# Power Sorting

Anastasios Kagkadis[†] Harald Lohre[†§] Ingmar Nolte[†]

Sandra Nolte[†] Nikolaos Vasilas[†]

[†]*Lancaster University Management School*

[§]*Robeco Quantitative Investments*

October 6, 2023

**Abstract**

We propose a novel approach for constructing characteristic-based equity factors, termed "power sorting". Power sorting exploits the non-linearities and asymmetries inherent in the characteristic-return relations, while it remains computationally simple and avoids excessive weights. We demonstrate that power sorting achieves consistently superior out-of-sample performance compared to traditional quantile sorting and other factor portfolio construction methods. Our results are pervasive across factors, robust through time, cannot be attributed to increased turnover or tail risk, and can be extended to multi-factor strategies. Finally, we show that power sorted versions of well-known asset pricing factor models outperform the original ones.

*Keywords: Factor Investing, Portfolio Construction, Power Sorting*

*JEL Classification: G10, G11, C52, C55*

We would like to thank David Blitz, Amit Goyal, Matthias Hanauer, Clint Howard, Laurens Swinkels, Steven Young, as well as the participants at the CQA 2023 Fall Conference, the Robeco research seminar and the LUMS research seminar for their valuable comments and suggestions. The views and opinions expressed in this paper are solely those of the authors and may not necessarily align with those of Robeco. This paper was awarded first place in the Chicago Quantitative Alliance 30th Annual Academic Competition. Send correspondence to Nikolaos Vasilas, n.vasilas@lancaster.ac.uk.

# 1 Introduction

When associating observable firm characteristics with equity returns, the classic approach in empirical asset pricing is to construct characteristic-sorted portfolios, commonly referred to as factor portfolios or simply, factors. Such factor portfolios have been widely used by academics to identify asset pricing anomalies and construct asset pricing models. They are also used by investment practitioners who look for systematic exposure to rewarded factors, provided these are investable. The conventional procedure for constructing factor portfolios invovles ranking the stock universe by a characteristic, creating quantile portfolios, and analyzing the long-short portfolio of the two extreme quantile portfolios. Despite its popularity and intuitive appeal, the procedure has its limitations. First, it lacks an objective criterion for choosing the number of quantile portfolios, with that number usually remaining invariant for the characteristic at hand. Usually, ten portfolios are considered, even though there is little motivation behind such a construction choice, apart from ensuring a decent characteristic spread. Second, the method cannot address variation in characteristics within quantile portfolios, as these are usually either equal or value-weighted.[1] In that respect, it also cannot account for potential non-linearities in the characteristic-return relations.[2] Third, the conventional weighting scheme is symmetric, implicitly assuming equal pricing ability of the characteristic on the long and the short side, while disregarding stocks in the middle. In that sense, it fails to explore the existence of monotonic patterns between returns and economic variables that are implied by finance theories (Patton and Timmermann, 2010).

To illustrate the limitations of conventional factor portfolio construction, Figure 1 plots the return for selected factors across the full spectrum of the respective characteristic, using 100 quantile bins. The main insight from Figure 1 is that characteristics can relate to average returns in non-trivial ways and decile sorting provides a simplistic perspective to a more complex set of patterns between the two. One such pattern is the inverted "smile" shape, where both stocks with very high and very low characteristics underperform (e.g., beta), resulting in insignificant return differences across the two legs. In this case, investing in the corner decile portfolios delivers an insignificant long-short spread, implicitly declaring the characteristic as an

[1] Such fixed weighting schemes introduce other factor exposures and can thus have a confounding effect on factor return inference (Swade et al., 2023).

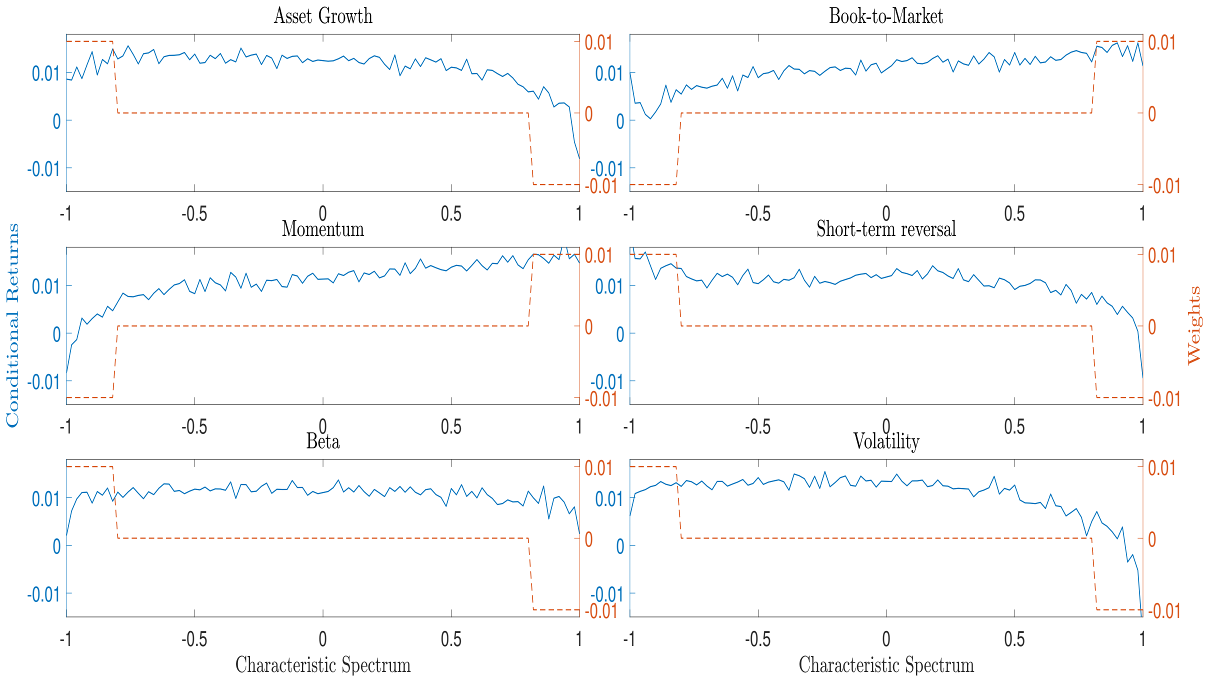
[2] Note that numerous leading finance theories predict that expected returns are highly non-linear functions of the underlying characteristics or state variables (e.g., Campbell and Shiller, 1988; Campbell and Cochrane, 1999; Bansal and Yaron, 2004; He and Krishnamurthy, 2013).

unimportant return predictor. Another common pattern is the inverted "smirk" shape, where stocks on the short side underperform, but stocks on the long side display no significant outperformance (e.g., asset growth and volatility). In that case, the factor portfolio spread is primarily driven by the short side. Lastly, average returns in the tails of the characteristics can drift in opposite directions. That is, they might drift in the intended direction as implied by the overall relationship, leading to an amplified effect in the extreme quantile portfolios (e.g., short-term reversal), but they might also turn in the opposite direction, reducing the return spread (e.g., book-to-market). Regardless of the underlying pattern, Figure 1 suggests that the extreme quantiles shall be treated differently and stocks in between the two extremes are also worth of consideration in the construction of factors. Nonetheless, any potential weighting scheme should be economically sound and theoretically motivated to ensure that the resulting portfolios retain the underlying factor structure and avoid overfitting and data mining concerns. Put differently, allowing the weight vector to vary freely without imposing any structure or economic prior could lead to overfitted factor portfolios that are based on return patterns alone and therefore unable to capture the underlying economic driver.

In this paper, we develop a data-oriented power sorting procedure to directly model factor portfolio weights as a function of firm characteristics. This procedure extends to conventional long-short factor portfolios by allocating some weight to all assets, while still allowing to tilt more towards stocks with extreme characteristics if deemed appropriate. Unlike conventional sorting, power sorting does not require manual selection of quantile breakpoints and seeks to exploit variation in characteristics across the full characteristic spectrum rather than overlaying fixed-weighting schemes that could mask the factor's nature. Importantly, power sorting can capture asymmetries and non-linearities from characteristics to returns, allowing for tailored treatment on the long and the short side and a deeper understanding of the behavior of the two complementary drivers of factor premia.

The power sorting procedure is based on the assumption of monotonicity between characteristic and return and is flexible enough to extract optimal performance from the underlying characteristic, while still creating portfolios that are theoretically guided and economically meaningful. Specifically, the cross-sectional weight vector for any given factor is obtained by expressing portfolio weights as a power series of the underlying characteristic rank. This formulation presents a tightly parameterized problem that accommodates a variety of monotonic

**Figure 1: Conditional monthly returns and conventional equal-weighted decile-sorted factor portfolio weights for six characteristics.** Characteristics are standardized in the $[-1, 1]$ range. The conditional returns are estimated by ranking stocks based on their underlying characteristic in the previous period and splitting them into 100 equal-weighted quantile portfolios. The blue line shows the average monthly return across portfolio groups. The dashed orange line shows the weight function for the factor portfolio that invests in the corner decile portfolios based on the underlying characteristic. The sample includes all common shares on NYSE, AMEX, and NASDAQ exchanges and covers the period from January 1980 to December 2021.




weighting schemes based on just two parameters, one for the long and one for the short leg of the factor portfolio. The two parameters determine the concentration of the power factor portfolio on stocks with extreme characteristic scores and can be estimated based on any arbitrary objective function, such as maximizing the Sharpe ratio of the factor portfolio. Importantly, unlike conventional optimization routines, our approach achieves this without explicitly requiring the use of a variance-covariance matrix, thus avoiding associated estimation challenges. Higher parameter values lead to portfolios that are more concentrated in stocks with characteristic extremes, while lower values lead to a more diversified factor exposure by spreading portfolio weights more evenly across stocks. Additionally, differences in the two parameters allow for capturing asymmetries and non-linearities in the weight function; for instance, one may construct factors that take an aggressive stance on one leg and a more passive stance on the other.

Our primary objective is to establish a framework for factor portfolio construction that accommodates characteristic-specific treatment of the various characteristics with clear interpretability of the underlying model parameters. Regarding the characteristics, several studies have consistently emphasized the asymmetric impact of the long and the short side on factor portfolio performance (Ang et al., 2006; Stambaugh, Yu, and Yuan, 2012; Blitz, Baltussen, and van Vliet, 2020; Leung et al., 2021). Furthermore, many characteristics documented in the literature have been found to yield insignificant performance, when the portfolio construction method is taken as given (Hou, Xue, and Zhang, 2015; Green, Hand, and Zhang, 2017; Hou, Xue, and Zhang, 2020). It is worth noting, however, that slight modifications in factor construction can lead to significantly different conclusions about factor significance (Jensen, Kelly, and Pedersen, 2023; Soebhag, van Vliet, and Verwijmeren, 2023). Consequently, the conventional portfolio construction technique cannot efficiently extract the underlying risk premium for the vast majority of characteristics, and can yield misleading conclusions about their economic and statistical significance.

For example, Hou, Xue, and Zhang (2020) find that factor portfolios based on market friction proxies exhibit insignificant performance under a conventional long-short quantile approach. This finding resonates with an inverted smirk pattern where the relationship between market friction proxies and next-period returns is highly asymmetric and factor performance is driven by the short side that contains the most illiquid stocks. Similar conclusions can be drawn for inverted smile patterns observed in many accounting variables. Power sorting proves particularly effective in modeling such patterns and producing weighting schemes that can exploit variation in the short leg while maintaining a more diversified stance in the long leg. Furthermore, power sorting can enhance the performance of already successful monotonic factors by leveraging the variation on both sides.

With regard to the model parameters, several degrees of freedom are involved in the construction of factor portfolios. For example, increasing the number of quantile portfolios — from terciles to quintiles, deciles, or beyond — produces portfolios that are concentrated in stocks with extreme characteristics. Additionally, researchers can affect the weighting scheme through other construction choices, such as value- or equal-weighting stocks in the selected quantile. Both schemes can introduce unwanted factor exposures that may unduly confound the targeted characteristic. Equal-weighting amplifies the effect of small stocks, while value-weighting results in portfolios that are heavily skewed towards very large stocks, thereby masking factor behavior via size effects. To address this issue, researchers can use NYSE breakpoints and winsorize market capitalizations. Such choices are often framed as data pre-processing steps and their implicit effect on portfolio performance is usually overlooked.

Ultimately, our approach constitutes a sample-efficient solution for deriving portfolio weights in an objectively optimal way, thereby alleviating p-hacking concerns related to subjective portfolio construction choices. Additionally, by explicitly parameterizing weight concentration in the tails, our framework enables clear interpretability of the underlying model parameters, thus bridging the gap between ad-hoc portfolio sorts and black-box machine learning methods. Finally, one distinctive feature of our method is the introduction of a hyper-parameter that controls for the impact of size in the construction of factor portfolios. This parameter is determined in a robust and transparent manner based on specific criteria, such as the maximum weight assigned to any individual stock. As a result, power factor portfolios are sufficiently diversified, easily interpretable, and practically relevant, establishing a data-driven and discretion-free framework for constructing factor portfolios.

Our results demonstrate that power sorting outperforms conventional sorting in terms of various portfolio metrics, using a set of 85 well-established characteristics in an out-of-sample period from March 1980 to December 2021. For many factor portfolios, the outperformance arises from adopting a more aggressive stance on the short leg and a more conservative stance on the long leg. This in turn implies that the characteristic signal is strong for underperforming stocks but it tends to be weaker for outperforming stocks. In the case of equal-weighted portfolios, the average factor portfolio Sharpe ratio increases by 57%, while for value-weighted portfolios, it doubles. Importantly, the observed performance enhancement is highly statistically significant and cannot be attributed to increased turnover or tail risk considerations. Furthermore, these economic gains also carry important asset pricing implications, as they lead to the resurrection of many documented factors that were previously deemed insignificant. Specifically, the factor significance rate rises from 40% to 75.3% for equal-weighted portfolios, and from 18% to 55.3% for value-weighted portfolios, even when employing a strict t-stat threshold of three (Harvey, Liu, and Zhu, 2016).

Lastly, despite its univariate nature, we provide evidence of the effectiveness of power sorting in a multi-factor context. Adopting an asset pricing perspective, we demonstrate that the incorporation of power factors into existing asset pricing models consistently improves their pricing ability as evidenced by a significantly higher model squared Sharpe ratio (Barillas et al., 2020). From an investment perspective, we highlight the empirical relevance of power sorting in combining individual factors into multi-factor portfolios. Our approach explicitly considers the asymmetric pricing abilities of different characteristics when combining signals, resulting in multi-factor portfolios with improved investment performance. The performance enhancement achieved through power sorting is substantial compared to single-characteristic strategies or equal weighted multi-factor approaches, particularly after accounting for size effects.

The remainder of the paper is structured as follows: Section 2 introduces the power sorting procedure and relates it to the conventional procedure and prior literature on characteristic-based portfolio construction. Section 3 explores power factor portfolio construction for a large set of characteristics and examines their out-of-sample performance on an individual and aggregate factor level. Section 4 compares power sorting to alternative methods proposed in the literature for factor construction and performs a variety of robustness tests to corroborate the validity of power-sorted factor portfolios. Section 5 concludes.

## 2 Power Sorting Methodology

The goal is to construct portfolios with exposure to some characteristic but in a way that one can best exploit its relationship to future returns. We begin by explaining the conventional portfolio construction technique, followed by the power sorting approach. The conventional sorting procedure is to rank the cross-section of stock returns according to a characteristic. The cross-sectional vector of characteristics, observable at the beginning of month $t$ is denoted by $x_t := (x_{t,1}, \dots x_{t,N_t})'$, where $N_t$ is the number of stocks available at time $t$. Equally, the vector of stock returns at the beginning of month $t+1$ is denoted by $r_{t+1} := (r_{t+1}, \dots, r_{t+1,N_t})'$. Finally, let $\{(1), (2), \dots, (N_t)\}$ be a permutation of $\{1, 2, \dots, N_t\}$ that results in ordered factor scores (from smallest to largest):[3]

$$x_{t,(1)} \leq x_{t,(2)} \leq \dots \leq x_{t,(N_t)}. \tag{1}$$

The essence of factor investing is the estimation of a weight vector $w_t := \left(w_{t,(1)}, \dots, w_{t,(N_t)}\right)'$ for $r_{t+1}$ based on $x_t$. A typical long-short portfolio satisfies,

$$\sum_{w_{t,(n)}<0} \left|w_{t,(n)}\right| = \sum_{w_{t,(n)}>0} \left|w_{t,(n)}\right| = 1, \tag{2}$$

[3]This is assuming a positive correlation between the characteristic and returns; otherwise the characteristic is inverted.

so that we have a unit dollar short leg and a unit dollar long leg.

## 2.1 Conventional long-short quantile factor portfolios

The standard procedure is based on partitioning the characteristic space into equal-sized quantile bins. Let $B$ be the number of quantile groups considered; for example, $B$ equals 5 for quintiles, or $B$ equals 10 for deciles. Let $d$ be the largest integer that is smaller than or equal to $N/B$ (the number of stocks in each quantile group). The $k$-th quantile of $x_t$, $q_{t,k} = 1, \dots, B-1$, is equal to $x_{t,(k \cdot d)}$. The weighting vector under a standard long-short decile portfolio scheme (i.e., $B$ equals 10) is denoted as $w_t^{LS}$. The standard long-short portfolio allocates equal weight to stocks belonging to the corner portfolios and disregards the rest,

$$\begin{aligned} w_{t,(1)}^{\mathrm{LS}} = \ldots = w_{t,(d)}^{\mathrm{LS}} &= -1/d, \\ w_{t,(d+1)}^{\mathrm{LS}} = \ldots = w_{t,(\mathrm{N_t}-d)}^{\mathrm{LS}} &= 0, \\ w_{t,(\mathrm{N_t}-d+1)}^{\mathrm{LS}} = \ldots = w_{t,(\mathrm{N_t})}^{\mathrm{LS}} &= 1/d. \end{aligned} \tag{3}$$

The resulting portfolio return at time $t+1$ is denoted by $r_{t+1}^{\mathrm{LS}} = r_{t+1}' w_t^{\mathrm{LS}}$. Value-weighted versions of those portfolios can be constructed by weighting stocks within each group based on their market capitalization:

$$w_{t,(n)}^{\mathrm{LS}} = \begin{cases} -mcap_{t,(n)} / \sum_{i=1}^{d} mcap_{t,(i)}, & \text{for } n \leq d \\ 0, & \text{for } d < n \leq N_t - d \\ mcap_{t,(n)} / \sum_{i=N_{t-d+1}}^{N_t} mcap_{t,(i)} & \text{for } n > N_t - d, \end{cases} \tag{4}$$

where $mcap_{t,(n)}$ is the market capitalization of stock $n$ at time $t$. Specifically, we construct capped value-weighted versions of the factor portfolios, following Jensen, Kelly, and Pedersen (2023), such that we assign weights to stocks based on their market capitalization winsorized at the $80^{th}$ percentile of the NYSE.

Regardless of the underlying weighting scheme, this approach has some important implications. First, the process is dependent on the specific choice of quantile breakpoints (e.g., terciles, quantiles or deciles). In essence, $B$ is a hyper-parameter that dictates the concentration of the long-short factor portfolio. Although deciles are commonly used, it is ultimately a choice parameter that can significantly affect inferences about the significance of factor premia (Soebhag, van

Vliet, and Verwijmeren, 2023). High values for $B$ can potentially improve return performance but lead to undiversified portfolios that are less practical as they over-concentrate in a small number of stocks. Second, employing equal- and value-weighted weighting schemes in portfolio construction introduces ad-hoc variation that may obscure the underlying return signal. Third, the approach places equal emphasis on the long and the short leg, while it disregards the information about mid-rank stocks. This attribute renders the method inadequate to effectively capture non-linearities and asymmetries in the underlying characteristic-return relationship.

## 2.2 Power sorting

### 2.2.1 Pure power-sorted portfolios

We propose power sorting that uses the underlying characteristic rank to determine factor portfolio weights, but the weighting vector is directly derived for the whole cross-section of stocks without requiring any grouping. In each period, we cross-sectionally rank all stock characteristics and map them onto the interval $[-1, 1]$ centered around the median rank. As shown below, the standardized characteristic rank vector $\tilde{s}_{t,(n)}$ is obtained as:

$$s_{t,(n)} = \left\lfloor \text{rank}(x_{t,(n)}) - \frac{N_t + 1}{2} \right\rceil \tag{5}$$

$$\tilde{s}_{t,(n)} = \begin{cases} -\frac{s_{t,(n)}}{s_{t,(1)}} & \text{for } s_{t,(n)} < 0 \\ 0 & \text{for } s_{t,(n)} = 0 \\ \frac{s_{t,(n)}}{s_{t,(N_t)}} & \text{for } s_{t,(n)} > 0, \end{cases} \tag{6}$$

where $\text{rank}(\cdot)$ is the rank function and $\lfloor \cdot \rceil$ is the function rounding to the nearest integer.

One advantage of using characteristic ranks rather than raw scores to derive the weighting vector is that the former is unaffected by the distribution of the characteristics. Next, we translate scores into weights by normalizing them based on the respective sums of scores as outlined below in equation (7). Stocks with below median characteristic rank are assigned negative weights and stocks with above median characteristic rank are assigned positive weights. Specifically, positive scores are divided by the sum of all positive scores and negative scores are divided by the sum of all negative scores, ensuring a unit dollar investment for the long and the short side. Non-linearities and asymmetries in the weight function are incorporated by introducing two parameters, one for the long ($p$) and one for the short side ($q$). These two parameters are exponents that are applied to positive and negative characteristic ranks before transforming them into portfolio weights. For exposition purposes, we assume $p$ and $q$ to be constant across time, while in our empirical investigation, we demonstrate how time-variability in $p$ and $q$ can impact the shape of the factor portfolio weight function over time. Hence, we express positive and negative scores as two independent power series and their scaling factors as their power sums. The resulting weighting vector for the power sorting portfolio is given by:

$$w_t^{PS}(\tilde{s}_{t,(n)}; p, q) = w_{t,(n)}^{PS} = \begin{cases} -\frac{|\tilde{s}_{t,(n)}|^q}{\sum\limits_{\tilde{s}_{t,(n)}<0} |\tilde{s}_{t,(n)}|^q} & \text{for } \tilde{s}_{t,(n)} < 0 \\ 0 & \text{for } \tilde{s}_{t,(n)} = 0 \\ \frac{\tilde{s}_{t,(n)}^p}{\sum\limits_{\tilde{s}_{t,(n)}>0} \tilde{s}_{t,(n)}^p} & \text{for } \tilde{s}_{t,(n)} > 0. \end{cases} \tag{7}$$

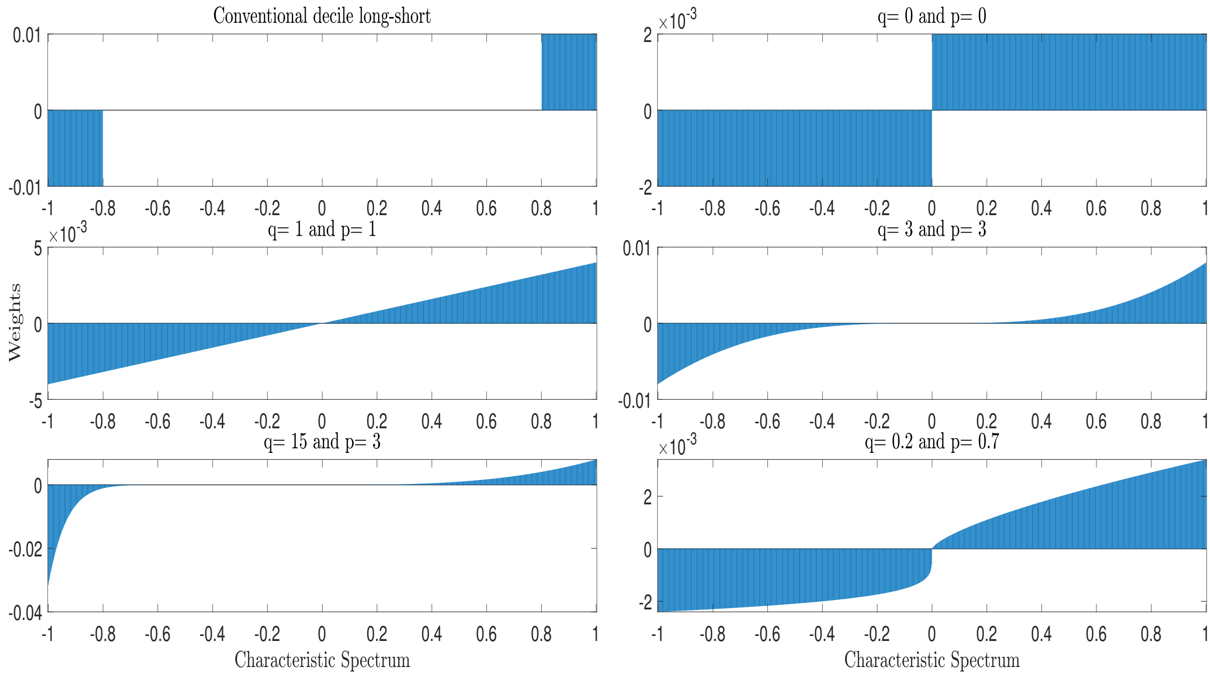
The two hyper-parameters $p$ and $q$ govern the concentration of the power sorting portfolio weights. Higher parameter values lead to portfolios that are more concentrated in the extreme ranks, as stocks with characteristic rank closer to the median shrink towards zero faster due to the function's exponential nature. Given that $\lim_{p \wedge q \to \infty} w_t^{PS} = (-1, 0, \ldots, 0, 1)$, all capital is allocated to the two stocks with the most extreme characteristics. This formulation provides a natural way to capture weight concentration in the tails and offers ample flexibility in modeling the underlying weighting function.

To illustrate, Figure 2 presents the resulting weighting function for various combinations of $p$ and $q$, alongside the conventional long-short weighting scheme. When $p \wedge q = 0$, the function evenly distributes weights between stocks above and below the median, resembling a conventional long-short portfolio with two groups and reflecting a passive factor approach. When $p \wedge q = 1$, the function aligns with a linear rank weighting scheme, where absolute weights increase linearly for stocks with characteristic ranks further from the median.

For values between 0 and 1 in $p \wedge q$, absolute weights increase at a marginally decreasing rate around the median, while for $p \wedge q > 1$ the weights increase at a marginally increasing rate, over-weighting the extreme ranks. Notice that when $p \wedge q < 0$, portfolio weights concentrate towards the centre, resulting in lack of monotonicity. Therefore $p \wedge q = 0$ constitutes a natural lower bound for the parameter space in the context of factor portfolio construction.

In general, high values for $p$ and $q$ correspond to an aggressive factor stance, where stocks with the most extreme characteristic ranks are expected to contribute most to factor performance, reflecting the existence of premium in the tails. Importantly, differences in $p$ and $q$ introduce asymmetries in the weighting scheme, allowing one leg to be more concentrated/less diversified.

**Figure 2: The opportunity set of power sorting.** The top-left chart displays a conventional equal-weighted decile-sorted long-short weighting scheme for a characteristic positively related to returns. The remaining charts display stock weights for different values of $p$ and $q$ under the power sorting scheme.




### 2.2.2 Value-weighted power factor portfolios

We next discuss the construction of value-weighted versions of the power factor portfolios. The rationale to use market capitalization for weighting stocks within each factor portfolio is to reflect the relative size of companies. Put differently, market capitalization weighting is likely to give sector and industry exposures similar to the overall market. However, under a conventional approach, such value-weighted factor portfolios tend to overweight mega-cap stocks, resulting in less diversified portfolios that cannot robustly capture the underlying factor premium. To this end, scholars have put forward ways to control the effect of market capitalization on portfolio composition. A well-known example is the Fama and French (1993) construction methodology, which gives half the weight to small stocks and the other half to big stocks. Jensen, Kelly, and Pedersen (2023) winsorize market capitalizations at the NYSE $80^{th}$ percentile before calculating factor portfolio weights, which avoids excessive weights on mega-cap stocks while still emphasizing large stocks. These approaches, although masked as data pre-processing steps, allow for different degrees of freedom in the estimation of value-weighted portfolios and can have a significant impact on portfolio outcomes (Soebhag, van Vliet, and Verwijmeren, 2023). Furthermore, the effect of such modifications on portfolio composition is usually unassessed. For this reason, we directly incorporate and parameterize the effect of size on the estimation of portfolio weights by computing the capitalization-adjusted versions of the power portfolio as:

$$w_{t,(n)}^{PS,cap} = \begin{cases} -\frac{|\tilde{s}_{t,(n)}|^q \cdot mcap_{t,(n)}^h}{\sum\limits_{\tilde{s}_{t,(n)}<0} |\tilde{s}_{t,(n)}|^q \cdot mcap_{t,(n)}^h} & \text{for } \tilde{s}_{t,(n)} < 0 \\ 0 & \text{for } \tilde{s}_{t,n} = 0 \\ \frac{\tilde{s}_{t,(n)}^p \cdot mcap_{t,(n)}^h}{\sum\limits_{\tilde{s}_{t,(n)}>0} \tilde{s}_{t,(n)}^p \cdot mcap_{t,(n)}^h} & \text{for } \tilde{s}_{t,(n)} > 0, \end{cases} \tag{8}$$

where parameter $h \in [0, 1]$ controls the concentration in mega-cap stocks. A value of $h = 1$ corresponds to the uncapped value-weighted versions, while a value of 0 corresponds to the pure characteristic-weighted power portfolios. Values between 0 and 1 regulate the effect of size in the estimation of weights and are crucial to avoiding corner allocations in mega-cap stocks. The reason is that the vector of ordered market capitalizations behaves as a power series with high exponential growth, as it is dominated by a handful of stocks of exponentially larger size than their peers. Hence, presuming no shrinkage on market caps ($h = 1$) means that the weighting vector of the value-weighted versions is the product of two power curves. This can lead to extreme concentrations in mega-cap stocks in cases where mega-cap stocks have extreme characteristic ranks and factor concentration ($p \vee q$) is high. As such, it is key to moderate the market capitalization component to avoid extreme mega-cap stock allocations.

Our approach allows for an efficient formulation of the weighting function and does not require any data pre-processing/manipulation step to avoid overconcentration, such as winsorization, NYSE breakpoints, grouping, or similar. The value of $h$ can be either calibrated based on the desired maximum portfolio weight or prespecified as a constant value. To mitigate data mining concerns, we opt for a constant value of $h = 0.5$ for all power portfolios, which is equivalent to taking the square root of market capitalization. As a benchmark, we estimate the value-weighted long-short portfolios using winsorized market caps at the $80^{th}$ NYSE breakpoints, as in Jensen, Kelly, and Pedersen (2023).

### 2.2.3 Managing weight concentration over time

The presented power sorting framework can naturally be extended to deal with extreme corner case allocations and account for time variation in $p$ and $q$ (i.e., consider $p_t$ and $q_t$). First, the maximum weight for each leg in each period is always allocated to the stock with the maximum absolute standardized characteristic rank. To illustrate, the maximum weight of the long leg portfolio is given by:

$$w_{max,t}^{PS} = w_{N_t}^{PS} = \frac{1}{F(\tilde{s}_{t,(N_t)}, p_t)}, \tag{9}$$

where $F(\tilde{s}_{t,(N_t)}, p_t) = \sum_{\tilde{s}_{t,(n)}>0} \tilde{s}_{t,(n)}^{p_t}$ is a power sum that can be efficiently computed using Faulhaber's formula (Knuth, 1993). Hence, $w_{N_t}^{PS}$ is decreasing in the number of available assets and increasing in the value of the power $p_t$, meaning that the effect of $p_t$ on $w_{N_t}^{PS}$ is conditional on the number of available assets and therefore characteristic- and time-dependent. Equation (9) highlights that finding a single optimal combination of $p$ and $q$ over time would lead to inconsistent weight distributions due to the variation in the size of the equity cross-section. In other words, imposing a single optimal power exponent would yield a variety of maximum weights over time. This inconsistency poses challenges when comparing different characteristics and determining the parameter values that maximize in-sample performance.[4] As the number of available assets is known at time $t$, the maximum weight can be constrained by setting an upper threshold to the maximum power. The threshold is calculated by solving:

$$\frac{1}{F(\tilde{s}_{t,(N_t)}, p_t)} - w^{ceil} = 0, \tag{10}$$

where $w^{ceil}$ is the targeted maximum weight. Depending on characteristic availability, the value of $p_t$ solving equation (10) will vary. We opt for a maximum portfolio weight of 2% when estimating the upper threshold for $p_t$ and $q_t$ (labeled $p_t^{max}$ and $q_t^{max}$, $t = 1, \ldots, T$) to ensure healthy portfolio diversification.

To ensure consistency in maximum weight concentration over time while still optimizing with respect to a single set of parameters, we define the concentration ratios for the two sides as $\tilde{p}_t = p_t/p_t^{max} \in [0,1]$ and $\tilde{q}_t = q_t/q_t^{max} \in [0,1]$, respectively. The concentration ratios

[4] It should be noted that the optimal power exponents cannot be directly compared across characteristics since they correspond to different availabilities of characteristics data.

correspond to the densities associated with the truncated distributions of $p_t$ and $q_t$ and are essentially standardized metrics that allow for a clear and intuitive interpretation. Specifically, a concentration ratio equal to one for either leg indicates that power factor portfolio performance is optimized when the weights are concentrated in the tail(s), with the maximum weight being no larger than $w^{ceil}$. Conversely, a value of zero implies that factor performance is optimized when a diversified stance is taken, equal-weighting stocks away from the median. This standardization allows for uniformity in the behavior of the weight distribution across time and characteristics and, hence, in the calibration of a single set of parameters. These optimal densities can then be mapped out to every period based on $p_t^{max}$ and $q_t^{max}$, allowing for the optimal $p_t$ and $q_t$ to be time-varying.

### 2.3 Power sorting and related literature

Our paper contributes to the literature on characteristic-based portfolio choice for asset pricing and investment applications. Conventional characteristic sorting has been a workhorse in empirical asset pricing due to its simplicity and intuitive interpretation. Early empirical contributors of portfolio characteristic sorts include Basu (1977) and Banz (1981), while the approach was popularised by Fama and French (1992) and Jegadeesh and Titman (1993). Despite its popularity, prior literature has identified some practical and theoretical limitations of the conventional portfolio construction. Jacobs and Levy (1993) raise various practical concerns that underline long-short strategies, while Patton and Timmermann (2010) highlight the inability of long-short strategies to test for monotonicity between characteristics and returns. In contrast to standard portfolio sorts, power sorting imposes monotonicity in the characteristic-return relationship and leverages variation across the characteristic spectrum to derive factor portfolio weights. Therefore, it promises to align more closely with economic theory.

Alternative approaches to portfolio construction like Frazzini and Pedersen (2014) or Koijen et al. (2018) utilize rank portfolios. Rank portfolios assign progressively higher weights to stocks as they deviate further from the characteristic median in a linear manner. This method proves effective for characteristics that demonstrate a monotonic relationship with returns, particularly when the effect is more pronounced for extreme values.[5] Notably, power sorting encompasses rank portfolios, allowing for a linear weighting function based on the characteristic rank when

[5] Novy-Marx and Velikov (2022) propose a rank- and capitalization-weighted scheme to account for market capitalizations in rank portfolios.

the underlying relationship is linear. However, power sorting goes beyond linear, enabling the incorporation of non-linear and asymmetric weights, thus offering greater flexibility in portfolio construction. A comprehensive comparison between rank portfolios and power portfolios is provided in Section 4.1.1.

The limitations of conventional portfolio sorting have prompted researchers to seek improvements in the construction of characteristic-based portfolios. Cattaneo et al. (2020) approach portfolio sorting as a non-parametric estimator, where the number of portfolios serves as a hyper-parameter. They demonstrate that the optimal number of portfolios, or the weighting function, should be time-varying and guided by the dynamics of the underlying characteristic. Zhang, Wu, and Chen (2022) propose a listwise learn-to-rank loss function that sequentially selects pairs of stocks for the long and the short leg. Closer to our study, Ledoit, Wolf, and Zhao (2019) utilize the DCC-NL estimator developed by Engle, Ledoit, and Wolf (2019) to estimate "efficient" factor portfolios. These portfolios aim to minimize variance while maintaining the overall factor exposure of traditional long-short portfolios. To assess the impact of parameter shrinkage resulting from power sorting compared to a more direct optimization approach, we compare our method with the efficient sorting methodology of Ledoit, Wolf, and Zhao (2019) in Section 4.1.2.

Our study is also related to a strand of the literature that models portfolio weights as a function of underlying firm characteristics and employs optimization-based approaches for portfolio construction. Notable examples of the former approach are Brandt (1999), Brandt and Santa-Clara (2006), Aït-Sahalia and Brandt (2001), and Brandt, Santa-Clara, and Valkanov (2009). Building upon the parametric portfolio policy framework of Brandt, Santa-Clara, and Valkanov (2009), Ammann, Coqueret, and Schade (2016) introduce leverage constraints, DeMiguel et al. (2020) incorporate transaction costs, and Simon, Weibels, and Zimmermann (2023) integrate feed-forward neural networks to capture non-linear and interaction characteristic effects. Hjalmarsson and Manchev (2012) demonstrate that within a mean-variance framework, the use of firm characteristics enables the reduction of the asset space to a set of characteristic-based portfolios. In an alternative approach, McGee and Olmo (2022) use non-parametric kernel methods to estimate the conditional moments of stock returns based on stock characteristics in a cross-sectional setting. These estimated moments are then used within a mean-variance objective function for portfolio construction.

Finally, our work relates to recent studies that construct characteristic-driven portfolios but with different objectives compared to ours. For instance, Fama and French (2020) utilize the cross-sectional regression approach of Fama and MacBeth (1973) to construct factors based on standardized characteristics.[6] Their findings reveal that these cross-sectional factors are more effective at explaining average returns compared to the original Fama-French-type factors. In a different context, Kim, Korajczyk, and Neuhierl (2021) introduce portfolios that aim to exploit mispricing information in the characteristics while hedging out systematic variation related to those characteristics. Similarly, Daniel et al. (2020) construct "characteristic efficient portfolios" by hedging away variation associated with unpriced risk using a hedge portfolio.

# 3 Optimal Power Sorting Portfolios

## 3.1 Characteristics and power thresholds

We replicate a large set of 85 characteristics that have been considered by Green, Hand, and Zhang (2017). The characteristics are calculated using data from the Center of Research on Securities (CRSP), Compustat, and the Institutional Brokers' Estimate System (I/B/E/S), covering the period from January 1980 to December 2021. The stock universe includes common stocks listed on NYSE, AMEX, and NASDAQ that have a record of month-end market capitalization on CRSP and a non-missing and non-negative common value of equity on Compustat. Additional information about the characteristics, including origination and characteristic description, can be found in Section A.1 of the Internet Appendix.

For every month, stock returns for month *t+1* are matched against their respective characteristics in month $t$. For accounting data, we allow at least six months to pass from the firms' fiscal year-end before they become available and at least four months to pass for quarterly data. To mitigate the effect of microcaps, we remove stocks with a market capitalization below the $10^{th}$ percentile at the portfolio formation period.

For constructing conventional benchmark factor portfolios, we first group stocks into equal-weighted deciles based on their characteristic scores in the previous month and then go long and short in the two extreme deciles, depending on the prevailing characteristic-return relationship. For value-weighted results, we use a "capped value-weighting" scheme following Jensen, Kelly,

[6] The regression slopes correspond to the returns of the zero investment factor portfolios with unit exposure to their characteristic and zero exposure to all the other characteristics.

and Pedersen (2023). Factors are categorized into six groups based on economic rationale, (Hou, Xue, and Zhang, 2015), namely: Momentum, Value, Investment, Profitability, Intangibles, and Trading Frictions.

A subtle but important detail is setting the maximum threshold for the hyper-parameters $p_t$ and $q_t$. This threshold depends on the targeted maximum portfolio weight and the number of available assets. To ensure that the power portfolios are sufficiently diversified, we set the maximum weight to 2% and solve equation (9) for the values of $p_t^{max}$ and $q_t^{max}$. Figure A.1 in the Internet Appendix displays the time-variation in the maximum powers for the long and the short leg of the different characteristics. Evidently, those thresholds vary significantly across characteristics and time, further stressing the importance of using a standardized measure for optimization purposes and for conducting comparisons.

### 3.2 Estimation procedure

The construction of the power sorting portfolio for a given characteristic $x_t$ requires an estimate of the powers $p_t \in [0, p_t^{max}]$ and $q_t \in [0, q_t^{max}]$ for each period $t$. To this end, we solve for the respective concentration ratios $\tilde{p}_t$ and $\tilde{q}_t$ that maximize the power sorting factor portfolio Sharpe ratio in the in-sample period and estimate the powers for the most recent cross-section by multiplying the ratios with the maximum power thresholds for the most recent period $p_t^{max}$ and $q_t^{max}$. To mitigate data-mining concerns regarding the selection of the estimation window, we adopt an expanding window approach and consider the longest out-of-sample period possible. In particular, the out-of-sample period covers March 1980 to December 2021, while different estimation windows are explored in Section 4.2. To illustrate, assuming a Sharpe ratio maximization objective and based on an underlying rank standardized characteristic $\tilde{s}_t := \tilde{s}_{t,1}, \ldots \tilde{s}_{t,N_t}$, the estimation problem at each investment date can be formulated as follows:

$$\{\hat{\tilde{p}}_t, \hat{\tilde{q}}_t\} = \underset{\tilde{p}_t \wedge \tilde{q}_t \in [0,1]}{\arg\max} \frac{\overline{r_t}^{PS}}{\sqrt{var(r_t^{PS})}}, \tag{11}$$

$$\overline{r_t}^{PS} = \frac{1}{t-1} \sum_{i=1}^{t-1} \sum_{j=1}^{N_i} r_{i+1,j} \cdot w_i^{PS}(\tilde{s}_{i,(j)}; p_i^{max} \cdot \tilde{p}_t, q_i^{max} \cdot \tilde{q}_t), \tag{12}$$

$$var(r_t^{PS}) = \frac{1}{t-2} \sum_{i=1}^{t-1} \left( \sum_{j=1}^{N_i} \left(r_{i+1,j} \cdot w_i^{PS}(\tilde{s}_{i,(j)}; p_i^{max} \cdot \tilde{p}_t, q_i^{max} \cdot \tilde{q}_t)\right) - \overline{r}^{PS} \right)^2, \tag{13}$$

which is a constrained optimization problem that can be solved numerically.[7] Notice that under this formulation there is no need to estimate the variance-covariance matrix (VCV) for individual stocks. Each combination of $\hat{\tilde{p}}_t$ and $\hat{\tilde{q}}_t$ practically corresponds to a set of cross-sectional weight vectors, and hence to a power portfolio return time-series for which the first and the second moments are computed directly. The out-of-sample power sorting portfolio return at time $t+1$ is then estimated as:

$$r_{t+1}^{PS} = r_{t+1}' \times w_t^{PS}(\tilde{s}_t; p_t^{max} \cdot \hat{\tilde{p}}_t, q_t^{max} \cdot \hat{\tilde{q}}_t). \tag{14}$$

Value-weighted results for the power versions of each factor are estimated as in equation (8), using the same maximum powers as in the pure characteristic weighted versions and a value of $h = 0.5$.

### 3.3 Power-sorted portfolios and concentration ratios

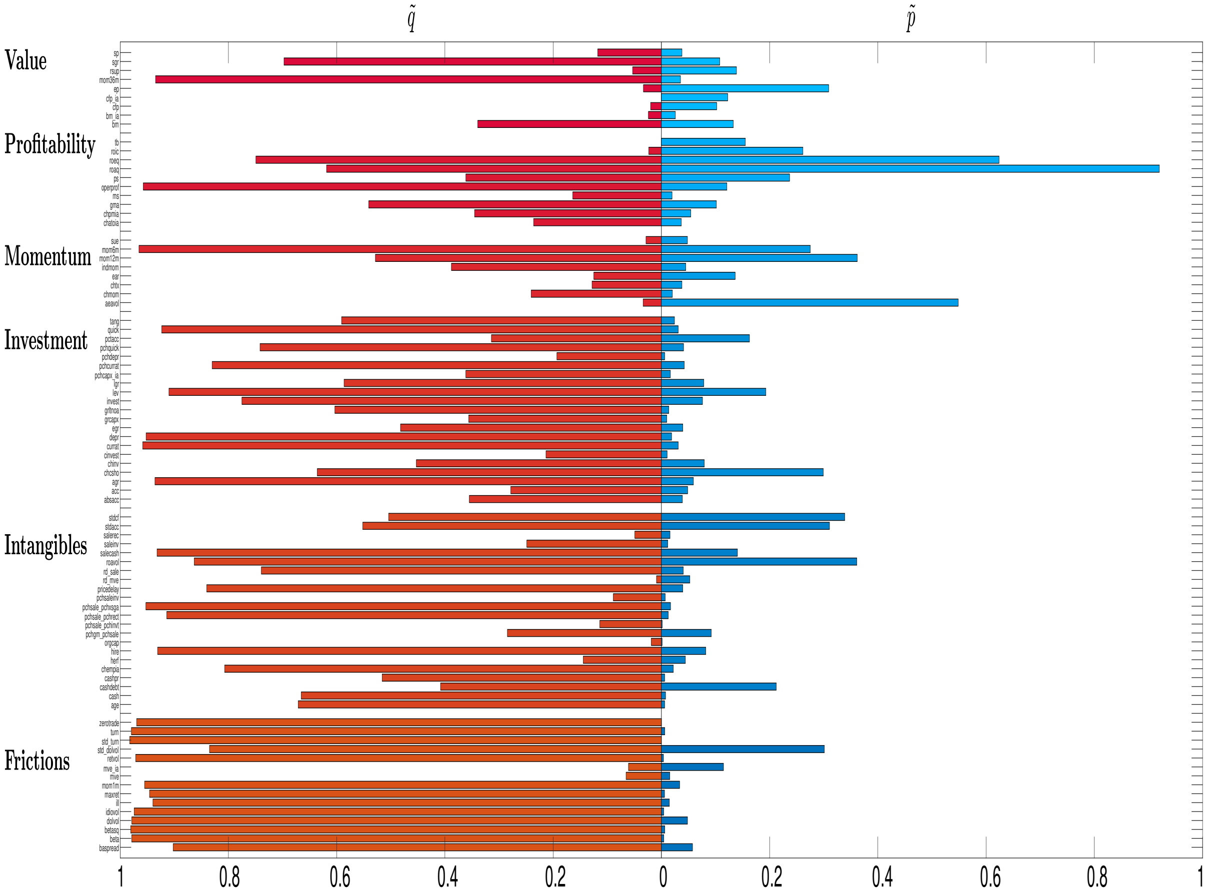
First, we examine the underlying form of the weight function for various characteristics. Power-sorted portfolios assign a portfolio weight to every stock that is uniquely determined by the $\tilde{p}_t$ and $\tilde{q}_t$ parameters. The use of an expanding estimation window implies that the out-of-sample parameters should gradually stabilize and converge to the optimal in-sample parameters as the sample expands. To foster intuition with respect to the underlying weight function, we present the average concentration ratios of each factor in Figure 3. Blue-shaded bars represent the average concentration ratio for the long side and red-shaded bars represent the average concentration ratio for the short side.

Figure 3 clearly illustrates that the optimal degree of concentration is highly asymmetric and skewed towards the short side for the majority of characteristics. That is the factor portfolio Sharpe ratio is maximized by adopting an aggressive stance on the short side and a more conservative stance on the long side. This finding indicates that stocks at the lower end of the conditional return distribution tend to perform very poorly, while stocks' outperformance at the extreme upper end is less extreme. Nonetheless, lower values for $\tilde{p}_t$ compared to $\tilde{q}_t$ do not imply that the long leg is an insignificant contributor to factor portfolio performance. In fact, as we show later in the analysis, the long leg of the power sorting portfolios delivers positive and significant returns. Nevertheless, this asymmetry suggests that conditional returns in the long

[7] In Section A.3 of the Internet Appendix we report results under a return-spread maximization objective. These results are similar to the ones presented in the main paper.

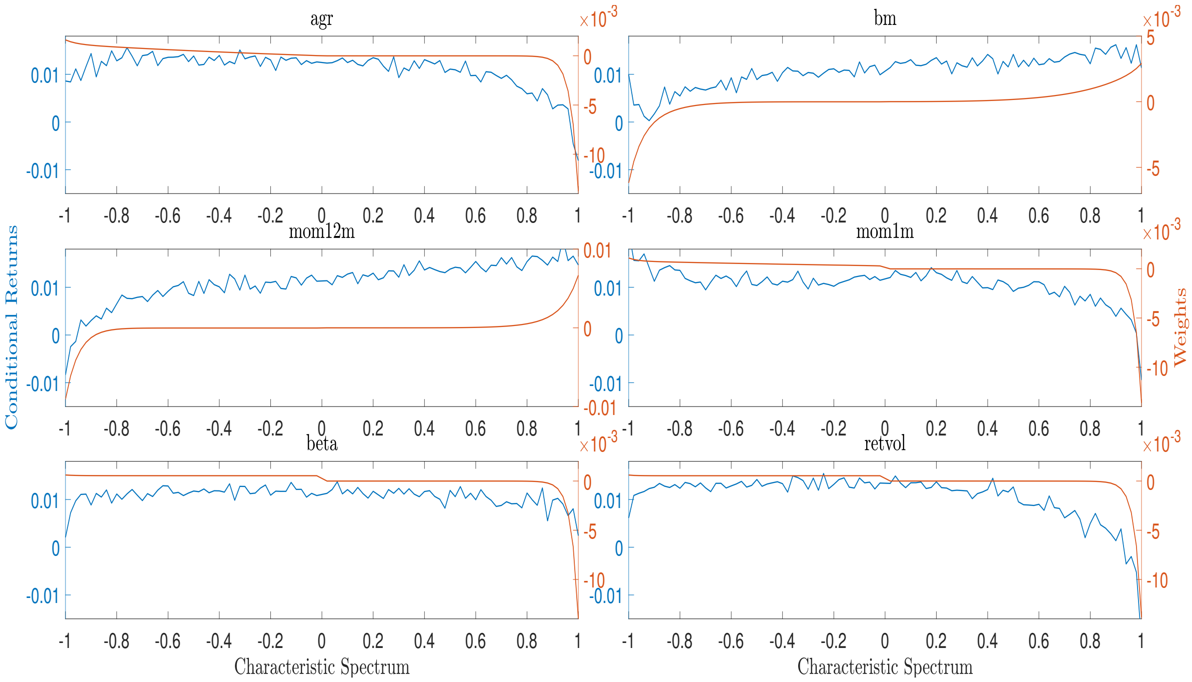
tail either remain relatively flat or that stocks in the corner of the long leg tend to underperform. Consequently, a lower concentration ratio in the long leg helps to avoid overinvesting in corner stocks that are likely to underperform when compared to their peers.

**Figure 3: Average estimated concentration ratios $\hat{\tilde{p}}$ and $\hat{\tilde{q}}$ of each factor.** Blue-shaded bars show the average concentration ratio for the long side and red-shaded bars show the average concentration ratio for the short side. Factors are sorted into six groups based on their economic rationale. The sample period is from March 1980 to December 2021.




To visualize the implications of the estimated concentration ratios in terms of portfolio weights, Figure 4 depicts the average weight function resulting from the selected values for $\hat{\tilde{p}}_t$ and $\hat{\tilde{q}}_t$ for the six factors presented in Figure 1. When compared to the conventional weighting scheme, power sorting is able to capture the underlying return patterns more accurately. In particular, our approach proves highly effective in dealing with characteristics that exhibit inverted "smirk" or inverted "smile" shapes, such as asset growth, return volatility, and beta. In such cases, power portfolios combine a high value for $\tilde{q}_t$ and a low value for $\tilde{p}_t$, thus producing inverted smirk weight schemes that increase exposure on stocks in the extremes of the short side, and reduce exposures on stocks in the extremes of the long side. Furthermore, for some characteristics, like momentum, the effect is amplified in the extremes and the algorithm opts for high values of $\tilde{p}_t$ and $\tilde{q}_t$, resonating with an aggressive stance in both the long and the short side to exploit variation in both tails.

**Figure 4: Conditional monthly returns and power factor portfolio weight function for six characteristic-based factor portfolios.** Characteristics are standardized in the $[-1, 1]$ range. Conditional returns are estimated by ranking stocks based on their underlying characteristic in the previous period and splitting them into 100 equal-weighted quantile portfolios. The blue line shows the average monthly return across portfolio groups. The orange line shows the average weight function for the factor portfolio as implied by the selected values for $\tilde{p}_t$ and $\tilde{q}_t$ across periods. The sample includes all common shares on NYSE, AMEX, and NASDAQ exchanges and covers the period from March 1980 to December 2021.




## 3.4 Power sorting versus conventional decile sorting

Table 1 compares the portfolio performance of power sorting against that of conventional decile sorting, presenting average portfolio statistics across factors for both equal-weighted and value-weighted cases. The results demonstrate the superiority of power sorting over the conventional approach across all portfolio metrics. Specifically, power sorted portfolios consistently exhibit significantly higher average returns and Sharpe ratios, with an average t-statistic above three for both value and equally weighted factors. Notably, for the value-weighted case, power sorting leads to an average Sharpe ratio that is twice as high as that achieved through the conventional approach (0.52 versus 0.26). Importantly, these results are not likely to be driven by increased trading costs or tail risk, as power factors exhibit on average a lower turnover and maximum drawdown compared to the conventional long-short portfolios. Finally, the resulting portfolios are more diversified, encompassing a higher number of effective names on both the long and short sides. The asymmetrically higher number of effective names for the long leg of the average power portfolio corroborates the patterns depicted in Figure 3, reflecting higher values of $\tilde{q}$ and a more aggressive stance for the short side.

**Table 1: Average performance measures across factors for power and conventional long-short portfolios**. Return: Average monthly return, Stand. dev.: Monthly standard deviation, Sharpe ratio: Annualized Sharpe ratio, t-stat: t-statistic on $H_0$: Return=0, Hit rate: Percentage frequency of positive returns, Turnover: Average monthly turnover (bounded by 200%), # of effective names long: Number of effective names for the long leg, # of effective names short: Number of effective names for the short leg. The sample period is from March 1980 to December 2021.

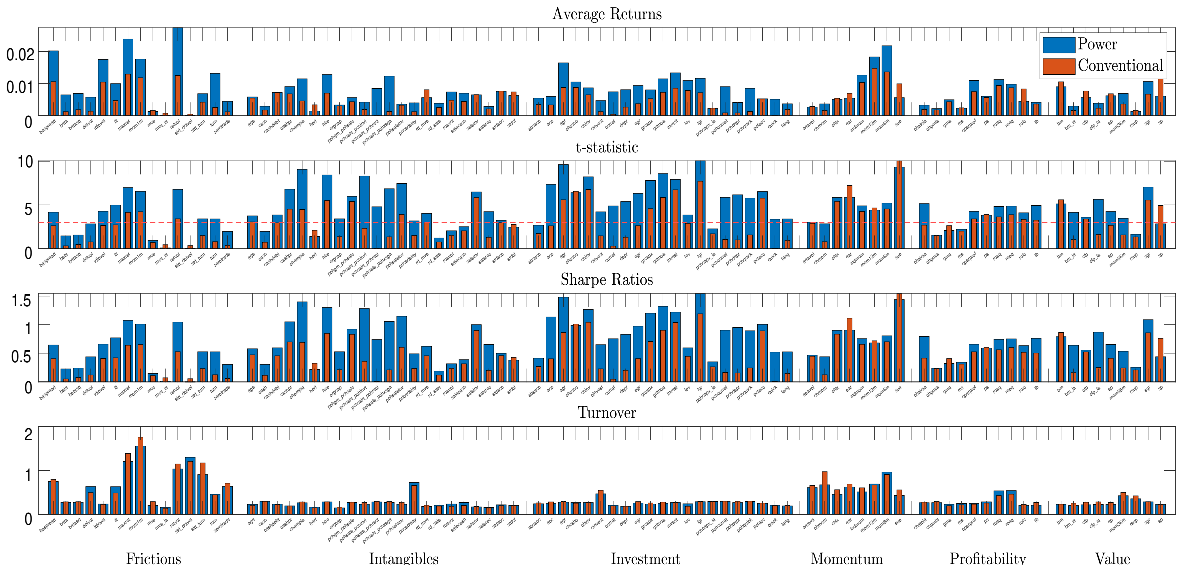
| | Equal-weighted | | Value-weighted | |
|---|---|---|---|---|
| | Power | Conventional | Power | Conventional |
| Return (%) | 0.77 | 0.51 | 0.62 | 0.32 |
| Standard deviation (%) | 4.04 | 4.21 | 4.17 | 4.39 |
| Sharpe ratio | 0.72 | 0.46 | 0.52 | 0.26 |
| t-stat | 4.64 | 2.96 | 3.40 | 1.71 |
| Maximum drawdown (%) | −45.79 | −55.35 | −48.96 | −59.22 |
| Hit rate (%) | 61.88 | 57.39 | 58.27 | 53.44 |
| Turnover (%) | 37.57 | 39.39 | 33.09 | 35.44 |
| # of effective names long | 1315.41 | 369.24 | 460.01 | 107.00 |
| # of effective names short | 535.60 | 370.33 | 229.79 | 98.42 |

Figure 5 presents selected out-of-sample portfolio evaluation measures for individual factors, comparing the power-sorted versions (blue) with the conventional long-short decile versions (orange). Panel A compares the pure power-sorted portfolios with equal-weighted decile benchmarks, while Panel B compares the capitalization-adjusted power versions with the capped value-weighted versions of the conventional long-short approach. Power sorting consistently leads to substantial gains in average returns and Sharpe ratios across the majority of factors, and these improvements cannot be attributed to increased turnover. Specifically, 75.3% of power versions have higher average returns, and 86% have higher Sharpe ratio. For value-weighted results, the respective numbers are 85.9% for returns and 96.5% for Sharpe ratios.

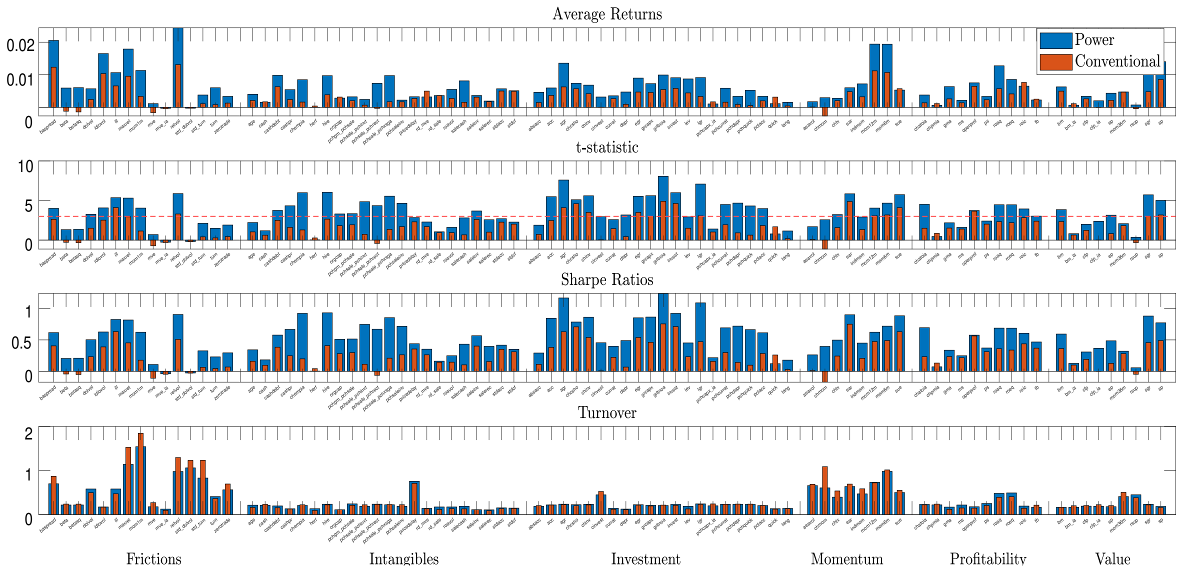
In addition, power sorting achieves a significantly higher significance rate for the average returns of factor portfolios, as indicated by a t-statistic above three (75.3% versus 40% for equal-weighted portfolios and 55.3% versus 18% for value-weighted portfolios). Hence, several factors deemed insignificant under the conventional weighting scheme become significant when the power weighting scheme is applied, even when using the stricter t-value threshold of three, as advocated by Harvey, Liu, and Zhu (2016). These results raise questions about the ability of decile sorting to efficiently extract the underlying signal from many characteristics, potentially leading to false rejections of factors.

**Figure 5: Portfolio evaluation measures for conventional long-sort and power versions.** Panel A displays equal-weighted results and Panel B displays value-weighted results. (1) Average monthly return, (2) t-statistics on average monthly return, (3) Annualized Sharpe ratio, (4) Monthly turnover. The optimal powers are selected using an in-sample expanding window starting from January 1980 to December 2021. Factors are sorted into six groups based on their economic rationale.

*A. Equal-weighted portfolios*




*B. Value-weighted portfolios*




It is worth noting that results are fairly consistent across the pure power-sorted and value-adjusted versions of the power portfolios, while the benchmark results considerably deteriorate under value-weighting. In fact, in some cases adjusting for market capitalization leads to value-weighted portfolios with negative average returns under the conventional method. Conversely, returns remain positive under a power sorting approach. Hence, the incorporation and parameterization of the size effect into the factor weighting procedure preserves the underlying factor behavior and controls for any confounding effects that might otherwise arise in naive value-weighted decile sorts.

Finally, investment gains from power sorting are evenly distributed across factors, yet notable significance is observed for factors associated with Frictions, Investment, and Intangibles. These factor themes are recognized for their asymmetric nature (i.e., Ang et al. (2006), Cooper, Gulen, and Schill (2008)), confirming the effectiveness of power sorting in capitalizing on the specific patterns inherent to these factors.

### 3.5 Dissecting long and short factor portfolio legs

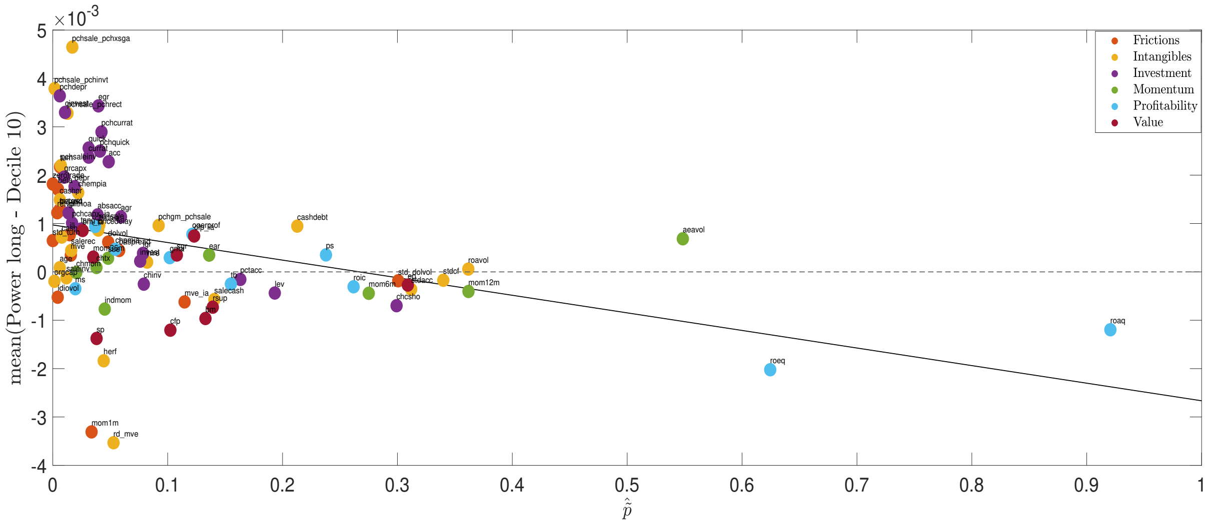
So far, we have established the presence of asymmetric and non-linear return effects of characteristics and demonstrated the effectiveness of power sorting in identifying and capitalizing on these patterns. To gain a deeper understanding, we further explore the implications of $\tilde{p}_t$ and $\tilde{q}_t$ on the long and short sides of each factor.

Figure 6 provides a visual depiction of the improvement in return performance for the long (Panel A) and short (Panel B) sides of the different factors, based on their corresponding values of $\tilde{p}_t$ and $\tilde{q}_t$. Each subfigure includes the line of best fit and the zero line. Data points positioned below the zero line indicate instances where the power leg underperforms its conventional decile leg. Panel A shows the increase in average return for the power long leg compared to decile ten, using the average optimal value of $\tilde{p}_t$. The relationship between the two is negative, suggesting that a more diversified approach that spreads weights across the long leg is preferable over concentrating solely on stocks in the extreme decile. Characteristics associated with Investment, Intangibles, and Market Friction proxies exhibit the most significant benefits from a low value of $\tilde{p}_t$. As already discussed, these variables demonstrate inverted smile and smirk patterns, indicating that conditional returns in the long tail either decrease or remain relatively flat. Consequently, a low $\tilde{p}_t$ reduces portfolio exposure to underperforming corner stocks in the long tail, enhancing diversification benefits and investment performance in the long leg.

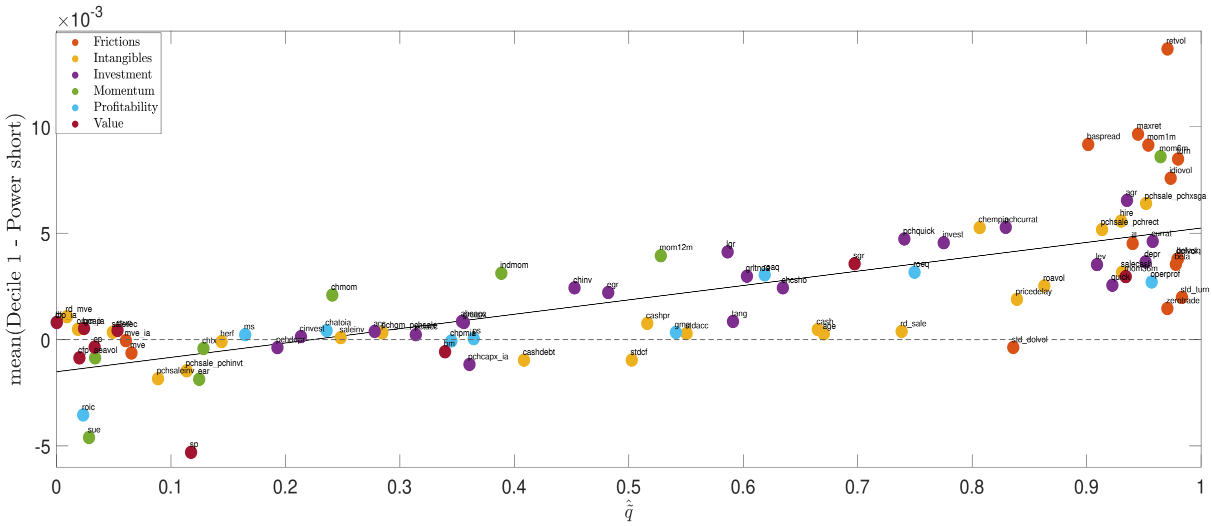
Panel B of Figure 6 shows the decrease in average return for the power short leg over decile one for the average optimal value of $\tilde{q}_t$. Higher values of $\tilde{q}_t$ are associated with lower average returns for the short side, increasing the long-short spread. Power sorting remains particularly effective for factors related to market frictions, such as maxret or retvol, as it capitalizes on the sharp decline in conditional returns on the short side. Still, intangibles and investment factors are now more spread across the line, implying heterogeneity in terms of optional concentration levels for the short leg. Intuitively, this result suggests that the different investment and intangible proxies agree on the long side but disagree on the short side.

**Figure 6: Concentration ratios and excess returns.** Panel A shows the increase in average returns for the long leg given the estimated value of $\tilde{p}_t$. Panel B shows the decrease in average return for the short leg given the estimated value of $\tilde{q}_t$. Each subfigure includes the line of best fit and the zero line. The sample period is from March 1980 to January 2021.

*A. Long leg*




*B. Short leg*



Note that stocks with the most extreme characteristics have a significant impact on the determination of optimal powers, as their weights increase exponentially. When using the median rank as the cutoff, the algorithm favors a low power in the long leg to avoid overinvestment in underperforming corner stocks relative to their peers. Due to the monotonic nature of the function, it has limited capacity to capture inflection points in the tails of the conditional return distribution. As a result, it adopts a passive approach by equally weighting stocks above the median to compensate. This pattern emerges also for the short leg of many value characteristics.

Overall, both sides of power portfolios outperform their conventional counterparts. The outperformance on the long side is driven by adopting a more balanced approach, spreading weights across a broader range of stocks. On the other hand, the outperformance on the short side is attributed to adopting a more aggressive stance, capitalizing on the specific patterns identified through power sorting.

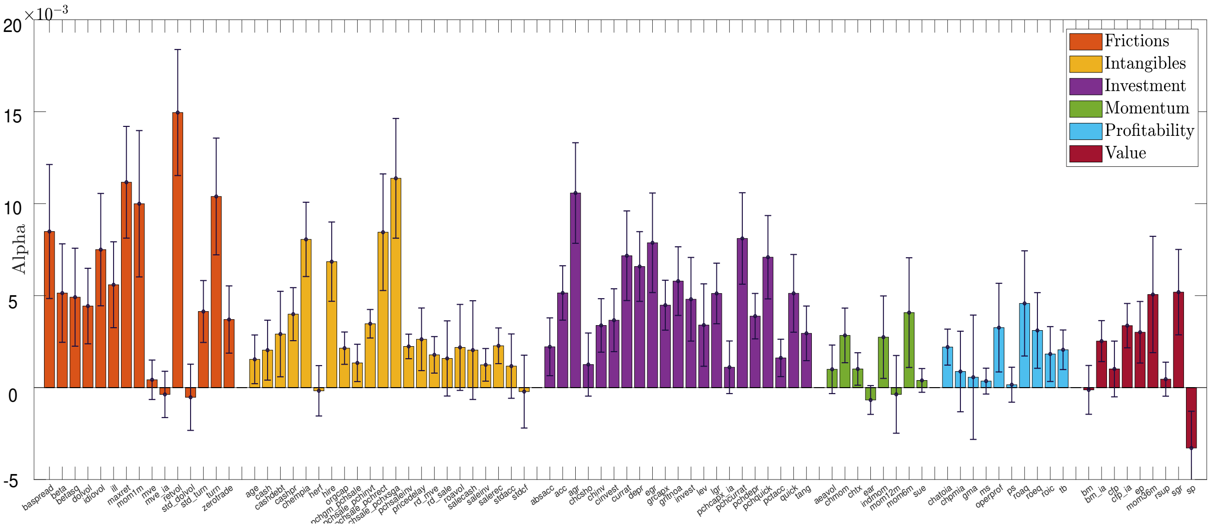
### 3.6 Spanning regressions

In Figure 7, we further report the monthly alphas from regressing power portfolio returns on those of their conventional long-short counterparts. The subfigures correspond to equal-weighted and value-weighted results and all estimates include their 95% confidence bounds. An interval that excludes (includes) zero indicates statistical significance (insignificance) at the 5% level.

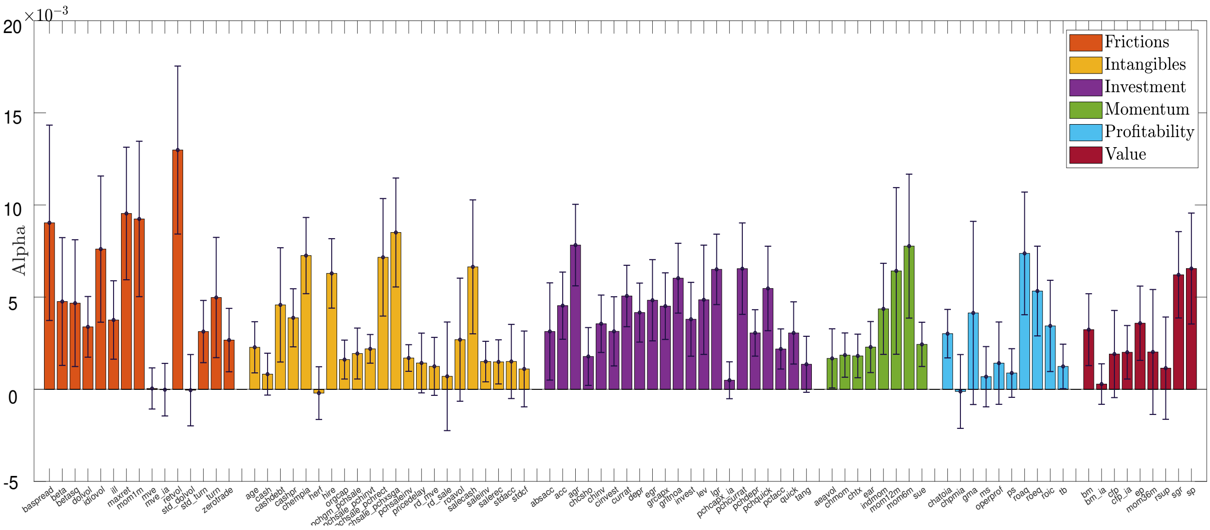
Out of the 85 alphas, 77 are positive and 62 are statistically significant for the equal-weighted case. In the value-weighted case, the corresponding numbers are 81 and 63, respectively. In fact, even factors that did not exhibit any significant improvement under power sorting in the equal-weighted case now exhibit alphas that are positive and significant. This improvement is particularly noticeable for momentum factors, which are infamous for experiencing a sharp decline in profitability with market capitalization (Hong, Lim, and Stein, 2000). Hence, results underscore the importance of incorporating size considerations in the factor construction process to mitigate performance deterioration or factor dilution in the value-weighted case. Finally, several of the alphas are also economically significant, with the 18% annualized alpha for the past month's volatility (retvol) being particularly noteworthy.

**Figure 7: Spanning regression alphas.** Intercepts from univariate regressions of power portfolio returns on conventional long-short portfolio returns for the sample period March 1980 to December 2021. Panel A shows equal-weighted results and Panel B shows value-weighted results. Power portfolios are constructed using an expanding window and conventional long-short using equal-weighted decile sorting. Factors are sorted into six groups based on their economic rationale.

*A. Equal-weighted portfolios*




*B. Value-weighted portfolios*




## 3.7 Multi-factor portfolios

Power sorting can be utilized to construct multi-factor portfolios through either an averaging or a combination approach. The averaging approach (AVP) involves aggregating factor exposures into a single weight vector by averaging the weights of each stock across different power factor portfolios. Let $W_t^{PS} = (w_{t,1}^{PS}, w_{t,2}^{PS}, \ldots, w_{t,M}^{PS})$ be a $(N_t \times M)$ matrix of weights, where $M$ is the number of available characteristics and $w_{t,m}^{PS}$ is the $m^{th}$ column of $W_t^{PS}$ based on characteristic $m = 1, \ldots, M$. Each $w_{t,m}^{PS}$ is estimated based on a set of values for $\hat{\tilde{p}}_t$ and $\hat{\tilde{q}}_t$, specific to the underlying characteristic. The average weight vector is then obtained as:

$$\overline{w}_t^P = \frac{1}{M} W_t^{PS} 1_M', \tag{15}$$

where $1_M$ is a $(1 \times M)$ vector of ones. To ensure a unit sum for the long and short sides, the weights are re-standardized:

$$w_{t,(n)}^{AVP} = \begin{cases} -\frac{\overline{w}_{t,(n)}^{PS}}{\sum\limits_{\overline{w}_{t,(n)}^{PS}<0} \left|\overline{w}_{t,(n)}^{PS}\right|} & \text{for } \overline{w}_{t,(n)}^{PS} < 0 \\ 0 & \text{for } \overline{w}_{t,(n)}^{PS} = 0 \\ \frac{\overline{w}_{t,(n)}^{PS}}{\sum\limits_{\overline{w}_{t,(n)}^{PS}>0} \left|\overline{w}_{t,(n)}^{PS}\right|} & \text{for } \overline{w}_{t,(n)}^{PS} > 0. \end{cases} \tag{16}$$

This approach promises significant diversification benefits by mixing factor exposures and allows for the cut-off point for the long and the short side to deviate from the characteristic median rank. As a benchmark, we repeat the same procedure using equal-weighted decile weights (AVD).

In the second approach, we combine standardized characteristic ranks into an equal-weighted composite score, which serves as a signal for constructing a power-sorted multi-factor portfolio. This approach is called Power-sorted Multi-factor Equal-weight (PME) since each characteristic contributes equally to the combined signal. To illustrate, let $\tilde{S}_t = (\tilde{s}_{\cdot,1}, \ldots \tilde{s}_{\cdot,M})$ be an $(N_t \times M)$ matrix of standardized characteristic ranks for $N_t$ stocks at time $t$. The next step is to use the average standardized characteristic rank as the underlying signal to obtain the weight vector for the composite power portfolio:

$$w_t^{PME} = w_t^{PS} \left( \frac{1}{M} \tilde{S}_t 1_M'; \tilde{p}_t, \tilde{q}_t \right). \tag{17}$$

Again, $\tilde{p}_t$ and $\tilde{q}_t$ are estimated based on the Sharpe ratio maximization objective to derive $w_t^{PME}$. As a benchmark, we use the average characteristic rank in conventional decile sorting, which is referred as Decile Mutli-factor Equal-weight (DME).

In our third approach, we construct the power multi-factor portfolio by using the sum of weights across power-sorted factors as the underlying signal. This approach considers not only the underlying characteristic scores but also the values of the characteristic powers in determining the contribution of each characteristic to the composite score.[8] We name this approach Power-sorted Multi-factor Power portfolio (PMP), and its weights are derived as:

$$w_t^{PMP} = w_t^{PS}\left(\frac{1}{M}W_t^{PS}1_M'; \tilde{p}_t, \tilde{q}_t\right). \tag{18}$$

Similarly to the previous case, the values of $\tilde{p}_t$ and $\tilde{q}_t$ are calibrated to maximize the portfolio Sharpe ratio. This approach assigns higher weights to characteristics that have a better ability to identify the extreme ends of the conditional return distribution. As a benchmark, we construct a decile-sorted multi-factor portfolio using the sum of weights from decile sorting as a ranking variable. We refer to this benchmark as Decile-sorted Multi-factor Decile-weighted (DMD).

Table 2 presents the out-of-sample performance of the three multi-factor strategies, comparing the utilization of power sorting to the conventional benchmark. One might anticipate a reduced opportunity set for power sorting in multi-factor portfolios due to the inclusion of multiple signals. However, our findings show that power sorting consistently outperforms the standard procedure across all construction schemes. This outperformance holds true for both equal-weighted and value-weighted portfolios, highlighting the robustness of our approach.

Regarding specific strategies, AVP and AVD exhibit a similar risk-return profile in the equal-weighted case. However, the strength and robustness of the power sorting procedure become evident when market capitalization is incorporated into the construction of the multi-factor portfolio. In this case, power sorting experiences significantly lower performance deterioration compared to the conventional approach, while also maintaining lower turnover and drawdown risk. This discrepancy in value-weighted results for the two approaches further emphasizes the importance of effectively incorporating the size effect within the factor weighting procedure, demonstrating its positive impact on the risk-return profile of the multi-factor portfolio.

Moving on to the combination approaches, both PME and PMP display significant outperformance in terms of average returns and Sharpe ratios compared to their respective benchmarks. The notable performance advantage of PME over DME demonstrates the ability of power sorting in generating superior portfolios utilizing the same information source, emphasiz-

[8] Consider as an example a hypothetical factor with $\tilde{q} = 1$ and $\tilde{p} = 0$. The concentration ratios indicate that a stock with a low characteristic should be allocated a highly negative weight, while a stock with a high characteristic should be allocated a moderately positive weight, reflecting the varying importance of the characteristic rank.

ing its effectiveness in extracting optimal performance from informative signals. Similarly, the superior performance of PMP over DMD highlights the advantages of combining power sorting with multiple characteristics and emphasizes the effectiveness of this approach in aggregating and integrating various characteristics into a composite signal.

**Table 2: Portfolio evaluation measures for multi-factor portfolios.** AVP: Multi-factor portfolio based on the average portfolio weight from individual power portfolios. AVD: Multi-factor portfolio based on average portfolio weight from individual decile long-short portfolios. PME: Power portfolio on based the average characteristic rank. DME: Decile long-short portfolio based on the average characteristic rank. PMP: Power portfolio based on the rank implied by average power portfolio weights. DMD: Decile long-short portfolio based on the rank implied by the average decile long-short portfolio weights. The sample period covers March 1980 to December 2021. Panel A shows equal-weighted results and Panel B shows value-weighted results.

| | AVP | AVD | PME | DME | PMP | DMD |
|---|---|---|---|---|---|---|
| *A. Equal-weighted portfolios* | | | | | | |
| Return (%) | 1.82 | 1.51 | 3.05 | 2.04 | 3.16 | 1.98 |
| Standard deviation (%) | 4.96 | 4.32 | 7.27 | 6.39 | 7.30 | 5.62 |
| Sharpe Ratio | 1.27 | 1.21 | 1.45 | 1.11 | 1.50 | 1.22 |
| t-stat | 8.28 | 7.89 | 9.44 | 7.19 | 9.72 | 7.92 |
| Maximum drawdown (%) | −46.08 | −41.80 | −55.51 | −56.59 | −55.40 | −53.10 |
| Hit rate (%) | 71.60 | 70.61 | 73.72 | 68.77 | 72.73 | 70.75 |
| Turnover (%) | 37.33 | 41.40 | 48.57 | 52.34 | 66.73 | 59.18 |
| # of effective names long | 1814.32 | 1425.83 | 1608.76 | 390.38 | 146.34 | 390.38 |
| # of effective names short | 600.71 | 935.10 | 101.38 | 390.38 | 141.27 | 390.38 |
| *B. Value-weighted portfolios* | | | | | | |
| Return (%) | 1.36 | 0.84 | 2.80 | 1.72 | 2.50 | 1.32 |
| Standard deviation (%) | 4.57 | 3.54 | 7.40 | 6.42 | 7.07 | 5.39 |
| Sharpe Ratio | 1.03 | 0.82 | 1.31 | 0.93 | 1.23 | 0.85 |
| t-stat | 6.68 | 5.32 | 8.52 | 6.01 | 7.96 | 5.51 |
| Maximum drawdown (%) | −41.21 | −42.01 | −61.54 | −57.63 | −56.61 | −43.26 |
| Hit rate (%) | 67.26 | 63.31 | 69.37 | 64.03 | 68.77 | 61.46 |
| Turnover (%) | 34.00 | 42.89 | 44.64 | 47.84 | 52.56 | 51.64 |
| # of effective names long | 641.23 | 420.89 | 578.14 | 211.20 | 166.51 | 191.04 |
| # of effective names short | 510.43 | 433.92 | 79.87 | 67.56 | 89.25 | 69.45 |

Overall, across the different portfolio weighting methods, PMP stands out with the highest overall return performance, followed by PME. It is worth noting that the key factor driving the performance difference between PMP and PME lies in their long legs. PME takes a diversified approach in the long leg, as indicated by a high number of effective names, suggesting that the combined characteristic rank does not strongly differentiate returns in the long tail. On the other hand, PMP adopts an aggressive long stance, indicating that the combined power portfolio weights can effectively identify strong performers. This outcome highlights the significant effectiveness of power sorting, as it allows for characteristic-specific treatment of weights. By assigning more weight to characteristics with higher concentration ratios, PMP leverages power sorting to identify and capitalize on assets with robust performance potential.

### 3.8 Asset pricing tests

Finally, we examine the asset pricing implications of power sorting across existing asset pricing models using the squared Sharpe ratio test of Barillas et al. (2020). This test enables direct model comparison by quantifying the difference in squared Sharpe ratios between two models, eliminating the need for test assets. Our objective is to assess whether incorporating power-sorted factors into predetermined models enhances the squared Sharpe ratio and, consequently, the pricing ability of these models beyond what is achieved by conventional factors.

We consider three asset pricing models that can be constructed from our characteristic universe. The first model is the 5-factor model (FF5) introduced by Fama and French (2015), which extends the previous 3-factor model by adding profitability and investment factors. The second model, FF5M, follows the framework proposed by Fama and French (2018) augmented by the momentum factor. Our final model is the 4-factor model suggested by Hou, Xue, and Zhang (2015), which includes size, investment, profitability, and the market factor.

To ensure a meaningful comparison, we employ value-adjusted power-sorted factors (using $h = 0.5$) and compare them to the factors provided in the original studies, which we obtained from the authors' websites. It is worth noting that while the proposed models concentrate on similar economic drivers—namely, market, size, profitability, and investment—they diverge in their approaches to constructing the underlying variables. For instance, while both models utilize the percentage change in total assets (agr) as a proxy for investment, Fama and French (2015) emphasize operating profitability (operprof), whereas Hou, Xue, and Zhang (2015) examine return on equity (roeq) as a measure of profitability. Moreover, the original papers have adopted distinct methodologies for constructing these factors. For instance, Fama and French (2015) used independent $2 \times 3$ sorts based on size, although they acknowledge that this choice is quite arbitrary. The motivation behind the $2 \times 3$ sorting methodology is to capture the factor effect across different size groups, ensuring a balanced representation of small and large stocks. By implementing value-adjusted power sorting with a parameter value of $h = 0.5$, we effectively replicate this effect, as it guarantees the inclusion of smaller capitalization stocks in the factor, provided they possess a sufficient characteristic rank. On the other hand, Hou, Xue, and Zhang (2015) conducted a triple $2\times3\times3$ sort on their characteristics to achieve orthogonality among the predictors. This sorting method helps reduce the covariance among factors, thereby decreasing the variance component of the squared Sharpe ratio. Under a power-sorting framework, a similar effect could be achieved by fine-tuning the powers of the factors to minimize factor covariance or even directly maximize the model squared Sharpe ratio.

**Table 3: Asset pricing models based power-sorted versus original factors.** $\theta_P^2$: Squared Sharpe ratio of factor model utilizing power-sorted factors. $\theta_O^2$: Squared Sharpe ratio of factor model utilizing original factors. $\theta_P^2 - \theta_O^2$: Difference in squared Sharpe ratio. We conduct nonnested pairwise model comparisons with traded factors using sequential testing. We first reject the null-hypothesis that the difference between the market factor, which is the only overlapping factor, and a model that includes all the non-overlapping factors from both competing model versions is different from zero. We then test whether the squared Sharpe ratios of the nonnested models are different by computing the p-value as in Barillas et al. (2020).

| | FF5 | FF5M | HXZ |
|---|---|---|---|
| $\theta_P^2$ | 0.236 | 0.281 | 0.238 |
| $\theta_O^2$ | 0.127 | 0.150 | 0.156 |
| $\theta_P^2 - \theta_O^2$ | 0.097 | 0.114 | 0.085 |
| p-value | 0.006 | 0.002 | 0.058 |

Table 3 evidences that models incorporating power-sorted factors consistently outperform conventional models in terms of squared Sharpe ratio across all scenarios. These results are statistically significant at a 1% level for two out of three cases, with the q-theory model proposed by Hou, Xue, and Zhang (2015) exhibiting significance at the 10% level. These findings underscore the significant asset pricing implications of power sorting, demonstrating its capacity to enhance the performance of asset pricing models.

# 4 Benchmarking and Robustness

## 4.1 Alternative Benchmarks

### 4.1.1 Rank portfolios

It is natural to investigate how power sorting compares to alternative factor portfolio weighting schemes.[9] As shown in Figure 2, rank sorting constitutes a special case of power sorting. Hence, it is important to examine whether the incorporation of non-linearities and asymmetries through the use of powers adds value beyond the use of simple rank portfolios. To

[9] In this section, we compare the performance of alternative approaches at a univariate level, while Section A.4 of the Internet Appendix presents results for multi-factor strategies applied to the alternative benchmarks.

this end, Table 4 presents the average portfolio results for power portfolios and rank portfolios, encompassing both value- and equal-weighted cases.

On average, power portfolios deliver a considerably higher annualized Sharpe ratio by providing more than double the return without doubling the risk. In contrast, rank portfolios generally demonstrate lower risk and turnover given broadly diversified positions. Specifically, their weight function corresponds to the rank in a linear manner, resulting in minor weight adjustments. Conversely, power portfolios can adopt more concentrated positions and vary their level of concentration over time, thus introducing an additional layer of turnover. This effect is more noticeable on the short side, where power portfolios tend to exhibit higher concentration, resulting also in a lower number of effective names.

**Table 4: Power sorting versus rank sorting**. Return: Average monthly return, Standard deviation: Monthly standard deviation, Sharpe ratio: Annualized Sharpe ratio, t-stat: t-statistic on $H_0$: Return=0, Hit rate: Percentage frequency of positive returns, Turnover: Average monthly turnover bounded by 200%, # of effective names long: Number of effective names for the long leg, # of effective names short: Number of effective names for the short leg. The sample period is from March 1980 to December 2021. In the value-weighted case we use a value $h = 0.5$ for both rank and power portfolios.

| | Equal-weighted | | Value-weighted | |
|---|---|---|---|---|
| | Power | Rank | Power | Rank |
| Return (%) | 0.77 | 0.32 | 0.62 | 0.22 |
| Standard deviation (%) | 4.04 | 2.78 | 4.17 | 2.68 |
| Sharpe ratio | 0.72 | 0.45 | 0.52 | 0.30 |
| t-stat | 4.64 | 2.91 | 3.40 | 1.97 |
| Maximum drawdown (%) | −45.79 | −41.20 | −48.96 | −41.80 |
| Hit rate (%) | 61.88 | 57.35 | 58.27 | 54.57 |
| Turnover (%) | 37.57 | 28.47 | 33.09 | 23.28 |
| # of effective names long | 1315.41 | 1329.93 | 460.01 | 543.98 |
| # of effective names short | 535.57 | 1331.95 | 229.79 | 622.76 |

#### 4.1.2 Efficient sorting portfolios

Next, we evaluate the performance of power sorting in comparison to the "efficient sorting" approach proposed by Ledoit, Wolf, and Zhao (2019). The term "efficient" refers to minimum variance-optimized factor portfolios that preserve the characteristic spread of the original long-short decile portfolio. Specifically, the weight vector $w_t^{EF}$ at each point in time is estimated as:

$$\min_{w_t^{EF}} w_t^{EF\prime} \hat{H}_t w_t^{EF} \tag{19}$$

$$\text{subject to} \quad x_t' w_t^{EF} = x_t' w_t^{LS} \text{ and} \tag{20}$$

$$\sum_{w_{t,i}^{EF}<0} \left| w_{t,i}^{EF} \right| = \sum_{w_{t,i}^{EF}>0} \left| w_{t,i}^{EF} \right| = 1 \tag{21}$$

where $\hat{H}_t$ is an estimator of the (conditional) VCV. The resulting portfolio is supposed to have the same exposure to the underlying characteristics as the original long-short portfolio because of (20), but smaller variance, and therefore a higher Sharpe ratio because of (19).[10] For the estimation of $H_t$, we employ the Quadratic-Inverse Shrinkage estimator proposed by Ledoit and Wolf (2022). Specifically, at each investment date, we estimate $\hat{H}_t$ for stocks with available return history over the most recent five years (i.e., 1260 days), which considerably reduces the viable investment universe in the comparison.[11] Finally, we winsorize the cross-sectional characteristic vector $m_t$ at each period $t$, following the methodology outlined in Ledoit, Wolf, and Zhao (2019).

A virtue of power sorting is that there is no need for computing a VCV at an individual stock level. We though investigate whether enriching the power sorting procedure by $\hat{H}_t$ is beneficial, thus making efficient sorting and power sorting more comparable. In particular, we modify equation (13) as:

$$var(r^{PS}) = w_t^{PS\prime} \hat{H}_t w_t^{PS}. \tag{22}$$

Figure 8 illustrates the average weight function for selected factors under power sorting and efficient sorting, alongside the conditional volatility across quantile groups. Evidently, the two approaches differ in terms of portfolio construction, reflecting their distinct underlying objective functions. Specifically, power sorting aims to exploit variations in conditional returns to maximize the factor portfolio Sharpe ratio, while efficient sorting focuses on minimizing variance while maintaining the same characteristic spread.

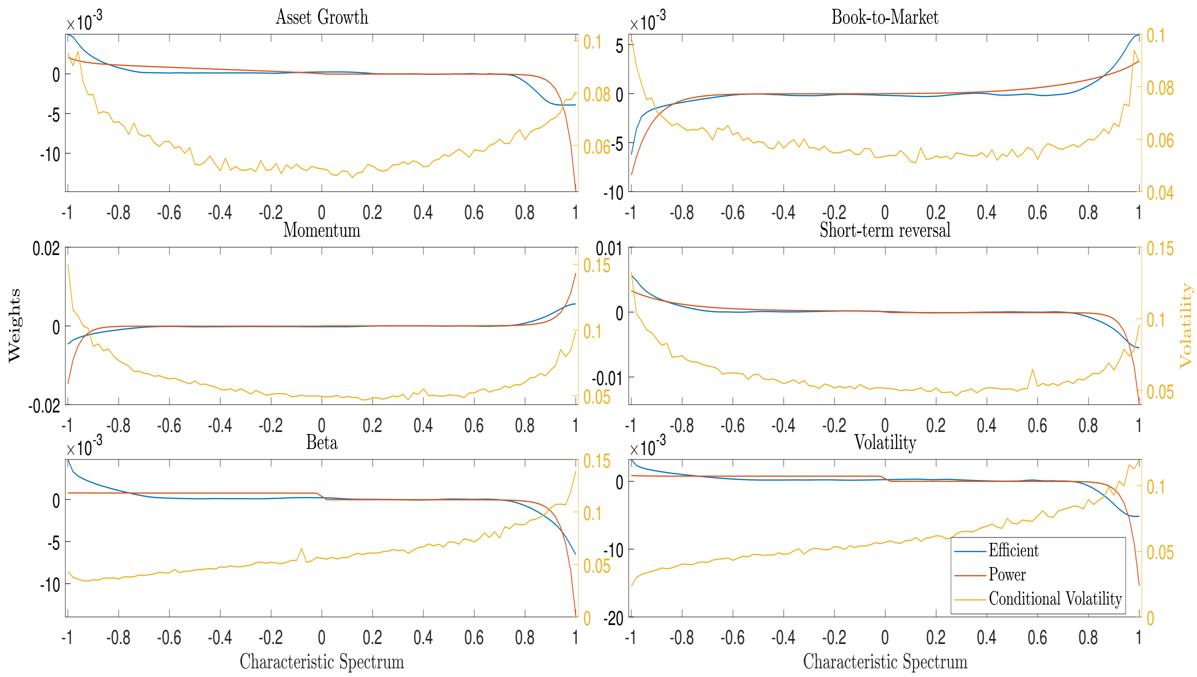
To gain insights into how these different objectives translate into portfolio decisions, consider the volatility factor as an illustrative example. Recall from Figure 1 that stocks with the highest volatility exhibit relatively lower average returns, while those with the lowest volatility do not demonstrate significant outperformance. Consequently, the power-sorted portfolio adopts

[10]Note that we additionally incorporate a maximum weight constraint of 2% to align with the power sorting framework and to prevent the minimum-variance optimizer from generating excessively large and imbalanced positions for the long and short side.

[11]The effect of this constraint on the sample size is illustrated in Figure A.2 in the Internet Appendix.

an aggressive stance on the short side and a more diversified one on the long side to capitalize on this pattern. In contrast, the efficient sorting portfolio aims to minimize variance by reducing exposure to stocks with the highest volatility on the short side and increasing exposure to those with the lowest volatility on the long side. Similar conclusions can also be drawn for beta, while in other cases efficient sorting tends to take a more passive stance, particularly on the short side. Only in the case of book-to-market ratio, efficient sorting adopts a more aggressive stance than power sorting, even though this behavior does not align with conditional returns. This result lies in the fact that as we move towards the extremes (high and low book-to-market ratios), the covariance between the long and short positions increases, leading to a reduction in the overall long-short variance. Hence, efficient sorting falls short of fully capturing the relationship between characteristics and returns beyond what is implied by covariance alone. On the other hand, power sorting integrates characteristic, return, and variance information, directly targeting the Sharpe ratio, while preserving the factor structure through the imposition of monotonicity.

**Figure 8: Average weight function for efficient sorting and power sorting.** Characteristics are standardized in the $[-1, 1]$ range. The conditional volatilities (orange lines) are estimated by ranking stocks based on their underlying characteristic in the previous period and splitting them into 100 equal-weighted quantile portfolios. The blue lines represent the weight function under efficient sorting, while the dashed blue lines depict the weight function under power sorting. The sample includes all common shares on NYSE, AMEX, and NASDAQ exchanges and covers the period from January 1980 to December 2021.



Next, Table 5 presents the average portfolio evaluation measures using the viable sample (that only includes stocks that have five years of return history at a given point in time) for the four approaches: the conventional approach, the efficient sorting approach, the original power sorting approach, and the modified power sorting approach utilizing the stock-level VCV, which we label as Power-VCV. Consistent with the findings of Ledoit, Wolf, and Zhao (2019), efficient sorting consistently reduces factor portfolio variance (2.90% for efficient sorting vs 3.96% for conventional sorting). However, this reduction in variance comes at a slight cost of lower average returns. This result aligns with the notion that characteristics are intertwined with the covariance structure of returns (Kelly, Pruitt, and Su, 2019). Consequently, when attempting to limit the underlying factor portfolio variance, there is an unavoidable trade-off with the underlying risk premia. Additionally, we note that the number of effective names for the efficient sorting approach implies a more symmetric stance that focuses on both extremes, even though the characteristic-return relationship is often asymmetric.

**Table 5: Power sorting versus efficient sorting.** Return: Average monthly return, Stand. dev.: Monthly standard deviation, Sharpe ratio: Annualized Sharpe ratio, t-stat: t-statistic on $H_0$: Return=0, Hit rate: Percentage frequency of positive returns, Turnover: Average monthly turnover bounded by 200%, # of effective names long: Number of effective names for the long leg, # of effective names short: Number of effective names for the short leg. The sample includes stocks with an available return history of five years at each investment date through the sample period from March 1980 to December 2021.

| | Power | Power-VCV | Efficient | Conventional |
|---|---|---|---|---|
| Return (%) | 0.50 | 0.44 | 0.31 | 0.36 |
| Standard deviation (%) | 3.73 | 3.46 | 2.90 | 3.96 |
| Sharpe ratio | 0.52 | 0.50 | 0.37 | 0.34 |
| t-stat | 3.35 | 3.24 | 2.43 | 2.19 |
| Maximum drawdown (%) | −45.50 | −43.63 | −44.71 | −55.62 |
| Hit rate (%) | 58.89 | 58.87 | 56.07 | 55.78 |
| Turnover (%) | 35.72 | 35.96 | 40.83 | 38.46 |
| # of effective names long | 940.97 | 1031.10 | 529.46 | 270.23 |
| # of effective names short | 434.85 | 478.56 | 493.05 | 264.10 |

On the other hand, power sorting effectively captures the inherent asymmetries in many characteristics, leading to a significant increase in average factor portfolio returns, along with a slight decrease in portfolio variance compared to the conventional method. As a result, the average Sharpe ratios and t-statistics show notable enhancements. Specifically, our findings demonstrate a 53% increase in the average t-stat through power sorting, compared to an 11% increase with efficient sorting. Importantly, this result remains consistent regardless of whether the variance is estimated directly from the power portfolio time-series or using a VCV approach.

Finally, note that while power sorting portfolios may exhibit higher volatility than efficient sorting portfolios, this increased volatility does not translate into higher drawdown risk, with average turnover being also lower.

## 4.2 Robustness Tests

Next, we analyze whether the presented results generalize to different sub-periods and methodological alternations. We show that power sorting generates performance that is robust to the choice of maximum weight thresholds, different size adjustment levels, and different sub-periods.

### 4.2.1 Lookback window

In the base case we employ an expanding window for estimating the optimal concentration ratios. Here, we explore the out-of-sample power using different rolling windows ranging from 12 months up to 10 years. With the expanding window, the estimated concentration ratios converge toward the values that were most effective through the whole sample, while a rolling window is more adaptive. Shorter windows adapt more dynamically to recent information, potentially introducing higher variation in the concentration ratios and resulting in more pro-cyclical strategies.

Table 6 presents the average portfolio evaluation measures across power portfolios for different lookback windows. To ensure consistency regarding the length of the evaluation period, results are assessed for the out-of-sample period from January 1990 onward. Our findings reveal that both short and long windows yield similar return performances, with the expanding window showing a slight advantage on average. Generally, longer lookback windows achieve comparable investment performance while maintaining significantly lower turnover, making them more desirable from a practical standpoint. Overall, results remain consistent across different formation periods, with power portfolios consistently outperforming the conventional benchmark, and the results not being driven by higher turnover or tail risk.

**Table 6: Robustness with respect to different lookback windows.** Return: average monthly return, Standard deviation: monthly standard deviation, Sharpe ratio: Annualized Sharpe ratio, t-stat: t-statistic on $H_0$: return=0, Hit rate: percentage frequency of positive returns, Turnover: monthly turnover bounded by 200%, # of effective names long: Number of effective names for the long leg, # of effective names short: Number of effective names for the short leg. The sample period is from January 1990 to December 2021. Panel A shows equal-weighted and Panel B shows value-weighted results.

| Lookback window | 12 Months | 36 Months | 60 Months | 120 Months | Expanding Window | Conventional |
|---|---|---|---|---|---|---|
| *A. Equal-weighted portfolios* | | | | | | |
| Return (%) | 0.65 | 0.60 | 0.62 | 0.63 | 0.65 | 0.43 |
| Standard deviation (%) | 4.50 | 4.53 | 4.49 | 4.41 | 4.30 | 4.55 |
| Sharpe ratio | 0.54 | 0.50 | 0.53 | 0.55 | 0.58 | 0.38 |
| t-stat | 3.04 | 2.86 | 3.00 | 3.11 | 3.31 | 2.14 |
| Maximum drawdown (%) | −46.14 | −49.75 | −49.44 | −48.93 | −45.70 | −54.04 |
| Hit rate (%) | 58.61 | 58.81 | 59.35 | 59.79 | 60.05 | 55.99 |
| Turnover (%) | 50.16 | 42.54 | 39.97 | 37.84 | 36.39 | 39.03 |
| # of effective names long | 1133.67 | 1201.18 | 1261.32 | 1311.78 | 1369.61 | 374.18 |
| # of effective names short | 709.94 | 594.08 | 564.05 | 523.48 | 518.53 | 377.85 |
| *B. Value-weighted portfolios* | | | | | | |
| Return (%) | 0.48 | 0.45 | 0.47 | 0.48 | 0.53 | 0.26 |
| Standard deviation (%) | 4.46 | 4.51 | 4.52 | 4.50 | 4.37 | 4.66 |
| Sharpe ratio | 0.39 | 0.36 | 0.38 | 0.39 | 0.44 | 0.21 |
| t-stat | 2.20 | 2.05 | 2.17 | 2.24 | 2.53 | 1.16 |
| Maximum drawdown (%) | −47.60 | −51.77 | −51.28 | −50.95 | −48.18 | −59.06 |
| Hit rate (%) | 55.77 | 55.74 | 56.34 | 56.45 | 56.82 | 52.65 |
| Turnover (%) | 48.23 | 39.27 | 36.24 | 33.81 | 32.61 | 35.44 |
| # of effective names long | 412.18 | 432.73 | 448.25 | 465.18 | 454.03 | 109.18 |
| # of effective names short | 306.44 | 260.62 | 239.16 | 216.16 | 187.25 | 100.92 |

#### 4.2.2 Maximum weights

In the main analysis, we opted for a maximum portfolio weight of 2% to ensure that the power portfolios are properly diversified. Table 7 shows how average results change for alternative choices of maximum portfolio weight, ranging from 0.5% to 10%. Higher maximum weights lead to higher values for $p_t^{max}$ and $q_t^{max}$ and hence to weight distributions that are potentially more concentrated in the tails, delivering higher returns at the expense of higher risk and turnover.

Increasing the upper weight threshold up to 10% can result in higher average return gains for portfolio performance, with the Sharpe ratio remaining practically unchanged. In the equal-weighted case, the maximum Sharpe ratio is achieved at a 3% weight threshold (0.73). For value-weighted data, the maximum Sharpe ratio is achieved with a weight concentration of 5% (0.54), suggesting that a higher power threshold is required to extract optimal performance after accounting for market capitalization. Importantly, power sorting does not appear to excessively increase concentration in both legs, even when higher maximum weight thresholds are allowed.

This can be attributed, in part, to the objective of maximizing the power portfolio Sharpe ratio, which helps maintain a balance between concentration and diversification.

**Table 7: Robustness with respect to the maximum stock weight.** Panel A shows equal-weighted results and Panel B shows value-weighted results. Return: average monthly return, Standard deviation: monthly standard deviation Sharpe ratio: Annualized Sharpe ratio, t-stat: t-statistic on $H_0$: return=0, Hit rate: percentage frequency of positive returns Turnover: monthly turnover bounded by 200%, # of effective names long: Number of effective names for the long leg, # of effective names short: Number of effective names for the short leg. The sample covers the period from March 1980 to December 2021. Panel A shows equal-weighted results and Panel B shows value-weighted results.

| $w^{ceil}$ | 0.5% | 1% | 2% | 3% | 4% | 5% | 10% |
|---|---|---|---|---|---|---|---|
| *A. Equal-weighted portfolios* | | | | | | | |
| Return (%) | 0.53 | 0.67 | 0.77 | 0.83 | 0.85 | 0.86 | 0.87 |
| Standard deviation (%) | 3.26 | 3.72 | 4.04 | 4.21 | 4.33 | 4.40 | 4.49 |
| Sharpe ratio | 0.65 | 0.69 | 0.72 | 0.73 | 0.72 | 0.72 | 0.71 |
| t-stat | 4.19 | 4.50 | 4.64 | 4.72 | 4.71 | 4.70 | 4.64 |
| Maximum drawdown (%) | −41.07 | −44.28 | −45.79 | −46.10 | −47.34 | −47.28 | −48.11 |
| Hit rate (%) | 60.68 | 61.65 | 61.88 | 62.09 | 62.21 | 62.11 | 62.18 |
| Turnover (%) | 32.49 | 35.55 | 37.57 | 38.62 | 39.29 | 39.69 | 40.01 |
| # of effective names long | 1413.22 | 1342.05 | 1315.41 | 1298.95 | 1290.90 | 1288.57 | 1287.28 |
| # of effective names short | 816.46 | 612.07 | 535.57 | 508.47 | 493.30 | 487.75 | 491.10 |
| *B. Value-weighted portfolios* | | | | | | | |
| Return (%) | 0.39 | 0.52 | 0.62 | 0.67 | 0.70 | 0.71 | 0.72 |
| Standard deviation (%) | 3.34 | 3.81 | 4.17 | 4.35 | 4.48 | 4.57 | 4.67 |
| Sharpe ratio | 0.44 | 0.49 | 0.52 | 0.53 | 0.54 | 0.54 | 0.53 |
| t-stat | 2.84 | 3.20 | 3.40 | 3.47 | 3.49 | 3.49 | 3.46 |
| Maximum drawdown (%) | −44.83 | −47.47 | −48.96 | −49.62 | −50.27 | −50.84 | −52.02 |
| Hit rate (%) | 56.78 | 57.74 | 58.27 | 58.34 | 58.34 | 58.53 | 58.60 |
| Turnover (%) | 31.12 | 34.08 | 36.02 | 37.24 | 38.01 | 38.41 | 38.70 |
| # of effective names long | 1306.86 | 1197.04 | 1144.98 | 1128.32 | 1119.88 | 1121.61 | 1117.87 |
| # of effective names short | 371.61 | 275.21 | 229.79 | 214.08 | 202.74 | 197.20 | 196.74 |

Conversely, the enforcement of high diversification via a low weight threshold may moderate the effectiveness of power sorting in exploiting return-relevant characteristic variation (i.e., setting $w^{ceil} = 0.5\%$). Nevertheless, values below 1% can lead to maximum power thresholds below one for the different factors ($p_t^{max} \wedge q_t^{max} < 1$), rendering them insufficient upper bounds for examining concentration in the tails.

### 4.2.3 Concentration in mega-cap stocks

In the base case, we employed $h = 0.5$ to address extreme concentration in mega-cap stocks when evaluating value-weighted results. Here, we examine the implications of different values of $h$ on the performance of value-weighted power portfolios. Additionally, we consider different variations of the conventional decile sorts to assess the sensitivity of the conventional approach with regard to the treatment of size effects. First, we compute "pure" value-weighted portfolios without winsorizing market-caps of individual stocks, effectively setting $h = 1$ in the power sorting framework. Second, since $h = 0.5$ is equivalent to using the square root, we analyze the effect of employing the square root of the market cap within the conventional decile sorting approach.

Table 8 displays the average portfolio evaluation results for values of $h$ ranging from 0 (equal-weighted) to 1 (pure value-weighted), along with the different benchmark variations. Lower values yield better portfolio performance as they minimize the effect of size on portfolio composition. Nonetheless, even when there is no size adjustment for the value-weighted power portfolios ($h = 1$), power sorting outperforms capped and unadjusted value-weighted decile sorting. Moreover, the performance of $h = 0.5$ significantly surpasses that of the square root approach in the conventional framework, reflecting the ability of the method to effectively incorporate size effects in portfolio construction.

**Table 8: Robustness with respect to concentration on mega-cap stocks**. Conv. VW (Capped): Stocks are weighted by their market cap winsorized at the NYSE $80^{th}$ percentile. Conv. VW (No Adj.): Stocks are weighted by their market cap without any adjustment. Conv. VW (Square root): Stocks are weighted by the square root of their market cap. Return: average monthly return, Standard deviation: monthly standard deviation Sharpe ratio: Annualized Sharpe ratio, t-stat: t-statistic on $H_0$: return=0, Hit rate: percentage frequency of positive returns, Turnover: monthly turnover bounded by 200%, # of effective names long: Number of effective names for the long leg, # of effective names short: Number of effective names for the short leg. The sample period covers March 1980 to December 2021.

| | Power Sorting | | | | | Conventional VW | | |
|---|---|---|---|---|---|---|---|---|
| | h=0 | h=0.25 | h=0.5 | h=0.75 | h=1 | Capped | No Adj. | Square root |
| Return (%) | 0.77 | 0.71 | 0.62 | 0.52 | 0.43 | 0.32 | 0.29 | 0.39 |
| Standard deviation (%) | 4.04 | 4.09 | 4.17 | 4.41 | 4.74 | 4.39 | 4.78 | 4.20 |
| Sharpe Ratio | 0.72 | 0.64 | 0.52 | 0.39 | 0.29 | 0.26 | 0.20 | 0.33 |
| t-stat | 4.64 | 4.15 | 3.40 | 2.54 | 1.85 | 1.71 | 1.31 | 2.19 |
| Maximum drawdown (%) | −45.79 | −47.10 | −48.96 | −52.91 | −58.50 | −59.22 | −65.05 | 56.81 |
| Hit rate (%) | 61.88 | 60.53 | 58.27 | 55.77 | 54.20 | 53.44 | 52.82 | 55.19 |
| Turnover (%) | 37.57 | 35.05 | 33.09 | 32.54 | 33.11 | 35.44 | 35.85 | 34.71 |
| # of effective names long | 1315.41 | 975.46 | 460.01 | 180.73 | 73.60 | 107.00 | 37.09 | 165.09 |
| # of effective names short | 535.57 | 409.18 | 229.79 | 116.87 | 67.47 | 98.42 | 37.21 | 158.65 |

### 4.2.4 Sub-period analysis

Finally, we conduct a decade-by-decade analysis in Table 9 which shows the average portfolio evaluation measures for power portfolios and the conventional approach for the four subperiods. The magnitude of the difference between power and conventional sorting covaries with the efficacy of factor investing as a whole, corroborating that results are driven by extracting optimal performance from the underlying factors rather than introducing other effects on the portfolio construction procedure. Confirming our full-sample analysis, the added value of power sorting is consistently positive within the chosen sub-period and is not driven by specific time periods. However, it is important to highlight that the performance of factor investing as a whole exhibits a noticeable decline in later years, as observed in previous studies (McLean and Pontiff, 2016).

**Table 9: Robustness across different sub-periods.** Return: average monthly return, Standard deviation: monthly standard deviation Sharpe ratio: Annualized Sharpe ratio, t-stat: t-statistic on $H_0$: return=0, Hit rate: percentage frequency of positive returns, Turnover: monthly turnover bounded by 200%, # of effective names long: Number of effective names for the long leg, # of effective names short: Number of effective names for the short leg. The sample period covers March 1980 to December 2021. Panel A shows equal-weighted results and Panel B shows value-weighted results.

| | 1980/1989 | 1990/1999 | 2000/2010 | 2011/2021 |
|---|---|---|---|---|
| *A. Equal-weighted portfolios* | | | | |
| *A.1 Power sorting* | | | | |
| Return (%) | 1.19 | 0.90 | 0.68 | 0.43 |
| Standard deviation (%) | 2.81 | 3.53 | 5.64 | 3.39 |
| Sharpe ratio | 1.40 | 0.97 | 0.50 | 0.44 |
| t-stat | 4.42 | 3.07 | 1.58 | 1.52 |
| Maximum drawdown (%) | −16.01 | −25.14 | −33.87 | −31.97 |
| Hit rate (%) | 67.81 | 63.75 | 60.21 | 57.01 |
| Turnover (%) | 41.49 | 38.50 | 38.52 | 32.92 |
| # of effective names long | 1138.13 | 1595.78 | 1461.20 | 1115.79 |
| # of effective names short | 590.50 | 610.45 | 548.69 | 421.27 |
| *A.2 Conventional* | | | | |
| Return (%) | 0.83 | 0.60 | 0.43 | 0.24 |
| Standard deviation (%) | 2.76 | 3.68 | 5.94 | 3.70 |
| Sharpe ratio | 1.00 | 0.66 | 0.31 | 0.23 |
| t-stat | 3.16 | 2.09 | 0.99 | 0.79 |
| Maximum drawdown (%) | −20.68 | −30.80 | −42.56 | −38.83 |
| Hit rate (%) | 62.33 | 59.25 | 55.86 | 53.13 |
| Turnover (%) | 40.59 | 40.67 | 40.11 | 36.81 |
| # of effective names long | 352.98 | 450.04 | 405.37 | 288.44 |
| # of effective names short | 345.68 | 455.04 | 411.05 | 289.38 |
| *B. Value-weighted portfolios* | | | | |
| *B.1 Power sorting* | | | | |
| Return (%) | 0.95 | 0.77 | 0.56 | 0.29 |
| Standard deviation (%) | 2.95 | 3.84 | 5.79 | 3.38 |
| Sharpe ratio | 1.02 | 0.72 | 0.37 | 0.29 |
| t-stat | 3.21 | 2.27 | 1.18 | 1.03 |
| Maximum drawdown (%) | −19.49 | −28.47 | −36.20 | −31.21 |
| Hit rate (%) | 62.71 | 60.50 | 56.27 | 54.60 |
| Turnover (%) | 38.33 | 33.06 | 33.58 | 28.54 |
| # of effective names long | 448.13 | 541.06 | 463.29 | 402.63 |
| # of effective names short | 288.98 | 257.16 | 213.02 | 174.07 |
| *B.2 Conventional* | | | | |
| Return (%) | 0.46 | 0.33 | 0.43 | 0.13 |
| Standard deviation (%) | 3.28 | 3.98 | 6.10 | 3.59 |
| Sharpe ratio | 0.48 | 0.29 | 0.28 | 0.11 |
| t-stat | 1.51 | 0.92 | 0.87 | 0.40 |
| Maximum drawdown (%) | −28.29 | −37.71 | −42.26 | −36.64 |
| Hit rate (%) | 55.60 | 54.86 | 53.33 | 50.63 |
| Turnover (%) | 37.33 | 35.34 | 35.47 | 34.00 |
| # of effective names long | 100.05 | 115.65 | 121.79 | 93.90 |
| # of effective names short | 90.42 | 109.48 | 108.19 | 88.25 |

## 5 Conclusion

We propose power sorting as a framework for constructing equity factors to improve upon conventional quantile sorting. Our method hinges on the assumption of a monotonic relationship between factor characteristics and returns. It is geared at creating refined versions of the factors while facilitating the construction of economically meaningful and sufficiently diversified portfolios. We deem the power sorting procedure as an effective compromise between conventional portfolio sorts and machine learning methods. While the former easily fails to account for characteristic-specific information, the latter is usually criticized for its lack of interpretability and its black box character. By striking a balance between interpretability and computational efficiency, our framework offers practical advantages. Under our modeling procedure, concentration ratios directly translate to weight concentration levels, allowing for a simple and intuitive interpretation of the model parameters.

We present several important empirical findings. First, we document the existence of asymmetric and non-linear patterns between characteristics and returns. Such patterns contradict the notion that the return signal is always amplified at the extremes and motivate separate treatment of the long and the short side of factor portfolios. As a consequence, off-the-shelf procedures may struggle to harvest the underlying factor premiums, which, in turn, can lead to false rejections of individual characteristics. The limitations of the conventional approach become more evident when dealing with value-weighted portfolios, as it fails to adequately account for confounding size effects. Unlike standard approaches, our method is designed to extract optimal performance from the vast majority of characteristics by allowing the weight function to be characteristic-specific and effectively incorporating size-effects in the construction of factors.

Building on these insights, we investigate the performance gains resulting from power sorting compared to the conventional quantile approach. Power sorting can generate average returns and Sharpe ratios that are up to double those achieved through conventional quantile sorting. These gains are both economically and statistically significant, survive size-adjustments, and are not driven by increased turnover or tail risk. Furthermore, the benefits persist when considering alternative optimization-based portfolio formulation approaches, suggesting that the use of exponential functions to model factor portfolio weights introduces structure to the weight vector that is beneficial in terms of out-of-sample performance.

The outperformance of power-sorted factor portfolios primarily stems from taking an aggressive stance on the short leg and adopting a more diversified one on the long leg. Hence, our results demonstrate that various characteristics are effective in identifying underperforming stocks, although they may provide mixed signals for outperforming stocks. Nonetheless, power sorting boosts performance in both the long and short leg of the various factor portfolios.

Lastly, the benefits of power sorting extend to a multi-factor level. For instance, by adopting power-sorted factors in existing asset pricing models, we can enhance the squared Sharpe ratio of the underlying model, thus increasing its ability to capture the cross-section of stock returns. In the context of multi-factor strategies, power sorting implicitly accounts for the informativeness of characteristics across the characteristic spectrum, yielding multi-factor portfolios with improved risk-return properties compared to simple equal-weighted schemes and individual factors.

## References


Aït-Sahalia, Y., and M. W. Brandt. 2001. Variable selection for portfolio choice. *Journal of Finance* 56(4):1297–1351.

Ammann, M., G. Coqueret, and J.-P. Schade. 2016. Characteristics-based portfolio choice with leverage constraints. *Journal of Banking & Finance* 70:23–37.

Ang, A., R. J. Hodrick, Y. Xing, and X. Zhang. 2006. The cross-section of volatility and expected returns. *Journal of Finance* 61(1):259–299.

Bansal, R., and A. Yaron. 2004. Risks for the long run: A potential resolution of asset pricing puzzles. *Journal of Finance* 59(4):1481–1509.

Banz, R. W. 1981. The relationship between return and market value of common stocks. *Journal of Financial Economics* 9(1):3–18.

Barillas, F., R. Kan, C. Robotti, and J. Shanken. 2020. Model comparison with Sharpe ratios. *Journal of Financial and Quantitative Analysis* 55(6):1840–1874.

Basu, S. 1977. Investment performance of common stocks in relation to their price-earnings ratios: A test of the efficient market hypothesis. *Journal of Finance* 32(3):663–682.

Blitz, D., G. Baltussen, and P. van Vliet. 2020. When equity factors drop their shorts. *Financial Analysts Journal* 76(4):73–99.

Brandt, M. W. 1999. Estimating portfolio and consumption choice: A conditional Euler equations approach. *Journal of Finance* 54(5):1609–1645.

Brandt, M. W., and P. Santa-Clara. 2006. Dynamic portfolio selection by augmenting the asset space. *Journal of Finance* 61(5):2187–2217.

Brandt, M. W., P. Santa-Clara, and R. Valkanov. 2009. Parametric portfolio policies: Exploiting characteristics in the cross-section of equity returns. *Review of Financial Studies* 22(9):3411–3447.

Campbell, J. Y., and J. H. Cochrane. 1999. By force of habit: A consumption-based explanation of aggregate stock market behavior. *Journal of Political Economy* 107(2):205–251.

Campbell, J. Y., and R. J. Shiller. 1988. The dividend-price ratio and expectations of future dividends and discount factors. *Review of Financial Studies* 1(3):195–228.

Cattaneo, M. D., R. K. Crump, M. H. Farrell, and E. Schaumburg. 2020. Characteristic-sorted portfolios: Estimation and inference. *Review of Economics and Statistics* 102(3):531–551.

Cooper, M. J., H. Gulen, and M. J. Schill. 2008. Asset growth and the cross-section of stock returns. *Journal of Finance* 63(4):1609–1651.

Daniel, K., L. Mota, S. Rottke, and T. Santos. 2020. The cross-section of risk and returns. *Review of Financial Studies* 33(5):1927–1979.

DeMiguel, V., A. Martin-Utrera, F. J. Nogales, and R. Uppal. 2020. A transaction-cost perspective on the multitude of firm characteristics. *Review of Financial Studies* 33(5):2180–2222.

Engle, R. F., O. Ledoit, and M. Wolf. 2019. Large dynamic covariance matrices. *Journal of Business & Economic Statistics* 37(2):363–375.

Fama, E. F., and K. R. French. 1992. The cross-section of expected stock returns. *Journal of Finance* 47(2):427–465.

———. 1993. Common risk factors in the returns on stocks and bonds. *Journal of Financial Economics* 33(1):3–56.

———. 2015. A five-factor asset pricing model. *Journal of Financial Economics* 116(1):1–22.

———. 2018. Choosing factors. *Journal of Financial Economics* 128(2):234–252.

———. 2020. Comparing cross-section and time-series factor models. *Review of Financial Studies* 33(5):1891–1926.

Fama, E. F., and J. D. MacBeth. 1973. Risk, return, and equilibrium: Empirical tests. *Journal of Political Economy* 81(3):607–636.

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

Green, J., J. R. Hand, and F. Zhang. 2017. The characteristics that provide independent information about average U.S. monthly stock returns. *Review of Financial Studies* 30(12):4389–4436.

Harvey, C. R., Y. Liu, and H. Zhu. 2016. . . . and the cross-section of expected returns. *Review of Financial Studies* 29(1):5–68.

He, Z., and A. Krishnamurthy. 2013. Intermediary asset pricing. *American Economic Review* 103(2):732–770.

Hjalmarsson, E., and P. Manchev. 2012. Characteristic-based mean-variance portfolio choice. *Journal of Banking & Finance* 36(5):1392–1401.

Hong, H., T. Lim, and J. C. Stein. 2000. Bad news travels slowly: Size, analyst coverage, and the profitability of momentum strategies. *Journal of Finance* 55(1):265–295.

Hou, K., C. Xue, and L. Zhang. 2015. Digesting anomalies: An investment approach. *Review of Financial Studies* 28(3):650–705.

———. 2020. Replicating anomalies. *Review of Financial Studies* 33(5):2019–2133.

Jacobs, B. I., and K. N. Levy. 1993. Long/short equity investing. *Journal of Portfolio Management* 20(1):52.

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

Jensen, T. I., B. T. Kelly, and L. H. Pedersen. 2023. Is there a replication crisis in finance? *Journal of Finance,* forthcoming.

Kelly, B. T., S. Pruitt, and Y. Su. 2019. Characteristics are covariances: A unified model of risk and return. *Journal of Financial Economics* 134(3):501–524.

Kim, S., R. A. Korajczyk, and A. Neuhierl. 2021. Arbitrage portfolios. *Review of Financial Studies* 34(6):2813–2856.

Knuth, D. E. 1993. Johann Faulhaber and sums of powers. *Mathematics of Computation* 61(203):2 77–294.

Koijen, R. S., T. J. Moskowitz, L. H. Pedersen, and E. B. Vrugt. 2018. Carry. *Journal of Financial Economics* 127(2):197–225.

Ledoit, O., and M. Wolf. 2022. Quadratic shrinkage for large covariance matrices. *Bernoulli* 28(3):1519–1547.

Ledoit, O., M. Wolf, and Z. Zhao. 2019. Efficient sorting: A more powerful test for cross-sectional anomalies. *Journal of Financial Econometrics* 17(4):645–686.

Leung, E., H. Lohre, D. Mischlich, Y. Shea, and M. Stroh. 2021. The promises and pitfalls of machine learning for predicting stock returns. *Journal of Financial Data Science* 3(2):21–50.

McGee, R. J., and J. Olmo. 2022. Optimal characteristic portfolios. *Quantitative Finance* 22(10):1 853–1870.

McLean, R. D., and J. Pontiff. 2016. Does academic research destroy stock return predictability? *Journal of Finance* 71(1):5–32.

Novy-Marx, R., and M. Velikov. 2022. Betting against betting against beta. *Journal of Financial Economics* 143(1):80–106.

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

Simon, F., S. Weibels, and T. Zimmermann. 2023. *Deep parametric portfolio policies.* Tech. rep. CFR Working Paper.

Soebhag, A., B. van Vliet, and P. Verwijmeren. 2023. Non-standard errors in asset pricing: Mind your sorts. *Available at SSRN.*

Stambaugh, R. F., J. Yu, and Y. Yuan. 2012. The short of it: Investor sentiment and anomalies. *Journal of Financial Economics* 104(2):288–302.

Swade, A., S. Nolte, M. Shackleton, and H. Lohre. 2023. Why do equally weighted portfolios beat value-weighted ones? *Journal of Portfolio Management* 49(5):167–187.

Zhang, X., L. Wu, and Z. Chen. 2022. Constructing long-short stock portfolio with a new listwise learn-to-rank algorithm. *Quantitative Finance* 22(2):321–331.

# Power Sorting

# Internet Appendix

Anastasios Kagkadis[†] Harald Lohre [†§] Ingmar Nolte[†]

Sandra Nolte[†] Nikolaos Vasilas[†]

[†] *Lancaster University Management School*

[§] *Robeco Quantitative Investments*

October 6, 2023

**Abstract**

We report characteristic sample information and supplementary results for "Power Sorting."

Send correspondence to Nikolaos Vasilas, n.vasilas@lancaster.ac.uk

## A.1 List of characteristics

**Table A1:** Listing of firm characteristics used in the study, including the source and the exact definition.

| Acronym | Author(s) | Journal | Definition | Group |
|---|---|---|---|---|
| absacc | Bandyopadhyay, Huang, & Wirjanto | 2010, WP | Absolute value of acc. | Investment |
| acc | Sloan | 1996, TAR | Annual income before extraordinary items (ib) minus operating cash flows (oancf) divided by average total assets (at); if oancf is missing then set to change in act – change in che – change in lct + change in dlc + change in txp–dp. | Investment |
| aeavol | Lerman, Livnat, and Mendenhall | 2008, WP | Average daily trading volume (vol) for 3 days around earnings announcement minus average daily volume for 1-month ending 2 weeks before earnings announcement divided by 1-month average daily volume. Earnings announcement day from Compustat quarterly (rdq) | Momentum |
| age | Jiang, Lee, & Zhang | 2005, RAS | Number of years since first Compustat coverage. | Intangibles |
| agr | Cooper, Gulen & Schill | 2008, JF | Annual percentage change in total assets (at). | Investment |
| baspread | Amihud & Mendelson | 1989, JF | Monthly average of daily bid-ask spread divided by average of daily spread. | Frictions |
| beta | Fama & MacBeth | 1973, JPE | Estimated market beta from weekly returns and equal weighted market returns for 3 years ending month t-1 with at least 52 weeks of returns. | Frictions |
| betasq | Fama & MacBeth | 1973, JPE | Market beta squared. | Frictions |
| bm | Rosenberg, Reid, & Lanstein | 1985, JPM | Book value of equity (ceq) divided by fiscal year end market capitalization. | Value |
| bm_ia | Asness, Porter & Stevens | 2000, WP | Industry adjusted book-to-market ratio. | Value |
| cash | Palazzo | 2012, JFE | Cash and cash equivalents divided by average total assets | Intangibles |
| cashdebt | Ou & Penman | 1989, JAE | Earnings before depreciation and extraordinary items (ib+dp) divided by avg. total liabilities (lt). | Intangibles |

| Acronym | Author(s) | Journal | Definition | Group |
|---|---|---|---|---|
| cashpr | Chandrashekar & Rao | 2009, WP | Fiscal year end market capitalization plus long-term debt (dltt) minus total assets (at) divided by cash and equivalents (che). | Intangibles |
| cfp | Desai, Rajgopal & Venkatachalam | 2004, TAR | Operating cash flows divided by fiscal year end market capitalization. | Value |
| cfp_ia | Asness, Porter & Stevens | 2000, WP | Industry adjusted cfp. | Value |
| chatoia | Soliman | 2008, TAR | 2-digit SIC fiscal year mean adjusted change in sales (sale) divided by average total assets (at). | Profitability |
| chcsho | Pontiff & Woodgate | 2008, JF | Annual percentage change in shares outstanding (csho). | Investment |
| chempia | Asness, Porter & Stevens | 1994, WP | Industry-adjusted change in number of employees. | Intangibles |
| chinv | Thomas & Zhang | 2002, RAS | Change in inventory (inv) scaled by average total assets (at). | Investment |
| chmom | Gettleman & Marks | 2006, WP | Cumulative returns from months t-6 to t-1 minus months t-12 to t-7. | Momentum |
| chpmia | Soliman | 2008, TAR | 2-digit SIC fiscal year mean adjusted change in income before extraordinary items (ib) divided by sales (sale). | Profitability |
| chtx | Thomas and Zhang | 2011, JAR | Percent change in total taxes (txtq) from quarter t-4 to t | Momentum |
| cinvest | Titman, Wei, and Xie | 2004, JFQA | Change over one quarter in net PP&E (ppentq) divided by sales (saleq) - average of this variable for prior 3 quarters; if saleq = 0, then scale by 0.01 | Investment |
| currat | Ou & Penman | 1989, JAE | Current assets / current liabilities. | Investment |
| depr | Holthausen & Larcker | 1992, JAE | Depreciation over PPE. | Investment |
| dolvol | Chordia, Subrahmanyam, & Anshuman | 2001, JFE | Natural log of trading volume times price per share from month t-2. | Frictions |
| ear | Kishore et al. | 2008, WP | Sum of daily returns in three days around earnings announcement. Earnings announcement from Compustat quarterly file (rdq) | Momentum |
| egr | Richardson, Sloan, Soliman & Tuna | 2005, JAE | Annual percentage change in book value of equity (ceq). | Investment |

| Acronym | Author(s) | Journal | Definition | Group |
|---|---|---|---|---|
| ep | Basu | 1977, JF | Annual income before extraordinary items (ib) divided by end of fiscal year market capitalization. | Value |
| gma | Novy-Marx | 2013, JFE | Revenues (revt) minus cost of goods sold (cogs) divided by lagged total assets (at). | Profitability |
| grcapx | Anderson & Garcia-Feijoo | 2006, JF | Percentage change in capital expenditures from year t-2 to year t. | Investment |
| grltnoa | Fairfield, Whisenant & Yohn | 2003, TAR | Growth in long term net operating assets. | Investment |
| herf | Hou & Robinson | 2006, JF | 2-digit SIC fiscal year sales concentration (sum of squared percentage of sales in industry for each company). | Intangibles |
| hire | Bazdresch, Belo & Lin | 2014, JPE | Percentage change in number of employees (emp). | Intangibles |
| idiovol | Ali, Hwang, & Trombley | 2003, JFE | Standard deviation of residuals of weekly returns on weekly equal weighted market returns for 3 years prior to month end. | Frictions |
| ill | Amihud | 2002, JFM | Average of daily (absolute return/dollar volume). | Frictions |
| indmom | Moskowitz & Grinblatt | 1999, JF | Equal weighted average industry 12-month returns. | Momentum |
| invest | Chen & Zhang | 2010, JF | Annual change in gross property, plant, and equipment (ppegt) + annual change in inventories (invt) all scaled by lagged total assets (at). | Investment |
| lev | Bhandari | 1988, JF | Total liabilities (lt) divided by fiscal year end market capitalization. | Investment |
| lgr | Richardson, Sloan, Soliman & Tuna | 2005, JAE | Annual percentage change in total liabilities (lt). | Investment |
| maxret | Bali, Cakici & Whitelaw | 2011, JFE | Maximum daily return from returns during calendar month t-1. | Frictions |
| mom12m | Jegadeesh | 1990, JF | 11-month cumulative returns ending one month before month end. | Momentum |
| mom1m | Jegadeesh & Titman | 1993, JF | 1-month cumulative return. | Frictions |
| mom36m | Jegadeesh & Titman | 1993, JF | Cumulative returns from months t-36 to t-13. | Value |
| mom6m | Jegadeesh & Titman | 1993, JF | 5-month cumulative returns ending one month before month end. | Momentum |
| ms | Mohanram | 2005, RAS | Sum of 8 indicator variables for fundamental performance | Profitability |

| Acronym | Author(s) | Journal | Definition | Group |
|---|---|---|---|---|
| mve | Banz | 1981, JFE | Natural log of market capitalization at end of month t-1. | Frictions |
| mve_ia | Asness, Porter, & Stevens | 2000, WP | 2-digit SIC industry-adjusted fiscal year end market capitalization. | Frictions |
| operprof | Fama & French | 2015, JFE | Revenue minus cost of goods sold - SG&A expense - interest expense divided by lagged common shareholders' equity. | Profitability |
| orgcap | Eisfeldt and Papanikolaou | 2013, JF | Capitalized SG&A expenses | Intangibles |
| pchcapx_ia | Abarbanell & Bushee | 1998, TAR | 2-digit SIC fiscal year mean adjusted percentage change in capital expenditures (capx). | Investment |
| pchcurrat | Ou & Penman | 1989, JAE | Percentage change in currat. | Investment |
| pchdepr | Holthausen & Larcker | 1992, JAE | Percentage change in depreciation. | Investment |
| pchgm_pchsale | Abarbanell & Bushee | 1998, TAR | Percentage change in gross margin (sale-cogs) minus percentage change in sales (sale). | Intangibles |
| pchquick | Ou & Penman | 1989, JAE | Percentage change in quick. | Investment |
| pchsale_pchinvt | Abarbanell & Bushee | 1998, TAR | Annual percentage change in sales (sale) minus annual percentage change in inventory (invt). | Intangibles |
| pchsale_pchrect | Abarbanell & Bushee | 1998, TAR | Annual percentage change in sales (sale) minus annual percentage change in receivables (rect). | Intangibles |
| pchsale_pchxsga | Abarbanell & Bushee | 1998, TAR | Annual percentage change in sales (sale) minus annual percentage change in SG&A (xsga). | Intangibles |
| pchsaleinv | Ou & Penman | 1989, JAE | Percentage change in sales-to-inventory ratio. | Intangibles |
| pctacc | Hafzalla, Lundholm & Van Winkle | 2011, TAR | Same as acc except that the numerator is divided by the absolute value of ib; if ib = 0 then ib set to 0.01 for denominator. | Investment |
| pricedelay | Hou & Moskowitz | 2005, RFS | The proportion of variation in weekly returns for 36 months ending in month t explained by 4 lags of weekly market returns incremental to contemporaneous market return. | Intangibles |
| ps | Piotroski | 2000, JAR | Sum of 9 indicator variables to form fundamental health score. | Profitability |
| quick | Ou & Penman | 1989, JAE | (current assets – inventory) / current liabilities. | Investment |

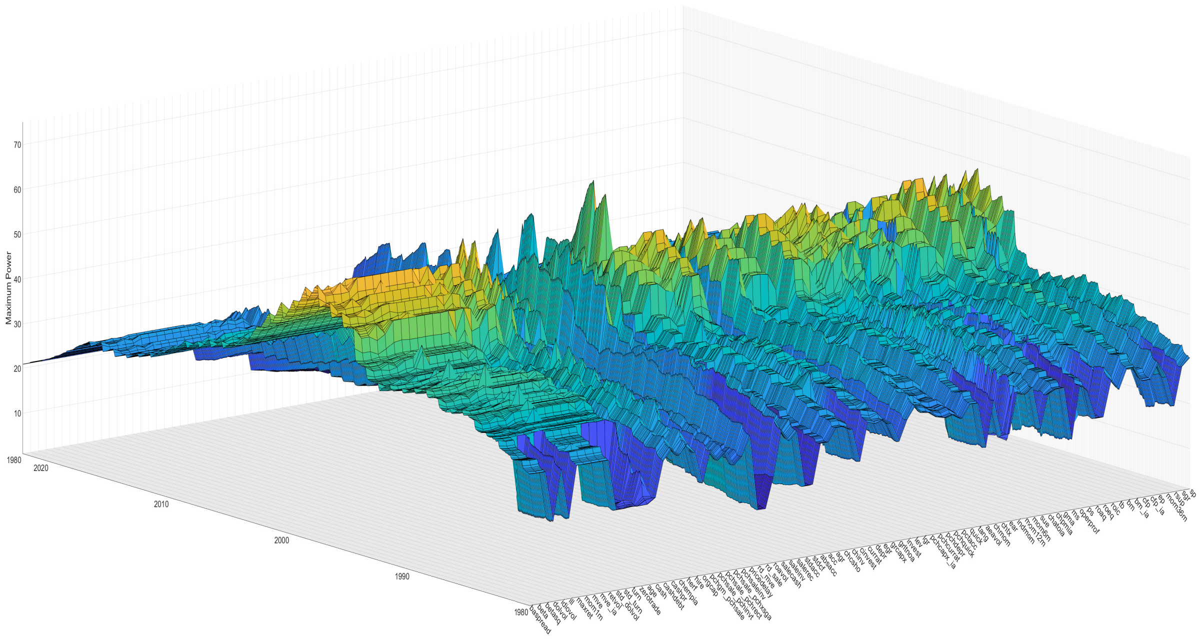
| Acronym | Author(s) | Journal | Definition | Group |
|---|---|---|---|---|
| rd_mve | Guo, Lev & Shi | 2006, JBFA | R&D expense divided by end of fiscal year market capitalization. | Intangibles |
| rd_sale | Guo, Lev & Shi | 2006, JBFA | R&D expense divided by sales (xrd/sale). | Intangibles |
| retvol | Ang et al. | 2006, JF | Standard deviation of daily returns from month t-1. | Frictions |
| roaq | Balakrishnan, Bartov, and Faurel | 2010, JAE | Income before extraordinary items (ibq) divided by one quarter lagged total assets (atq) | Profitability |
| roavol | Francis et al. | 2004, TAR | Standard deviation for 16 quarters of income before extraordinary items (ibq) divided by average total assets (atq) | Intangibles |
| roeq | Hou, Xue, and Zhang | 2015 RFS | Earnings before extraordinary items divided by lagged common shareholders equity | Profitability |
| roic | Brown & Rowe | 2007, WP | Annual earnings before interest and taxes (ebit) minus non-operating income (nopi) divided by non-cash enterprise value (ceq+lt–che). | Profitability |
| rsup | Kama | 2009, JBFA | Sales from quarter t minus sales from quarter t-4 (saleq) divided by fiscal-quarter-end market capitalization (cshoq * prccq) | Value |
| salecash | Ou& Penman | 1989, JAE | Annual sales divided by cash and cash equivalents. | Intangibles |
| saleinv | Ou& Penman | 1989, JAE | Annual sales divided by total inventory. | Intangibles |
| salerec | Ou& Penman | 1989, JAE | Annual sales divided by accounts receivable. | Intangibles |
| sgr | Lakonishok, Shleifer & Vishny | 1994, JF | Annual percentage change in sales (sale). | Value |
| sp | Barbee, Mukherji, & Raines | 1996, FAJ | Annual revenue (sale) divided by fiscal year end market capitalization. | Value |
| std_dolvol | Chordia, Subrahmanyam, & Anshuman | 2001, JFE | Monthly standard deviation of daily dollar trading volume. | Frictions |
| std_turn | Chordia, Subrahmanyam, & Anshuman | 2001, JFE | Monthly standard deviation of daily share turnover. | Frictions |
| stdacc | Bandyopadhyay, Huang, and Wirjanto | 2010, WP | Standard deviation for 16 quarters of accruals (acc measured with quarterly Compustat) scaled by sales; if saleq = 0, then scale by 0.01 | Intangibles |
| stdcf | Huang | 2009, JEF | Standard deviation for 16 quarters of cash flows divided by sales (saleq); if saleq = 0, then scale by 0.01. Cash flows defined as ibq minus quarterly accrual | Intangibles |

| **Acronym** | **Author(s)** | **Journal** | **Definition** | **Group** |
| --- | --- | --- | --- | --- |
| sue | Rendelman, Jones, and Latane | 1982, JFE | Unexpected quarterly earnings divided by fiscal-quarter-end market cap. Unexpected earnings is I/B/E/S actual earnings minus median forecasted earnings if available, else it is the seasonally differenced quarterly earnings before extraordinary | Momentum |
| tang | Almeida & Campello | 2007, RFS | Cash holdings + 0.715 × receivables + 0.547 × inventory + 0.535 × PPE/total assets. | Investment |
| tb | Lev & Nissim | 2004, TAR | Tax income, calculated from current tax expense divided by maximum federal tax rate, divided by income before extraordinary items. | Profitability |
| turn | Datar, Naik, & Radcliffe | 1998, JFM | Average monthly trading volume for most recent 3 months scaled by number of shares outstanding in current month. | Frictions |
| zerotrade | Liu | 2006, JFE | Turnover weighted number of zero trading days for most recent 1 month. | Frictions |

## A.2 Supplementary Figures

**Figure A.1: Upper Thresholds for $q_t^{max}$ (a) and $p_t^{max}$ (b) for the 85 characteristics using a Maximum Weight Constraint of ($w^{ceil}$) 2%.** The figures illustrate the time variability in the maximum threshold due to the varying number of the cross-sections across different characteristics. The maximum power thresholds vary with characteristics, reflecting the different characteristic variabilities within each characteristic due to the time-varying size of the cross-sections and between the long and short legs of the same characteristic due to the presence of ties in the underlying characteristic distribution. The sample period is from January 1980 to December 2021.

**(a)** $q_t^{max}$




**(b)** $p_t^{max}$

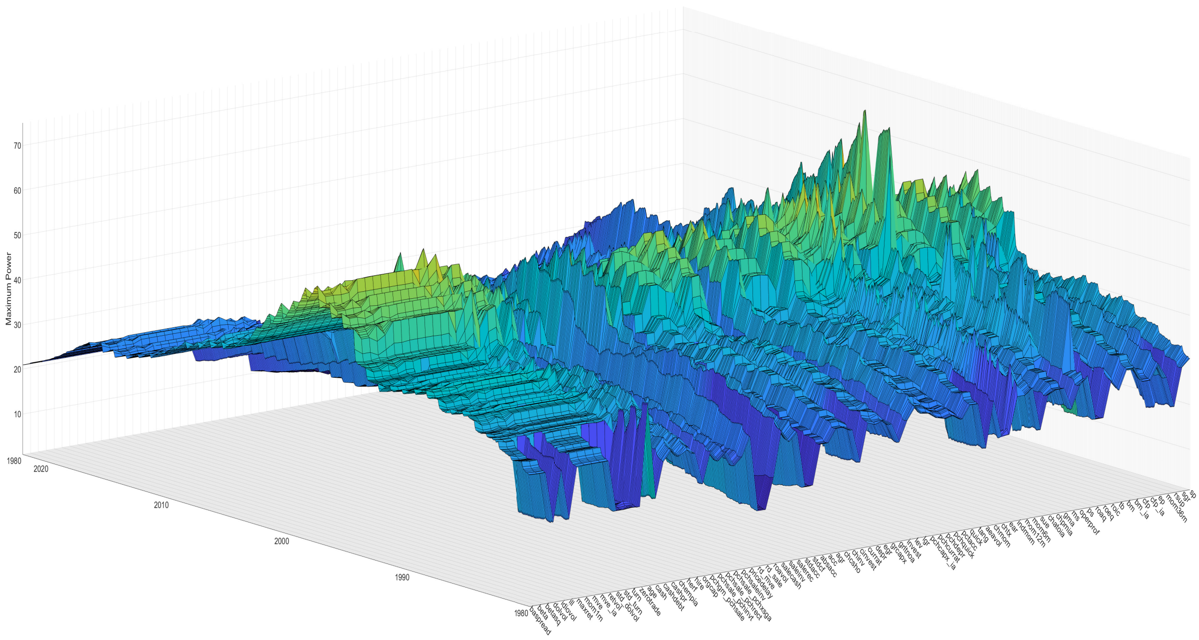
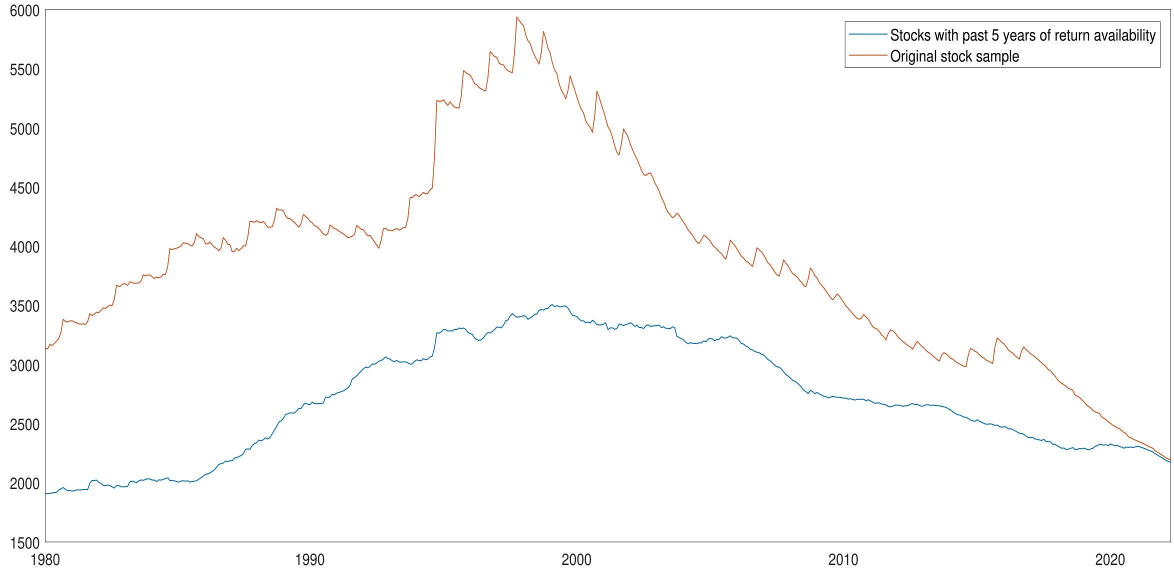


**Figure A.2: Cross-sectional stock size for original sample and sub-sample with five years of daily return data**. The chart illustrates the difference in the number of stocks in the sample when five years of past daily data are required. The disparity is particularly pronounced in earlier periods, notably during the build-up of the dot com bubble, but becomes less noticeable in later years.




## A.3 Return-spread maximization objective

In the base case, power portfolios are constructed based on a Sharpe ratio maximization objective. To explore the sensitivity of the approach to the underlying objective, we here construct power portfolios under a return maximization objective. Table A.2 presents the average portfolio statistics for these power portfolios, alongside the decile-sorted benchmarks. The results show strong consistency with those in Table 1, confirming the significant outperformance of power sorting over the conventional benchmark.

As expected, the switch from Sharpe ratio to returns leads to power portfolios with a higher average return but a lower Sharpe ratio, in line with the new objective. This is achieved by slightly increasing concentration in the tails, as evident from a lower number of effective names for both the long and short sides.

**Table A.2: Power portfolios under a return maximization objective**. Return: Average monthly return, Standard deviation.: Monthly standard deviation, Sharpe ratio: Annualized Sharpe ratio, t-stat: t-statistic on $H_0$: Return=0, Maximum drawdown: Maximum drawdown, Hit rate: Percentage frequency of positive returns, Turnover: Average monthly turnover (bounded by 200%), # of effective names long: Number of effective names for the long leg, # of effective names short: Number of effective names for the short leg. The sample period is from March 1980 to December 2021.

| | Equal-weighted | | Value-weighted | |
|---|---|---|---|---|
| | Power | Conventional | Power | Conventional |
| Return (%) | 0.91 | 0.51 | 0.69 | 0.32 |
| Standard deviation (%) | 5.03 | 4.21 | 4.95 | 4.39 |
| Sharpe ratio | 0.65 | 0.46 | 0.48 | 0.26 |
| t-stat | 4.24 | 2.96 | 3.12 | 1.71 |
| Maximum drawdown (%) | −56.22 | −55.35 | −57.05 | −59.22 |
| Hit rate (%) | 61.21 | 57.39 | 57.83 | 53.44 |
| Turnover (%) | 39.39 | 39.93 | 35.05 | 35.44 |
| # of effective names long | 1168.99 | 369.24 | 425.66 | 107.00 |
| # of effective names short | 267.78 | 370.33 | 155.07 | 98.42 |

## A.4 Multi-factor strategies for alternative benchmarks

We briefly examine how the results reported in Section 5 generalize for multi-factor strategies. Table A.3 reports the results for the three mutli-factor strategies, comparing rank portfolios to power portfolios. Evidently, multi-factor portfolios based on rank portfolios demonstrate very similar performance across the three strategies, indicating the method's inability to effectively combine signals in an objective-oriented fashion. In contrast, power portfolio-based strategies clearly showcase the method's capability to combine individual factors into multi-factor portfolios with maximum Sharpe ratio.

When comparing averaging strategies, AVP outperforms AVR in terms of average returns and Sharpe ratios, while also exhibiting less tail risk. In the case of the combination approaches, which are more aggressive in nature, PME and PMP deliver more than double the return compared to the rank-based strategies, leading to significantly higher Sharpe ratios and t-statistics. Overall, our results demonstrate that the strict enforcement of a linear weighting scheme hinders performance at both a univariate and a multivariate level. Furthermore, a simple rank-based approach leads to portfolios that inherit a passive stance, limiting the effective extraction of underlying signals or the combination of different signals.

**Table A.3: Portfolio evaluation measures for multi-factor power and rank portfolios.** AVP: Multi-factor portfolio based on the average portfolio weight from individual power portfolios. AVR: Mutli-factor portfolio based on the average portfolio weight from individual rank portfolios. PME: Power portfolio based on the average characteristic rank. RME: Rank portfolio based on the average characteristic rank. PMP: Power portfolio based on the rank implied by average power portfolio weights. RMR: Rank portfolio based on the rank implied by average rank portfolio weights. Panel A shows equal-weighted results and Panel B shows value-weighted results. The sample period is from March 1980 to December 2021.

| | AVP | AVR | PME | RME | PMP | RMR |
|---|---|---|---|---|---|---|
| *A. Equal-weighted portfolios* | | | | | | |
| Return (%) | 1.82 | 1.45 | 3.05 | 1.27 | 3.16 | 1.26 |
| Standard deviation (%) | 4.96 | 4.60 | 7.27 | 4.29 | 7.30 | 4.16 |
| Sharpe ratio | 1.27 | 1.09 | 1.45 | 1.03 | 1.50 | 1.05 |
| t-stat | 8.28 | 7.09 | 9.44 | 6.66 | 9.72 | 6.82 |
| Maximum drawdown (%) | −46.08 | −41.81 | −55.51 | −39.17 | −55.40 | −38.11 |
| Hit rate (%) | 71.60 | 69.03 | 73.72 | 67.19 | 72.73 | 68.77 |
| Turnover (%) | 37.33 | 32.82 | 48.57 | 32.16 | 66.73 | 31.47 |
| # of effective names long | 1814.32 | 1392.70 | 1608.76 | 1463.92 | 146.34 | 1463.92 |
| # of effective names short | 600.71 | 1150.37 | 101.38 | 1463.92 | 141.27 | 1463.92 |
| *B. Value-weighted portfolios* | | | | | | |
| Return (%) | 1.36 | 0.97 | 2.80 | 1.00 | 2.50 | 0.98 |
| Standard deviation (%) | 4.57 | 4.23 | 7.40 | 4.33 | 7.07 | 4.20 |
| Sharpe ratio | 1.03 | 0.80 | 1.31 | 0.80 | 1.23 | 0.81 |
| t-stat | 6.68 | 5.17 | 8.52 | 5.22 | 7.96 | 5.25 |
| Maximum drawdown (%) | −41.21 | −42.03 | −61.54 | −43.12 | −56.61 | −42.29 |
| Hit rate (%) | 67.26 | 63.31 | 69.37 | 62.45 | 68.77 | 62.45 |
| Turnover (%) | 34.00 | 29.57 | 44.64 | 27.04 | 52.56 | 26.30 |
| # of effective names long | 641.23 | 404.75 | 578.14 | 562.56 | 166.51 | 556.31 |
| # of effective names short | 510.43 | 1062.73 | 79.87 | 849.04 | 89.25 | 821.85 |

Next, we see how efficient and power sorting generalize to a multi-factor level using the updated sample. In Panel A of Table A.4, we begin by analyzing the averaging strategy. Power sorting exhibits the highest average return, while efficient sorting exhibits the smallest volatility and highest Sharpe ratio. Nonetheless, the performance differences across all variations, including the conventional approach, are relatively small. This can be attributed to the averaging strategy, which blends exposures without adequately differentiating the strength of the underlying signal.

**Table A.4: Portfolio evaluation measures for multi-factor power and efficient portfolios.** Return: Average monthly return, Standard deviation: Monthly standard deviation, Sharpe ratio: Annualized Sharpe ratio, t-stat: t-statistic on $H_0$: Return=0, Maximum drawdown: Maximum drawdown, Hit rate: Percentage frequency of positive returns, Turnover: Average monthly turnover bounded by 200%, # of effective names long: Number of effective names for the long leg, # of effective names short: Number of effective names for the short leg. The sample includes stocks with an available return history of five years at each investment date through the period from March 1980 to December 2021. Panel A shows the Weight averaging strategy, Panel B shows the average characteristic rank strategy, and Panel C shows the average weight rank strategy.

| | Power | Power VCV | Efficient | Conventional |
|---|---|---|---|---|
| *A. Weight Averaging Strategy* | | | | |
| Return (%) | 1.38 | 1.28 | 1.19 | 1.14 |
| Standard deviation (%) | 4.58 | 4.67 | 3.52 | 3.81 |
| Sharpe ratio | 1.04 | 0.95 | 1.17 | 1.04 |
| t-stat | 6.78 | 6.15 | 7.59 | 6.75 |
| Maximum drawdown (%) | −43.85 | −45.96 | −38.96 | −41.82 |
| Hit rate (%) | 68.97 | 66.01 | 70.36 | 70.36 |
| Turnover (%) | 36.63 | 35.76 | 39.96 | 42.14 |
| # of effective names long | 1244.12 | 1242.10 | 995.91 | 990.89 |
| # of effective names short | 434.08 | 488.97 | 601.37 | 632.24 |
| *B. Average Characteristic Rank Strategy* | | | | |
| Return (%) | 2.25 | 1.93 | 1.52 | 1.53 |
| Standard deviation (%) | 6.41 | 5.51 | 4.31 | 5.61 |
| Sharpe ratio | 1.22 | 1.22 | 1.22 | 0.95 |
| t-stat | 7.90 | 7.90 | 7.92 | 6.14 |
| Maximum drawdown (%) | −55.28 | −55.28 | −50.02 | −57.66 |
| Hit rate (%) | 70.16 | 69.96 | 71.34 | 68.18 |
| Turnover (%) | 47.26 | 46.71 | 66.76 | 52.68 |
| # of effective names long | 1051.80 | 1195.59 | 440.34 | 270.94 |
| # of effective names short | 103.19 | 294.13 | 271.34 | 270.94 |
| *C. Average Weight Rank Strategy* | | | | |
| Return (%) | 2.39 | 2.29 | 1.23 | 1.55 |
| Standard deviation (%) | 6.99 | 7.37 | 3.97 | 5.03 |
| Sharpe ratio | 1.18 | 1.08 | 1.07 | 1.06 |
| t-stat | 7.68 | 6.99 | 6.96 | 6.91 |
| Maximum drawdown (%) | −50.76 | −59.18 | −45.41 | −51.68 |
| Hit rate (%) | 70.95 | 69.37 | 70.16 | 70.55 |
| Turnover (%) | 61.99 | 63.35 | 71.56 | 59.73 |
| # of effective names long | 99.17 | 98.74 | 443.71 | 270.94 |
| # of effective names short | 134.26 | 128.99 | 413.40 | 270.94 |

Panel B presents the results for the strategy that utilizes the average characteristic rank. In the case of efficient sorting, the characteristics are cross-sectionally standardized and added together, rather than using their ranks, to maintain consistency with the original framework. The optimization-based approaches clearly outperform the conventional approach when the underlying signal is informative, resulting in a significant 30% increase in the Sharpe ratio. While the power sorting approaches primarily increase average returns through asymmetric concentration in the tails, the efficient sorting approach achieves similar results by reducing variance. Moreover, the efficient sorting approach demonstrates the lowest drawdown, albeit at a cost of 20% higher turnover per month compared to power sorting.

Finally, Panel C displays the findings for the strategy that utilizes the sum of weights as the underlying signal. Once again, power sorting emerges with the highest Sharpe ratio by maximizing average returns, despite exhibiting the highest variance. In contrast, efficient sorting demonstrates the lowest volatility, although it comes with the highest turnover. The conventional approach falls in the middle, achieving a Sharpe ratio comparable to that of the sophisticated approaches.

Notably, the sophisticated approaches inherently differ in their approach to factor portfolio construction. Specifically, when there is a strong underlying signal, as is the case with the sum of power portfolio weights, power sorting adopts an aggressive stance by increasing the concentration ratios to maximize performance. This results in a significant improvement in average returns, albeit at the expense of higher volatility. In contrast, efficient sorting does not distinguish between weak and strong signals, consistently striving to minimize variance, even if it slightly reduces the underlying premia. However, it is worth noting that a variance reduction objective can also be achieved through a power sorting framework.