratio N_S/N_L of the Fama-French Value Factor

This table presents the results of an OLS regression. The dependent variable is the ratio N_S/N_L across the FF portfolios, i.e., the quotient between the average number of firms in the short leg (S) portfolios and that in the long leg (L) portfolios. The independent variables are (i) the cyclically adjusted price earnings ratio (CAPE), (ii) the long-term interest rate (Rate GS10), which are all obtained from Robert Shiller's website, and (iv) the diversity measure (Diversity) for the US markets (NYSE-AMEX-NASDAQ) computed using equation (13) as well as the cross-sectional volatility the diversity measures of the portfolios formed in the independent (2x3) sort is denoted $\sigma_{diversity}^{Ind}$. Diversity measures are computed at the portfolio formation date (beginning of July of each year *y*). Model (1) uses the ratio N_S/N_L as a dependent variable; model (2) shows the results when the ratio N_S/N_L is higher than 1, and model (3) presents the results when N_S/N_L is lower than 1. *, **, and *** indicate statistical significance at the 0.1, 0.05 and 0.01 levels, respectively. The *t*-statistics of the coefficients are reported in parentheses. The sample period ranges from July 1963 to December 2014.

Model	(1)	(2)	(3)
Intercept	0.084	-1.995***	2.080***
	(0.449)	(-3.904)	(5.613)
CAPE	0.020***	0.047***	-0.027**
	(3.549)	(3.079)	(-2.452)
Rate GS10	0.074***	0.184***	-0.110***
	(5.687)	(5.213)	(-4.314)
Diversity	0.001***	0.002***	-0.001***
	(3.963)	(3.784)	(-3.214)
$\sigma_{diversity}^{Ind}$	-0.003***	-0.008***	0.005***
	(-4.169)	(-3.684)	(2.973)
R^2	0.511	0.490	0.414
R^2-Adj	0.470	0.446	0.365

Electronic copy available at: https://ssrn.com/abstract=3179655

Table 6

Size Factors: Summary Statistics

The table shows the annualized arithmetic mean, volatility (SD) (both in %) as well as the Sharpe Ratio, Square Sharpe Ratio (SSR) and *t*-stats of the size factors constructed under seven sorting configurations (Models 1 to 7) using value-weighted basis portfolios. We consider a full sample period of daily returns from the 1st of July 1963 to the 31st December 2014 and two sub-periods corresponding to pre- and post-NASDAQ introduction in 1973.

	Model 1: Independent, Asymmetric (2x3), NYSE-BPs	Model 2: Independent, Asymmetric (2x3), Name-BPs	Model 3: Independent, Symmetric (3x3), NYSE-BPs	Model 4: Independent, Symmetric (3x3), Name-BPs	Model 5: Dependent, Asymmetric (2x3), Name-BPs	Model 6: Dependent, Symmetric (3x3), Name-BPs	Model 7: Dependent, Symmetric (3x3x3), Name-BPs
Full Sample							
Mean	0.65	-3.66	-0.04	-7.73	-4.14	-8.29	-6.32
SD	8.85	10.63	10.72	13.84	10.22	12.75	12.34
SR	0.07	-0.34	0.00	-0.56	-0.41	-0.65	-0.51
SSR	0.01	0.12	0.00	0.31	0.16	0.42	0.26
t-stat	0.53	-2.55	-0.03	-4.23	-3.01	-4.94	-3.85
Pre-1973							
Mean	-0.25	-1.39	-0.59	-2.70	-2.57	-3.20	-2.02
SD	5.70	7.28	7.41	9.44	6.76	8.97	8.00
SR	-0.04	-0.19	-0.08	-0.29	-0.38	-0.36	-0.25
SSR	0.00	0.04	0.01	0.08	0.14	0.13	0.06
t-stat	-0.14	-0.60	-0.25	-0.91	-1.21	-1.14	-0.80
Post-1973							
Mean	0.89	-4.19	0.16	-8.85	-4.45	-9.43	-7.24
SD	9.48	11.32	11.41	14.75	10.91	13.55	13.20
SR	0.09	-0.37	0.01	-0.60	-0.41	-0.70	-0.55
SSR	0.01	0.14	0.00	0.36	0.17	0.48	0.30
t-stat	0.61	-2.45	0.09	-4.08	-2.70	-4.74	-3.69

Electronic copy available at: https://ssrn.com/abstract=3179655

Table 7

"Calendar effect" and the Size Premium

T-statistics on the daily size premium are presented per month and per sorting method. The sample period ranges from 1st of July 1963 to 31st December 2014.

Month	Model 1: Independent, Asymmetric (2x3), NYSE-BPs	Model 2: Independent, Asymmetric (2x3), Name-BPs	Model 3: Independent, Symmetric (3x3), NYSE-BPs	Model 4: Independent, Symmetric (3x3), Name-BPs	Model 5: Dependent, Symmetric (3x3x3), Name-BPs
January	4.00	5.10	6.66	8.37	6.97
February	2.27	2.07	1.24	1.22	1.36
March	-0.11	-0.20	-0.81	-0.70	-0.45
April	-0.71	-1.10	-1.93	-1.69	-2.19
May	-0.25	-0.27	-0.44	-1.02	0.17
June	2.14	2.01	0.80	-0.06	-0.19
July	-1.51	-1.59	-1.16	-0.93	-0.42
August	-1.21	-2.16	-3.66	-4.48	-4.59
September	0.27	0.14	-0.91	-1.82	-1.24
October	-3.04	-3.43	-4.29	-5.00	-4.52
November	-0.18	-0.67	-2.65	-3.97	-3.79
December	1.73	1.33	-1.29	-3.52	-3.29

Electronic copy available at: https://ssrn.com/abstract=3179655

Table 8

Spanning Basis Portfolios: The Sequential Test of Harvey and Liu (2019)

The table reports the sequential test developed by Harvey and Liu (2019). It evaluates the incremental contribution of a set of candidate factors (SMB and HML constructed according to the rules described Columns (1) to (3)) to a baseline model for explaining the cross-section of expected returns of test assets. The table displays the value of the scaled intercept (SI) over the first sequence of the test as well as the p-value of the selected size and value candidates (bottom line) when controlling for multiple testing. Test assets are as follows: (1) the FF 5x5 Size and Book-to-Market portfolios, (2) the DSN 5x5 Size and (then) Book-to-Market portfolios, (3) the DSN 5x5 Book-to-Market and (then) Size portfolios, and (4) the 12 Industry Portfolios. The sample period ranges from July 1963 to December 2014.

Baseline Model: Rm-Rf + WML						
Sort	Splits	Breakpoints	SI (Scaled Intercept)			
			(1)	(2)	(3)	(4)
Independent	2x3	NYSE	-0.805	-0.742	-0.752	-0.100
Independent	3x3	NYSE	-0.785	-0.716	-0.742	-0.051
Independent	3x3x3	NYSE	-0.614	-0.570	-0.687	**-0.212**
Independent	2x3	Name	-0.825	-0.715	-0.707	0.079
Independent	3x3	Name	-0.782	-0.727	-0.759	-0.008
Independent	3x3x3	Name	-0.788	-0.624	-0.680	0.094
S-Post	2x3	NYSE	-0.801	-0.662	-0.749	-0.184
S-Post	3x3	NYSE	-0.779	-0.683	-0.789	-0.086
S-Post	3x3x3	NYSE	-0.762	-0.655	-0.783	-0.190
S-Post	2x3	Name	-0.679	-0.739	**-0.814**	0.043
S-Post	3x3	Name	-0.690	-0.680	-0.758	-0.179
S-Post	3x3x3	Name	-0.711	-0.718	-0.796	-0.164
S-Pre	2x3	NYSE	-0.761	-0.675	-0.765	-0.204
S-Pre	3x3	NYSE	-0.776	-0.705	-0.775	-0.140
S-Pre	3x3x3	NYSE	-0.735	-0.682	-0.728	-0.169
S-Pre	2x3	Name	-0.788	-0.756	-0.767	0.098
S-Pre	3x3	Name	-0.799	**-0.826**	-0.803	0.099
S-Pre	3x3x3	Name	**-0.827**	-0.694	-0.781	0.018
multiple test p-value			0.006	0.000	0.000	0.581

Electronic copy available at: https://ssrn.com/abstract=3179655

Table 9

Mean-Variance Spanning Test

The table presents the results of the step-down mean-variance spanning test from Kan and Zhou (2012) (MATLAB code available on Prof. Zhou's website). We report the intercept (α) from the OLS-regression model (monthly and in percent) as well as the F-test (F_1), which tests the null hypothesis that the test asset (R_2^t) does not improve the ex-post tangency portfolio composed of the benchmark assets (R_1^t). Here, R_1^t is the Carhart (1997) four-factor model, i.e., Rm-Rf, SMB, HML, and WML, obtained from Kenneth French's website. R_2^t corresponds to one candidate among the set of size and value factors. SMB and HML are constructed according to the rules described Columns (1) to (3). Basis portfolios are value-weighted. We report in bold the *p*-values that are rejected with a confidence interval of 90%. The sample period ranges from July 1963 to December 2014.

Sort	Splits	Breakpoints	α (in %)	F_1	p-val	α (in %)	F_1	p-val
			Panel A: All sample					
			SMB Factors			HML Factors		
Independent	3x3	NYSE	-0.001	0.000	0.986	0.036	2.163	0.142
Independent	3x3x3	NYSE	0.001	0.001	0.982	0.006	0.019	0.890
Independent	2x3	Name	-0.041	0.331	0.565	0.226	23.545	**0.000**
Independent	3x3	Name	0.047	0.172	0.679	0.285	34.349	**0.000**
Independent	3x3x3	Name	0.078	0.458	0.499	0.285	33.453	**0.000**
S-Post	2x3	NYSE	0.016	1.028	0.311	0.036	1.551	0.213
S-Post	3x3	NYSE	0.003	0.008	0.927	0.071	6.411	**0.012**
S-Post	3x3x3	NYSE	0.013	0.126	0.722	0.065	3.901	**0.049**
S-Post	2x3	Name	-0.025	0.133	0.716	0.332	38.387	**0.000**
S-Post	3x3	Name	0.089	0.628	0.428	0.402	53.614	**0.000**
S-Post	3x3x3	Name	0.096	0.669	0.414	0.370	41.232	**0.000**
S-Pre	2x3	NYSE	0.017	0.783	0.376	0.007	0.059	0.808
S-Pre	3x3	NYSE	-0.002	0.006	0.936	0.043	2.033	0.154
S-Pre	3x3x3	NYSE	0.003	0.006	0.940	0.022	0.483	0.487
S-Pre	2x3	Name	-0.018	0.077	0.782	0.196	16.587	**0.000**
S-Pre	3x3	Name	0.014	0.017	0.896	0.261	29.779	**0.000**
S-Pre	3x3x3	Name	0.075	0.506	0.477	0.236	22.950	**0.000**

Electronic copy available at: https://ssrn.com/abstract=3179655

Table 10
Mean-Variance Spanning Test: Stambaugh-Yuan (2017) Four-Factor model

The table presents the results of the step-down mean-variance spanning test from Kan and Zhou (2012) (MATLAB code available on Prof. Zhou's website). We report the intercept (α) from the OLS-regression model (monthly and in percent) as well as the F-test (F_1), which tests the null hypothesis that the test asset (R_2^t) does not improve the ex-post tangency portfolio composed of the benchmark assets (R_1^t). Here, R_1^t is the Stambaugh-Yuan (2017) Four-Factor model, i.e., Rm-Rf, SMB, MNGMT, and PERF, obtained from Robert Stambaugh's website. R_2^t corresponds to one candidate among the set of size and value factors. SMB and HML are constructed according to the rules described Columns (1) to (3). Basis portfolios are value-weighted. We report in bold the *p*-values that are rejected with a confidence interval of 90%. The sample period ranges from July 1963 to December 2014.

Sort	Splits	Breakpoints	α (in %)	F_1	p-val	α (in %)	F_1	p-val
			SMB Factors			HML Factors		
Independent	2x3	NYSE	-0.039	0.662	0.416	-0.003	0.002	0.969
Independent	3x3	NYSE	-0.040	0.343	0.558	-0.012	0.021	0.886
Independent	3x3x3	NYSE	-0.058	0.591	0.442	0.010	0.014	0.907
Independent	2x3	Name	-0.022	0.045	0.832	0.146	2.471	0.117
Independent	3x3	Name	0.121	0.721	0.396	0.200	4.992	**0.026**
Independent	3x3x3	Name	0.118	0.665	0.415	0.219	6.865	**0.009**
S-Post	2x3	NYSE	-0.041	0.800	0.371	0.027	0.110	0.740
S-Post	3x3	NYSE	-0.054	0.699	0.403	0.032	0.162	0.688
S-Post	3x3x3	NYSE	-0.057	0.695	0.405	0.044	0.314	0.575
S-Post	2x3	Name	-0.036	0.134	0.714	0.315	11.254	**0.001**
S-Post	3x3	Name	0.130	0.908	0.341	0.327	12.697	**0.000**
S-Post	3x3x3	Name	0.115	0.652	0.420	0.283	9.446	**0.002**
S-Pre	2x3	NYSE	-0.016	0.111	0.739	0.024	0.093	0.761
S-Pre	3x3	NYSE	-0.042	0.419	0.518	0.044	0.300	0.584
S-Pre	3x3x3	NYSE	-0.050	0.553	0.457	0.040	0.291	0.590
S-Pre	2x3	Name	0.010	0.010	0.920	0.155	2.882	**0.090**
S-Pre	3x3	Name	0.094	0.475	0.491	0.213	6.436	**0.011**
S-Pre	3x3x3	Name	0.148	1.263	0.262	0.174	4.587	**0.033**

Electronic copy available at: https://ssrn.com/abstract=3179655

Table 11

Mean-Variance Spanning Test: Hou, Xue, and Zhang (2015) *q*-Factor model

The table presents the results of the step-down regression-based mean-variance spanning test from Kan and Zhou (2012), for which the MATLAB code is made available on Prof. Zhou's website. We display the results for the set of size (SMB) and value (HML) factors. We report the intercept (α) from the OLS-regression model (monthly and in percent) as well as the F-test (F_1), which tests the null hypothesis that the test asset (R_2^t) does not improve the ex-post tangency portfolio composed of the benchmark assets (R_1^t). Here, R_1^t is the Hou, Xue, and Zhang (2015) q-Factor model, i.e., Rm-Rf, SMB, IA, and ROE, obtained from http://global-q.org/. R_2^t corresponds to one candidate among the set of size and value factors. Basis portfolios are value-weighted. We report in bold the *p*-values that are rejected with a confidence interval of 90%. The sample period ranges from January 1967 to December 2014.

Sort	Splits	Breakpoints	α (in %)	F_1	p-val	α (in %)	F_1	p-val
			SMB Factors			HML Factors		
Independent	2x3	NYSE	0.087	5.507	**0.019**	0.035	0.134	0.714
Independent	3x3	NYSE	0.153	7.263	**0.007**	0.015	0.024	0.878
Independent	3x3x3	NYSE	0.163	6.706	**0.010**	0.061	0.370	0.543
Independent	2x3	Name	0.206	4.999	**0.026**	0.176	2.825	**0.093**
Independent	3x3	Name	0.358	6.749	**0.010**	0.248	5.927	**0.015**
Independent	3x3x3	Name	0.384	7.567	**0.006**	0.257	7.548	**0.006**
S-Post	2x3	NYSE	0.106	9.448	**0.002**	0.114	1.613	0.205
S-Post	3x3	NYSE	0.150	8.226	**0.004**	0.133	2.211	0.138
S-Post	3x3x3	NYSE	0.169	9.831	**0.002**	0.159	3.366	**0.067**
S-Post	2x3	Name	0.201	5.631	**0.018**	0.401	15.907	**0.000**
S-Post	3x3	Name	0.377	8.313	**0.004**	0.426	18.032	**0.000**
S-Post	3x3x3	Name	0.378	7.860	**0.005**	0.387	14.508	**0.000**
S-Pre	2x3	NYSE	0.093	5.359	**0.021**	0.038	0.173	0.678
S-Pre	3x3	NYSE	0.128	5.538	**0.019**	0.043	0.230	0.631
S-Pre	3x3x3	NYSE	0.124	4.496	**0.034**	0.031	0.142	0.706
S-Pre	2x3	Name	0.212	6.098	**0.014**	0.137	1.761	0.185
S-Pre	3x3	Name	0.339	6.828	**0.009**	0.229	5.557	**0.019**
S-Pre	3x3x3	Name	0.353	8.081	**0.005**	0.180	4.352	**0.037**

Electronic copy available at: https://ssrn.com/abstract=3179655

Table 12

Mean-Variance Spanning Test: Equal-Weighted Portfolios

The table presents the results for the step-down mean-variance spanning test from Kan and Zhou (2012) (MATLAB code available on Prof. Zhou's website). We report the intercept (α) from the OLS-regression model (monthly and in percent) as well as the F-test (F_1) which tests the null hypothesis that the test asset (R_2^t) do not improve the ex-post tangency portfolio composed of the benchmark assets (R_1^t). Here, R_1^t is the Carhart (1997) four-factor model, i.e., Rm-Rf, SMB, HML, and WML, obtained from Kenneth French's website. R_2^t corresponds to one candidate among the set of size and value factors. SMB and HML are constructed according to the rules described Columns (1) to (3). Basis portfolios are equally weighted. We report in bold the *p*-values that are rejected with a confidence interval of 90%. The sample period ranges from July 1963 to December 2014.

Sort	Splits	Breakpoints	α (in %)	F_1	p-val	α (in %)	F_1	p-val
			SMB Factors			HML Factors		
Independent	3x3	NYSE	0.026	1.000	0.318	-0.046	3.025	**0.083**
Independent	3x3x3	NYSE	0.047	1.818	0.178	-0.120	5.971	**0.015**
Independent	2x3	Name	0.096	2.989	**0.084**	0.103	7.356	**0.007**
Independent	3x3	Name	0.196	6.851	**0.009**	0.081	4.261	**0.039**
Independent	3x3x3	Name	0.192	5.840	**0.016**	-0.008	0.031	0.860
S-Post	2x3	NYSE	-0.029	2.102	0.148	0.158	17.076	**0.000**
S-Post	3x3	NYSE	0.030	0.982	0.322	-0.031	1.150	0.284
S-Post	3x3x3	NYSE	0.044	2.089	0.149	-0.006	0.033	0.856
S-Post	2x3	Name	0.066	1.503	0.221	0.217	28.459	**0.000**
S-Post	3x3	Name	0.167	5.279	**0.022**	0.217	30.058	**0.000**
S-Post	3x3x3	Name	0.190	6.563	**0.011**	0.206	24.503	**0.000**
S-Pre	2x3	NYSE	0.109	4.688	**0.031**	-0.002	0.004	0.951
S-Pre	3x3	NYSE	0.042	2.401	0.122	-0.035	1.489	0.223
S-Pre	3x3x3	NYSE	0.041	1.849	0.174	-0.019	0.384	0.536
S-Pre	2x3	Name	0.099	3.582	**0.059**	0.144	15.064	**0.000**
S-Pre	3x3	Name	0.206	8.602	**0.003**	0.099	7.728	**0.006**
S-Pre	3x3x3	Name	0.241	10.685	**0.001**	0.128	10.869	**0.001**

Electronic copy available at: https://ssrn.com/abstract=3179655

Table 13

Equally Weighted Value Factors: Summary Statistics

The table shows the annualized arithmetic mean, volatility (SD) (both in %) as well as the Sharpe Ratio (SR), Square Sharpe Ratio (SSR) and *t*-stats of the value factors constructed under seven sorting configurations (Models 1 to 7) using equally weighted basis portfolios. We consider a full sample period of daily returns from the 1st of July 1963 to the 31st December 2014 and two sub-periods corresponding to pre- and post-NASDAQ introduction in 1973.

	Model 1: Independent, Asymmetric (2x3), NYSE-BPs	Model 2: Independent, Asymmetric (2x3), Name-BPs	Model 3: Independent, Symmetric (3x3), NYSE-BPs	Model 4: Independent, Symmetric (3x3), Name-BPs	Model 5: Dependent, Asymmetric (2x3), Name-BPs	Model 6: Dependent, Symmetric (3x3), Name-BPs	Model 7: Dependent, Symmetric (3x3x3), Name-BPs
Full Sample							
Mean	8.07	9.27	6.66	7.98	9.83	8.31	7.64
Std	7.54	7.43	7.57	7.34	7.39	7.27	6.48
SR	1.07	1.25	0.88	1.09	1.33	1.14	1.18
SSR	1.14	1.56	0.77	1.18	1.77	1.31	1.39
t-stat	7.49	8.69	6.20	7.61	9.24	7.99	8.26
Pre-1973							
Mean	6.91	7.35	6.08	6.36	7.23	5.24	6.06
Std	5.05	6.02	5.13	6.01	5.99	5.93	5.48
SR	1.37	1.22	1.19	1.06	1.21	0.88	1.11
SSR	1.87	1.49	1.41	1.12	1.45	0.78	1.23
t-stat	4.16	3.71	3.62	3.22	3.66	2.71	3.38
Post-1973							
Mean	8.07	9.55	6.51	8.15	10.35	8.92	7.89
Std	8.04	7.73	8.08	7.64	7.67	7.55	6.71
SR	1.00	1.24	0.81	1.07	1.35	1.18	1.18
SSR	1.01	1.53	0.65	1.14	1.82	1.40	1.38
t-stat	6.26	7.66	5.07	6.65	8.33	7.34	7.34

Electronic copy available at: https://ssrn.com/abstract=3179655

Table 14

Mean-Variance Spanning Test

The table replicates the results displayed in Table 9 except that it excludes stocks belonging to the 20% quantile based on NYSE breakpoints. The sample period ranges from July 1963 to December 2014.

Sort	Splits	Breakpoints	α (in %)	F_1	p-val	α (in %)	F_1	p-val
			Replication of Table 9 without the 20% lowest cap stocks					
			SMB Factors			HML Factors		
Independent	3x3	NYSE	-0.015	0.283	0.595	0.029	2.401	0.122
Independent	3x3x3	NYSE	-0.010	0.067	0.796	0.000	0.000	0.992
Independent	2x3	Name	-0.068	1.868	0.172	0.130	12.432	**0.000**
Independent	3x3	Name	-0.084	1.039	0.309	0.216	24.661	**0.000**
Independent	3x3x3	Name	-0.042	0.225	0.635	0.081	4.103	**0.043**
S-Post	2x3	NYSE	-0.037	7.161	**0.008**	0.149	22.360	**0.000**
S-Post	3x3	NYSE	-0.009	0.079	0.779	0.050	4.621	**0.032**
S-Post	3x3x3	NYSE	-0.003	0.006	0.938	0.053	3.188	**0.075**
S-Post	2x3	Name	-0.064	1.805	0.180	0.180	26.128	**0.000**
S-Post	3x3	Name	-0.070	0.783	0.377	0.271	43.041	**0.000**
S-Post	3x3x3	Name	-0.049	0.365	0.546	0.245	34.409	**0.000**
S-Pre	2x3	NYSE	-0.020	0.216	0.642	0.017	0.377	0.539
S-Pre	3x3	NYSE	-0.018	0.544	0.461	0.031	1.056	0.305
S-Pre	3x3x3	NYSE	-0.012	0.132	0.716	0.013	0.160	0.689
S-Pre	2x3	Name	-0.061	1.869	0.172	0.114	8.347	**0.004**
S-Pre	3x3	Name	-0.071	0.858	0.355	0.215	24.851	**0.000**
S-Pre	3x3x3	Name	-0.066	0.664	0.415	0.169	17.773	**0.000**

Electronic copy available at: https://ssrn.com/abstract=3179655

Table 15

Mean-Variance Spanning Test: The Definition of Breakpoints

The table presents the results of the step-down mean-variance spanning test from Kan and Zhou (2012) (MATLAB code is made available on Prof. Zhou's website). We report the intercept (α) from the OLS-regression model (monthly and in percent) as well as the F-test (F_1), which tests the null hypothesis that the test asset (R_2^t) does not improve the ex-post tangency portfolio composed of the benchmark assets (R_1^t). Here, R_1^t is the Carhart (1997) four-factor model, i.e., Rm-Rf, SMB, HML, and WML, obtained from Kenneth French's website. R_2^t corresponds to one candidate among the set of size and value factors. SMB and HML are constructed according to the rules described Columns (1) to (3). Basis portfolios are value-weighted. We report in bold the *p*-values that are rejected with a confidence interval of 90%. The sample period ranges from July 1963 to December 2014.

Sort	Splits	Breakpoints	L - H	α (in %)	F_1	p-val	α (in %)	F_1	p-val
				SMB Factors			HML Factors		
S-Pre	3x3x3	Name	33 - 66	0.073	0.540	0.463	0.216	25.496	**0.000**
S-Pre	3x3x3	Name	30 - 70	0.075	0.506	0.477	0.236	22.950	**0.000**
S-Pre	3x3x3	Name	20 - 80	0.081	0.303	0.582	0.374	26.222	**0.000**
S-Pre	3x3x3	Name	10 - 90	0.598	6.054	**0.014**	0.246	1.409	0.236

Electronic copy available at: https://ssrn.com/abstract=3179655

Table 16

Investment (CMA) and Profitability (RMW) factors under the FF and DSN Methods

The table reports the monthly mean and volatility (SD) (both in percentages), t-statistics of the investment (CMA) and profitability (RMWo) factors of Fama and French (2015)[22] as well as the profitability (RMWgp) factor of Novy-Marx (2013)[23] derived from two different portfolio construction configurations: (i) the Fama and French (1993)'s independent and asymmetric (2x3) sort with NYSE breakpoints; (ii) a dependent, symmetric (3x3x3) sort with all-name breakpoints. The dependent sort first starts with the firm's size, then by the book-to-market equity ratio, and the final sort is on the price variable, i.e., investment for CMA and profitability for RMWo and RMWgp. The sample period is July 1963 - December 2014.

	CMA		RMWo		RMWgp	
	Independent, Asymmetric (2x3), NYSE-BPs	Dependent, Symmetric (3x3x3), Name-BPs	Independent, Asymmetric (2x3), NYSE-BPs	Dependent, Symmetric (3x3x3), Name-BPs	Independent, Asymmetric (2x3), NYSE-BPs	Dependent, Symmetric (3x3x3), Name-BPs
Mean	0.32	0.41	0.26	0.32	0.19	0.43
SD	2.02	1.52	2.22	2.92	2.31	1.77
# Obs	618	618	618	618	618	618
t-stat	3.95	6.68	2.96	2.71	2.01	5.98

[22] Investment is the change in total assets ($TA_t - TA_{t-1}$) over the last fiscal year divided by the total assets of the latest fiscal year (TA_{t-1}). Profitability variable in June of year t is the annual revenues (REVT) minus the cost of goods sold (COGS), interest expense (XINST+XINTD), and selling, general, and administrative expenses (XSGA) divided by the book equity (BE) for the last fiscal year-end in t-1.

[23] Profitability in June of year t is measured as the gross profits-to-assets (GP/AT), where gross profits are the annual revenues (REVT) minus the cost of goods sold (COGS).

Electronic copy available at: https://ssrn.com/abstract=3179655