# Looking for the Tangent Portfolio: Risk Optimization Techniques on Equity Style Buckets*

Boris Fays[1], Marie Lambert[1], and Nicolas Papageorgiou[2]

[1]HEC Liege, Management School of the University of Liege

[2]HEC Montreal

January 30, 2019

**Abstract**

The inflation of multifactor models has questioned the existence of a mean-variance efficient portfolio. We examine the risk-return properties of low-risk portfolios obtained by the risk optimization of characteristic-sorted equity portfolios. We group stocks based on size, value and momentum characteristics through several sorting procedures. We claim that low-risk portfolios constructed using long-only positions on basis portfolios formed using a dependent sorting scale with whole-sample breakpoints deliver superior Sharpe ratio to an optimal portfolio spanning the single or the 3-factor models. Besides, while the three-factor empirical model spans an opportunity set augmented with low-risk portfolios formed on other sorting procedures, the pricing performance of low-risk investment strategies based on these dependent basis portfolios is shown to be superior to that of the single-index factor model, the three-factor model and its competing low-risk benchmarks. Our testing framework is based on bootstrapped mean-variance spanning tests and shows valid conclusions out-of-sample and net of transaction costs. The results still hold when controlling for data snooping biases through multiple testing and luck. Economically, our results are supported by diversification-based properties.

*Boris Fays is a PhD candidate at HEC Liège, ULiège. Rue Louvrex 14, Bldg. N1, 4000 Liège, Belgium. E-mail: boris.fays@uliege.be. Phone number: +32 4 232 73 47. Marie Lambert is a Professor of Finance at HEC Liège, ULiège. She is also a Research Associate at EDHEC Risk Institute. Address: HEC Liège, Rue Louvrex 14, Bldg. N1, 4000 Liège, Belgium. E-mail: marie.lambert@uliege.be. Phone number: +32 4 232 74 32. Nicolas Papageorgiou is a Full Professor of Finance at HEC Montreal, and Vice-President and Head of R&D, Fiera Capital Corporation. Mailing address: 3000 Cote Ste Catherine, Montreal, Quebec. E-mail: Nicolas.Papageorgiou@hec.ca Phone number: +1 514 340 7309.

# I. Introduction

For more than fifty years, passive investors have considered capitalization-weighted (CW) indices to be a suitable proxy for the tangency portfolio, namely the maximum Sharpe ratio (MSR) portfolio. The inflation of multifactor models in the recent literature has raised the question of the existence of a mean-variance efficient (MVE) portfolio. If it exists, the MVE portfolio should span the available factor portfolios and display the MSR achievable with the factor portfolios alone (Daniel et al. (2017) and Grinblatt and Saxena (2018)). Upon a stratification of the equity universe into style buckets, we implement long-only, low-risk optimization techniques. We claim that the portfolios constructed this way exhibit interesting mean-variance properties and outperform the current CW indices used as a proxy for the market portfolio and even outperform a multifactor model based on the related long-short style factors.

Our research is grounded in the recent work of Grinblatt and Saxena (2018) and Ao, Li, and Zheng (2018). Similar to their work, we consider the underperformance of the current proxies for the market portfolio. First, although CW indices provide a simple, cost-effective and intuitive manner to allocate stocks, they are also exposed to certain inherent weaknesses, notably their embedded momentum bias and their concentration in large capitalization stocks. Second, performing mean-variance optimization on individual assets induces large estimation errors of the variance-covariance matrix. Our approach will simplify the allocation and reduce these estimation errors by considering low-risk portfolios built on a limited number of basis portfolios. More specifically, Ao, Li, and Zheng (2018) compare the properties of the minimum variance (MV) and mean-variance efficient portfolios for a large set of individual assets augmented with risk factors using both sample and robust estimates of the variance-covariance matrix. The authors design a new statistical approach to reduce estimation error and show that considering risk factors together with individual assets manages to deliver optimal risk-return properties. Similar to their work, we compare various portfolio optimization techniques including the MV portfolio. We do not, however, consider MSR portfolios due to the empirical challenge of estimating expected returns and focus on low-risk optimizations, namely, minimum variance (MV), maximum diversification (MD) and risk parity (RP). In addition, Grinblatt and Saxena (2018) propose a MVE candidate that relies on a statistical technique to infer the weights of basis portfolios formed by sorting stocks according to their style characteristics. This portfolio is shown to span the opportunity set formed from a 3-factor (Fama and French (1993)) model. Their work induces long and short positions into extreme size and value portfolios, which might constitute an unfeasible outcome for common investors. On the contrary, our work provides a long-only investment solution.

We demonstrate that superior performance is attached to portfolios obtained by the optimization of "smart" characteristic-sorted equity portfolios. In our empirical exercise, we consider size, value and momentum basis portfolios based on the US equity market. We first compare the standard way to allocate stocks into these basis portfolios (using an independent scale with NYSE breakpoints) to the more recent dependent technique, which works in successive subportfolios. Whole-sample breakpoints are jointly used with a dependent sorting. Hereafter, we refer to these two sets of basis portfolios as the dependent and independent basis portfolios. We show that strategic beta portfolios formed on the latter dependent basis portfolios outperform (in terms of Sharpe ratio and alpha) low-risk portfolios formed on traditional independent basis portfolios. We also show that the risk-based portfolios formed on dependent basis portfolios span the traditional (Fama and French (1993)) three-factor model. Our empirical approach relies on the mean-variance spanning test of Kan and Zhou (2012) augmented with a bootstrap approach similar to Fama and French (2010) and Harvey and Liu (2016) to ensure the robustness of our results. We also perform the factor selection technique of Harvey and Liu (2016) to conduct a horse race between the different configurations of the low-risk portfolios. We finally build on the "diversification return" from Booth and Fama (1992) and the extensions of Willenbrock (2011) and Erb and Harvey (2006) to infer the diversification properties of our strategic beta portfolios.

The rest of the paper is organized as follows: Section II presents a literature review. Section III describes and compares the opportunity sets used; i.e., the data and methodology used to construct the characteristic-based portfolios. Section IV presents the smart investment strategies and their diversification properties. In Section V, mean-variance spanning tests are used to compare smart investment strategies against single-index and multifactor models. In Section VI, we test the significance of our smart investment strategies to complement a multifactor model and explain the cross section of characteristic-sorted portfolios. Section VII concludes the paper.

## II. Literature Review

The work of Markowitz (1952) establishes the foundations of modern portfolio theory (MPT). Under this framework, investors are assumed to be individual agents with homogeneous preferences for the first two moments of the distribution of financial assets. In such circumstances, investors optimally invest in mean-variance efficient portfolios lying on a so-called mean-variance efficient frontier. Under several additional assumptions (unlimited risk-free borrowing and short selling, no frictions (taxes, transaction costs) and nontradable assets (social security claims, housing, human capital)), investors are concerned about only the tangency portfolio to this frontier; i.e., the MSR portfolio, commonly referred to as the "market" portfolio (Sharpe (1964), Lintner (1965), and Mossin (1966)). CW indices have long been a popular proxy for this portfolio.

Under real-world conditions, however, the market portfolio may not be efficient (Sharpe (1991) and Markowitz (2005)). The recent literature has proposed non-CW strategies to circumvent the drawbacks of CW allocation schemes.

An issue potentially preventing the mean-variance efficiency of the CW portfolio could be the use of market capitalization as a measure of fair value. If stock prices do not fully reflect firm fundamentals, then the CW portfolio might be suboptimal because it over- (under) weights over- (under) priced stocks (Hsu (2006)). This drawback has led to fundamental indexing and the creation of characteristic-based indices that weight stocks according to their economic footprints (such as revenues, book values, and earnings). According to Arnott (2005), compared to traditional CW indices, this new heuristic scheme provides consistently superior mean-variance performance. Academics and practitioners have recently explored scientific diversification. Looking for the MSR portfolio faces the challenge of estimating robust inputs. Sophisticated statistical techniques have been proposed (see, for instance, Demiguel and Nogales (2009), Ledoit et al. (2016), and Ao, Li, and Zheng (2018)). Risk-based optimization techniques have also been examined as they simplify the mean-variance estimation process by "giving up" on the estimates of expected returns (Clarke, Silva, and Thorley (2013), Choueifaty and Coignard (2008), Maillard, Roncalli, and Teïletche (2010), and Ardia, Boudt, and Nguyen (2018)). All these risk optimizations assume that the expected return of an asset increases in proportion to its risks. Amenc, Goltz, and Martellini (2013) and Frazzini and Pedersen (2014) ground these risk-based optimization techniques into the theory of the recently discovered low-beta anomaly. The low-beta anomaly contradicts the MPT theory in the sense that stocks with low volatility (low beta) are empirically shown to deliver higher returns than high-volatility (high beta) stocks (Baker, Bradley, and Wurgler (2011), Baker, Bradley, and Taliaferro (2014), and Cederburg and O'Doherty (2016)). The techniques of MV, MD and RP fall under this category of risk-based optimizations.

In recent years, an equally weighted (EW) investment strategy has gained popularity. As CW allocations suffer from their concentration in large capitalization stocks and from exposure to uncontrolled sources of risk, such as momentum due to a price-based weighting scheme, one simple or naive way to ensure good diversification and low idiosyncratic risk would be to equally weight all $N$ constituents of the portfolio. An EW scheme, referred to as "$1/N$", is a heuristic method that approximates mean-variance efficiency only when the assets have the same expected return and covariance (Chaves et al. (2012)). DeMiguel, Garlappi, and Uppal (2009) demonstrated that none of the "optimal" allocation schemes they reviewed (Bayesian methods as well as the CW portfolio) significantly outperform out-of-sample the "$1/N$" portfolio in terms of the Sharpe ratio and certainty equivalent return. Plyakha, Uppal, and Vilkov (2015) attribute the sources of the outperformance of $1/N$ portfolios on CW and $1/N$ portfolios to rebalancing and the embedded reversal strategy of the strategy. Regarding the simplicity of the strategy, DeMiguel, Garlappi, and Uppal (2009) claim that the "$1/N$" portfolio should be defined as a benchmark to evaluate alternative weighting schemes.

Thus, smart beta ranges from scientific diversification (such as the MV portfolio or risk efficient indexing) and risk-based heuristic methods (MD indexing, diversity-weighted indexing or RP indexing) to fundamental indexing (e.g., using dividend yield as a proxy for asset market value). Hsu and Kalesnik (2014) show that compare to other allocation strategies (i.e., strategies based on fundamental weight, MV, and $1/N$), traditional CW indexing exhibits (by construction) a drag in their expected returns because the strategy involves buying stocks when prices are high and selling stocks when prices are low. Nevertheless, Perold (2007) and Graham (2012) conclude that there is no evidence thatthis return drag of cap-weighted indices is valid regardless of the period. In reality, fundamental indexing is another method used to implement style investing: it produces a significant bias toward distressed stocks (Jun and Malkiel (2007) and Perold (2007)). This method, therefore, has a risk of concentration equivalent to that of traditional CW portfolios.

These strategies have mostly been implemented at the individual stock level as the equity building block to construct portfolios that satisfy specific investor objectives or gain exposure to specific systematic risk factors (e.g., Arnott et al. (2013), Clarke, Silva, and Thorley (2013)).

Investors can also find benefits in performing strategic beta allocations at the portfolio level (Boudt and Peeters (2013)) or even at the asset class level (Ardia, Boudt, and Nguyen (2018)). In fact, recent studies have recognized the use of asset or factor portfolios as the new opportunity set (Idzorek and Kowara (2013), Roncalli and Weisang (2016), Ao, Li, and Zheng (2018), and Grinblatt and Saxena (2018)). The value added by working on equity style buckets rather than asset classes or individual assets when implementing risk-based optimizations and the performance of the so-constructed efficient portfolios have only recently been studied (see Grinblatt and Saxena (2018)). Our paper contributes to this recent literature.

# III. Investment Opportunity Set

This section describes our opportunity set; i.e., the set of portfolios that constitute our basis assets. Our approach consists of stratifying the US stocks' universe in investment style portfolios; namely, size, book-to-market and momentum portfolios.

Grouping stocks into portfolios offers several advantages. First, forming groups of stocks into style portfolios circumvents the burden of estimating a large covariance matrix of returns (Berk (2000) and Ao, Li, and Zheng (2018)). In addition, our framework is consistent with the stylized facts of Barberis and Shleifer (2003), who demonstrate the natural tendency of investors to allocate funds according to asset categories, and Froot and Teo (2008), who also observe that institutional investors tend to reallocate their funds across style groupings. Our objective to perform risk optimization techniques on investment style portfolios is therefore in line with the reallocation practice of institutional investors and avoids the implementation costs of working with a wide variety of individual securities.

Our stratification relies on two sorting methodologies. The first construction methodology is based on an independent double or triple sort of stocks into portfolios and has become a standard in the asset-pricing literature for constructing characteristic-sorted portfolios (Fama and French (1993), Fama and French (1995), and Fama and French (2015)). The second sorting methodology follows Lambert, Fays, and Hübner (2016) and applies a double or triple dependent sort using whole-sample breakpoints; this strategy implies the sorting of stocks in successive subportfolios according to characteristics. We stratify the US stock universe into six (2×3), nine (3×3) or twenty-seven (3×3×3) groups. The double sort is performed on size and book-to-market characteristics, while the 3×3×3 split is constructed on the momentum, firm size, book-to-market characteristics. More details of the two methodologies can be found below.

## A. Data

The data are obtained by merging data from the Center for Research in Security Prices (CRSP) and Compustat. The CRSP database contains historical price information, whereas Compustat provides accounting information for all stocks listed on the major US stock exchanges. The sample period ranges from July 1963 to December 2015 and covers all stocks listed on the NYSE, AMEX, and NASDAQ.[1] For stocks listed on the NASDAQ, the data collection starts in 1973. The analysis covers a total of 618 monthly observations. Following Fama and French (1993) to filter the database and construct cross-sectional portfolios, we keep stocks with a CRSP share[2] 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 $t$, at least 2 years of listing on Compustat to avoid survival bias and a positive book-equity value at the end of December of year $y-1$. We define the book value of equity as the Compustat book value of stockholders' equity (SEQ) plus the balance-sheet deferred taxes and investment tax credit (TXDITC). If available, we decrease this amount by the book value of the preferred stock (PSTK). If the book value of stockholders' equity (SEQ) plus the balance-sheet deferred taxes and investment tax credit (TXDITC) is not available, we use the firm's total assets (AT) minus its total liabilities (LT).

Book-to-market equity (B/M) is the ratio of the book value of equity for the fiscal year ending in calendar year $y-1$ to market equity. Market equity is defined as the price (PRC) of the stock times the number of shares outstanding (SHROUT) at the end of June y to construct the size characteristic and at the end of December of year $y-1$ to construct the B/M ratio. Momentum is defined as in Carhart (1997); i.e., based on a $t-2$ until $t-12$ cumulative prior return.

[1]Data regarding Compustat and CRSP are available from January 1950 and January 1926, respectively. After correcting the databases for survival and backfill biases the sample starts in July 1953. For comparison purpose, we start our empirical analyses from July 1963 onwards as in Fama and French (1993).

[2]see Hasbrouck (2009, p. 1455):"*restricted to ordinary common shares (CRSP share code 10 or 11) that had a valid price for the last trading day of the year and had no changes of listing venue or large splits within the last 3 months of the year*".

## B. Sorting Out Stocks

In the original Fama–French approach, portfolios are constructed using a 2×3 independent sorting procedure: two-way sorting (small and large) on market capitalization and three-way sorting (low, medium, high) on the book-to-market equity ratio. Six portfolios are constructed at the intersection of the 2×3 classifications and are rebalanced on a yearly basis at the end of June. These style classifications are defined according to the NYSE stock exchange only and are then applied to the whole sample (AMEX, NASDAQ and NYSE).[3] and motivate the use of NYSE breakpoints by the need to have approximately the same market capitalization across portfolios and the same number of NYSE firms in each portfolio.

The second sorting methodology is an extension of the Fama–French sorting methodology. Lambert, Fays, and Hübner (2018) sort stocks in successive subportfolios according to various characteristics; moreover, they define sorting breakpoints based on the whole sample rather than considering only the NYSE. The authors indeed uncover that these NYSE breakpoints create an imbalance in the (total) number of stocks between small- and large-cap portfolios such that independent sorting leads to a higher number of stocks in small-value portfolios. Using independent sorting on negatively (or positively) correlated variables can induce, by design, a strong tilt toward the extreme categories of inverse ranks (low-high and high-low). From January 1963 to December 2014, the market equity and book-to-market equity of a firm were, on average, negatively correlated (−5%). In Figure 1, we illustrate the implications of the choice of sorting methodology on stratifying the US equity universe into (2×3) characteristic-sorted portfolios. The independent sorting methodology results in a large part of the universe falling into the small-value (28.1%) category, whereas dependent sorting delivers a well-balanced distribution of stocks in all portfolios (approximately 16%).

[Figure 1 near here]

The practical consequence when sorting stocks into portfolios, as already stated by Chan, Dimmock, and Lakonishok (2009), is that the original independent sorting procedure could induce large value stocks to be categorized as growth stocks. Supportive evidence can be found in the recent work of Lettau, Ludvigson, and Mandel (2018) who characterize the holdings of value mutual funds using

[3]The NYSE is represented by stocks that account for the largest capitalization in the CRSP database. The exchange codes 1, 2 and 3 represent the NYSE, NASDAQ and AMEX, respectively

Daniel et al. (1997) methodology.[4] Lettau, Ludvigson, and Mandel (2018) show that value mutual funds tend to hold a large proportion of their investments in growth stocks. However, the ranking into quantiles relies on NYSE breakpoints . Lambert, Fays, and Hübner (2018) document that the choices underlying the sorting methodology are important to draw robust inference on firm style characteristics. In particular, the standard procedure of NYSE breakpoints and the sequence of the dependent sort matter. If the sorting methodology is responsible for these empirical results, we claim that forming basis portfolios using this procedure will lead to a biased allocation of stocks into style portfolios and stratification of the US equity universe, and therefore to a misleading optimization exercise.

To better understand the problem, we compare the Morningstar style classification of 8,739 mutual funds (focused on the US equity market) to the ones implied by the dependent on all breakpoints and independent on NYSE breakpoints sorting procedures. For the dependent sort, the classification of stocks under growth and value characteristics is obtained by applying a first sort on the size characteristic of a firm and then performing a the second sort on the book-to-equity market of firm. We construct a matrix of 5×5 portfolios along the size and value characteristics of a firm. For an independent sort, the output is similar to the 5×5 size and value portfolios available on Kenneth French's website.

The sample of mutual funds is obtained from Morningstar and CRSP databases over the period April 2002 to December 2005. Databases are merged according to the funds' CUSIP and a phrase matching techniques applied on funds' name. Monthly performance and quarterly holdings are obtained from CRSP Mutual Fund Database. Style classification is obtained from Morningstar. Next, we match the information of funds' holdings with the value-growth classification from the independent (with NYSE breakpoints) and dependent (with whole sample breakpoints) sorting methodologies. The stock universe is then split according to a 1–5 scale: 1 represents a growth tilt, 3 represents a blend/neutral style, and 5 represents a value tilt. The classification is applied according to accounting information obtained from Compustat at the end of June for each stock.

Figure 2 illustrates the distribution of funds along the dependent-name breakpoints (hereafter referred simply to dependent sort) and independent-NYSE (hereafter referred simply to independent

[4] Daniel et al. (1997) sort stocks at the end of June of each year to form 125 portfolios along a triple dependent sort with a first sort on firms size, a second sort on firms' industry adjusted book-to-market and a final sort on firms' momentum (cumulative return from t-2 to t-12).

sort) frameworks for the following Morningstar categories: growth (left), blend (middle), value (right). The distributions across the growth and value styles demonstrate that the BM-score of mutual funds computed using a dependent scale instead of an independent scale is better aligned with the style classification of the fund. Indeed, the distribution for growth funds is more skewed to the left for the dependent sort as shown by the 21.32% (dependent sort) vs 9.17% (independent sort) of the observations falling under the first quintile of the distribution. Similarly for value funds, the distribution is more skewed to the right for the dependent sort given that 7.30% (dependent sort) vs 4.81% (independent sort) of the observations falling under the last quintile of the distribution. Lastly, the mode of the distribution of blend mutual funds under a dependent scale falls around the third quintile as 49.42% of the observations are found in the quintile 2 and 4. Using an independent scale, the mode is shifted to values below 3, which are representative of growth stocks. This would wrongly indicate that these funds hold more growth than value stocks.

In summary, value (growth) mutual funds have a higher probability of being categorized as value (growth) funds under a dependent sorting procedure than an independent sorting procedure. Blend mutual funds also show a better neutrality to the value-growth categorization using a dependent sort.

[Figure 2 near here]

Next, we break down the Morningstar value-growth categories across three levels of size attributed to the holdings style of mutual funds: large (top), mid (middle), small (bottom) capitalization. Figure 3 display the breakdown across the size categories. Results confirm that under an independent sort, the mutual funds holdings style would be biased due to the intrinsic asymmetric relationship between the size and book-to-market stock classification. More precisely, an independent sort will systematically attribute a greater value (growth) score to funds holding small (large) cap stocks. Conversely, the dependent sort does not exhibit this systematic size-related bias as the distributions of the funds' BM-score appear unconditional to funds' size subsample.

[Figure 3 near here]

Finally, we focus in Table 1 on the 15 mutual funds with the highest discrepancy in book-to-market related score between the independent and dependent framework. All of them are marketed as value funds and achieve a score higher than 3 under the dependent framework. We can see that, although they declare a value focus, some funds are wrongly categorized as blended funds under the independent framework with a score below but near to 3.

[Table 1 near here]

## C. Pair-Wise Correlation of Style Portfolios

Figure 4 illustrates the stock distribution when the number of portfolios is increased either by a larger split of the sample (from a 2×3 to a 3×3 split) or by adding a new characteristic (3×3×3). The 3×3×3 splits are constructed based on the size, value and momentum characteristics of a firm. We observe that using an independent sort results in an imbalance of stocks across the portfolios, and this effect becomes larger when more groups are constructed.

[Figure 4 near here]

We expect the better diversification induced by the dependent sorting and by the higher dimensional space representation of the US equity market to deliver additional diversification benefits for risk-based optimizations with regard to independent basis portfolios.

To verify this hypothesis, in Table 2, we compute the average correlation between the investment style portfolios. It can be shown that the correlation is lower when stocks are sorted dependently and are split into a larger number of groups (i.e., 3×3×3).

[Table 2 near here]

We form 6, 9, and 27 investment style portfolios and use 60 daily returns to estimate the covariance matrix.[5]

In the most extreme case (27 portfolios), we are left with 0.17 data points per parameter, which might create large sampling errors if we only consider the sample covariance matrix in our optimizations. In our applications, we use a traditional shrinkage methodology developed by Ledoit and Wolf (2004) to estimate the covariance with lower sampling errors. Further details on the shrinkage method used can be found in the Appendix A.1.

[5]We use a range of 60-day to estimate variance-covariance matrices for two reasons; first, Fama and French (2018) use 60 days of lagged returns to estimate the monthly variance of stocks, and second, real-life applications on tradable assets would also impose practical constraints over the length available for time-series (Idzorek and Kowara (2013)).

## IV. Smart Investment Strategies: Diversification Properties

We consider three low-risk investment strategies (MV, MD, RP) for the various opportunity sets (dependent versus independent portfolios; 2×3, 3×3 and 3×3×3 sort). Our basis portfolios are all cap-weighted portfolios to mitigate the impact of small cap stocks.

Table 3 recalls the analytic forms of the risk-based allocations that serve as a practical base in our empirical analysis; namely, *minimum variance*, *maximum diversification*, and *risk parity.* Following Ardia, Boudt, and Nguyen (2018), Roncalli and Weisang (2016), Grinblatt and Saxena (2018), and Ao, Li, and Zheng (2018) among others, these risk-based allocations are rebalanced on a monthly basis.

[Table 3 near here]

This section compares the diversification returns achieved through implementing risk-based optimization based on dependent and independent basis portfolios and further decomposes the diversification return into its two components and performs a paired difference test.

The diversification return according to Booth and Fama (1992) is defined as the difference between the compound return of a portfolio and the weighted average of the compound return of its constituent assets. This relationship assumes that the portfolios are rebalanced so that the weights are held constant. In this situation, the diversification return increased with the spread between the individual asset variance and its covariance with the portfolio.

Denoting the geometric average return as $g$, the volatility as $\sigma$, and the arithmetic average return as $\mu$, the geometric return of a portfolio $p$ can expressed mathematically as follows:

$$g_p = \mu_p - \frac{\sigma_p^2}{2} \tag{1}$$

The diversification return can be written as follows (Booth and Fama (1992) and Willenbrock (2011)):

$$\mathrm{DR} = g_p - \sum_{i}^{N} w_i g_i \tag{2}$$

where $i$ stands for the $i^{th}$ security in the portfolio $p$, and $g$ refers to the geometric return. Weights ($w_i$) are assumed to be constant over the estimation period. We refer to fixed-weight diversification return using the superscript (FW).

Substituting (1) in (2), we obtain

$$\text{DR}^{FW} = \mu_p - \frac{\sigma_p^2}{2} - \sum_i^N w_i \left( \mu_i - \frac{\sigma_i^2}{2} \right) \tag{3}$$

Rearranging the terms,

$$\text{DR}^{FW} = \underbrace{\mu_p - \sum_i^N w_i \mu_i}_{\text{DR}_1^{FW} = 0 \text{ if weights are constant}} + \underbrace{\frac{1}{2}\left( \sum_i^N w_i \sigma_i^2 - \sigma_p^2 \right)}_{\text{DR}_2^{FW} = \text{variance reduction benefit}} \tag{4}$$

In the last part of the equation, we retrieve the variance reduction benefit ($\text{DR}_2$) of Booth and Fama (1992) and Willenbrock (2011). Note that in theory, $w_i$ should be determined at inception and remain constant over the life of the strategy. To implement equation (4) for rebalancing strategies (non fixed weight), Erb and Harvey (2006) use the average of the weights over the sample period ($\overline{w_i} = \frac{1}{T} \sum_1^T w_i^t$). Thus, we follow this approach to evaluate the diversification return of our strategies for which weights are not held constant over time.

To test the statistical difference in diversification return brought by a pair of strategies performed on two opportunity sets, we follow the indirect bootstrap framework of Ledoit and Wolf (2008). Their model is initially constructed to compare whether a pair of strategies have statistically equivalent Sharpe ratio. In their conclusion, the authors suggests to extend their model to other mean-variance performance measures. We thus revisit their framework to a spread in diversification return between a pair of strategies. We provide more details on our extension in the Appendix A.2. In short, we aim to compare the spread in diversification return ($\Delta$ Dep-Ind) estimated in the original sample to an empirical distribution of spreads constructed from bootstrapped samples and infer the level of significance of this spread.

We report, in Table 4, the results of the diversification return for the low-risk investment strategies based on 2×3, 3×3, and 3×3×3 basis portfolios. We observe that the spread in variance reduction benefit ($\text{DR}_2^{FW}$) is in 8 out of 9 times statistically greater for the set of dependent basis portfolios at the usual significance level.

[Table 4 near here]

The total spreads ($\Delta$ Dep-Ind) are however close to zero due to the first negative DR component induced by the artificial fixed weight benchmark. As the computation of the diversification return induces a comparison with a static portfolio endogenous to each strategic, it is difficult to compare the total diversification gains across a pair of strategies. We therefore extend equation (4) to consider a rebalanced portfolio $p$ and its diversification return with regard to an EW benchmark. We chose the equal-weighted strategy because this is the only allocation for which we know ex-ante the value of $w_i$, that is (1/N), as long as the amount of securities ($N$) remains constant in the portfolio. In this alternative framework, we make sure that two smart beta strategies constructed on a equivalent number of basis portfolios ($N$) share the same benchmark ($1/N$). We denoted the principle that the diversification return is compared to an EW strategy using the superscript (EW) as follow,

$$\mathrm{DR}^{EW} = \underbrace{\mu_p - \frac{1}{N}\sum_{i}^{N}\mu_i}_{\mathrm{DR}_1^{EW}} + \underbrace{\frac{1}{2}\left(\frac{1}{N}\sum_{i}^{N}\sigma_i^2 - \sigma_p^2\right)}_{\mathrm{DR}_2^{EW}} \tag{5}$$

Table 5 presents the results when the benchmark is equal-weighted and the strategies on the independent and dependent portfolios share, at least, a benchmark with equivalent weighting scheme.For 7 out of 9 risk-based optimizations, the dependent opportunity set offers significantly higher diversification returns than the independent sort. Consistent with Grinblatt and Saxena (2018), risk-return improvement can be achieved by using optimization techniques for constructing an MVE benchmark rather than constructing risk factors by using equal weights on the long and short legs. Again, the dependent opportunity sets systematically outperform the independent opportunity set. Also, the variance reduction benefit remains higher in 8 out of 9 times for the dependent-sorted opportunity set compared to the independent-sorted opportunity set.

[Table 5 near here]

The next section is dedicated to providing a methodological analysis on the mean-variance performance of the smart beta strategies.

## V. Mean-Variance Spanning Test

Mean-variance spanning *à la* Huberman and Kandel (1987) means that a set of $K$ risky assets spans a larger set of $K + N$ assets if the efficient frontier made of the $K$ assets is identical to the efficient frontier comprising the $K + N$ assets. We initially set $R_1$ to a $K$-vector of the returns on

$K$ benchmark assets, $R_2$ to a $N$-vector of the returns on $N$ test assets, and R to the raw returns on $K + N$ assets. Huberman and Kandel (1987) define the following regression test:

$$R_2^t = \alpha + \beta R_1^t + e^t \tag{6}$$

The null hypothesis $H_0$ sets $\alpha = 0$ and $\delta = 1 - \beta = 0$.

The null hypothesis implies mean-variance spanning as the benchmark assets dominate the test assets; both assets have the same mean, but the $K$ benchmarks have a lower variance than the test assets.

Considering an efficient frontier comprising $K + N$ assets, the following two formulas express the optimal weights of the $N$ assets into the tangent ($Qw_1$) and GMV[6] ($Qw_2$) portfolios:

$$\begin{aligned} Qw_1 &= \frac{QV^{-1}\mu}{1'_{N+K}V^{-1}\mu} = \frac{\Sigma^{-1}\alpha}{1'_{N+K}V^{-1}\mu} \\ Qw_2 &= \frac{QV^{-1}1_{N+K}}{1'_{N+K}V^{-1}1_{N+K}} = \frac{\Sigma^{-1}\delta}{1'_{N+K}V^{-1}1_{N+K}} \end{aligned} \tag{7}$$

where $Q = [0_{N\times K}, I_N]$ with $I_N$, an $N \times N$ identity matrix, $\Sigma = V_{22} - V_{21}V_{11}^{-1}V_{12}$ which comes from $V$ the variance-covariance matrix of the $K$ benchmark assets ($R_1$) plus the $N$ test assets ($R_2$) that is,

$$V = Var[R_1^t, R_2^t] = \begin{bmatrix} V_{11} & V_{12} \\ V_{21} & V_{22} \end{bmatrix} \tag{8}$$

The value of alpha will determine whether the tangency portfolio is improved by the introduction of the $N$ assets, while testing beta will determine whether a significant change is induced in the GMV portfolio by the addition of the $N$ assets. Huberman and Kandel (1987) jointly test these two conditions. The rejection of mean-variance spanning could thus find two sources: an improvement in the slope of the tangency portfolio or an improvement in the risk-return properties of the GMV portfolio. However, beta can be estimated more accurately than alpha, as it does not depend on the expected returns of the assets (see equation 7). Therefore, the statistical significance of the change

[6] To make a clear distinction between the risk-optimization that minimizes the portfolio variance and the ex-post global minimum variance portfolio, we denote the former MV and the latter GMV in the rest of the paper.

in the composition of the GMV portfolio can be reached without implying economic significance. To circumvent this problem, Kan and Zhou (2012, hereafter KZ) propose to separately test the two conditions and to adjust the significance threshold of the two tests to economic significance. If the GMV condition is rejected more easily, the significance threshold should be reduced.

The KZ step-down test proceeds as follows: The first test defines the null hypothesis for the tangent portfolio such that $\alpha = 0_N$ using the OLS regression. The tangency portfolio is improved when the null hypothesis is rejected.

$$H_0^1 : \alpha = 0_N \tag{9}$$

Kan and Zhou (2012) perform a test for the statistical significance of the hypothesis similar to a GRS $F$-test. The $F$-test for the first hypothesis ($H_0^1$) is

$$F_1 = \frac{T-K-N}{N} \frac{\hat{a} - \hat{a}_1}{1+\hat{a}_1} \tag{10}$$

where $T$ is the number of observations; $K$ is the number of benchmark assets; $N$ is the number of test assets; $\hat{a}_1 = \hat{\mu}_1' \hat{V}_{11}^{-1} \hat{\mu}_1$ represents the squared Sharpe ratio of the K benchmark assets ($R_1$), with $\hat{V}_{11}$ denoting the variance and $\hat{\mu}_1$ the vector of mean return of the benchmark assets; and $\hat{a}$ takes the same notation as $\hat{a}$ but refers to the benchmark assets plus the new test asset ($R$).

The second test of the step-down procedure defines the null hypothesis for the GMV portfolio. This second test is conditional on the first test, $\alpha = 0_N$, and verifies whether $\delta = 1_N - \beta 1_K = 0_N$. Only when both conditions are rejected does the test suggest that the GMV portfolio is improved by adding $N$ assets to the $K$ benchmark assets.

$$H_0^2 : \delta = 1_N - \beta 1_K = 0_N | \alpha = 0_N \tag{11}$$

The $F$-test for the second hypothesis ($H_0^2$) is

$$F_2 = \frac{T-K-N+1}{N} \left[ \frac{\hat{c}+\hat{d}}{\hat{c}_1+\hat{d}_1} \frac{1+\hat{a}_1}{1+\hat{a}} - 1 \right] \tag{12}$$

where $\hat{c}_1 = 1_K' \hat{V}_{11}^{-1} 1_K$ and $\hat{d}_1 = \hat{a}_1 \hat{c}_1 - \hat{b}_1^2$ are the efficient set (hyperbola) constants with $\hat{a}_1 = \hat{\mu}_1' \hat{V}_{11}^{-1} \hat{\mu}_1$ and $\hat{b}_1 = \hat{\mu}_1' \hat{V}_{11}^{-1} 1_K$ for the benchmark assets ($R_1$). $\hat{\mu}_1$ and $\hat{V}_{11}$ denote the vector of mean return and the variance of the benchmark assets. $\hat{a}$, $\hat{b}$, $\hat{c}$ and $\hat{d}$ are the equivalent notations for the benchmark assets plus the new test assets ($R$).

In Panel A of Figure 6, we graphically illustrate a significant improvement in the tangency portfolio when a test asset ($R_2$) is added to the benchmark assets ($R_1$). Panel B indicates a significant improvement in the GMV portfolio when a test asset ($R_2$) is added to the benchmark assets ($R_1$).

[Figure 6 near here]

Mean-variance spanning implies that both null hypotheses hold ($H_0^1$ and $H_0^2$). The benchmark assets $R_1$ are said to span the test assets $R_2$ if the weight attributed to the $N$ test assets within the efficient frontier comprising $K + N$ assets is trivial. Put differently, discarding the $N$ test assets does not significantly change the efficient frontier of the $K$ benchmark assets from a statistical standpoint. By testing the two hypotheses separately, we gain understanding of the reason for mean-spanning rejection. If the mean-variance test is rejected, the test assets improve either the slope of the tangency portfolio or the risk-return properties of the GMV portfolio. Assuming the existence of a risk-free rate, investors are mostly concerned by the difference in the tangency portfolios.

Our application of mean-variance spanning tests whether smart investment strategies span existing benchmarks, such as the single-factor model or the expanded universe comprising the multifactor model of Fama and French (1993) (Section V.B). Spanning tests between the different configurations of low-risk portfolios are also performed to investigate the consequences of the use of different opportunity sets (Section V.C).[7]

## A. Out-of-Sample Test: A Bootstrap Approach

To test the robustness of our results, we extend the mean-variance spanning tests to address out-of-sample and data dredging concerns. We implement the bootstrap method used in Harvey and Liu (2016). It does not only allows to validate our results out-of-sample but also to control for multiple testing.

To analyze the mean-variance properties of a portfolio $P$ over the sample ranging from July 1963 to December 2015 (630 months), the method proceeds in 3 steps:

**Step 1:** Orthogonalization Under the Null

[7] The MATLAB code is available on Prof. Zhou's website.

The goal of this step is to modify the original times series of $R_2$ such that the null hypothesis appears to be true in-sample (Harvey and Liu (2016) and White (2000)). To do this, we perform the following regression,

$$R_2^t = \alpha + \beta R_1^t + R_2^{t,e} \tag{13}$$

Forcing the F-tests of $H_0^1$ and $H_0^2$ to be zero under the null (i.e. forcing the tangent and GMV portfolio to attribute zero weight to the test asset $R_2$) is achieved by adding the time series of one benchmark assets present in $R_1$ to the time-series of the residuals ($R_2^{t,e}$). In fact, the residuals ($R_2^{t,e}$) have by construction a zero mean but preserve the original time-series variation of $R_2$. In our applications, we always use the asset present in the benchmark asset ($R_1$) which proxies for the MVE market portfolio ($R_{1,MVE}^t$). The transformation equals to $R_2^{orth} = [R_2^{t,e} + R_{1,MVE}^t]$ and represents in-sample the orthogonal asset with zero weight in both the MVE and GMV portfolios. We use this artificial time-series $R_2^{orth}$ in the bootstrap samples to find the F-values of both hypothesis tests ($H_0^1$ and $H_0^2$).

**Step 2:** Bootstrap

The bootstrap procedure is a random selection of monthly observations of the strategy returns with replacement (i.e. $R_1$ and $R_2^{orth}$). We jointly resample the monthly strategies returns to ensure that we preserve the cross-sectional correlations between the strategies returns in our sample (i.e. $R_1$ and $R_2^{orth}$). Also, we make sure that the resample time-series have the same size as the original time frame (630 months). As in Fama and French (2010) and Harvey and Liu (2016), the bootstrap preserves cross-sectional and time-series dependence.

**Step 3:** MVE spanning test

We apply the mean-variance spanning from Kan and Zhou (2012) on the bootstrapped samples according to the benchmark assets ($R_1$) and the test asset ($R_2^{orth}$). We repeat the operation 1,000 times to construct an empirical distribution of the performance measures which are true (in-sample) under the null for $H_0^1$ and $H_0^2$. The empirical distribution serve as a threshold for the critical value of the F-tests. Each bootstrap contains two F-tests: one for the tangent

($F^b_{1,ind}$ and $F^b_{1,dep}$) and one for the GMV ($F^b_{2,ind}$ and $F^b_{2,dep}$), with the subscript $b$ denoting the b-th bootstrapped sample while *ind* and *dep* denote independent and dependent-sorted portfolios. In these bootstrap samples, the null is valid in-sample and a significant value for the F-tests simply arise from the resampling (or luck). Besides, we force the bootstrap sampling to have the same size as the original time series (630 observations) to ensure that the degree of freedom in the measurements of the F-tests remain equal to the F-test from the original sample.

Following the multiple testing framework developed by Harvey and Liu (2016), we compute the maximal measure of each hypothesis (tangent and GMV) that serve as a reference point for the collection of statistics from the bootstrapped samples and control for data snooping. For instance for the tangent hypothesis, $F^b_1 = max(F^b_{1,ind}, F^b_{1,dep})$. We can then compare these $B$ statistic measures to the one find in the original sample $F^o_{1,ind}$ and $F^o_{1,dep}$.

The frequency of observations in the bootstrap sample that are greater than the F-test under the original sample define the bootstrap p-value. Thus, the p-value is the sum of $I\{F^o_1 < F^b_1\}$ divided by the total number of bootstraps $B$.

## B. Mean-Variance Spanning Test of the Traditional (Multi) Factor Models

We consider three initial conditions. We assume the market model of Sharpe (1964) and the Fama and French (1993) 3-factor empirical model as our initial conditions; i.e., the initial benchmark portfolio $R_1$. In Scenario 1, $R_1$ comprises 2 assets: an investment in a 30-year US treasury bond (B30) and the market portfolio, which are both given in excess of the risk-free rate.[8] The market portfolio and the risk-free rate are obtained from Kenneth French's website while the 30-year US treasury bond (B30) is obtained from CRSP US Treasury and Inflation Indexes. $R_2$ is the gross return of one smart beta strategy. When implementing the bootstrap, $R_{1,MVE}$ is proxied by the cap-weighted market portfolio (Mkt) and $R^{orth}_2$ is equal to the residuals of the spanning regression (defined in the Step 1 of the Boostrap approach) plus $R_{1,MVE}$. In Scenario 2, $R_1$ comprises 4 assets: a 30-year US treasury bond (B30), the cap-weighted market portfolio (Mkt), the size (SMB) and value (HML) factors of Fama and French.[9] $R_2$ is the gross return of one smart beta strategy. When

[8]The risk-free rate refers to the one-month T-bill from Ibbotson.

[9]Data are obtained from Ken French's data library.

implementing the bootstrap, $R_{1,MVE}$ is also proxied by the cap-weighted market portfolio (Mkt) and $R_2^{orth}$ is equal to the residuals of the spanning regression plus $R_{1,MVE}$. In Scenario 3, $R_1$ comprises 2 assets: an investment in a 30-year US treasury bond (B30) and the gross return of one smart beta strategy. $R_2$ is the return on the market portfolio. When implementing the bootstrap, $R_{1,MVE}$ is now proxied by the gross return on the smart beta strategy and $R_2^{orth}$ is equal to the residuals of the spanning regression plus $R_{1,MVE}$. In all scenarios, Mkt, B30 and the smart beta strategies are taken in excess of the risk-free rate. However, we do not consider the smart beta strategies net of transaction costs because the risk factors (i.e., Mkt, B30, SMB, HML) used as explanatory variables are also gross of transaction costs. We provide further evidence on the performance of the strategies net of transactions latter in the next sub-section of the paper.

The step-down spanning test proceeds as follows: We first test the null hypothesis $H_0^1$ that the $\alpha$ is equal to 0, meaning that no improvement is obtained by adding the smart beta strategy to the initial portfolio. We consider the usual significance thresholds; i.e., 1%, 5% and 10%. The results will be further split into two sub-periods: the period for the full sample and the period after the publication date of the seminal Fama and French (1993) paper.

We report in Table 6 the results for the dependent and independent opportunity sets and for standard $p$-values (p-val) and bootstrapped $p$-values (p-val$^b$) on the full sample. Panel A shows that the traditional CAPM model does not span an expanded set augmented with risk-optimization strategies. The tangent portfolio level is significantly improved when adding the smart beta strategies using both the dependent and independent basis portfolios (all p-val$^b$ are significant with a 99% confidence level). Results of Panel B indicate however that a three-factor model spans the larger set comprising the original assets supplemented by a smart beta factor defined using *independent* basis portfolios. However, several smart beta strategies performed on a *dependent* opportunity set improve the tangency portfolio implied by the three-factor model: 4 strategies out of 9 improve the initial 3-factor portfolio at the 90% confidence level. Finally, Panel C shows that smart beta strategies performed on dependent basis portfolios span (7 cases out of 9) the tangent portfolio made of the traditional cap-weighted market portfolio. These results do not hold for independent basis portfolios. This evidence makes the latter sub-optimal with regard to low risk strategies implemented on *dependent* basis portfolios (i.e., 3×3 and 3×3×3).

[Table 6 near here]

Next, we present in Table 7 the results on the post-publication period of the Fama and French 3-factor model. Findings suggest that all low-risk strategies performed on a *dependent* opportunity set reject the mean-variance spanning hypothesis of the CAPM and 3-factor model as both sub-hypotheses (on alpha and delta) are statistically different from 0. This means that two portfolios of the mean-variance frontier (the MSR and the GMV) are improved under a *dependent* framework. However, the three-factor model continues to span 4 low-risk portfolios that are formed on the *independent* opportunity, especially the strategies aiming at maximizing the portfolio diversification (MD). This last evidence is particularly important as it confirms that the traditional *independent* sorting can not compete with a dependent sort when forming basis portfolios which offer sufficient cross-sectional variation.

[Table 7 near here]

In summary, our results on the post-publication period clearly highlights the improvement brought by considering low-risk portfolios constructed on style basis portfolios with regard to the related multi-factor model. These results might be explained by the increasing market diversity offering a higher potential for diversification and the increase in volumes traded on the US stock exchanges; this necessitates performing the optimization exercise on basis portfolios or factors rather than individual stocks. Findings also support the outperformance of low-risk strategies performed on the dependent opportunity set. A horse race between the two competing sets of basis portfolios will be performed in the next subsection.

## C. Horse Race Between Dependent and Independent Basis Portfolios

The previous subsection suggests that dependent basis portfolios offer interesting properties to perform risk-based optimization. We therefore carry out a horse race between the opportunity sets made of basis portfolios formed after dependent and independent sorting. Our spanning test considers whether a portfolio $R_1$ comprising US government 30-year bonds and one smart beta portfolio (defined with a dependent or independent opportunity set, respectively) span another set of portfolios made of $R_1$ and $R_2$ which comprises a US government bond and the equivalent smart beta portfolios but performed on other basis portfolios. The test is performed in both directions to test the superiority of one opportunity set over the other. In Scenario 1, $R_{1,MVE}$ is proxied by the smart beta on the dependent-sorted portfolios ($SB_{dep}$) and $R_2$ is the same smart beta strategy on the independent-sorted portfolios ($SB_{ind}$).[10] In Scenario 2, $R_{1,MVE}$ is proxied by the smart beta on the independent-sorted portfolios ($SB_{ind}$) and $R_2$ is the same smart beta strategy on the dependent-sorted portfolios ($SB_{dep}$).[11] Smart beta strategies are taken in excess of the risk-free rate and account for transaction costs as denoted by the super-script *net*. Both $H_1^0$ (the test on the tangency portfolio) and $H_2^0$ (the test on the GMV) are tested. Bootstrapped $p$-values are reported.

To consider transaction costs, Plyakha, Uppal, and Vilkov (2015) implement a decreasing function of transaction costs from 1% in 1978 to 0.5% in 1993 for their S&P 500 sample. However, in our paper, we trade stocks on NYSE-NASDAQ-AMEX exchanges and consequently have to differentiate between transaction costs for small and large-cap stocks. We thus follow an approach similar to that of Novy-Marx and Velikov (2016) and use the individual stock estimates from the Gibbs sampling developed in Hasbrouck (2009). Further details of this method can be found in the Appendix A.3. In Figure 5, we show the annual box-and-whisker plot for the CRSP/Gibbs estimates of transaction costs (variable $c$ from equation (A.16)) from 1963 to 2015.

[Figure 5 near here]

Novy-Marx and Velikov (2016) uncover a minor drawback to Hasbrouck's estimation technique, which requires relatively long series of daily prices to perform the estimation (250 days), resulting in a number of missing observations (mostly for non-NYSE stocks), for which the authors perform a nonparametric matching method and attribute equivalent transaction costs to the stock with a missing value according to its closest match to a stock with nonmissing value according to their size and idiosyncratic volatility. However according to the authors, these missing observations represent only 4% of the total market capitalization universe. Thus, instead, we replace the missing values with a transaction cost of 0.50%. We employ this value because (1) we see from Figure 5 that none of the estimates from Hasbrouck's algorithm have breached a trading cost of 50 bps since 1963, (2) this choice will more strongly impact illiquid stocks with a small amount of daily observations (small-

[10] In Scenario 1, $R_1$ is composed of B30 and a smart beta on dependent-sorted portfolios, $R_2$ refers to the same smart beta constructed on independent-sorted portfolios and $R_2^{orth}$ is equal to the residuals of the spanning regression - defined in Step 1 of the Boostrap procedure - plus $R_{1,MVE}$.

[11] In Scenario 2, $R_1$ is composed of B30 and a smart beta on independent-sorted portfolios while $R_2$ refers to the same smart beta constructed on dependent-sorted portfolios and $R_2^{orth}$ is equal to the residuals of the spanning regression plus $R_{1,MVE}$.

capitalization stocks), and (3) Plyakha, Uppal, and Vilkov (2015) also choose to set this threshold for transaction costs from 1993 onwards.

Table 8 presents the results for CW basis portfolios. Both panels demonstrate that the dependent opportunity sets outperform the independent set. In Panel A, we test whether the low-risk strategies formed on a dependent opportunity set span a larger universe augmented with independent sets. For all low-risk strategies, we cannot reject mean-variance spanning at the 5% confidence level. This means that the efficient frontier comprising a low-risk optimization of dependent portfolios and an investment in a long-term US government bond cannot be improved using an independent opportunity set. However, Panel B indicates that the MD (2×3, 3×3, and 3×3×3) and MV (2×3, 3×3) strategies performed on a dependent opportunity set improve both the tangency and the GMV portfolios formed on an independent opportunity set. This is evidenced by the levels of $p$-values attached to $F$-tests on $H_1^0$ and $H_2^0$ when the dependent portfolio is used as $R_2$.

[Table 8 near here]

# VI. MVE Benchmark Selection

We follow the method of Harvey and Liu (2016) to select the most appropriate (without luck) MVE benchmark among the low-risk portfolios and the original CW portfolio for explaining the cross-section of expected returns. Our test assets are the 2×3 and 3×3 portfolios sorted on size and book-to-market or the 3×3×3 when the sorting procedure first pre-condition on a firm's momentum. The MVE benchmark should best complement a basis multi-factor model comprising a long-term US Government rate as a proxy for the risk-free rate (B30) and the size (SMB) and the value (HML) factors of Fama and French (1993).

The method is an alternative to the test developed by Gibbons, Ross, and Shanken (1989). It departs however from the GRS test as it allows the initial or basis model to be sub-optimal and tests the incremental contribution of the additional factor.

To measure the incremental contribution of the selected candidate, Harvey and Liu (2016) define a scaled intercept (SI) measure and look at the spread between the scaled (by the standard error of the estimated intercept) intercept of the augmented and initial model. Using equivalent notations as the authors, the measure is defined as follow,

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

where $median(.)$ is the median value of the ratio $|a_i^g|/s_i^b$ or $|a_i^b|/s_i^b$. Here the superscript $b$ is for the baseline model and $g$ 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 negative value of the SI means that the augmented model outperforms the baseline model to explain the variations of the $J$ test assets returns. To define a statistical level of confidence to the measure, Harvey and Liu (2016) use the bootstrapping method presented in Step 2 of Section V.A.

To orthogonalize the MVE candidates, the authors regress the returns of $R_2^i$, where $i$ denotes the i-th candidate among the list of $K$ candidates, against the baseline benchmark $R_1$ and then subtract the intercept from the time-series $R_2^i$, as follows:

$$\begin{aligned} R_2^i &= \alpha^i + \beta^i R_1 + e^i \\ R_2^{\alpha,i} &= R_2^i - \alpha^i = \beta^i R_1 + e^i \end{aligned} \tag{15}$$

In our applications $R_1$ is composed of the risk-free rate (B30), the size (SMB) and the value (HML) factors. $R_2^{\alpha,i}$ is defined as a linear combination of the benchmark assets ($R_1$), i.e. the risk-free rate (B30), the size (SMB) and the value (HML) factors such that it does not bring any additional information to the baseline model.

Then in each sample of the $B$ bootstrap, a score for the scaled intercept $SI_{ew}^{med}$ can be obtained for the $K$ number of orthogonalized candidates (i.e, $R_2^{\alpha,i}$ with the $i = \{1, 2, ..., K\}$ candidate). Hence, the single test p-value for the i-th candidate is given by,

$$\text{p-val} = \frac{\#\{SI^o > SI^b\}}{B} \tag{16}$$

To control for multiple testing, the authors suggest to take the minimum value among $K$ estimates of $SI$ in the b-th bootstrap as follow,

$$SI^{b,*} = \underbrace{min}_{i \in \{1,2,...,K\}} \{SI^{b,i}\} \tag{17}$$

Hence, the multiple test p-value for the i-th candidate is written as,[12]

$$\text{p-val} = \frac{\#\{SI^o > SI^{b,*}\}}{B} \tag{18}$$

Next, our objective is to apply the method to a multiple set of MVE candidates and filter them to find the best candidate. For example, the first natural candidate to consider is the traditional cap-weighted market portfolio but smart beta strategies on the Fama-French's independent 2×3 size and value portfolios or on the dependent 2×3 size and value portfolios can also constitute MVE candidates to augment the baseline model. We also extend the 2×3 size and value grid, to a 3×3 or 3×3×3 with an additional sort on firms' momentum and end up with a set of 7 candidates as MVE portfolio. We run the test sequentially, as in Harvey and Liu (2016), until the single test p-value of each candidate is greater than a pre-specified threshold. In our application, we set the threshold to 10%. In each run, the selected candidate have a single test p-value but will also be attributed with a multiple test p-value to control for data snooping. The candidate is only accepted if the multiple test p-value is significant at a 90% confidence level.

Table 9 presents the results for the different types of basis assets and strategic beta portfolios. The table shows the single-test $p$-value for each MVE candidate as well as the final joint $p$-value for the selected candidates considering the multiple testing framework. Note that except for the cap-weighted market portfolios, all strategic portfolios candidates are net of transaction costs as computed in Appendix A.3. For 2×3 portfolios, the optimal MVE candidate comes from the same family as the set of basis portfolios to be explained. However, as soon as we increase the dimension of the sort and therefore the dispersion between portfolios, the dependent candidates win the horse race. The results of Panels A and B are explained by two facts; first, the space of 2×3 and 3×3 test assets is too low-dimensional to provide a robust statistical framework (Lewellen, Nagel, and

[12]Note that the sign of the indicator function is important. Here, we want to count the number of bootstrapped scaled intercepts ($SI^b$) that have lower values (improvement of the model) than the scaled intercept from the original sample($SI^o$). In other words, when the test is performed on the time-series of $R_2$ from the original data.

Shanken (2010)), also coined as "rank deficiency" by Grinblatt and Saxena (2018), and second, the imbalance of the distribution of stocks in portfolios under an independent sort is too sensitive to a number of macro-economic factors (Daniel and Titman (2012)). Consequently, these lead to the construction of sub-optimal basis portfolios whereas smart beta strategies benefit from a dependent sort which overcome these issues.

[Table 9 near here]

## VII. Conclusion

We currently observe a dual paradigm shift to so called "smart beta strategies" including low-risk portfolios and "strategic beta factors" (also called "style investing"). On the one hand, smart beta strategies provide an alternative weighting scheme for stocks; i.e., an alternative way to diversify risk. On the other hand, strategic beta investing looks to allocate "investment style" portfolios more efficiently to better capture systemic sources of market risk premiums. We build a bridge between these two paradigms as we show that the opportunity sets and the way to stratify the universe into style portfolios is important when performing smart beta techniques.

Our paper proposes new proxies for tangent/well-diversified (US equity) market portfolios by applying simple, long-only risk-based strategies to characteristic-sorted equity portfolios (i.e., an opportunity set sorted by market capitalization, book-to-market ratio and momentum characteristics). This method ensures that the risk properties of style portfolios are taken into account and simplifies the allocation by reducing the errors in the covariance matrix of returns.

We claim that the question is economically important for two reasons. First, the recent inflation of discovered risk factors challenges the candidate of the value-weighted market portfolio in terms of being mean-variance efficient. Second, there is a common practice among institutional investors to reallocate funds across style groupings (e.g., Froot and Teo (2008)). We show that the methodology for grouping stocks in different style buckets has strong implications for the performance of the final strategy. To categorize stocks in investment style portfolios, we stratify the universe along the common dimensions of size, value and momentum characteristics. We implement two sorting methodologies to construct characteristic-based portfolios: independent sorting and dependent sorting. To demonstrate the implications of the sorting methodologies for the strategies' performance, we apply a bootstrap version of the mean-variance spanning tests from Kan and Zhou (2012) to the risk-oriented strategies that use characteristic-based portfolios as opportunity sets. The results show that a dependent stratification with whole sample breakpoints of stocks in portfolios provides significantly higher Sharpe ratios for risk-oriented strategic beta strategies. Because dependent sorting controls for correlated variables and stratifies the stock universe in well-diversified portfolios (Lambert, Fays, and Hübner (2016)), this sorting methodology delivers better diversification benefit potential for smart beta strategies. To demonstrate this point, we follow the approach of Erb and Harvey (2006) and Willenbrock (2011)) and provide an extended decomposition of the diversification return from Booth and Fama (1992). We find that the diversification return is higher for risk-based strategies implemented on these dependent portfolios than for those implemented on independent portfolios and higher than a naive equally weighted strategy. Our results hold out-of-sample and are robust to multiple testing.

## References


Amenc, N., Goltz, F., and Martellini, L. (2013) Smart Beta 2.0, *Working Paper - EDHEC Risk Institute*, 83–99.

Ao, M., Li, Y., and Zheng, X. (2018) Approaching Mean-Variance Efficiency for LArge Portfolios, *Review of Financial Studies* **00**, 1–30.

Ardia, D., Boudt, K., and Nguyen, G. (2018) Beyond Risk-Based Portfolios : Balancing Performance and Risk Contributions in Asset Allocation, *Quantitative Finance* **18**, 1249–1259.

Arnott, R., Hsu, J., Kalesnik, V., and Tindall, P. (2013) The Surprising Alpha from Malkiel's Monkey and Upside-Down Strategies, *The Journal of Portfolio Management* **39**, 91–105.

Arnott, R. D. (2005) Disentangling Size and Value, *Financial Analysts Journal* **61**, 12–15.

Baker, M., Bradley, B., and Taliaferro, R. (2014) The Low-Risk Anomaly: A Decomposition into Micro and Macro Effects, *Financial Analysts Journal* **70**, 43–58.

Baker, M., Bradley, B., and Wurgler, J. (2011) Benchmarks as Limits to Arbitrage : Understanding the Low-Volatility Anomaly, *Financial Analysts Journal* **67**, 40–54.

Barberis, N. and Shleifer, A. (2003) Style investing, *Journal of Financial Economics* **68**, 161–199.

Berk, J. B. (2000) Sorting Out Sorts, *Journal of Finance* **55**, 407–427.

Booth, D. G. and Fama, E. F. (1992) Diversification Returns and Asset Contributions, *Financial Analysts Journal* **48**, 26–32.

Boudt, K. and Peeters, B. (2013) Smart Harvesting of Equity Style Premia, *Working Paper*, 1–18.

Carhart, M. M. (1997) On Persistence in Mutual Fund Performance, *Journal of Finance* **52**, 57–82.

Cederburg, S. and O'Doherty, M. S. (2016) Does It Pay to Bet Against Beta ? On the Conditional Performance of the Beta Anomaly, *Journal of Finance* **71**, 737–774.

Chan, L. K. C., Dimmock, S. G., and Lakonishok, J. (2009) Performance : Issues and Evidence, *Review of Financial Studies* **22**, 4554–4599.

Chaves, D., Hsu, J., Li, F., and Shakernia, O. (2012) Efficient Algorithms for Computing Risk Parity Portfolio Weights, *Journal of Investing* **21**, 150–163.

Choueifaty, Y. and Coignard, Y. (2008) Toward Maximum Diversification, *The Journal of Portfolio Management* **35**, 40–51.

Clarke, R., Silva, H. de, and Thorley, S. (2013) Risk Parity, Maximum Diversification, and Minimum Variance: An Analytic Perspective, *The Journal of Portfolio Management* **39**, 39–53.

Daniel, K. D. and Titman, S. (2012) Testing Factor-Model Explanations of Market Anomalies, *Critical Finance Review* **1**, 103–139.

Daniel, K., Grinblatt, M., Titman, S., and Wermers, R. (1997) Measuring Mutual Fund Performance with Characteristic-Based Benchmarks, *Journal of Finance* **52**, 1035–1058.

Daniel, K., Mota, L., Rottke, S., and Santos, T. (2017) The Cross-Section of Risk and Return, *Working Paper*.

DeMiguel, V., Garlappi, L., and Uppal, R. (2009) Optimal versus naive diversification: How inefficient is the 1/N portfolio strategy?, *The Review of Financial Studies* **22**, 1915–1953.

Demiguel, V. and Nogales, F. J. (2009) Portfolio Selection with Robust Estimation, *Operations Research* **57**, 560–577.

Elton, E. J. and Gruber, M. J. (1973) Estimating the Dependence Structure of Share Prices, *Journal of Finance* **28**, 1203–1232.

Erb, C. B. and Harvey, C. R. (2006) The Strategic and Tactical Value of Commodity Futures, *Financial Analysts Journal* **62**, 69–97.

Fama, E. F. and French, K. R. (1993) Common Risk Factors in the Returns Stocks and Bonds, *Journal of Financial Economics* **33**, 3–56.

Fama, E. F. and French, K. R. (1995) Size and Book-to-Market Factors in Earnings and Returns, *Journal of Finance* **50**, 131–155.

Fama, E. F. and French, K. R. (2010) Luck versus Skill in the Cross-Section of Mutual Fund Returns, *Journal of Finance* **65**, 1915–1947.

Fama, E. F. and French, K. R. (2015) A Five-Factor Asset Pricing Model, *Journal of Financial Economics* **116**, 1–22.

Fama, E. F. and French, K. R. (2018) Choosing factors, *Journal of Financial Economics* **128**, 234–252.

Frazzini, A. and Pedersen, L. H. (2014) Betting against Beta, *Journal of Financial Economics* **111**, 1–25.

Froot, K. and Teo, M. (2008) Style Investing and Institutional Investors, *The Journal of Financial and Quantitative Analysis* **43**, 883–906.

Gibbons, M. R., Ross, S. A., and Shanken, J. (1989) A Test of the Efficiency of a Given Portfolio, *Econometrica* **57**, 1121.

Graham, J. (2012) Comment on the Theoretical and Empirical Evidence of Fundamental Indexing, *Journal of Investment Management* **10**, 90–98.

Grinblatt, M. and Saxena, K. (2018) When Factors Don't Span Their Basis Portfolios, *Journal of Financial and Quantitative Analysis.*

Harvey, C. R. and Liu, Y. (2016) Lucky Factors, *Working Paper*, Duke University.

Hasbrouck, J. (2009) Trading Costs and Returns for U . S . Equities: Estimating Effective Costs from Daily Data, *Journal of Finance* **64**, 1445–1477.

Hsu, J. (2006) Cap-weighted portfolios are sub-optimal portfolios, *Journal of Investment Management* **4**, 44–53.

Hsu, J. and Kalesnik, V. (2014) Measuring the "Skill" of Index Portfolios, *White Paper.*

Huberman, G. and Kandel, S. (1987) Mean-Variance Spanning, *The Journal of Finance* **42**, 873–888.

Idzorek, T. M. and Kowara, M. (2013) Factor-Based Asset Allocation vs. Asset-Class-Based Asset Allocation, *Financial Analysts Journal* **69**, 19–29.

Jun, D. and Malkiel, B. G. (2007) New Paradigms in Stock Market Indexing, *European Financial Management* **14**, 118–126.

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

Lambert, M., Fays, B., and Hübner, G. (2016) Size and Value Matter , But Not The Way You Thought, *Working Paper*, 1–67.

Lambert, M., Fays, B., and Hübner, G. (2018) Factoring Characteristics into Returns : A Clinical Study on the SMB and HML Portfolio Construction Methods, *Working Paper.*

Ledoit, O. and Wolf, M. (2004) A well-conditioned estimator for large-dimensional covariance matrices, *Journal of Multivariate Analysis* **88**, 365–411.

Ledoit, O. and Wolf, M. (2008) Robust performance hypothesis testing with the Sharpe ratio, *Journal of Empirical Finance* **15**, 850–859.

Ledoit, O., Wolf, M., Zhao, Z., and Ledoit, O. (2016) Beyond Sorting : A More Powerful Test for Cross-Sectional Anomalies Beyond Sorting : A More Powerful Test for Cross-Sectional Anomalies, *Working Paper.*

Lettau, M., Ludvigson, S., and Mandel, P. (2018) Characteristics of Mutual Fund Portfolios: Where Are the Value Funds?, *Working paper.*

Lewellen, J., Nagel, S., and Shanken, J. (2010) A skeptical appraisal of asset pricing tests, *Journal of Financial Economics* **96**, 175–194.

Lintner, J. (1965) Security Prices, Risk, and Maximal Gains From Diversification, *Journal of Finance* **20**, 587–615.

Maillard, S., Roncalli, T., and Teïletche, J. (2010) The Properties of Equally Weighted Risk Contribution Portfolios, *The Journal of Portfolio Management* **36**, 60–70.

Markowitz, H. (1952) Portfolio Selection, *Journal of Finance* **7**, 77–91.

Markowitz, H. (2005) Market Efficiency : A Theoretical Distinction and So What ?, *Financial Analysts Journal* **61**, 17–30.

Marshall, B. R., Nguyen, N. H., and Visaltanachoti, N. (2011) Commodity Liquidity Measurement and Transaction Costs, *Review of Financial Studies* **25**, 599–638.

Mossin, J. (1966) Equilibrium in a Capital Asset Market, *Econometrica* **34**, 768–783.

Novy-Marx, R. and Velikov, M. (2016) A Taxonomy of Anomalies and their Trading Costs, *The Review of Financial Studies* **29**, 104–147.

Perold, A. F. (2007) Fundamentally Flawed Indexing, *Financial Analysts Journal* **63**, 31–37.

Plyakha, Y., Uppal, R., and Vilkov, G. (2015) Why Do Equal-Weighted Portfolios Outperform Value-Weighted Portfolios?, *Working Paper*, 1–25.

Politis, D. and Romano, J. (1992) A circular block-resampling procedure for stationary data. In: LePage, R., Billard, L. (Eds.), Exploring the Limits of Bootstrap. In: *John wiley.* New York, 263–270.

Roll, R. (1984) A simple measure of the bid/ask spread in an efficient market, *The Journal of Finance* **39**, 1127–1139.

Roncalli, T. and Weisang, G. (2016) Risk Parity Portfolios with Risk Factors, *Quantitative Finance* **16**, 377–388.

Sharpe, W. F. (1964) Capital Asset Prices : A Theory of Market Equilibrium under Conditions of Risk, *Journal of Finance* **19**, 425–442.

Sharpe, W. F. (1991) The Arithmetic of Active Management, *Financial Analysts Journal* **47**, 7–9.

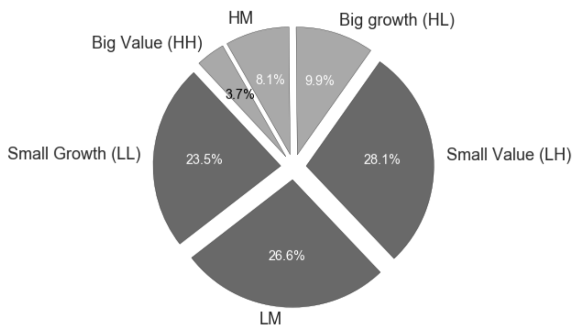
White, H. (2000) A Reality Check for Data Snooping, *Econometrica* **68**, 1097–1126.

Willenbrock, S. (2011) Diversification Return, Portfolio Rebalancing, and the Commodity Return Puzzle, *Financial Analysts Journal* **67**, 42–49.

**Figure 1:** Stock Distribution with Independent vs Dependent Sorting

The figure displays the stock distribution into the 2×3 characteristic-sorted portfolios on size (low and high) and book-to-market equity ratio (low, medium and high) for the independent (left) and dependent (right) sorting methodologies. The independent sorting uses the NYSE as a reference for breakpoints, while the dependent sorting uses all name breakpoints (NYSE, NASDAQ, and AMEX). The period is the interval from July 1963 to December 2015.

**(a)** Independent (2×3)




**(b)** Dependent (2×3)

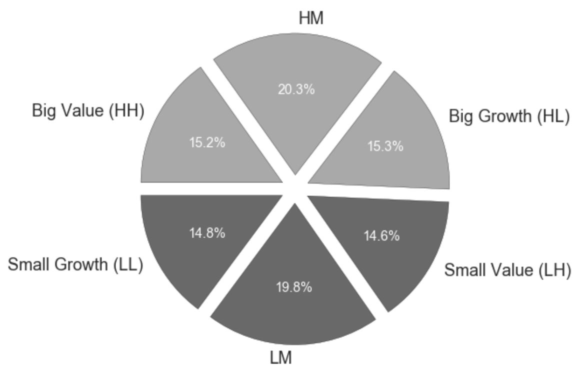


**Figure 2:** Distribution of BM-scores of Mutual Funds: Independent vs Dependent Sorts

The figure shows the kernel distribution in BM-score of mutual funds with a focused on the US equity market for which Morningstar attributes a value-growth classification. The value-growth classification applied to the mutual funds present in the CRSP mutual funds database. For each point in time where a fund reports its holdings, we associate a BM-score from a 1–5 scale according the Fama–French's 5x5 size and value *independent* sorting methodology or a 5x5 size and value *dependent* sorting methodology. The fund's BM-score is then calculated as the percentage of Total Net Assets (TNA) weighted average of the 1–5 scale of the securities the fund holds. Distributions are displayed for 3 Morningstar classifications of funds: growth (left), blend (middle), value (right). The sample period ranges from April 2002 to December 2015.

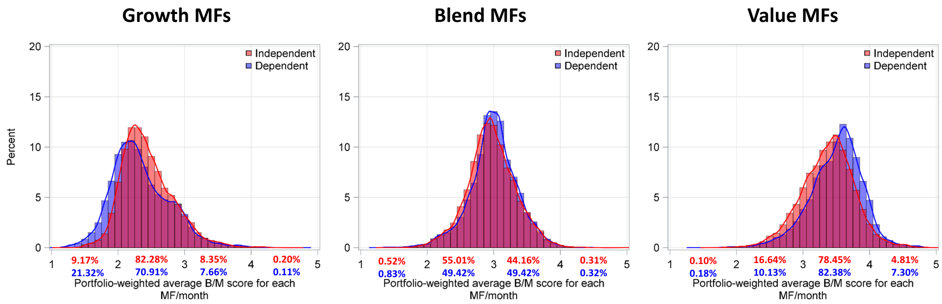


**Figure 3:** Distribution of BM-scores of Mutual Funds: Morningstar Category Breakdown

The figure presents the kernel distribution in BM-score of mutual funds with a focused on the US equity market for which Morningstar attributes size and value classifications. The value-growth classification applied to the mutual funds present in the CRSP mutual funds database. For each point in time where a fund reports its holdings, we associate a BM-score from a 1–5 scale according the Fama–French's 5x5 size and value *independent* sorting methodology or a 5x5 size and value *dependent* sorting methodology. The fund's BM-score is then calculated as the percentage of Total Net Assets (TNA) weighted average of the 1–5 scale of the securities the fund holds. Distributions are displayed for 3 Morningstar's value classifications of funds: growth (left), blend (middle), value (right) and 3 size classifications of funds: small (bottom), mid (middle), large (top). The sample period ranges from April 2002 to December 2015.

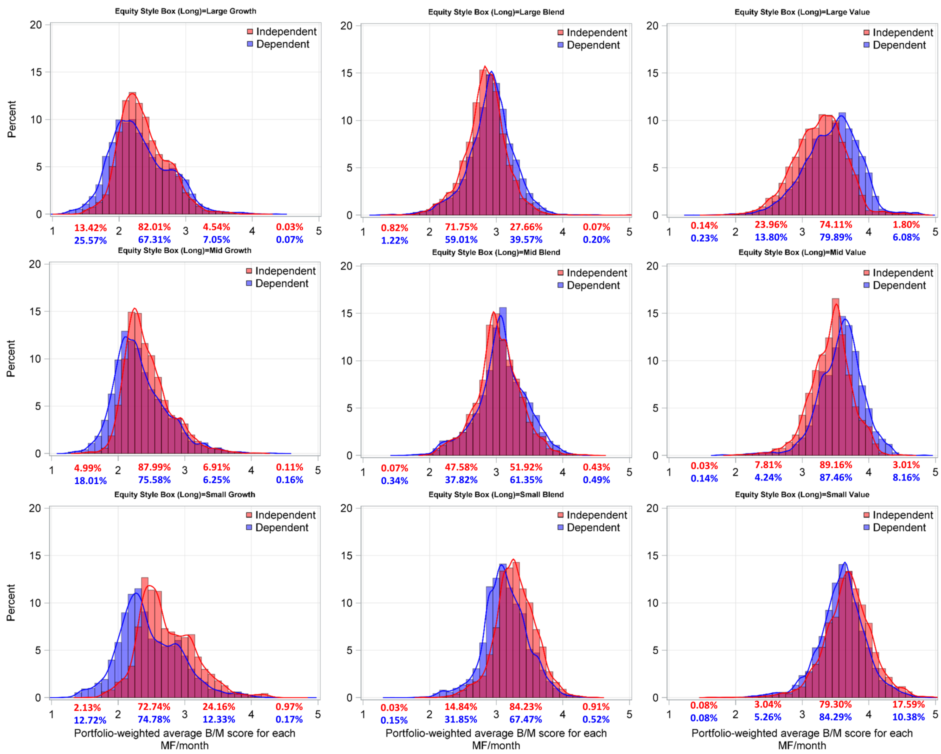


**Figure 4:** Stock Distribution with Independent vs Dependent Sorting

These plots show the stock distribution among the 2×3 and 3×3 characteristic-sorted portfolios based on size (low, medium and high) and the book-to-market equity ratio (low, medium and high) for the independent- and dependent-sorting methodologies. We also report the average percentage of stock distribution among the 3×3×3 characteristic-sorted portfolios when momentum is added as a third variable. For clarity, we group the 27 portfolios according to their size classifications (small, medium, and large). The period is the interval from July 1963 to December 2015.

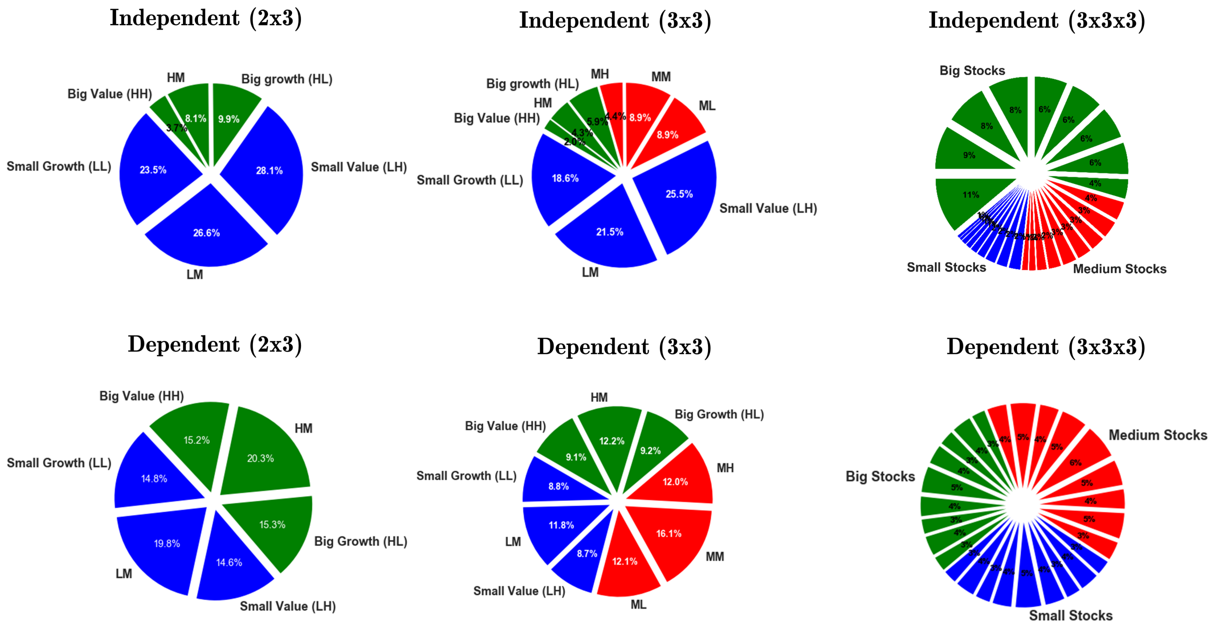


**Figure 5:** Variation of Transaction Cost Estimates

The figure presents a boxplot of the distribution of individual stock transaction costs estimated as in Hasbrouck (2009). The sample period is the interval from 1963 to 2015. The whiskers represent the distribution of the $5^{th}$ to $95^{th}$ percentile, and the upper and lower edges of the boxes correspond to the $25^{th}$ and $75^{th}$ percentiles. The gray dots represent outliers.

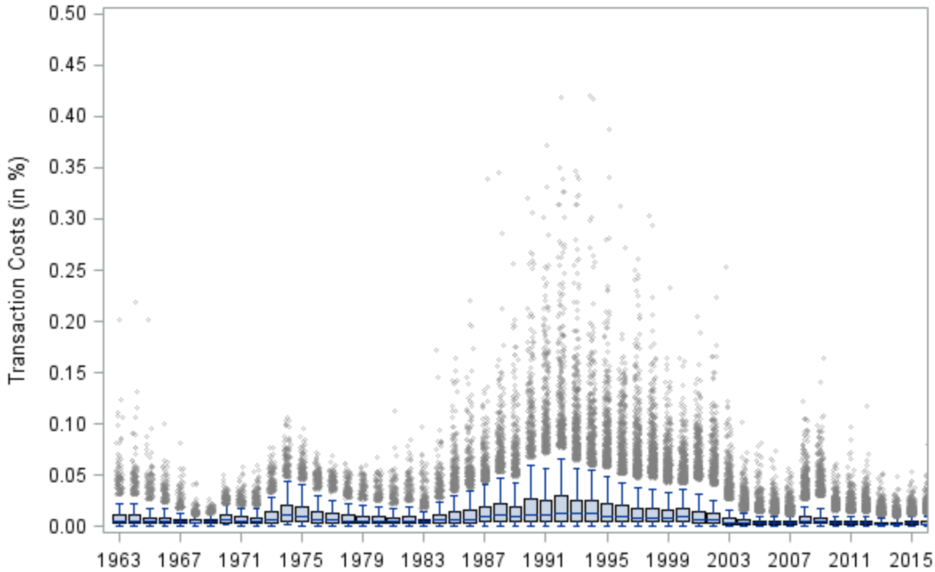
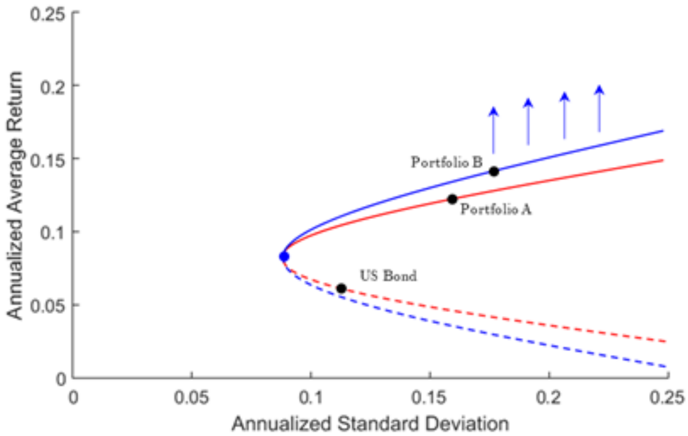


**Figure 6:** Improving the Tangency (**a**) and GMV (**b**) Portfolios

The figure displays the spanning illustration for opportunity sets comprising the benchmark assets ($R_1$), i.e., the 30-Year US Treasury Bond and Portfolio A, in the color red. The benchmark assets plus a test asset ($R_2$), i.e., Portfolio B, are displayed in the color blue. The $x$-axis reports the annualized standard deviation (in %), and the $y$-axis reports the annualized average return (in %). This example is fictitious but illustrates in Panel A (Panel B) an improvement of the tangency (GMV) portfolio after Portfolio B is added to the benchmark assets.

**(a)** Tangency Portfolio




**(b)** GMV Portfolio

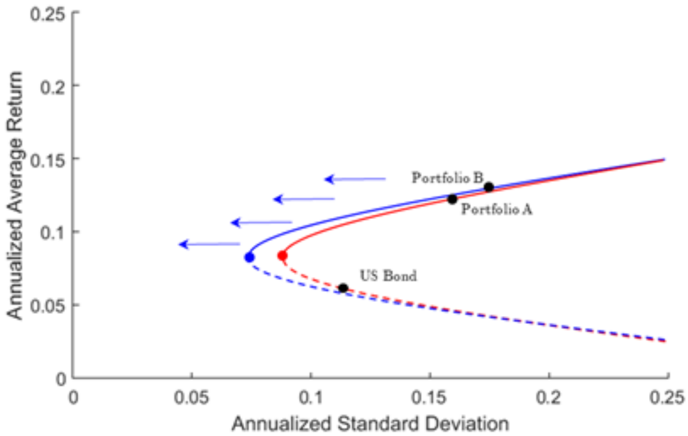


**Table 1:** Mutual Funds Score BM: 15 Largest Discrepancies

The table reports the 15 mutual funds with a focused on the US equity market for which the discrepancies between the BM-scores from the dependent and independent sorts are larger. We report in the first column the attributed Morningstar style classification of the funds; in the second column the ticker; in the third column displays the funds' name; columns four and five respectively report the BM-score under an independent sort and dependent sort; column six is the difference between $\text{BM}_{dep}$ and $\text{BM}_{ind}$; the last column report the average number of stocks of funds across the period going from April 2002 to December 2015.

| Morningstar Equity Style Box | Ticker | Fund Name | Score $\text{BM}_{Ind}$ | Score $\text{BM}_{Dep}$ | Score $\Delta$ BM | Average # of stocks |
|---|---|---|---|---|---|---|
| Large Value | NPLVX | NorthPointe Large Cap Value Investor | 2.713 | 3.232 | 0.519 | 23 |
| Large Growth | IMETX | ING Legg Mason ClearBridge Agrsv Gr S2 | 2.880 | 3.325 | 0.445 | 23 |
| Large Value | LLPFX | Longleaf Partners | 3.178 | 3.614 | 0.435 | 12 |
| Mid Growth | LSHAX | Kinetics Spin-Off and Corp Rest Adv A | 3.249 | 3.683 | 0.435 | 22 |
| Large Value | MVRAX | Monteagle Value A | 2.874 | 3.295 | 0.421 | 24 |
| Mid Blend | PVUCX | Principal MidCap Value II C | 2.528 | 2.944 | 0.416 | 61 |
| Large Value | LOPBX | DWS Dreman Concentrated Value B | 3.392 | 3.796 | 0.403 | 23 |
| Large Value | QALVX | Federated MDT Large Cap Value A | 3.360 | 3.758 | 0.398 | 90 |
| Large Value | LSVVX | LSV Conservative Value Equity | 3.152 | 3.547 | 0.395 | 76 |
| Large Value | FLVSX | Fidelity Series Large Cap Value | 3.130 | 3.522 | 0.393 | 122 |
| Large Value | ISVAX | ING Amer Cent Large Comp Value Adv | 2.866 | 3.240 | 0.374 | 82 |
| Large Value | DENVX | MassMutual Premier Disciplined Value Svc | 3.203 | 3.572 | 0.369 | 432 |
| Large Value | DFLVX | DFA US Large Cap Value I | 4.193 | 4.561 | 0.369 | 186 |
| Large Value | MSLVX | BlackRock Advantage Large Cap Val Svc | 3.312 | 3.670 | 0.358 | 91 |
| Large Value | MRLVX | BlackRock Advantage Large Cap Val R | 3.312 | 3.670 | 0.358 | 91 |

**Table 2:** Correlation Between Characteristic-Sorted Portfolios

The table reports the average correlation (in %) for the characteristic-sorted portfolios constructed using independent and dependent sorting methodologies. The third column specifies the difference in the average correlation between the independent and dependent sorting results. Correlations are estimated based on daily returns, and the sample period extends from 01/07/1963 to 31/12/2015.

| #Number of portfolios | Independent Sort (1) | Dependent Sort (2) | Difference (1)-(2) |
|---|---|---|---|
| | Panel A: Cap-weighted Portfolios | | |
| 2×3 | 84.99 | 78.00 | 6.99 |
| 3×3 | 84.99 | 75.81 | 9.18 |
| 3×3×3 | 78.38 | 66.8 | 11.58 |

**Table 3:** List of the Smart Beta Strategies' Objective Functions

The table decomposes the smart beta strategies' objective function and the constraints applied on the constituents' weights. The first column refers to the common name of the strategy. The second column specifies the main authors who have analyzed the strategy. The third column reports the objective function for minimization or maximization, and the last column shows the unleveraged long-only constraint applied to the constituents' weight. In the objective function, $w$ refers to the weights, $N$ is the total amount of assets introduced in the optimization, $i$ and $j$ denote the i-th asset and the j-th asset, $\sigma_{ij}$ is the covariance between the i-th asset and j-th asset, $p$ refers to portfolio, and $(\Sigma w)_i$ is the risk contribution of the i-th asset.

| Strategy | Referenced Authors | Objective Function | Constraints |
|---|---|---|---|
| Minimum Variance (MV) | Clarke, Silva, and Thorley (2013) | $\min f(w) = \sum_i^N \sum_j^N w_i \sigma_{ij} w_j$ | |
| Maximum Diversification (MD) | Choueifaty and Coignard (2008) | $\max f(w) = \frac{\sum_i^N w_i \sigma_i}{\sqrt{\sum_i^N \sum_j^N w_i \sigma_{ij} w_j}}$ | $w_i \in [0,1]$ and $\sum_{i=1}^N w_i = 1$ |
| Risk parity (RP) | Maillard, Roncalli, and Teïletche (2010) | $\min f(w) = \sum_i^N \sum_j^N (w_i \times (\Sigma w)_i - w_j \times (\Sigma w)_j)^2$ | |

**Table 4:** Diversification Returns: Fixed-Weight Benchmark

The table reports the spread of diversification return obtained from equation (4) for the three different strategic beta strategies: MD, MV, and RP. These strategies are applied to portfolios that are sorted independently (ind) or dependently (dep). These portfolios are rebalanced on a monthly basis and the number of portfolios is either six (2×3), nine (3×3) or twenty-seven (3×3×3). The components of diversification return (i.e., $DR_1$, $DR_2$, and DR) are reported in percentage and on a monthly basis. The sample period extends from July 1963 to December 2015. We then provide the p-value of the hypothesis that the spread in the component of diversification are equivalent for a pair of strategies applied on independent or dependent portfolios. To extract estimate a p-value for this static measures, we use the framework on hypothesis testing with the Sharpe ratio from Ledoit and Wolf (2008) and substitute the Sharpe ratio by the measures of diversification return.

Fixed-Weight (FW) Benchmark

| | $DR_1^{FW}$ | | | | $DR_2^{FW}$ | | | | $DR^{FW}$ | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Ind | Dep | $\Delta$ Dep-Ind | p-val | Ind | Dep | $\Delta$ Dep-Ind | p-val | Ind | Dep | $\Delta$ Dep-Ind | p-val |
| $MD_{2x3}$ | -0.005 | -0.003 | 0.002 | 0.936 | 0.025 | 0.038 | 0.012 | **0.000** | 0.020 | 0.034 | 0.014 | 0.524 |
| $MD_{3x3}$ | 0.004 | -0.032 | -0.035 | 0.257 | 0.031 | 0.050 | 0.019 | **0.000** | 0.034 | 0.018 | -0.016 | 0.594 |
| $MD_{3x3x3}$ | -0.041 | -0.075 | -0.034 | 0.428 | 0.046 | 0.080 | 0.034 | **0.000** | 0.005 | 0.005 | 0.000 | 0.993 |
| $MV_{2x3}$ | 0.013 | -0.025 | -0.038 | 0.554 | 0.024 | 0.031 | 0.007 | 0.165 | 0.037 | 0.006 | -0.031 | 0.635 |
| $MV_{3x3}$ | -0.001 | -0.036 | -0.035 | 0.581 | 0.024 | 0.045 | 0.020 | **0.003** | 0.023 | 0.009 | -0.014 | 0.807 |
| $MV_{3x3x3}$ | -0.018 | -0.115 | -0.097 | **0.093** | 0.047 | 0.070 | 0.023 | **0.001** | 0.029 | -0.045 | -0.074 | 0.191 |
| $RP_{2x3}$ | 0.005 | 0.005 | 0.000 | 0.992 | 0.024 | 0.036 | 0.012 | **0.000** | 0.030 | 0.041 | 0.012 | 0.308 |
| $RP_{3x3}$ | 0.005 | 0.000 | -0.005 | 0.679 | 0.027 | 0.042 | 0.014 | **0.000** | 0.032 | 0.041 | 0.009 | 0.460 |
| $RP_{3x3x3}$ | -0.001 | -0.010 | -0.009 | 0.522 | 0.043 | 0.065 | 0.023 | **0.000** | 0.041 | 0.056 | 0.014 | 0.307 |

**Table 5:** Diversification Returns: Equal-Weight Benchmark

The table reports the spread of diversification return obtained from equation (5) for the three different strategic beta strategies: MD, MV, and RP. These strategies are applied to portfolios that are sorted independently (ind) or dependently (dep). These portfolios are rebalanced on a monthly basis and the number of portfolios is either six (2×3), nine (3×3) or twenty-seven (3×3×3). The components of diversification return (i.e., $DR_1$, $DR_2$, and DR) are reported in percentage and on a monthly basis. The sample period extends from July 1963 to December 2015. We then provide the p-value of the hypothesis that the spread in the component of diversification are equivalent for a pair of strategies applied on independent or dependent portfolios. To extract estimate a p-value for this static measures, we use the framework on hypothesis testing with the Sharpe ratio from Ledoit and Wolf (2008) and substitute the Sharpe ratio by the measures of diversification return.

Equal-Weight (EW) Benchmark

| | $DR_1^{EW}$ | | | | $DR_2^{EW}$ | | | | $DR^{EW}$ | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Ind | Dep | $\Delta$ Dep-Ind | p-val | Ind | Dep | $\Delta$ Dep-Ind | p-val | Ind | Dep | $\Delta$ Dep-Ind | p-val |
| $MD_{2x3}$ | 0.005 | 0.088 | 0.083 | **0.012** | 0.027 | 0.038 | 0.011 | **0.003** | 0.032 | 0.126 | 0.094 | **0.007** |
| $MD_{3x3}$ | 0.010 | 0.070 | 0.060 | 0.125 | 0.033 | 0.048 | 0.015 | **0.000** | 0.043 | 0.118 | 0.075 | **0.055** |
| $MD_{3x3x3}$ | -0.034 | 0.063 | 0.097 | 0.107 | 0.046 | 0.061 | 0.014 | **0.010** | 0.012 | 0.124 | 0.112 | **0.067** |
| $MV_{2x3}$ | 0.082 | 0.218 | 0.136 | **0.071** | 0.033 | 0.028 | -0.004 | 0.407 | 0.115 | 0.246 | 0.131 | **0.075** |
| $MV_{3x3}$ | 0.050 | 0.143 | 0.093 | 0.163 | 0.034 | 0.046 | 0.012 | **0.066** | 0.084 | 0.189 | 0.105 | 0.112 |
| $MV_{3x3x3}$ | 0.022 | 0.019 | -0.003 | 0.955 | 0.061 | 0.080 | 0.018 | **0.002** | 0.084 | 0.099 | 0.015 | 0.785 |
| $RP_{2x3}$ | 0.020 | 0.056 | 0.036 | **0.013** | 0.026 | 0.036 | 0.010 | **0.001** | 0.046 | 0.092 | 0.047 | **0.005** |
| $RP_{3x3}$ | 0.019 | 0.048 | 0.029 | **0.063** | 0.029 | 0.042 | 0.013 | **0.000** | 0.048 | 0.090 | 0.041 | **0.015** |
| $RP_{3x3x3}$ | 0.011 | 0.035 | 0.024 | 0.102 | 0.046 | 0.068 | 0.022 | **0.000** | 0.057 | 0.102 | 0.046 | **0.007** |

**Table 6:** Spanning Tests with Mutliple Factor: Full Sample

The table reports the results for the bootstrap mean-variance spanning test from Kan and Zhou (2012). For one bootstrap, the mean-variance test goes as follow: we test whether a benchmark portfolio $R_1$ have a significant improvement at the tangent ($F_1$), or at the GMV ($F_2$) portfolio level when a test asset, i.e. $R_2$, is added to the benchmark assets ($R_1$). The test is performed twice, given that we have two proxies for $R_1$, that is a smart beta strategy on independent-sorted portfolio and the same smart beta strategy on dependent-sorted portfolios, respectively denoted with the subscript ind and dep. The outcomes of the test are the following, (i) the F-tests p-value from the original sample and, (ii) the bootstrap p-value that control for multiple testing. Results presented below are composed of 1,000 simulations for each smart beta strategy, i.e. maximum diversification (MD), minimum variance (MV), risk parity (RP). All strategies and explanatory variables are taken in excess of the risk-free rate except for the long-short size and value factors. B30 refers to the 30-Year US Treasury Bonds in excess of the risk-free rate. We also report the estimate of $\alpha$ from the OLS-regression of $R_1$ against $R_2$. The sample period ranges from July 1963 to December 2015.

| MVE candidates | $\alpha_{ind}$ | $F_{1,ind}$ | p-val | p-val$^b$ | $F_{2,ind}$ | p-val | p-val$^b$ | $\alpha_{dep}$ | $F_{1,dep}$ | p-val | p-val$^b$ | $F_{2,dep}$ | p-val | p-val$^b$ |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Panel A: $R_1 = MKT + B30, R_2 = SB, R_2^{orth} = SB^e + Mkt$ | | | | | | | | | | | | | | |
| $MD_{2x3}$ | 0.0020 | 11.69 | 0.00 | **0.00** | 2.37 | 0.12 | 0.31 | 0.0031 | 13.00 | 0.00 | **0.00** | 13.52 | 0.00 | 0.01 |
| $MD_{3x3}$ | 0.0022 | 11.89 | 0.00 | **0.00** | 2.96 | 0.09 | 0.25 | 0.0034 | 10.66 | 0.00 | **0.00** | 21.30 | 0.00 | 0.00 |
| $MD_{3x3x3}$ | 0.0019 | 7.21 | 0.01 | **0.01** | 1.42 | 0.23 | 0.58 | 0.0038 | 9.45 | 0.00 | **0.00** | 19.39 | 0.00 | 0.00 |
| $MV_{2x3}$ | 0.0032 | 13.63 | 0.00 | **0.00** | 11.83 | 0.00 | 0.01 | 0.0049 | 15.55 | 0.00 | **0.00** | 23.12 | 0.00 | 0.00 |
| $MV_{3x3}$ | 0.0029 | 10.60 | 0.00 | **0.00** | 6.68 | 0.01 | 0.07 | 0.0044 | 13.39 | 0.00 | **0.00** | 28.21 | 0.00 | 0.00 |
| $MV_{3x3x3}$ | 0.0029 | 13.60 | 0.00 | **0.00** | 10.24 | 0.00 | 0.02 | 0.0034 | 10.02 | 0.00 | **0.00** | 19.87 | 0.00 | 0.00 |
| $RP_{2x3}$ | 0.0021 | 13.03 | 0.00 | **0.00** | 1.62 | 0.20 | 0.42 | 0.0027 | 10.60 | 0.00 | **0.00** | 7.48 | 0.01 | 0.04 |
| $RP_{3x3}$ | 0.0022 | 11.57 | 0.00 | **0.00** | 0.49 | 0.48 | 0.72 | 0.0029 | 8.70 | 0.00 | **0.00** | 8.49 | 0.00 | 0.02 |
| $RP_{3x3x3}$ | 0.0023 | 12.09 | 0.00 | **0.00** | 0.25 | 0.62 | 0.84 | 0.0031 | 10.28 | 0.00 | **0.00** | 7.34 | 0.01 | 0.04 |
| Panel B: $R_1 = MKT + B30 + SMB + HML, R_2 = SB, R_2^{orth} = SB^e + Mkt$ | | | | | | | | | | | | | | |
| $MD_{2x3}$ | 0.0001 | 0.21 | 0.65 | 0.85 | 1419.96 | 0.00 | 0.00 | 0.0011 | 4.89 | 0.03 | **0.05** | 578.95 | 0.00 | 0.00 |
| $MD_{3x3}$ | 0.0002 | 0.66 | 0.42 | 0.63 | 1392.77 | 0.00 | 0.00 | 0.0011 | 3.13 | 0.08 | 0.14 | 365.03 | 0.00 | 0.00 |
| $MD_{3x3x3}$ | -0.0003 | 0.65 | 0.42 | 0.63 | 824.06 | 0.00 | 0.00 | 0.0012 | 2.22 | 0.14 | 0.22 | 285.69 | 0.00 | 0.00 |
| $MV_{2x3}$ | 0.0009 | 2.22 | 0.14 | 0.22 | 317.78 | 0.00 | 0.00 | 0.0022 | 6.21 | 0.01 | **0.02** | 249.22 | 0.00 | 0.00 |
| $MV_{3x3}$ | 0.0006 | 1.02 | 0.31 | 0.47 | 438.93 | 0.00 | 0.00 | 0.0018 | 4.59 | 0.03 | **0.06** | 254.99 | 0.00 | 0.00 |
| $MV_{3x3x3}$ | 0.0007 | 2.22 | 0.14 | 0.23 | 564.35 | 0.00 | 0.00 | 0.0010 | 1.92 | 0.17 | 0.28 | 310.91 | 0.00 | 0.00 |
| $RP_{2x3}$ | 0.0002 | 0.92 | 0.34 | 0.47 | 1786.24 | 0.00 | 0.00 | 0.0007 | 3.21 | 0.07 | 0.11 | 882.21 | 0.00 | 0.00 |
| $RP_{3x3}$ | 0.0002 | 0.88 | 0.35 | 0.52 | 2015.93 | 0.00 | 0.00 | 0.0007 | 2.36 | 0.12 | 0.20 | 801.67 | 0.00 | 0.00 |
| $RP_{3x3x3}$ | 0.0003 | 1.16 | 0.28 | 0.44 | 1895.22 | 0.00 | 0.00 | 0.0010 | 4.27 | 0.04 | **0.06** | 811.01 | 0.00 | 0.00 |
| Panel C: $R_1 = SB + B30, R_2 = Mkt, R_2^{orth} = Mkt^e + SB$ | | | | | | | | | | | | | | |
| $MD_{2x3}$ | -0.0013 | 5.92 | 0.02 | **0.01** | 7.93 | 0.01 | 0.03 | -0.0016 | 4.20 | 0.04 | **0.05** | 8.11 | 0.00 | 0.03 |
| $MD_{3x3}$ | -0.0014 | 5.51 | 0.02 | **0.02** | 9.40 | 0.00 | 0.02 | -0.0013 | 1.99 | 0.16 | 0.22 | 10.40 | 0.00 | 0.01 |
| $MD_{3x3x3}$ | -0.0010 | 2.22 | 0.14 | 0.18 | 17.25 | 0.00 | 0.01 | -0.0010 | 0.92 | 0.34 | 0.49 | 19.97 | 0.00 | 0.01 |
| $MV_{2x3}$ | -0.0017 | 3.97 | 0.05 | **0.05** | 12.51 | 0.00 | 0.01 | -0.0016 | 2.31 | 0.13 | 0.17 | 21.01 | 0.00 | 0.00 |
| $MV_{3x3}$ | -0.0014 | 2.63 | 0.11 | 0.13 | 18.31 | 0.00 | 0.00 | -0.0014 | 1.87 | 0.17 | 0.22 | 15.30 | 0.00 | 0.00 |
| $MV_{3x3x3}$ | -0.0017 | 4.60 | 0.03 | **0.05** | 10.31 | 0.00 | 0.01 | -0.0012 | 1.45 | 0.23 | 0.30 | 14.54 | 0.00 | 0.00 |
| $RP_{2x3}$ | -0.0014 | 6.82 | 0.01 | **0.01** | 9.52 | 0.00 | 0.02 | -0.0014 | 3.44 | 0.06 | **0.07** | 10.80 | 0.00 | 0.01 |
| $RP_{3x3}$ | -0.0014 | 5.35 | 0.02 | **0.03** | 16.33 | 0.00 | 0.00 | -0.0011 | 1.81 | 0.18 | 0.21 | 16.07 | 0.00 | 0.00 |
| $RP_{3x3x3}$ | -0.0014 | 5.50 | 0.02 | **0.02** | 19.14 | 0.00 | 0.00 | -0.0013 | 2.46 | 0.12 | 0.14 | 17.64 | 0.00 | 0.00 |

**Table 7:** Spanning Tests with Mutliple Factor: Sub Sample

The table reports the results for the bootstrap mean-variance spanning test from Kan and Zhou (2012). For one bootstrap, the mean-variance test goes as follow: we test whether a benchmark portfolio $R_1$ have a significant improvement at the tangent ($F_1$), or at the GMV ($F_2$) portfolio level when a test asset, i.e. $R_2$, is added to the benchmark assets ($R_1$). The test is performed twice, given that we have two proxies for $R_1$, that is a smart beta strategy on independent-sorted portfolio and the same smart beta strategy on dependent-sorted portfolios, respectively denoted with the subscript ind and dep. The outcomes of the test are the following, (i) the F-tests p-value from the original sample and, (ii) the bootstrap p-value that control for multiple testing. Results presented below are composed of 1,000 simulations for each smart beta strategy, i.e. maximum diversification (MD), minimum variance (MV), risk parity (RP). All strategies and explanatory variables are taken in excess of the risk-free rate except for the long-short size and value factors. B30 refers to the 30-Year US Treasury Bonds in excess of the risk-free rate. We also report the estimate of $\alpha$ from the OLS-regression of $R_1$ against $R_2$. The sample period ranges from July 1993 to December 2015.

| MVE candidates | $\alpha_{ind}$ | $F_{1,ind}$ | p-val | p-val$^b$ | $F_{2,ind}$ | p-val | p-val$^b$ | $\alpha_{dep}$ | $F_{1,dep}$ | p-val | p-val$^b$ | $F_{2,dep}$ | p-val | p-val$^b$ |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Panel A: $R_1 = MKT + B30, R_2 = SB, R_2^{orth} = SB^e + Mkt$ | | | | | | | | | | | | | | |
| $MD_{2x3}$ | 0.0019 | 3.77 | 0.05 | 0.11 | 6.48 | 0.01 | 0.09 | 0.0048 | 12.70 | 0.00 | **0.00** | 16.12 | 0.00 | 0.00 |
| $MD_{3x3}$ | 0.0024 | 5.08 | 0.03 | **0.06** | 7.26 | 0.01 | 0.05 | 0.0049 | 9.14 | 0.00 | **0.01** | 16.55 | 0.00 | 0.00 |
| $MD_{3x3x3}$ | 0.0016 | 1.86 | 0.17 | 0.33 | 5.87 | 0.02 | 0.17 | 0.0053 | 9.38 | 0.00 | **0.01** | 16.66 | 0.00 | 0.02 |
| $MV_{2x3}$ | 0.0039 | 9.03 | 0.00 | **0.00** | 19.93 | 0.00 | 0.00 | 0.0071 | 14.13 | 0.00 | **0.00** | 17.80 | 0.00 | 0.00 |
| $MV_{3x3}$ | 0.0038 | 7.55 | 0.01 | **0.02** | 11.39 | 0.00 | 0.01 | 0.0066 | 13.10 | 0.00 | **0.00** | 21.37 | 0.00 | 0.00 |
| $MV_{3x3x3}$ | 0.0042 | 11.36 | 0.00 | **0.00** | 18.15 | 0.00 | 0.00 | 0.0050 | 9.92 | 0.00 | **0.00** | 19.77 | 0.00 | 0.00 |
| $RP_{2x3}$ | 0.0021 | 4.73 | 0.03 | **0.07** | 4.94 | 0.03 | 0.12 | 0.0039 | 9.39 | 0.00 | **0.01** | 9.07 | 0.00 | 0.03 |
| $RP_{3x3}$ | 0.0024 | 5.02 | 0.03 | **0.05** | 3.44 | 0.06 | 0.21 | 0.0039 | 6.64 | 0.01 | **0.02** | 7.98 | 0.01 | 0.05 |
| $RP_{3x3x3}$ | 0.0027 | 6.19 | 0.01 | **0.03** | 4.37 | 0.04 | 0.13 | 0.0043 | 8.69 | 0.00 | **0.01** | 9.42 | 0.00 | 0.02 |
| Panel B: $R_1 = MKT + B30 + SMB + HML, R_2 = SB, R_2^{orth} = SB^e + Mkt$ | | | | | | | | | | | | | | |
| $MD_{2x3}$ | 0.0006 | 1.52 | 0.22 | 0.40 | 384.37 | 0.00 | 0.00 | 0.0036 | 16.78 | 0.00 | **0.00** | 118.87 | 0.00 | 0.00 |
| $MD_{3x3}$ | 0.0010 | 3.71 | 0.06 | 0.11 | 382.82 | 0.00 | 0.00 | 0.0036 | 9.31 | 0.00 | **0.00** | 63.08 | 0.00 | 0.00 |
| $MD_{3x3x3}$ | 0.0001 | 0.02 | 0.88 | 0.99 | 249.23 | 0.00 | 0.00 | 0.0041 | 8.88 | 0.00 | **0.01** | 39.76 | 0.00 | 0.00 |
| $MV_{2x3}$ | 0.0027 | 6.52 | 0.01 | **0.03** | 42.73 | 0.00 | 0.00 | 0.0056 | 14.39 | 0.00 | **0.00** | 44.38 | 0.00 | 0.00 |
| $MV_{3x3}$ | 0.0025 | 5.89 | 0.02 | **0.02** | 92.31 | 0.00 | 0.00 | 0.0052 | 13.19 | 0.00 | **0.00** | 43.67 | 0.00 | 0.00 |
| $MV_{3x3x3}$ | 0.0028 | 11.43 | 0.00 | **0.00** | 124.93 | 0.00 | 0.00 | 0.0038 | 9.86 | 0.00 | **0.00** | 51.57 | 0.00 | 0.00 |
| $RP_{2x3}$ | 0.0008 | 3.50 | 0.06 | 0.11 | 563.63 | 0.00 | 0.00 | 0.0026 | 14.80 | 0.00 | **0.00** | 217.80 | 0.00 | 0.00 |
| $RP_{3x3}$ | 0.0010 | 5.07 | 0.03 | **0.05** | 652.92 | 0.00 | 0.00 | 0.0026 | 9.23 | 0.00 | **0.00** | 174.22 | 0.00 | 0.00 |
| $RP_{3x3x3}$ | 0.0013 | 7.72 | 0.01 | **0.02** | 618.11 | 0.00 | 0.00 | 0.0031 | 13.91 | 0.00 | **0.00** | 158.94 | 0.00 | 0.00 |
| Panel C: $R_1 = SB + B30, R_2 = Mkt, R_2^{orth} = Mkt^e + SB$ | | | | | | | | | | | | | | |
| $MD_{2x3}$ | -0.0008 | 0.74 | 0.39 | 0.56 | 3.25 | 0.07 | 0.16 | -0.0024 | 3.02 | 0.08 | 0.14 | 5.50 | 0.02 | 0.06 |
| $MD_{3x3}$ | -0.0011 | 1.10 | 0.30 | 0.44 | 4.15 | 0.04 | 0.11 | -0.0015 | 0.89 | 0.35 | 0.50 | 11.18 | 0.00 | 0.01 |
| $MD_{3x3x3}$ | -0.0002 | 0.02 | 0.89 | 0.99 | 7.24 | 0.01 | 0.07 | -0.0013 | 0.65 | 0.42 | 0.65 | 15.00 | 0.00 | 0.01 |
| $MV_{2x3}$ | -0.0016 | 1.35 | 0.25 | 0.40 | 5.11 | 0.02 | 0.07 | -0.0019 | 1.08 | 0.30 | 0.46 | 21.32 | 0.00 | 0.00 |
| $MV_{3x3}$ | -0.0013 | 0.88 | 0.35 | 0.51 | 10.74 | 0.00 | 0.01 | -0.0018 | 1.04 | 0.31 | 0.46 | 16.16 | 0.00 | 0.00 |
| $MV_{3x3x3}$ | -0.0021 | 2.54 | 0.11 | 0.16 | 4.27 | 0.04 | 0.07 | -0.0015 | 0.89 | 0.35 | 0.49 | 11.11 | 0.00 | 0.00 |
| $RP_{2x3}$ | -0.0010 | 1.15 | 0.28 | 0.39 | 4.53 | 0.03 | 0.07 | -0.0019 | 2.27 | 0.13 | 0.19 | 7.03 | 0.01 | 0.02 |
| $RP_{3x3}$ | -0.0011 | 1.10 | 0.30 | 0.40 | 7.88 | 0.01 | 0.02 | -0.0012 | 0.70 | 0.40 | 0.52 | 13.32 | 0.00 | 0.00 |
| $RP_{3x3x3}$ | -0.0013 | 1.53 | 0.22 | 0.30 | 7.30 | 0.01 | 0.03 | -0.0016 | 1.34 | 0.25 | 0.34 | 11.52 | 0.00 | 0.01 |

**Table 8:** Horse Race Boostrap Test

Smart beta strategies are constructed on CW portfolios. The table reports the results for the bootstrap mean-variance spanning test from Kan and Zhou (2012). For one bootstrap, the mean-variance test goes as follow: we test whether a benchmark portfolio $R_1$ have a significant improvement at the tangent ($F_1$), or at the GMV ($F_2$) portfolio level when a test asset, i.e. $R_2$, is added to the benchmark assets ($R_1$). The test is performed twice, given that we have two proxies for $R_1$, that is a smart beta strategy on independent-sorted portfolio and the same smart beta strategy on dependent-sorted portfolios, respectively denoted with the subscript *ind* and *dep*. Here, we only report the bootstrap p-values that control for multiple testing which are more conservative that the p-values of the original sample. Results presented below are composed of 1,000 simulations for each smart beta strategy, i.e. maximum diversification (MD), minimum variance (MV), risk parity (RP). All smart strategies are taken in excess of the risk-free rate and net of transactions costs. B30 refers to the 30-Year US Treasury Bonds in excess of the risk-free rate. In Panel A, $R_1$ is composed of B30 and a proxy for the MVE market portfolio ($R_{1,MVE}$) defined as a smart beta strategy constructed on dependent-sorted portfolios while $R_2$ is the same smart beta strategy constructed on independent-sorted portfolios. In Panel A, $R_1$ is composed of B30 and a proxy for the MVE market portfolio ($R_{1,MVE}$) defined as a smart beta strategy constructed on independent-sorted portfolios while $R_2$ is the same smart beta strategy constructed on dependent-sorted portfolios. The sample period ranges from July 1963 to December 2015.

| | Panel A: $\mathrm{R}_1 = B30 + SB^{net}_{dep}$ $\mathrm{R}_2 = SB^{net}_{ind}$ | | | | Panel B: $\mathrm{R}_1 = B30 + SB^{net}_{ind}$ $\mathrm{R}_2 = SB^{net}_{dep}$ | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | $\mathrm{F}^o_{1,ind}$ | p-val$^b$ | $\mathrm{F}^o_{2,ind}$ | p-val$^b$ | $\mathrm{F}^o_{1,dep}$ | p-val$^b$ | $\mathrm{F}^o_{2,dep}$ | p-val$^b$ | MVE Candidate | GMV Candidate |
| $\mathrm{MD}_{2x3}$ | 0.801 | 0.448 | 0.064 | 0.944 | 5.823 | 0.012 | 11.175 | 0.013 | Dependent | Dependent |
| $\mathrm{MD}_{3x3}$ | 0.016 | 0.973 | 0.121 | 0.914 | 4.160 | 0.049 | 17.576 | 0.000 | Dependent | Dependent |
| $\mathrm{MD}_{3x3x3}$ | 0.001 | 1.000 | 2.470 | 0.361 | 5.959 | 0.026 | 18.900 | 0.005 | Dependent | Dependent |
| $\mathrm{MV}_{2x3}$ | 0.118 | 0.871 | 6.515 | 0.054 | 5.254 | 0.015 | 10.170 | 0.015 | Dependent | Dependent |
| $\mathrm{MV}_{3x3}$ | 0.035 | 0.956 | 0.090 | 0.954 | 6.095 | 0.011 | 24.206 | 0.000 | Dependent | Dependent |
| $\mathrm{MV}_{3x3x3}$ | 1.490 | 0.285 | 3.210 | 0.178 | 0.788 | 0.471 | 9.723 | 0.010 | Dep $\approx$ Ind | Dependent |
| $\mathrm{RP}_{2x3}$ | 0.006 | 0.984 | 0.299 | 0.773 | 1.743 | 0.173 | 5.814 | 0.061 | Dep $\approx$ Ind | Dependent |
| $\mathrm{RP}_{3x3}$ | 0.035 | 0.931 | 0.116 | 0.885 | 1.497 | 0.204 | 9.125 | 0.019 | Dep $\approx$ Ind | Dependent |
| $\mathrm{RP}_{3x3x3}$ | 0.009 | 0.984 | 0.087 | 0.922 | 2.315 | 0.109 | 9.240 | 0.021 | Dep $\approx$ Ind | Dependent |

**Table 9:** Lucky MVE Candidates

The table reports sequential test developed by Harvey and Liu (2016) to select robust MVE candidates. As MVE candidates, we use the traditional cap-weighted market portfolio obtained from Kenneth French's website, and the smart beta strategies constructed on the two sets of basis portfolios: one from an (ind)ependent sort and the other from a (dep)endent sort. All strategies are taken in excess of the risk-free rate and net of transactions costs. The test assets used in the MVE test are the basis portfolios used to construct the smart beta strategies. These ones are indicated next to the referenced Panel. The baseline model used in our test is the 30-Year US Treasury Bonds in excess of the risk-free rate (B30), and the size (SMB) and value (HML) factors obtained from Kenneth French's website. We report in the first row of each panels the Gibbons, Ross, and Shanken (1989) test and in the second row its respective p-value; the third row presents the value of the scaled intercept (SI); the fourth row displays the single test p-value for SI; the fifth row shows the numbering sequence for which SI and its p-values are referring; the penultimate row reports the selected MVE candidates among the smart beta strategies; and the final row displays the p-value of the selected candidate when controlling for multiple testing. The sample period ranges from July 1963 to December 2015.

| MVE Candidates → | Mkt | MV$_{dep}$ | MD$_{dep}$ | RP$_{dep}$ | MV$_{ind}$ | MD$_{ind}$ | RP$_{ind}$ | Mkt | MV$_{dep}$ | MD$_{dep}$ | RP$_{dep}$ | MV$_{ind}$ | MD$_{ind}$ | RP$_{ind}$ |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Baseline = B30 + SMB + HML | | | | | | | Baseline = B30 + SMB + HML | | | | | | |
| | Panel A: 2x3 cap-weighted independent portfolios as test assets | | | | | | | Panel B: 2x3 cap-weighted dependent portfolios as test assets | | | | | | |
| GRS | 4.836 | 4.189 | 4.391 | 4.538 | 4.155 | 4.445 | 4.341 | 12.058 | 10.947 | 11.344 | 11.589 | 11.552 | 11.888 | 11.962 |
| p-value | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 |
| Scaled intercept (SI) | 0.042 | 0.070 | 0.041 | 0.036 | 0.066 | -0.012 | -0.842 | 0.049 | 0.061 | -0.838 | 0.023 | 0.009 | -0.001 | -0.008 |
| Single test p-value | 0.701 | 0.984 | 0.929 | 0.894 | 0.893 | 0.325 | 0.000 | 0.880 | 0.865 | 0.000 | 0.782 | 0.569 | 0.547 | 0.468 |
| SI sequence | 2 | 2 | 2 | 2 | 2 | 2 | 1 | 2 | 2 | 1 | 2 | 2 | 2 | 2 |
| Selected candidate(s) | | | | **RP**$_{ind}$ | | | | | | | **MD**$_{dep}$ | | | |
| Multiple test p-value | | | | [0.000] | | | | | | | [0.000] | | | |
| | Panel C: 3x3 cap-weighted independent portfolios as test assets | | | | | | | Panel D: 3x3 cap-weighted dependent portfolios as test assets | | | | | | |
| GRS | 3.403 | 2.792 | 2.892 | 2.975 | 2.909 | 2.918 | 2.929 | 8.228 | 7.757 | 7.959 | 7.950 | 8.116 | 8.108 | 8.108 |
| p-value | 0.000 | 0.003 | 0.002 | 0.002 | 0.002 | 0.002 | 0.002 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 |
| Scaled intercept (SI) | -0.006 | 0.041 | 0.255 | 0.373 | -0.876 | 0.100 | 0.169 | 0.210 | 0.249 | 0.221 | -0.806 | 0.001 | 0.150 | 0.107 |
| Single test p-value | 0.542 | 0.755 | 0.883 | 0.895 | 0.000 | 0.697 | 0.755 | 0.937 | 0.979 | 0.916 | 0.000 | 0.528 | 0.861 | 0.849 |
| SI sequence | 2 | 2 | 2 | 2 | 1 | 2 | 2 | 2 | 2 | 2 | 1 | 2 | 2 | 2 |
| Selected candidate(s) | | | | **MV**$_{ind}$ | | | | | | | **RP**$_{dep}$ | | | |
| Multiple test p-value | | | | [0.000] | | | | | | | [0.000] | | | |
| | Panel E: 3x3x3 cap-weighted independent portfolios as test assets | | | | | | | Panel F: 3x3x3 cap-weighted dependent portfolios as test assets | | | | | | |
| GRS | 2.262 | 2.135 | 2.164 | 2.213 | 2.085 | 2.332 | 2.110 | 4.591 | 4.366 | 4.335 | 4.283 | 4.335 | 4.608 | 4.379 |
| p-value | 0.000 | 0.001 | 0.001 | 0.000 | 0.001 | 0.000 | 0.001 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 |
| Scaled intercept (SI) | 0.245 | 0.030 | -0.790 | 0.301 | 0.283 | 0.111 | 0.225 | -0.015 | 0.007 | -0.668 | -0.039 | 0.006 | 0.013 | 0.020 |
| Single test p-value | 0.943 | 0.661 | 0.000 | 0.970 | 0.962 | 0.816 | 0.952 | 0.566 | 0.571 | 0.000 | 0.431 | 0.610 | 0.686 | 0.739 |
| SI sequence | 2 | 2 | 1 | 2 | 2 | 2 | 2 | 2 | 2 | 1 | 2 | 2 | 2 | 2 |
| Selected candidate(s) | | | | **MD**$_{dep}$ | | | | | | | **MD**$_{dep}$ | | | |
| Multiple test p-value | | | | [0.000] | | | | | | | [0.000] | | | |

# A. Appendices

## A.1. Estimation of the Covariance Matrix

In this section, we briefly describe a shrinkage methodology used in our applications to estimate the covariance with lower sampling errors following Ledoit and Wolf (2004). In their model, the authors build on Elton and Gruber (1973), who use a constant correlation coefficient to shrink the assets' covariance toward a global average correlation estimator.

The constant correlation coefficient is determined using

$$\hat{\rho} = \frac{1}{N(N-1)} \left( \sum_{i}^{N} \sum_{j}^{N} \hat{\rho}_{ij} - N \right) \tag{A.1}$$

where $N$ is the number of portfolios – in our applications, either 6, 9 or 27. The term $\hat{\rho}_{ij}$ is the historical correlation estimate between the $i^{th}$ portfolio and the $j^{th}$ portfolio. Ledoit and Wolf (2004) then obtain an optimal structure for the covariance matrix and reduce the sampling error of a traditional sample covariance matrix ($S$) as follows:

$$\Sigma = \delta F + (1 - \delta) S \tag{A.2}$$

where $\Sigma$ is the output covariance matrix obtained from the shrinkage estimation, and $\delta$ is the optimal shrinkage intensity.[13] $S$ is the sample covariance matrix from our 60 daily returns, and $F$ is the structured covariance matrix with the assets' covariance estimated via the constant correlation estimator in equation (A.1).[14] In our empirical study, the estimations of the sample and the structured covariance matrices are based on 60-day rolling windows to accommodate for gradual changes in the return distribution and short-term variations. A real-life application with tradable assets (Idzorek and Kowara (2013)) would impose constraints on the historical information available to replicate our results. For this reason - to stay as close as possible to what real-world applications may offer - we limit our optimizations on 60-day windows. This choice is also consistent with Fama and French (2018) who estimate the monthly variance of stocks using 60 days of lagged returns.

[13] Matlab code is available at Prof. Wolf's website.

[14] The covariance of the matrix $F$ is given by $\sigma_{ij} = \hat{\rho}\sigma_i\sigma_j$.

## A.2. Testing the Incremental Diversification Return

Ledoit and Wolf (2008) propose an 'indirect' bootstrap methodology to construct an empirical distribution of the spread in a function of the underlying first and second moments of two time series. They test the significance of the spread by considering whether a 1-$\alpha$ confidence interval (e.g. 90%) contains zero.

They first consider that the difference between the true first and second moments of the two series converge towards their sample estimate such that,

$$\sqrt{T}(\hat{u} - u) \xrightarrow{d} N(0, \Omega) \tag{A.3}$$

where $\hat{u} = (\hat{\mu_i}, \hat{\mu_j}, \hat{\sigma_i}^2, \hat{\sigma_j}^2)$ are the sample estimates of $u = (\mu_i, \mu_j, \sigma_i^2, \sigma_j^2)$, $\xrightarrow{d}$ refers to the convergence in distribution of the parameters, $T$ is the length of the time-series, and $\Omega$ refers to variance of the estimator distribution.

Considering the sample uncentered second moments instead of the sample estimated variances, i.e. $\hat{\gamma_i} = E(r_i^2)$ and $\hat{\gamma_i} = E(r_j^2)$, and taking into account non-normality and auto-correlation in returns, the relationship (A.3) becomes

$$\sqrt{T}(\hat{v} - v) \xrightarrow{d} N(0, \Psi) \tag{A.4}$$

where $\hat{v} = (\hat{\mu_i}, \hat{\mu_j}, \hat{\gamma_i}, \hat{\gamma_j})$ is the sample estimates of $v = (\mu_i, \mu_j, \gamma_i, \gamma_j)$.

The estimator $\Psi$ is estimated through an heteroskedasticity and autocorrelation (HAC) robust kernel method. We refer to the paper of Ledoit and Wolf (2008) for a more detailed discussion on the computation of this estimator.

The standard error of the spread $\hat{\Delta}$ in a function $f(\hat{v})$ can be defined as,

$$s(\hat{\Delta}) = \sqrt{\frac{\nabla' f(\hat{v}) \hat{\Psi} \nabla f(\hat{v})}{T}} \tag{A.5}$$

where $\nabla' f(\hat{v})$ is the gradient function of $f(\hat{v})$ and $T$ is the length of the time-series.

To obtain a confidence interval attached to $\hat{\Delta}$, we resample the original time-series using the block-bootstrap method of Politis and Romano (1992) and construct an empirical (bootstrap) distribution of a studentized test statistic ($d^b$) defined as

$$d^b = \frac{|\hat{\Delta}^b - \hat{\Delta}|}{s(\hat{\Delta}^b)} \tag{A.6}$$

where the superscript $b$ denotes the b-th bootstrap sample and where $s(\hat{\Delta^b})$, for the b-th bootstrap is obtained by using both the gradient of $f(\hat{v}^b)$ and the HAC kernel estimator $\hat{\Psi}^b$ and defined as follows,

$$s(\hat{\Delta}^b) = \sqrt{\frac{\nabla' f(\hat{v}^b)\hat{\Psi}^b \nabla f(\hat{v}^b)}{T}} \tag{A.7}$$

The boostrap 1-$\alpha$ confidence interval is defined as:

$$\left[\hat{\Delta} - z^b_{|.|,1-\alpha/2} s(\hat{\Delta}), \hat{\Delta} + z^b_{|.|,1-\alpha/2} s(\hat{\Delta})\right] \tag{A.8}$$

with $z^b_{|.|,1-\alpha}$ the quantile of the distribution function of the studentized statistic estimated from the bootstrap and denoted $\mathcal{L}(d^b)$.

In our applications, we use a block-bootstrap of 10 observations and runs 4999 simulations.[15] The bootstrap process works as follow: First, we set a length for the block of observations (e.g. 10) that we want to resample in order to capture serial autocorrelation. Second, we match the length of the original time-series in the bootstrap samples to preserve the uncertainty and the degree of freedom from the original data. Third, we randomly resample (with replacement) the sequence of time-series for the b-th bootstrap and keep the same sequence for resampling the time-series of the strategies and their underlying opportunity sets. This way we make sure to preserve the cross-sectional correlation across the assets (Fama and French (2010) and Harvey and Liu (2016)). Lastly, we repeat the operation B times (e.g. 4999) to construct an empirical distribution of centered studentized test statistics in which the standard error,

Defining a studentized test statistic ($d$) on the original time-series as follows,

$$d = \frac{|\hat{\Delta}|}{s(\hat{\Delta})} \tag{A.9}$$

[15]Our results are not sensitive to the choice of the block length. As a matter of fact, we run the test with blocks of length $\{2, 4, 6, 8, 10\}$ and found very similar results (available upon request). Also, we run 4999 simulations to stay aligned with the recommendations of Ledoit and Wolf (2008, p. 858).

The p-value attached to the test of the spread $\hat{\Delta}$ in a function $f(\hat{v})$ is computed as,

$$\text{p-val} = \frac{\#\{d^b \geq d\} + 1}{B + 1} \tag{A.10}$$

Ledoit and Wolf (2008) apply this framework to a test of Sharpe ratio for a pair of strategies. They consider the following function

$$f(a, b, c, d) = \frac{a}{\sqrt{c - a^2}} - \frac{b}{\sqrt{d - b^2}} \tag{A.11}$$

Where $a = \hat{\mu_i}$, $b = \hat{\mu_j}$, $c = \hat{\gamma_i}$, and $d = \hat{\gamma_j}$.

The gradient of the function is defined as $\nabla' f(\hat{v}) = (\frac{c}{(c-a^2)^{1.5}}, -\frac{d}{(d-b^2)^{1.5}}, -\frac{1}{2}\frac{a}{(c-a^2)^{1.5}}, \frac{1}{2}\frac{b}{(d-b^2)^{1.5}})$ is the gradient function of $f(\hat{v})$.

Our estimates of $\hat{\Delta}$ are defined as follows,

$$\begin{aligned} \hat{\Delta}(DR) &= DR^{Dep} - DR^{Ind} \\ \hat{\Delta}(DR_1) &= DR_1^{Dep} - DR_1^{Ind} \\ \hat{\Delta}(DR_2) &= DR_2^{Dep} - DR_2^{Ind} \end{aligned} \tag{A.12}$$

However obtaining the gradient for these functions is cumbersome because additionally to the pair of strategies, we are left with a number $N$ of dependent-sorted and independent-sorted portfolios. Given that in our applications this amount $N$ can take the value of 6, 9 or 27, finding the gradient for this large amount of parameters is difficult. Consequently, we take the assumptions that if there is large deviations in the spread of diversification return for a pair of strategies then there should also be large deviations in their spread of Sharpe ratio. This assumption is helpful as we can now only substitute the numerator in equations (A.6) and (A.9) by the spread in diversification return while keeping the standard error from the spread in Sharpe ratio derived in the initial framework of Ledoit and Wolf (2008). To test the sensitivity of this assumption, we also substitute the standard error from the spread in Sharpe ratio by the standard error from the spread in geometric return, and also by the standard error of the spread in geometric return scaled by the standard deviation. Their gradient function ($\nabla' f(\hat{v})$) are respectively given by $(a + 1, -b - 1, -0.5, 0.5)$ and $(\frac{0.5a^3-0.5ac-c}{a^2-c\sqrt{c-a^2}}, -\frac{0.5b^3+0.5bd+d}{b^2-d\sqrt{d-b^2}}, \frac{-0.25a^2+0.5a+0.25c}{a^2-c\sqrt{c-a^2}}, \frac{0.25b^2-0.5b-0.25d}{b^2-d\sqrt{d-b^2}})$. We obtained results qualitatively similar under all robustness tests. Results are available upon request.

## A.3. Transaction Costs: Gibbs Estimates

A traditional model to estimate trading costs of a security is documented by Roll (1984) and simply use the autocovariance of the change in trade price ($\Delta p_t$) to find an effective estimate of spread such that,

$$c^{Roll} = \begin{cases} \sqrt{-Cov(\Delta p_t, \Delta p_{t-1})} & \text{if } Cov(\Delta p_t, \Delta p_{t-1}) < 0 \\ 0 & \text{if } Cov(\Delta p_t, \Delta p_{t-1}) \geq 0 \end{cases} \tag{A.13}$$

In last equation, we see that when the autocovariance is positive the model fails to provide a fair estimate of effective costs. For this reason, Hasbrouck (2009) extend the measure under Roll (1984)'s framework on the price dynamics in a market with transaction costs. In this framework, the model only requires information about the daily trade price, the prior midpoint of the bid-ask prices and the sign of trade to perform the estimation. Formally, the price dynamic is written as follows:

$$\begin{aligned} m_t &= m_{(t-1)} + u_t \\ p_t &= m_t + cq_t \end{aligned} \tag{A.14}$$

where $m_t$ is the log midpoint of the prior bid-ask price (the efficient price), $p_t$ is the log trade price (the real price), $q_t$ is the sign of the last trade of the day ($+1$ for a buy and $-1$ for a sale), $c$ is the effective cost, and $u_t$ is assumed to be unrelated to the sign of the trade ($q_t$).

Since we use the logarithm for the price variables in equation (A.14), the daily change in price is given by

$$\begin{aligned} \Delta p_t &= p_t - p_{t-1} \\ &= m_t + cq_t - m_{t-1} - cq_{t-1} \\ &= c\Delta q_t + u_t \end{aligned} \tag{A.15}$$

Hasbrouck (2009) extends Roll (1984)'s model with a market factor to capture a larger part of the changes in prices not due to transaction costs. They estimate the effective trading costs using Bayesian Gibbs sampling applied to the daily prices of U.S. equities retrieved from CRSP data.[16] The market-factor model is presented as follows:

$$\Delta p_t = c\Delta q_t + \beta_{rm} rm_t + u_t \tag{A.16}$$

where $rm_t$ is the market return on day $t$ and $\beta_{rm}$ is the parameter estimate obtained from a Bayesian regression on the market return.

The Bayesian methodology estimates the effective costs ($c$) based on a sequence of iterations where the initial prior for $c$ is strictly positive and follows a normal distribution with mean of 0.01 and variance equal to $0.01^2$, denoted $N^+$ ($\mu = 0.01$, $\sigma^2 = 0.01^2$). This initial prior of $\beta_{rm}$ follows a normal distribution with mean and variance of 1, i.e. $N$ ($\mu = 1$, $\sigma^2 = 1$) and the prior of $\sigma_u^2$ follows an inverted Gamma distribution initiated at $IG$ ($\alpha = 10^{-12}$, $\beta = 10^{-12}$).[17] The objective of the Gibbs sampling is to estimate the value of the parameters $c$ and $\beta_{rm}$ conditional on the values drawn for $q_t$, which is based on the sign of trade ($\Delta p_t$), and the error term ($u_t$). Initially, $q_1$ is set to +1 and $\sigma_u^2$ is set to 0.001. Next, the sampler runs as follow,

**for 1 to 1,000 sweeps**

1. Perform a Bayesian OLS regression on a 250-day of lagged observations to estimate the new values of $c$ and $\beta_{rm}$, update the posterior distribution of the parameters and make a new draw of the coefficients.

2. Back out $u_t$ from the values of $c$, $\beta_{rm}$, $\Delta p_t$, $rm_t$, $\Delta q_t$ as follow,

$$u_t = \Delta p_t - \beta_{rm} rm_t - c\Delta q_t \tag{A.17}$$

3. Update the posterior $\sigma_u^2$ according to the series of $u_t$,

[16] The SAS code is available on Prof. Hasbrouck's website.

[17] These initial values of the priors are the ones found in the SAS code made available by Prof. Hasbrouck. According to Hasbrouck (2009), the initial values of the prior should not impact the final estimate of the effective cost of a stock because the first 200 iterations (of 1,000) are disregarded to compute the average of the estimated values for the trading cost ($c$).

4. Draw new series of $q_t$ knowing the new value of $\sigma_u^2$. Given that $u_t = \Delta p_t - \beta_{rm} rm_t - cq_t + cq_{t-1}$, estimate $u_t$ if $q_t = +1$ or $q_t = -1$. Find the probability of $u_t(q_t = +1)$ and $u_t(q_t = -1)$ given that $u_t \sim N(0, \sigma_u^2)$ and compute the odds ratio for a buy order as follow,

$$\text{Odds} = \frac{f(u_t(q_t = +1))}{f(u_t(q_t = -1))} \begin{cases} q_t = +1 & \text{if Odds} > 1 \\ q_t = -1 & \text{if Odds} < 1 \end{cases} \tag{A.18}$$

**end**

The process is repeated 1,000 times and the final value for $c$ is the average of the last 800 estimations of the procedure (”burn in” the 200 first observations). For more information on simulating the probability distributions of $q_t$ and $u_t$ as well as on the iterative process, interested readers should refer to Hasbrouck (2009, p. 1449-1951).[18]

[18]Further details regarding the application of the estimation technique can also be found in Marshall, Nguyen, and Visaltanachoti (2011) and Novy-Marx and Velikov (2016).