Multivariate Factors: Accounting for the Joint Dependence among Characteristics

Gustavo Freire[*a,b,c], **Rasmus Lönn**[†a,c], and **Anastasija Tetereva**[‡a,b]

[a]Erasmus University Rotterdam
[b]Tinbergen Institute
[c]Erasmus Research Institute of Management

December 15, 2025

Abstract

We propose a new approach for constructing characteristics-based factor portfolios. Instead of forming independent quantile or rank sorts, we allocate weights proportional to the conditional distribution of one characteristic given the other characteristics. We model this conditional distribution in a data-driven way using copulas. Empirically, we apply our method to the five Fama-French characteristics-based factor portfolios. Relative to the original construction, our factors increase the maximum attainable Sharpe ratio from 1.04 to 1.80 and decrease by half the number of anomalies with significant alphas in the cross-section of stock returns.

Keywords: Factor models, Cross-section of returns, Characteristic portfolios, Anomalies, Copulas

JEL Classification: C58, G11, G12

[*]freire@ese.eur.nl.
[†]lonn@ese.eur.nl.
[‡]tetereva@ese.eur.nl.

1 Introduction

The main goal of cross-sectional asset pricing research has been to find a factor model that can price stock returns. In theory, such a factor model exists and can be represented by the conditional mean-variance efficient portfolio of individual assets: $\Sigma_t^{-1}\mu_t$. However, this portfolio is essentially unfeasible in practice as it requires the inversion of a large covariance matrix and knowledge of conditional expected returns. Alternatively, researchers have turned to heuristic approaches for factor construction exploiting information about expected returns contained in firm characteristics. As discussed by Kozak and Nagel (2024), the most prominent examples are sorted factors (e.g., Fama and French, 1993), obtained by sorting stocks by characteristics and forming portfolios based on quantile cutoffs, and univariate factors (e.g., Kozak, Nagel, and Santosh, 2020), constructed with weights proportional to the cross-sectional rank for each characteristic.

When constructing characteristics-based factor portfolios, it is usually desirable to isolate the effect of one characteristic from the others. For instance, in the motivation of their five-factor model, Fama and French (2015) use the dividend discount model to determine how book-to-market, profitability or investment should relate to expected returns *fixing the other characteristics to be constant*. Empirically, they consider several approaches for factor construction based on forming different independent quantile sorts, recognizing that a particular choice of quantile cutoffs is inherently arbitrary. This includes the traditional 2×3 sorts on size and the characteristic of interest, and $2\times2\times2\times2$ sorts meant to provide allocations more accurately isolating the specific premium associated with a given characteristic.

In this paper, we propose a new approach for constructing characteristics-based factors that formally accounts for the joint dependence among characteristics while avoiding the need to pre-specify quantile cutoffs. We allocate portfolio weights proportional to the conditional distribution of each characteristic given the others, i.e., maximizing cross-sectional variation in the characteristic of interest while keeping the others fixed. This is a natural generalization of univariate factors, which allocate weights based on the marginal distribution of a given characteristic. We transform the marginal distributions into the joint multivariate distribution of the characteristics in a simple and data-driven way using copulas. By effectively neutralizing the effect of other characteristics, we can better assess the individual premium associated with a particular factor.

More specifically, at every rebalancing month, the marginal distributions of the characteristics in the cross-section of stocks are individually estimated by the empirical cumulative distribution functions. Their inverse functions, which are uniformly distributed between zero and one, are used as inputs to estimate the parameters of the copula by maximum likelihood. Using information criteria, we select the most suitable copula from the following candidates: Gaussian, Student-t and the most common Archimedean copulas (Clayton, Gumbel and Frank), each of which capture different asymmetries in the distribution of characteristics. Once the copula is estimated, the conditional distribution is computed as the derivative of the copula with respect to the characteristics we want to condition on and be neutral to. The availability of analytical derivatives is one of the reasons why we employ parametric copulas. Finally, the conditional distribution is scaled into portfolio weights to form a dollar-neutral portfolio.

We apply our method to construct a new version of the Fama and French (2015) five-factor model, for which the theoretical motivation strongly relies on making characteristics conditional on each other. Our benchmarks are: the original 2×3 sorted factors; univariate factors based on those characteristics, which our approach generalizes; the original $2 \times 2 \times 2 \times 2$ factors, which aim to make factors neutral with respect to each other with independent quantile sorts; and the recent Daniel, Mota, Rottke, and Santos (2020) hedged factors, which focus on hedging unpriced risk associated with characteristics sorts. We investigate whether our approach, which uses the same characteristics as the benchmarks, is able to generate factors with unspanned priced information and improved explanatory power for the cross-section of returns.

We start by comparing the individual factors in the models. Relative to the original 2×3 construction, the copula factors have a similar premium, but with smaller volatilities that can be even halved. This results in substantially larger Sharpe ratios and stronger statistical significance for the premium of each factor. Univariate factors, on the other hand, are more erratic, with value and investment factors with high t-statistics for their average returns, but a profitability factor that is statistically insignificant. The $2 \times 2 \times 2 \times 2$ and hedged factors generally improve upon the original construction, but are outperformed by our approach uniformly across the factors.

The individual factor comparison suggests that the copula factors isolate better the premium associated with each characteristic. We next analyze whether this enhanced premium can be explained by the benchmark approaches for factor construction. We consider spanning regressions where we regress each of the copula factors on the benchmark factor models. We find that the alphas from these regressions are positive and statistically significant in nearly all cases, confirming that our approach extracts priced variation from the characteristics that is not spanned by existing methods. The only exceptions are for the value factor when regressed on univariate factors and hedged factors. Interestingly, univariate factors explain a large fraction of the time-series variation of the copula factors, but not their expected returns, reinforcing the importance of accounting for the joint dependence among characteristics in isolating their premia.

The improvement associated with formally making characteristics neutral with respect to each other manifests itself not only in the individual Sharpe ratios of the factors, but also in the maximum Sharpe ratio that can be attained with the factors, i.e., the Sharpe ratio of the tangency portfolio of the factors. This is the relevant metric for comparing linear factor models in terms of their ability to span the mean-variance frontier (Barillas and Shanken, 2017). The copula factors combined yield an annualized Sharpe ratio of 1.80, which is higher than that obtained with the benchmarks and, in particular, the 1.04 Sharpe ratio of the tangency portfolio of the original 2×3 factors. In other words, our factors expand considerably the mean-variance frontier relative to the existing approaches. Relatedly, we also show that the copula tangency portfolio is not spanned by the tangency portfolios of the benchmarks.

Finally, we assess the performance of the competing five-factors models in pricing a large set of 150 anomalies from Hou, Xue, and Zhang (2020). Overall, the copula factors deliver the smallest average pricing errors and Gibbons, Ross, and Shanken (1989) statistics. More strikingly, the number of anomalies with statistically significant alphas – following the rule-of-thumb of Harvey, Liu, and Zhu (2016) – is more than halved when controlling for our factors relative to the benchmarks. Specifically, 22 anomalies survive the copula factor model, while 49 cannot be explained by the original 2×3 factors. Looking at specific classes of anomalies, we observe that most of this improvement comes from pricing momentum anomalies. That is, our Fama-French five-factor model accounting for the joint dependence among characteristics is more successful in explaining momentum, which Fama and French (2016) show is a challenge for the original specification.

The remainder of the paper is structured as follows. After a brief review of the related literature, Section 2 motivates and describes the portfolio sorting methodology we propose. In Section 3, we conduct a simulation exercise comparing different sorting approaches. Our empirical results are presented in Section 4, while Section 5 concludes the paper. Lastly, Appendix A provides the details of our simulation design, while Appendix B collects the figures and tables of the paper.

1.1 Related literature

Our paper is related to the vast literature exploiting the information content of firm characteristics about cross-sectional differences in stock returns to build portfolios. Based on the evidence from Fama and French (1992) that firm size and book-to-market are associated with spreads in stock returns, Fama and French (1993) introduce the size and value factors sorted on those characteristics. Their three-factor model is later extended in Fama and French (2015) to include factors based on investment and profitability. Fama and French (2020) compare their original factors with OLS factors extracted from cross-sectional regressions of returns on characteristics. Over the last decades, a large number of papers has proposed factors built from different characteristics (Harvey et al., 2016), while others have documented anomaly portfolios that are challenging to price with traditional factor models (Hou et al., 2020). More recently, Gu, Kelly, and Xiu (2020), Kozak, Nagel, and Santosh (2020), Freyberger, Neuhierl, and Weber (2020), Feng, He, Polson, and Xu (2023), among others, leverage machine learning techniques to study which characteristics provide incremental information for the cross-section of returns.

Our contribution is to provide a simple alternative to standard sorting methods to construct tradable factors that formally controls for the joint dependence among characteristics, such that the premium of each factor is more accurately isolated. Assuming that expected returns are linear in characteristics, Kozak and Nagel (2024) show under which conditions different methods for constructing characteristics-based factors approximate the unfeasible conditional mean-variance efficient portfolio of individual assets. Our results suggest that accounting for the joint dependence among characteristics and isolating their individual effects helps get closer to this unfeasible optimum relative to popular approaches for factor construction in the literature.

Bryzgalova, Pelger, and Zhu (2024) use decision trees to produce portfolios based on optimally chosen interactions of characteristics that can serve as test assets that are challenging to price. In a sense, our goal is the opposite: when sorting on a given characteristic, we aim to isolate its effect from that of other characteristics, instead of combining their information. This gives us factors that can be interpreted as capturing the independent effect of each characteristic. Kelly, Pruitt, and Su (2019) and Kim, Korajczyk, and Neuhierl (2021) propose methods for extracting statistical factors exploiting information from characteristics, which Kozak and Nagel (2024) show can be implemented using PCA on OLS factors and PCA on univariate factors constructed with orthonormalized characteristics, respectively. Our approach is complementary as it can be combined with these methods by using PCA on our copula factors instead.

The closest paper to ours is by Daniel et al. (2020), who also evaluate how to improve the construction of characteristics-based factors, but focus on hedging unpriced risk associated with characteristics sorts. We address a different challenge, in the sense that even in the absence of unpriced risk, we would still like to build factors that isolate the effect of each characteristic. While we compare our approach to theirs and show that our Fama-French factors achieve higher Sharpe ratios, the methods are complementary as both procedures could be combined.

2 Characteristics-based factor portfolios

Firm characteristics such as size, book-to-market, profitability and investment have been shown to explain cross-sectional differences in stock returns. As a result, a common approach in the asset pricing literature has been to form factors as managed zero-cost portfolios from these characteristics to explain the cross-section of excess returns. In this paper, we propose a new procedure to form these portfolios as to account for the joint dependence among characteristics and isolate the individual premium associated with a particular factor. In this section, we theoretically motivate and describe our approach.

2.1 A single characteristic

To motivate our analysis, we start by considering a single characteristic in which expected returns are assumed to increase. A traditional sort ranks assets according to the value of the characteristic and forms a long-short portfolio by taking long positions in the top quantile and short positions in the bottom quantile. For the sake of the exposition, we consider equally-weighted positions within each quantile. Let x_i denote the characteristic for asset i. Given a cross-section of N assets, we order them such that $x_1 \leq x_2 \cdots \leq x_N$. For a given quantile q, we let N_q be the number of assets in that quantile and take short (long) positions in assets with a characteristic below (above) that of x_{N_q} (x_{N-N_q}). Hence,

$$w_1 = w_2 = \cdots = w_{N_q} = -1/N_q \quad \text{and} \quad w_{N-N_q} = w_{N-N_q+1} = \cdots = w_N = 1/N_q, \tag{1}$$

where assets with characteristics greater than x_{N_q} but lower than x_{N-N_q} are assigned weights equal to zero.

With the portfolio weights $w = (w_1 \ldots w_N)$ and a model for excess returns, we can determine the return on the long-short portfolio. Consider a simple economy in which each asset's excess return r_i depends on a single risk factor f, such that:

$$r_i = \beta_i(f + \lambda) + \epsilon_i, \tag{2}$$

where $\mathrm{E}[f] = \mathrm{E}[\epsilon_i] = 0$ for all i, $Var[f] = \sigma_f^2$ and $Var[\epsilon_i] = \sigma_\epsilon^2$ for all i. Here, λ denotes the risk premium associated with the factor. Collecting the cross-section of assets returns and factor exposures into vectors $r = (r_1, r_2, \ldots, r_N)$ and $\beta = (\beta_1, \beta_2, \ldots, \beta_N)$, respectively, expected returns are given by $\mathrm{E}[r] = \mu = \beta\lambda$.

Introducing the vector of characteristics $x = (x_1, \ldots, x_N)$, we follow Daniel et al. (2020) and assume that expected returns line up with the characteristic, so that $\mu = \gamma x$, where γ represents the characteristic premium. For this specification to be consistent with the factor representation in (2), each return must satisfy:

$$r_i = x_i(\gamma/\lambda)(f + \lambda) + \epsilon_i. \tag{3}$$

This formulation provides economic meaning to the analysis of characteristic-sorted long-short portfolio returns, as it links cross-sectional variation in characteristics directly to systematic risk exposures.

To keep the notation concise, let $\bar{\beta}^{(Q)}(N_q)$ and $\bar{\epsilon}^{(Q)}(N_q)$ denote the average factor exposure and average idiosyncratic error within the top quantile (Q), where N_q is the number of assets in each quantile. Similarly, let $\bar{\beta}^{(1)}(N_q)$ and $\bar{\epsilon}^{(1)}(N_q)$ represent the corresponding averages in the bottom quantile.[1] The return on a characteristic-sorted, dollar-neutral long-short portfolio can then be expressed as:

$$R_x = R'w = (f + \lambda)(\bar{\beta}^{(Q)}(N_q) - \bar{\beta}^{(1)}(N_q)) + \bar{\epsilon}^{(Q)}(N_q) - \bar{\epsilon}^{(1)}(N_q). \tag{4}$$

This allocation is dollar neutral and highly exposed to the factor. If the portfolio is mean-variance efficient, the time-series intercept (alpha) from regressing its excess return on the factor is zero.

The maximum attainable Sharpe ratio in the economy is given by λ/σ_f. Taking the expectation and variance of (4), the Sharpe ratio of the characteristic-sorted portfolio becomes:

[1] More precisely, $\bar{\beta}^{(1)}(N_q) = N_q^{-1}\sum_{i\leq N_q}\beta_i$ and $\bar{\beta}^{(Q)}(N_q) = N_q^{-1}\sum_{i\geq N-N_q}\beta_i$ along with $\bar{\epsilon}^{(1)}(N_q) = N_q^{-1}\sum_{i\leq N_Q}\epsilon_i$ and $\bar{\epsilon}^{(Q)}(N_q) = N_q^{-1}\sum_{i\leq N-N_q}\epsilon_i$.

$$
\begin{aligned}
SR(R_x) = \frac{E[R'w]}{\sqrt{V[R'w]}} &= \frac{\lambda}{\sigma_f} \frac{(\bar{\beta}^{(Q)}(N_q) - \bar{\beta}^{(1)}(N_q))}{\sqrt{\left((\bar{\beta}^{(Q)}(N_q) - \bar{\beta}^{(1)}(N_q)^2 + \frac{2}{N_q}\sigma_\epsilon^2\right)}}, \\
&= \frac{\lambda}{\sigma_f} \frac{(\bar{x}_{N-N_q} - \bar{x}_{N_q})}{\sqrt{(\bar{x}_{N-N_q} - \bar{x}_{N_q})^2 + \frac{2}{\gamma^2 N_q}\sigma_\epsilon^2 \left(\frac{\lambda}{\sigma_f}\right)^2}},
\end{aligned}
\tag{5}
$$

where $\bar{x}_{(.)}$ denotes the average value of the characteristic in the respective quantiles, defined analogously to the averages of β and ϵ.

Equation (5) highlights that the Sharpe ratio of the characteristic-sorted portfolio depends both on the difference in average characteristics between the top and bottom quantiles and on the number of assets included in each quantile. As such, there exists a trade-off when constructing long-short portfolios: selecting more extreme quantiles (i.e., narrowing the tails) increases the characteristic spread between long and short positions, potentially enhancing expected returns. However, doing so reduces the number of assets N_q in each tail, thereby increasing idiosyncratic risk and lowering diversification. This trade-off implies that there is no universally optimal quantile size. Rather, the optimal choice balances return amplification from characteristic dispersion against increased portfolio volatility from reduced diversification. Importantly, as the total number of assets N grows, the impact of idiosyncratic noise diminishes, and the Sharpe ratio of the characteristic-based portfolio converges to that of the underlying factor, which represents the maximum attainable Sharpe ratio in the economy.

2.2 Multiple characteristics

While univariate portfolio analysis offers valuable insights into the relation between a single characteristic and asset returns, most empirical studies construct portfolios based on multiple characteristics. When these characteristics are correlated, it is desirable to form portfolios that are highly exposed to one target characteristic while remaining approximately neutral with respect to others. To illustrate this, extending equation (2) to a two-factor model yields:

$$r_i = \beta_{i,1}(f_1 + \lambda_1) + \beta_{i,2}(f_2 + \lambda_2) + \epsilon_i, \tag{6}$$

where exposures $\beta_{i,1}$ and $\beta_{i,2}$ may be associated with characteristics that are correlated. To take this into account, it is common to implement independent sorts for each characteristic. Breakpoints are chosen separately for each dimension, yielding two ordered sequences of assets: $x_{1,1} \leq x_{1,2} \leq \cdots \leq x_{1,N}$ and $x_{2,1} \leq x_{2,2} \leq \cdots \leq x_{2,N}$. Assets falling into the p-th and q-th quantiles of these sorts are assigned non-zero weights in the portfolio vector $w_{p,q}$. Portfolio returns are then computed from the intersections of these sorts. For example, Fama and French (1993) use $Q \times P = 2 \times 3$ sorts on size and value, forming value-weighted portfolios in each cell. Their size and value factors (SMB and HML) are constructed as:

$$SMB = \frac{1}{3}\Big(r'w_{1,1} + r'w_{1,2} + r'w_{1,3}\Big) - \frac{1}{3}\Big(r'w_{2,1} + r'w_{2,2} + r'w_{2,3}\Big), \tag{7}$$

$$HML = \frac{1}{2}\Big(r'w_{1,3} + r'w_{2,3}\Big) - \frac{1}{2}\Big(r'w_{1,3} + r'w_{2,3}.\Big) \tag{8}$$

The sorting methods described above are fully non-parametric, requiring no assumptions about the underlying distribution of characteristics. However, they depend on the selection of the corresponding quantiles for the breakpoints. Choosing an appropriate number of quantiles is important both to approximate the true Sharpe ratio and to neutralize characteristics with respect to each other. Alternatively, we propose a procedure that optimally neutralizes the effects of other characteristics while accurately isolating the premium associated with the characteristic of interest, without having to pre-specify quantile cutoffs. As described in the next subsection, we achieve this by estimating the conditional cumulative distribution function of a given characteristic while simultaneously controlling for the influence of all other characteristics using copulas.

2.3 Coupling characteristics with copula

Although quantile sorts remain the standard in the literature, recent studies have explored alternative methods to construct characteristic-based portfolios that are more diversified, less sensitive to outliers, and free from the need to choose arbitrary breakpoint thresholds. This can be done by applying a rank transformation to the characteristics. Specifically, assets are sorted based on the values of a given characteristic and then ranked cross-sectionally. This method has been employed by, among others, Frazzini and Pedersen (2014), Freyberger et al. (2020), and Kozak et al. (2020).

Importantly, there exists a one-to-one relation between a stock's rank and the empirical cumulative distribution function (ECDF) of a characteristic. Specifically, for characteristic x_j, the ECDF for stock i is given by:

$$\widehat{F}_N^j(x_{ij}) = \frac{\text{rank}(x_{ij})}{N+1}, \tag{9}$$

where $\text{rank}(x_{ij})$ denotes the cross-sectional rank of stock i, and N is the total number of stocks. Portfolio weights based on this approach are constructed by centering and scaling the ECDF-transformed characteristic:

$$w_{ij} = \frac{\widehat{F}_N^j(x_{ij}) - \frac{1}{N}\sum_{i=1}^N \widehat{F}_N^j(x_{ij})}{\sum_{i=1}^N \left|\widehat{F}_N^j(x_{ij}) - \frac{1}{N}\sum_{i=1}^N \widehat{F}_N^j(x_{ij})\right|}, \tag{10}$$

where the numerator centers the ECDF values around their cross-sectional mean, and the denominator scales the weights to ensure comparability across portfolios. This transformation maps each characteristic value to a point in the $[0, 1]$ interval, effectively mitigating the influence of extreme observations through the probability integral transform.

However, the univariate rank-based approach does not account for the fact that, in the presence of multiple characteristics, it is essential to neutralize the influence of other characteristics when constructing a sort based on a characteristic of interest. In fact, achieving this requires modeling the multivariate joint distribution of characteristics. To this end, we propose the use of copulas, which decompose the joint distribution into its marginal distributions and a copula that captures the dependence structure. By differentiating the joint distribution, one can derive conditional distributions for individual characteristics. This is a natural generalization of univariate rank-based portfolios, which allocate weights based on the marginal distribution of a given characteristic.

Copulas enable the separate modeling of marginal and joint distributions in two distinct steps. According to Sklar's theorem (Sklar, 1959), any p-dimensional joint distribution function $F(x_1, \ldots, x_p)$ with marginals $F_1(x_1), \ldots, F_p(x_p)$ can be represented as:

$$F(x_1, \ldots, x_p) = C\{F_1(x_1), \ldots, F_p(x_p)\}, \tag{11}$$

where $C(u_1, \ldots, u_p)$ is a p-dimensional copula that captures the dependence structure between the variables independently of their marginal distributions.

We consider both elliptical (Gaussian and Student-t) and Archimedean (Clayton, Gumbel, Frank) copula families. Among elliptical copulas, the Gaussian copula is one of the most widely used and is defined as:

$$C(u_1, u_2, \ldots, u_p; \rho) = \Phi_\rho\left(\Phi^{-1}(u_1), \Phi^{-1}(u_2), \ldots, \Phi^{-1}(u_p)\right),$$

where ρ is the correlation matrix of the variables, Φ_ρ is the cumulative distribution function of the p-dimensional standard Gaussian distribution with correlation matrix ρ, and $\Phi^{-1}(\cdot)$ denotes the inverse of the standard univariate Gaussian cumulative distribution function. While the Gaussian copula is popular due to its simplicity, it has notable limitations: it imposes symmetric dependence and exhibits asymptotic independence in the tails. In the context of factor construction, this implies that extreme combinations of characteristics (such as firms with both very high book-to-market ratios and very small size) are modeled as independent in the limit, which contradicts empirical evidence.

The assumption of asymptotic independence of extreme observations, inherent in the Gaussian copula, is relaxed in the Student-t copula. Analogous to the Gaussian copula, the Student-t copula is derived from the multivariate Student-t distribution. It is defined as:

$$C(u_1, u_2, \ldots, u_p; \rho, \nu) = t_{\rho,\nu}\left(t_\nu^{-1}(u_1), t_\nu^{-1}(u_2), \ldots, t_\nu^{-1}(u_p)\right),$$

where $t_{\rho,\nu}$ denotes the cumulative distribution function of the p-dimensional Student-t distribution with ν degrees of freedom and correlation matrix ρ, and t_ν^{-1} is the inverse cumulative distribution function of the univariate Student-t distribution with ν degrees of freedom. The presence of the additional parameter ν provides greater flexibility, allowing the Student-t copula to capture tail dependence. Smaller values of ν induce stronger tail dependence, approximating the behavior of a Cauchy distribution, while larger values make the distribution approach the Gaussian case with asymptotic independence. Despite its advantages, a limitation of the Student-t copula is that it only captures symmetric dependence between variables.

Another widely used class of copulas is the Archimedean family, which includes the Clayton, Gumbel, and Frank copulas, among others. Archimedean copulas are characterized by a generator function $\psi_\theta(t)$ and are parametrized by a dependence parameter θ. The general form of an Archimedean copula is:

$$C_d(u_1, \ldots, u_d; \theta) = \psi_\theta\left\{\psi_\theta^{[-1]}(u_1) + \ldots + \psi_\theta^{[-1]}(u_d)\right\},$$

where $\psi_\theta^{[-1]}$ denotes the pseudo-inverse of ψ_θ, defined as $\psi_\theta^{-1}(y)$ for $0 \leq y \leq \psi_\theta(0)$ and 0 otherwise, and $\psi_\theta(\infty) = 0$. The generator functions for the Clayton and Gumbel copulas are, respectively:

$$\psi_\theta(y) = \frac{1}{\theta}(y^{-\theta} - 1), \quad \text{and} \quad \psi_\theta(y) = \left(-\log(y)\right)^\theta.$$

The Frank copula has a different generator, given by:

$$\psi_\theta(y) = -\log\left(\frac{\exp(-\theta y) - 1}{\exp(-\theta) - 1}\right).$$

Archimedean copulas offer greater flexibility compared to elliptical copulas and are particularly useful for modeling asymmetric tail dependence. Specifically, the Clayton copula captures stronger dependence in the lower tail, while the Gumbel copula captures stronger dependence in the upper tail. The Frank copula, although symmetric, provides more flexibility in the tails than the Gaussian and Student-t copulas.

2.4 Copula estimation and portfolio construction

To estimate the copula, several procedures can be considered. One approach is to solve a joint maximum likelihood optimization problem, estimating the parameters of both the marginal distributions and the copula simultaneously. However, in practice, a two-step procedure is more commonly employed due to its computational simplicity and flexibility. In the two-step maximum likelihood method, the marginal distributions are estimated first, after which a cumulative distribution function transformation is applied to construct uniformly distributed pseudo-observations.

To maintain a fully flexible modeling framework, we adopt a two-step estimation procedure. Specifically, we use ECDFs to estimate the marginal distributions and to obtain pseudo-observations that are uniformly distributed on $[0, 1]$. This choice is consistent with the approach described in (10), where portfolio weights are constructed based on the ECDF or rank transformation. In the second step, the copula parameters are estimated based on the pseudo-observations, using the inversion of Kendall's tau method (for more details, see Durante and Sempi, 2010). Before estimating the copula parameters, we select the most suitable copula family using the cross-validation copula information criterion introduced by Grønneberg and Hjort (2014). This criterion, similar in spirit to the Akaike Information Criterion (AIC), is specifically designed for copula model selection.

Modeling the joint distribution of characteristics using copulas naturally extends the traditional sorting methodology, making it more robust and properly accounting for dependencies among characteristics. In the univariate rank procedure, portfolio weights are constructed to be proportional to the ECDF of the corresponding characteristic. Therefore, when constructing a portfolio based on a characteristic or interest while controlling for the influence of others, a natural extension is to consider the conditional distribution of that characteristic. Using copulas makes this construction straightforward. It is well known that:

$$c_{u_2}(u_1) = P(U_1 \leq u_1 \,|\, U_2 = u_2) = \frac{\partial C(u_1, u_2)}{\partial u_2}, \tag{12}$$

where $C(u_1, u_2)$ is the copula function linking the marginal distributions. In other words, the conditional distribution function of U_1 given $U_2 = u_2$ is equal to the partial derivative of the copula with respect to u_2.[2]

Thus, a natural generalization of the univariate weighting scheme in (10) to the multivariate case is to construct portfolio weights proportional to the conditional distribution given by (12). This approach also scales naturally to conditioning on multiple characteristics. Specifically, one takes the partial derivative of the copula with respect to all characteristics on which neutrality is desired. For example, when constructing a sort based on size, we first apply rank transformations to size, book-to-market, profitability, and investment characteristics. We then estimate the copula based on these transformed variables and take the derivative of the copula with respect to the transformed values of book-to-market, profitability, and investment. This procedure eliminates the need for selecting breakpoints for each characteristic, remains robust to outliers, and fully captures dependence among characteristics. Importantly, copulas model general forms of dependence, not just simple linear correlations, providing a more comprehensive adjustment for interactions between characteristics.

When applied to the portfolio construction problem, the standardized values of the conditional distribution of a characteristic can be used as the weights of a long-short portfolio. Analogous to (10), where the weights were defined as the demeaned ECDF values of a single characteristic, the copula-based vector of weights can be constructed as:

$$w_{i,j} = \frac{c_{u_1,\ldots,u_{-j},\ldots,u_p}(u_{i,j}) - \frac{1}{N}\sum_{i=1}^{N} c_{u_1,\ldots,u_{-j},\ldots,u_p}(u_{i,j})}{\sum_{i=1}^{N}\left| c_{u_1,\ldots,u_{-j},\ldots,u_p}(u_{i,j}) - \frac{1}{N}\sum_{i=1}^{N} c_{u_1,\ldots,u_{-j},\ldots,u_p}(u_{i,j})\right|}, \tag{13}$$

where $c_{u_1,\ldots,u_{-j},\ldots,u_p}(u_{i,j})$ denotes the derivative of the copula function with respect to all arguments except to characteristic j for stock i. This quantity is equivalent to the conditional distribution of the pseudo-observations u_j associated with the j-th characteristic,

[2]For a more detailed discussion on the relation between conditional distributions and copula derivatives, see Nelsen (2007).

conditional on all other characteristics.

As a consequence, constructing portfolio weights proportional to the conditional cumulative distribution function of a characteristic offers several important advantages. First, by assigning weights across all firms rather than restricting attention to extreme quantiles, the resulting portfolio is better diversified. Second, this methodology is robust to extreme values in the characteristics, as the probability integral transform maps outliers into a bounded domain. Third, by estimating the distribution of the characteristic of interest conditional on all other characteristics, it becomes possible to neutralize the influence of confounding factors. This, in turn, facilitates a more precise isolation and estimation of the risk premium associated with the characteristic of interest.

To illustrate the implications of accounting for joint dependence among characteristics when constructing factor portfolios, Figure 1 presents a scatterplot of the observed size and value ranks across stocks on a representative date from our sample, together with the size portfolio weights as denoted by the colors. We compare four different methods for portfolio construction: the Fama-French SMB portfolio (equally-weighted for ease of exposition), the univariate rank-based approach, and two conditional copula-based approaches, where one neutralizes size with respect to value, while the other makes size neutral with respect to value, operational profitability and investment.

As can be seen, the SMB portfolio essentially ignores the information from value, buying small stocks and selling large stocks in the same proportion. Similarly, the univariate rank sort buys small stocks and sells large stocks, but with weights changing continuously with the size rank. In contrast, when constructing a size factor that is neutral with respect to value using copulas, the threshold defining which small stocks are bought strongly depends on value: among high-value stocks, only the smallest are bought while most are sold, whereas among low-value stocks, only the largest are sold while most are bought. The conditional size sort only behaves similarly to the univariate rank sort for median-value stocks. Finally, constructing the copula size factor to be neutral not only to value, but also profitability and investment, introduces further heterogeneity in the portfolio weights that depends on the joint distribution of all these characteristics.

3 Simulation experiment

Having motivated and described our approach for accounting for joint dependence among characteristics in factor portfolio construction, we now turn to a simulation analysis to compare the performance of different portfolio construction methods. We consider two return-generating processes. The first case considers a market where returns are driven by two priced factors, following (6). The second considers a market with two priced factors and a third factor without a premium, such that:

$$r_i = \beta_{i,1}(f_1 + \lambda_1) + \beta_{i,2}(f_2 + \lambda_2) + \eta_i g + \epsilon_i. \tag{14}$$

We assume that all factors and the idiosyncratic error are uncorrelated. The unpriced factor affects the variances and covariances of returns but not average returns. This case provides an empirically relevant extension to the stylized model that solely contain priced factors, in light of the relevance of unpriced risk as discussed by Daniel et al. (2020).

We calibrate the factors, the characteristics, and the idiosyncratic error such that the optimal Sharpe ratio, the variance of the market index, the variance of expected returns, and the average variance of returns, remain the same in both data-generating processes (DGPs). All parameters are summarized in Table 1. The highest possible Sharpe ratio is restricted to 0.8. We let this Sharpe ratio through a linear pricing kernel of the priced factors. For simplicity, the risk prices of the factors in that kernel are set to one. Likewise, we let the expected value of the pricing kernel equal to one. This way, the Hansen and Jagannathan (1991) bound restricts the highest possible Sharpe ratio to the volatility of the pricing kernel. For both of the return-generating processes, we ensure that the volatility of the market index ($\sqrt{V(\bar{r})}$) is set to 20 percent. Also, the standard deviation of the cross section of expected returns ($\sqrt{V(\mu)}$) and average volatility of returns ($\sqrt{\bar{V}(r)}$) are 10 and 60 percent, respectively. We fix these values by calibrating the asset characteristics that proxy for the exposures to the priced risk factors. The design of the simulation is presented in detail in Appendix A.

Aside from the economic restrictions above, we make some restrictions out of convenience. We assume that characteristics only proxy for the exposures to one risk factor, such that $\beta_1 = x_1(\gamma_1/\lambda_1)$ and $\beta_2 = x_2(\gamma_2/\lambda_2)$. However, while the two factors are independent, we let the two characteristics be correlated. We restrict the means and variances of the two characteristics to be the same and set the correlation between the two equal to 0.5. Furthermore, we calibrate the design such that the two factors contribute equally to the optimal Sharpe ratio of the market.

3.1 Simulation results

To evaluate the performance of the sorting methodologies, we use them to form long-short portfolios and compute the out-of-sample Sharpe ratios of the respective allocations. The weights $w_k^{(m)}$ denote the weights for the kth factor obtained using method m. The methods are either equally weighted independent quantile sorts or a conditional cumulative distribution constructed using copulas. We set the independent sorts such that we obtain $P \times Q = 3 \times 3$ sorts.

Since we know the expected excess returns and covariance matrix by construction of the simulations, we can compute $SR(w_k^{(m)}) = \mu' w_k^{(m)} / \sqrt{w_k^{(m)\prime} \Sigma w_k^{(m)}}$. We then compare this Sharpe ratio to the theoretical Sharpe ratio of the factor:

$$Loss(w_k^{(m)}) = \left| \left(\frac{SR(w_k^{(m)})}{\lambda_k/\sigma_{f,k}} - 1 \right) \right| \times 100. \tag{15}$$

We obtain the error distributions by computing the deviations $Loss(w_k^{(m)})$ associated with the respective factors over 10000 draws from the two data-generating processes. We consider different numbers of assets in the simulation such that $N = \{500, 1500, 3500\}$.

The simulation results are presented in Table 2. Considering first the case with two priced factors, when including $N = 3500$ assets, we find that the average Sharpe ratio loss associated with the independent $Q \times P$ sorts is 9.83 percent. The corresponding average associated with the copula sorts is only 0.32 percent. Hence, the conditional cumulative distributions provide a more accurate assessment of the investment opportunities associated with the respective characteristics than the assessment obtained through independent sorts. Furthermore, the standard deviation of the loss distribution for the copula-sorted factor is significantly lower than what is obtained using independent sorts.

Similarly, when considering cross-sections of 500 and 1500 assets, factors formed using the copula obtain Sharpe ratios closer to the theoretical Sharpe ratios in expectation. The sampling variance of $Loss(w_k^{(m)})$ also remains lower using the copula sorts. Using independent sorts, the size of the average loss is smaller in the cross-sections with fewer assets. This is due to the presence of additional idiosyncratic risk in the factor sorts. While the positive correlation between characteristics inflates the Sharpe ratios of the sorts, the additional volatility attenuates the Sharpe ratios. The idiosyncratic risk is removed through diversification when the number of assets is large.

For the second data generating process, we include an unpriced factor in the return process. The bottom panel of Table 2 reports these results. Unpriced factors introduce additional risk for which there is no premium. The expected losses associated with both sorting methods are greater than in the case without unpriced risk, and the largest losses are more extreme than in the case where non-priced risk is excluded. Nonetheless, we find that using the conditional cumulative distribution continues to be more effective. The variance of the Sharpe ratio distribution obtained from the copula sorts is lower, and the mean of the distribution is much closer to the actual Sharpe ratio of the factors than we find using the independent $Q \times P$ sorts. These results highlight the importance of accounting for the dependence across characteristics when constructing factor portfolios, particularly in settings with unpriced risk.

4 Empirical analysis

Empirically, we apply our method to construct a new version of the Fama and French (2015) five-factor model, for which the theoretical motivation strongly relies on making characteristics conditional on each other. In this section, we first describe the data used in our analysis. Then, we compare our copula factors to four benchmarks (original 2×3 factors, univariate rank factors, original $2 \times 2 \times 2 \times 2$ factors and hedged factors) in terms of their individual risk premia and conduct several spanning regressions. Finally, we analyze the mean-variance frontier spanned by the tangency portfolio of each factor model and assess their ability to price a large set of cross-sectional anomaly portfolios.

4.1 Data

We use standard data from the Center for Research in Security Prices (CRSP) and Compustat. The historical data available in CRSP received a major update during 2022. In this paper, we make use of the older version of the data to simplify the comparison with prior work in the literature. In the cases where stocks delist from the market, we adjust returns according to common practice (see, e.g., Shumway, 1997) and returns below -100% are set equal to -100%. Data are compiled following the common practice and only include common stocks (share codes 10 and 11) listed on the NYSE, AMEX or NASDAQ exchanges. The data collection and construction of the financial characteristics and factor sorts closely follow Scheuch et al. (2023).

We focus on the factors proposed by Fama and French (2015), which in addition to the market factor, include Size, Value, Operating Profitability, and Investment. We contrast our method with four other variations of these factors. The first corresponds to 2×3 sorts, where assets are independently sorted into small and big stocks based on market capitalization, and into three groups based on each of the other characteristics. The second approach uses $2 \times 2 \times 2 \times 2$ sorts, where assets are assigned to factor portfolios depending on whether their values for each characteristic are above or below the median. We make use of value-weighted portfolios within the respective sorts. The third is the set of hedged factor sorts proposed by Daniel et al. (2020).[3] Finally, we consider factors formed through simple rank sorting based on individual characteristics. Except for the hedged factor sorts, the market factor is the same across all models, and is the market return in excess of the one-month Treasury bill from Kenneth French's data library.

The definitions of the characteristics follow those provided by Fama and French (2015). Size is measured by market capitalization, while the Value characteristic is defined as the

[3]We are grateful to the authors for making their factors available.

ratio of book equity to market equity. Operating profitability is calculated as revenue minus expenses, divided by book equity, and Investment is approximated by the ratio of current total assets to lagged total assets. Firms with negative lagged total assets or negative book equity are excluded. The sorting characteristics are updated annually in June, and are used to construct factors based on monthly returns over the same period.[4]

To assess the factors obtained using our methodology, we evaluate spanning regressions for individual factors and the ex post mean-variance efficient portfolios. These regressions use monthly returns from July 1964 to December 2023. In addition, we examine asset pricing errors across a large set of anomalies. These anomalies are constructed as in Hou et al. (2020), based on test assets available from Global-Q (www.global-q.org). The pricing errors are evaluated using monthly returns from January 1973 to March 2023.

4.2 Summary statistics of individual factors

Panel A of Table 3 reports the average monthly returns, volatilities and t-statistics for each of the factors in the different models we consider. Focusing first on the original 2×3 Fama-French sorts, we can see that all the factors yield a statistically significant risk premium at the 5% confidence level, with t-statistics for the average excess returns ranging from 1.99 (for the SMB) to 3.49 (for the CMA). Relative to the original approach, the copula factors provide a similar level of average returns, but decrease the volatility substantially. This is such that the t-statistics uniformly increase, from 3.32 (for the SMB) to 9.21 (for the CMA), being all significant at the 1% level. That is, by formally accounting for the joint dependence among characteristics, our method better isolates the risk premium associated with each characteristics-based factor portfolio.

As previously mentioned, our approach generalizes univariate rank factors, which allocate weights based on the marginal distribution of a given characteristic. Therefore, it is useful to compare both approaches. Table 3 shows that the behavior of univariate

[4] As mentioned in Section 2.3, the copula type employed in our method is selected every rebalancing period using the cross-validation copula information criterion. According to this criterion, the Student-t copula is always selected as the best fit for the characteristics we consider over our sample period. Although the best copula type does not change over time, its parameters do. Therefore, it is still important to re-estimate the model in every period.

factors is more erratic. The t-statistics of the SMB, HML and CMA factors increase substantially relative to the 2×3 approach, but that of RMW decreases to an insignificant level of 1.09. This indicates that failing to account for the size effect when conditioning based on profitability leads to an insignificant effect on stock returns. In contrast, each copula factor is appropriately neutralized with respect to every other characteristic, such that our RMW factor has a highly statistically significant average return.

The original 2×3 sorts make each characteristic neutral only to size. To build factors that are all neutral to each other, Fama and French (2015) heuristically propose to use $2\times2\times2\times2$ sorts to construct the factors. As can be seen in Panel A of Table 3, this method results in factors with generally smaller volatilities relative to the original approach, which translates into higher t-statistics for SMB, HML and RMW. However, the average return of CMA is greatly reduced, such that its t-statistics decreases substantially. In other words, while this heuristic conditioning approach mostly goes in the right direction, it is evidently sub-optimal relative to our copula approach, which delivers uniformly higher t-statistics.

The last benchmark we consider are the hedged factors from Daniel et al. (2020). By removing exposures to unpriced risk factors, this method essentially reduces the volatilities of the factors relative to the original 2×3 approach, at the natural expense of slightly lower average excess returns. The net result is that the t-statistics uniformly increase, that is, the factors offer a better risk-return trade-off than their unhedged counterparts. Compared to our copula factors, however, the hedged sorts still deliver smaller t-statistics. This highlights the relative importance of isolating the effect of each characteristic from the others. In any case, we emphasize that both approaches are complementary and could potentially be combined for even better performance.

Panel B of Table 3 reports the pairwise correlation structure between factors within each model. The correlation between the value factor and the investment factor is particularly high within the 2×3 portfolio sorts. Using our method, this correlation falls from 0.69 to 0.37. However, it is important to note that making characteristics neutral with respect to each other does not necessarily imply low correlation between factor returns, as seen by the high correlation between HML and RMW in the copula sorts, for example. As noted by Fama and French (2015), this is also evident from the $2 \times 2 \times 2 \times 2$ sorts, which generally have higher (absolute) correlations than the 2×3 sorts. Interestingly, the correlation structure of the copula factors is also substantially different from that of the univariate rank factors, which we generalize, and the hedged factors, which focus on hedging unpriced risk instead of imposing neutrality among characteristics.

We also assess, in Table 4, the pairwise correlation between the same factors using our approach and each of the benchmarks. The copula factors are positively related to all respective factors, as would be expected. Correlations with the 2×3 and $2 \times 2 \times 2 \times 2$ factors are similar and range from 55% for the SMB to 85% for the HML. The strongest correlations are observed for the univariate rank factors, around 90% for all factors, which makes sense given that the copula factors use the marginal distributions of the characteristics (which are the portfolios weights in the rank sorts) as inputs to compute their joint multivariate distribution. On the other hand, the weakest correlations appear for the hedged factors, around 40%. This highlights, again, how both approaches can be seen as complementary.

4.3 Factor spanning regressions

The individual factor comparisons suggest that the copula factors isolate better the risk premium associated with each characteristic. We next evaluate to which extent each of the factors we have constructed are spanned by the joint specification of all factors in the benchmark models. This allows us to address whether the enhanced premium associated with our factors can be explained by the other models, or, equivalently, whether each of the copula factors is able to improve the mean-variance efficiency of the benchmark five-factor specifications. Table 5 presents the results for the spanning regressions, which are time-series regressions of each of our factors onto the five factors of another model. As can be seen, the intercepts (or alphas) from these regressions are positive and statistically significant in nearly all cases. The only exceptions are for the value factor, when regressed onto the univariate rank factors, for which the alpha is significantly negative, and onto the hedged factors, for which the alpha is positive but insignificant.

Interestingly, univariate factors explain a large fraction of the time-series variation of the copula factors, given the large R^2s of the regressions, but not their expected returns, reinforcing the importance of accounting for the joint dependence among characteristics in isolating their premia. Relative to the original Fama-French 2×3 or $2 \times 2 \times 2 \times 2$ sorts and the hedged sorts, the largest alphas are observed for the SMB and CMA factors, suggesting that these characteristics benefit the most from isolating out the effect of the remaining characteristics. In sum, the individual copula factors carry unspanned priced information and individually improve the mean-variance profile of the five-factor models we consider as benchmarks. While these results are striking on their own, the important consideration for asset pricing that remains is how the joint specification of our characteristics sorts compares to the others, which we address in the next subsection.

4.4 Maximum attainable Sharpe ratios

Comparing linear factor models when factors are traded boils down to a comparison of the maximum attainable Sharpe ratio from the set of factors included in each competing model (Barillas and Shanken, 2017). The maximum Sharpe ratio, in turn, is simply the Sharpe ratio of the tangency portfolio of the factors in a given model. In Table 6, we report the monthly average returns, volatility and Sharpe ratio of the tangency portfolios implied by each model. The original 2×3 factors, when combined, yield an annualized Sharpe ratio of 1.04, which is substantially smaller than the 1.80 Sharpe ratio of the copula tangency portfolio. In fact, the copula factors yield the highest maximum Sharpe ratio, followed by univariate rank factors and hedged factors. The outperformance associated with our approach is statistically significant: as shown in the table, the alpha from regressing the copula tangency portfolio on the tangency portfolio of each of the benchmark models is always positive and highly statistically significant. That is, our method, by formally accounting for the joint dependence among characteristics when forming factor portfolios, greatly expands the mean-variance frontier.

Table 6 also reports the composition of the respective tangency portfolios. We find that the position in the value factor is relatively small compared to that of other factors for most of the sorts, which resonates the evidence from Fama and French (2015) that the HML factor is largely redundant. The exception is our method, for which the HML factor actually gets a weight that is slightly larger (in absolute terms) than the one for the market factor. Compared to the original 2×3 approach, the main difference in the composition of the copula tangency portfolio is a decrease in the weight of the RMW factor accompanied by an increase in the weight of the CMA factor. This is aligned with the strong individual performance of the copula investment factor observed in Table 3.

4.5 Pricing cross-sectional anomaly portfolios

While the previous results show that the copula factors outperform the benchmark factors in terms of relative pricing performance, researchers are often also interested in the absolute performance of factor models for pricing a set of test assets. In particular, factor models are commonly used to assess whether a given cross-sectional anomaly portfolio delivers alpha or simply reflects exposures to risk factors. For instance, Fama and French (2016) use their 2×3 five-factor model to price several anomalies, finding relative success with exception for explaining returns on accruals and momentum portfolios. In this subsection, we assess the performance of the competing five-factors models in our analysis in pricing a large set of 150 anomalies from Hou et al. (2020).

For each anomaly portfolio return $r_{i,t}$, we run the following time-series regression:

$$r_{i,t} = \alpha_i + \beta_i Mkt_t + h_i SMB_t^{(s)} + c_i HML_t^{(s)} + o_i RMW_t^{(s)} + v_i CMA_t^{(s)} + \varepsilon_{i,t}, \qquad (16)$$

where s denotes the sorting methodology $s \in \{CS, 2 \times 3, Rank, 2 \times 2 \times 2 \times 2, Hedged\}$.[5] Our main interest with respect to these regressions are the intercepts, i.e., the alphas. For a given factor model, we consider the following metrics of pricing performance: the average absolute α_i; the Gibbons et al. (1989) test statistic for the null hypothesis that the alphas are jointly equal to zero; and the number of anomalies for which the α_i is statistically significant following the rule-of-thumb of Harvey et al. (2016) that the t-statistics must be higher than 3 to account for multiple testing biases. For all these

[5]For the hedged factors, the market factor is different as they also modify it to hedge unpriced risk.

metrics, the lower the metric, the better the factor model prices the anomalies.

Table 7 reports the results of this pricing exercise. Focusing first on the panel aggregating performance across the entire set of anomalies, we can see that the copula factors deliver the smallest average pricing errors and GRS statistic. The original 2×3 and $2 \times 2 \times 2 \times 2$ factors are the second and third best in terms of average alphas, but the worst in terms of GRS statistics, which weight the alphas taking into account the covariance among the test assets returns. This means that the original factors are outperformed by the univariate rank factors and hedged factors in pricing anomalies that contribute more for the mean-variance efficient portfolio of the test assets. In terms of the explanatory power for the cross-section, the last column of the table shows that, after controlling for the copula factors, only 21 anomalies deliver statistically significant alphas. Strikingly, this number is at least twice as high for the benchmark models. For instance, 49 anomalies are significant if we use the original 2×3 five-factor model.

Figure 2 sheds further light by plotting the distribution of the (absolute) t-statistics associated with the alphas of the anomalies for each factor model. As can be seen, the right tail of the t-statistics distribution is much thinner for the copula factors compared to the benchmark models, with considerably less mass in the region given by $|t| > 3.00$. Importantly, the improved pricing performance comes without making use of any additional information relative to the original Fama-French specification. Rather, the improvements come solely from accounting for the joint multivariate distribution of the characteristics when making characteristics portfolios neutral to each other. By doing so, the five-factor model based on size, value, profitability and investment goes a long way in explaining the cross-section of anomaly portfolios.

The remaining panels of Table 7, which break down the pricing metrics by anomaly category, help understand for which anomalies the pricing improvements associated with our approach are coming from.[6] The copula factors are uniformly better across categories in terms of rendering alphas insignificant. The only exception is for the investment category, where univariate rank factors explain more anomalies. Overall, it is clear that the main improvement from employing the copula factors comes from pricing profitability

[6] We follow the anomaly category definitions of Hou et al. (2020).

anomalies, with half as significant anomalies as the benchmarks, and especially momentum anomalies, with twice to six times less significant anomalies relative to the competing models. In fact, from 32 anomalies in the momentum category, only 5 are able to deliver significant alphas with respect to the copula factor model. That is, after accounting for the joint dependence among characteristics, the five-factor model is mostly successful in explaining momentum. This is particularly remarkable given the evidence from Fama and French (2016) that their original 2×3 five-factor model has a hard time in pricing momentum portfolios.

5 Conclusion

In this paper, we propose a new approach for constructing characteristics-based factor portfolios. In contrast to existing methods, we formally account for the joint dependence among characteristics by allocating portfolio weights proportional to the conditional distribution of each characteristic given the others, i.e., maximizing cross-sectional variation in the characteristic of interest while keeping the others fixed. We are able to do so in a data-driven way by using copulae to model the joint multivariate distribution of characteristics. This generalizes univariate factors, avoids the need to arbitrarily specify quantile cutoffs and allows for a better assessment of the individual premium associated with a particular factor.

We apply our sorting methodology to construct a new version of the five Fama and French (2015) factors, for which the theoretical motivation relies on making characteristics neutral to each other. We show that, relative to the 2×3 sorted factors, univariate factors, $2 \times 2 \times 2 \times 2$ factors and the hedged factors of Daniel et al. (2020), our copula factors increase the individual Sharpe ratios, carry unspanned information about returns and expand considerably the mean-variance frontier in terms of higher maximum attainable Sharpe ratios. Importantly, we also find that our factors improve substantially the explanatory power for a large set of anomalies in the cross-section of returns. This improvement mainly comes from being able to price momentum anomalies, which is a well-known challenge for the original 2×3 factors.

Our paper contributes to a large literature assessing the role of firm characteristics in explaining cross-sectional differences in stock returns. Much of the recent work in this literature focuses on bringing discipline to the large "zoo" of characteristics that have been shown to explain returns. While we do not address this challenge explicitly, as our goal is to improve the construction of characteristics-based factor portfolios in low-dimensional models, our results suggest that extending our methodology to make sorts conditional on each other in a large set of characteristics could both expand the mean-variance frontier attainable from these sorts and help isolate the individual premium associated with each characteristic. We leave this interesting venue for future research.

References

Barillas, F. and J. Shanken (2017): "Which alpha?" *The Review of Financial Studies*, 30, 1316–1338.

Bryzgalova, S., M. Pelger, and J. Zhu (2024): "Forest Through the Trees: Building Cross-Sections of Stock Returns," *Journal of Finance, forthcoming.*

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

Durante, F. and C. Sempi (2010): "Copula theory: an introduction," in *Copula theory and its applications*, Springer, 3–31.

Fama, E. F. and K. R. French (1992): "The Cross-Section of Expected Stock Returns," *Journal of Finance*, 47, 427–65.

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

——— (2015): "A five-factor asset pricing model," *Journal of financial economics*, 116, 1–22.

——— (2016): "Dissecting Anomalies with a Five-Factor Model," *The Review of Financial Studies*, 29, 69–103.

——— (2020): "Comparing Cross-Section and Time-Series Factor Models," *The Review of Financial Studies*, 33, 1891–1926.

Feng, G., J. He, N. Polson, and J. Xu (2023): "Deep learning in characteristics-sorted factor models," *Journal of Financial and Quantitative Analysis.*

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

Freyberger, J., A. Neuhierl, and M. Weber (2020): "Dissecting characteristics nonparametrically," *The Review of Financial Studies*, 33, 2326–2377.

Gibbons, M. R., S. A. Ross, and J. Shanken (1989): "A test of the efficiency of a given portfolio," *Econometrica: Journal of the Econometric Society*, 1121–1152.

Grønneberg, S. and N. L. Hjort (2014): "The copula information criteria," *Scandinavian Journal of Statistics*, 41, 436–459.

Gu, S., B. Kelly, and D. Xiu (2020): "Empirical Asset Pricing via Machine Learning," *The Review of Financial Studies*, 33, 2223–2273.

Hansen, L. and R. Jagannathan (1991): "Implications of Security Market Data for Models of Dynamic Economies," *Journal of Political Economy*, 99, 225–62.

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

Hou, K., C. Xue, and L. Zhang (2020): "Replicating Anomalies," *The Review of Financial Studies*, 33, 2019–2133.

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

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

Kozak, S. and S. Nagel (2024): "When do cross-sectional asset pricing factors span the stochastic discount factor?" *Available at SSRN 4443643*.

Kozak, S., S. Nagel, and S. Santosh (2020): "Shrinking the cross-section," *Journal of Financial Economics*, 135, 271–292.

Nelsen, R. B. (2007): *An introduction to copulas*, Springer Science & Business Media.

Scheuch, C., S. Voigt, and P. Weiss (2023): *Tidy Finance with R*, Chapman and Hall/CRC, 1st ed.

Shumway, T. (1997): "The delisting bias in CRSP data," *The Journal of Finance*, 52, 327–340.

Sklar, M. (1959): "Fonctions de repartition an dimensions et leurs marges," *Publ. inst. statist. univ. Paris*, 8, 229–231.

A Simulation design

We calibrate a simple model for returns on risky assets to assess the factor constructs in a stylized setting. For the calibration of the critical parameters, we assume that the market is large, i.e., $N \to \infty$. We consider two uncorrelated factors with equal variances and premia, f_1 and f_2, and two characteristics, $x_{i,1}$ and $x_{i,2}$, that proxy for the exposures to the factors in asset i. Expected returns are set to be linear in the characteristics such that:

$$E(r_i) = \gamma_1 x_{i,1} + \gamma_2 x_{i,2}, \tag{17}$$

and we set the pricing kernel equal to $m = 1 - f_1 - f_2$.

The first step of the simulation design is to calibrate the the risk premia of the respective factors and the variances of the factors. The Hansen-Jagannathan bound controls the obtainable Sharpe ratio by the variance of the pricing kernel. We set the optimal annual Sharpe ratio to be 0.8, which gives us the variance of the factors $\sigma^2_{f,1} = \sigma^2_{f,2} = \sigma^2_f = 0.8^2/2$. The fundamental pricing conditions $E(mr) = 0 = \lambda(\beta_1 + \beta_2) - \sigma^2_f(\beta_1 + \beta_2)$ give premia of the factors $\lambda_1 = \lambda_2 = \lambda = \sigma^2_f$.

A.1 Calibration when there is no unpriced risk

We calibrate the expected value of the financial characteristics using the variance of a market index. If there are no unpriced factors in the market, the asset returns are:

$$r_i = \beta_{i,1}(f_1 + \lambda) + \beta_{i,2}(f_2 + \lambda) + \epsilon_i, \tag{18}$$

and the index is $\bar{r} = N^{-1}\sum_{i=1}^{N}(\beta_{i,1}(f_1+\lambda)+\beta_{i,2}(f_2+\lambda)+\epsilon_i)$. We let $N \to \infty$ and obtain the variance of the market index:

$$V(\bar{r}) = V(\bar{\beta}_1(f_1 + \lambda) + \bar{\beta}_2(f_2 + \lambda)) = \sigma^2_f(\bar{\beta}_1^2 + \bar{\beta}_2^2). \tag{19}$$

Using that $\beta_{i,1} = x_{i,1}(\gamma_1/\lambda)$, and setting $\gamma_1 = \gamma_2 = \gamma$ along with $\mathrm{E}(x_1) = \mathrm{E}(x_2) = \mathrm{E}(x)$ we obtain $V(\bar{r}) = \sigma^2_f((\bar{x}_1(\gamma/\lambda))^2 + (\bar{x}_2(\gamma/\lambda))^2)$, which gives us $E(x) = (\lambda/\gamma)\sqrt{V(\bar{r})/(2\sigma^2_f)}$.

To calibrate the variances, $V(x_1)$ and $V(x_2)$, we make use of a restriction on the cross-sectional variation in expected returns on the individual assets ($V(\mu_i)$), which is,

$$
\begin{aligned}
V(\mu_i) &= N^{-1}\sum_{i=1}^{N}(\mu_i - E(\mu_i))^2,\\
&= N^{-1}\sum_{i=1}^{N}(\lambda(\beta_{i,1}+\beta_{i,2}) - E(\beta_{i,1}+\beta_{i,2}))^2,\\
&= \lambda^2 N^{-1}\sum_{i=1}^{N}((\beta_{i,1}-\bar{\beta}_{i,2}) + (\beta_2 - \bar{\beta}_2))^2,\\
&= \lambda^2 N^{-1}\sum_{i=1}^{N}(\tilde{\beta}_{i,1}+\tilde{\beta}_{i,2})^2,\\
&= \lambda^2 N^{-1}\sum_{i=1}^{N}(\tilde{\beta}_{i,1}^2+\tilde{\beta}_{i,2}^2 + 2\tilde{\beta}_{i,1}\tilde{\beta}_{i,2}), \qquad (20)
\end{aligned}
$$

where $\tilde{\beta}_{i,1} = (\beta_{i,1} - \bar{\beta}_1)$. Again, using $\beta_{i,1} = x_{i,1}(\gamma_1/\lambda)$ we obtain the cross-sectional variance of expected returns as a function of the characteristics x_1 and x_2,

$$
\begin{aligned}
V(\mu_i) &= \lambda^2(\sigma_{x_1}^2(\gamma/\lambda)^2 + \sigma_{x_2}^2(\gamma/\lambda)^2 + 2(\gamma/\lambda)^2\sigma_{x_1,x_2})\\
&= 3\gamma^2\sigma_x^2.
\end{aligned}
$$

The last equality follows because we set the correlation between characteristics to be equal to $\sigma_{x_1,x_2}/(\sigma_{x_1}\sigma_{x_2}) = 0.5$ and $\sigma_{x_1} = \sigma_{x_2} = \sigma_x$. Hence, the variance of the respective characteristics is $\sigma_x^2 = V(\mu_i)/(3\gamma^2)$.

Lastly, we calibrate the variance of the idiosyncratic errors of returns, which we assume to be constant across assets and denote σ_ϵ^2. This parameter is calibrated from the average variance of asset returns,

$$
\begin{aligned}
\bar{V}(r) &= N^{-1}\sum_{i=1}^{N} V(\beta_{1,i}(f_1+\lambda) + \beta_{2,i}(f_2+\lambda) + \epsilon_i),\\
&= N^{-1}\sum_{i=1}^{N}(\beta_{1,i}^2\sigma_f^2 + \beta_{2,i}^2\sigma_f^2 + \sigma_\epsilon^2),\\
&= \sigma_f^2(\bar{\beta}_1^2 + \bar{\beta}_2^2) + \sigma_\epsilon^2. \qquad (21)
\end{aligned}
$$

Substituting $\bar{\beta_1^2} = V(\beta_1) + \bar{\beta_1}^2$ gives:

$$\bar{V}(r) = \sigma_f^2(V(\beta_1) + \bar{\beta_1}^2 + V(\beta_2) + \bar{\beta_2}^2) + \sigma_\epsilon^2, \tag{22}$$

and since $\beta_1 = x_1(\gamma/\lambda)$ along with $\mathrm{E}(x_1) = \mathrm{E}(x_2) = \mathrm{E}(x)$, and $\sigma_{x_1} = \sigma_{x_2} = \sigma_x$ we obtain:

$$\bar{V}(r) = 2\sigma_f^2(\sigma_x^2 + \mathrm{E}(x)^2)(\gamma/\lambda)^2 + \sigma_\epsilon^2. \tag{23}$$

This renders the idiosyncratic error volatility equal to $\sigma_\epsilon^2 = \bar{V}(r) - 2\sigma_f^2(\sigma_x^2 + \mathrm{E}(x)^2)(\gamma/\lambda)^2$.

A.2 Calibration when there is unpriced risk

The previous model excludes the case where financial returns are exposed to factors that do not carry a risk premium. These unpriced factors affect the variance and co-variances of returns but not average returns. Daniel et al. (2020) highlights that this is a particularly relevant case, as unpriced factors may attenuate the Sharpe ratio of the sorts. We make this our second model and take excess returns r_i from a model of three factors (f_1, f_2 and g). The factors f_1 and f_2 have a premium but g does not. The relevant factors, irrelevant factor, and the idiosyncratic error are orthogonal and give the following model for the returns:

$$r_i = \beta_{i,1}(f_1 + \lambda_1) + \beta_{i,2}(f_2 + \lambda_2) + \eta_i g + \epsilon_i,$$

where $\mathrm{E}[f_1] = \mathrm{E}[f_2] = \mathrm{E}[g] = \mathrm{E}[e] = 0$. We can still calibrate the priced factors and the risk premia by fundamental pricing conditions and the Hansen-Jagannathan bound, similar to before, which gives $\lambda_1 = \lambda_2 = \sigma_{f,1}^2 = \sigma_{f,2}^2 = \sigma_f^2 = 0.8^2/2$ since the factor g does not carry a premium. To calibrate the characteristics and remaining variances, we make one new assumption that the priced risk factors in the model account for 50 percent of the variance in the market index ($R^2 = 0.5$). The remaining variation of the index is due to the unpriced risk factor.

In this framework, the market index associated with this model is equal to:

$$N^{-1}\sum_{i}^{N} r_i = N^{-1}\sum_{i}^{N}(x_i(\gamma/\lambda)(f_1+\lambda) + x_i(\gamma/\lambda)(f_2+\lambda) + \eta_i g). \tag{24}$$

Following the same steps as before we obtain $V(\bar{r}) = \sigma_f^2(\bar{\beta}_1^{\,2}+\bar{\beta}_2^2)+\bar{\eta}^2\sigma_g^2 = 2\sigma_f^2(\gamma/\lambda)^2E(x)^2+\bar{\eta}^2\sigma_g^2$. Using the restriction on the variance stemming from the priced factors, we further have $R^2V(\bar{r}) = 2\sigma_f^2(\gamma/\lambda)^2E(x)^2$. In turn, this renders the expected value of the characteristics equal to:

$$E(x_1) = E(x_2) = (\lambda/\gamma)\sqrt{R^2V(\bar{r})/(2\sigma_f^2)}. \tag{25}$$

Likewise, $(1-R^2)V(\bar{r}) = \bar{\eta}^2\sigma_g^2$ results in $\bar{\eta} = \sqrt{(1-R^2)V(\bar{r})/\sigma_g^2}$.

Just as in the previous return generating process, we use the cross-sectional variance of the expected returns in this model to calibrate the variance of the characteristics, which are unaffected by the irrelevant factor. The cross-sectional variance of expected returns remains equal to:

$$V(\mu_i) = N^{-1}\sum_{i=1}^{N}(\mu_i - E(\mu_i))^2 = N^{-1}\sum_{i=1}^{N}\left(\lambda(\beta_1+\beta_2) - \lambda(\bar{\beta}_1+\bar{\beta}_2)\right)^2, \tag{26}$$

which gives use the variance of the characteristics $\sigma_x^2 = V(\mu_i)/(3\gamma^2)$.

The final step is to compute the idiosyncratic error volatility from the average volatility of returns. This is in the second return generating process,

$$\begin{aligned}\bar{V}(r_i) &= \bar{\beta}_1^2\sigma_f^2 + \bar{\beta}_2^2\sigma_f^2 + \bar{\eta}^2\sigma_g^2 + \sigma_\epsilon^2,\\ &= 2\sigma_f^2(V(\beta)+\mathrm{E}(x)^2)(\gamma/\lambda)^2 + \sigma_g^2(V(\eta)+\bar{\eta}^2) + \sigma_\epsilon^2.\end{aligned}$$

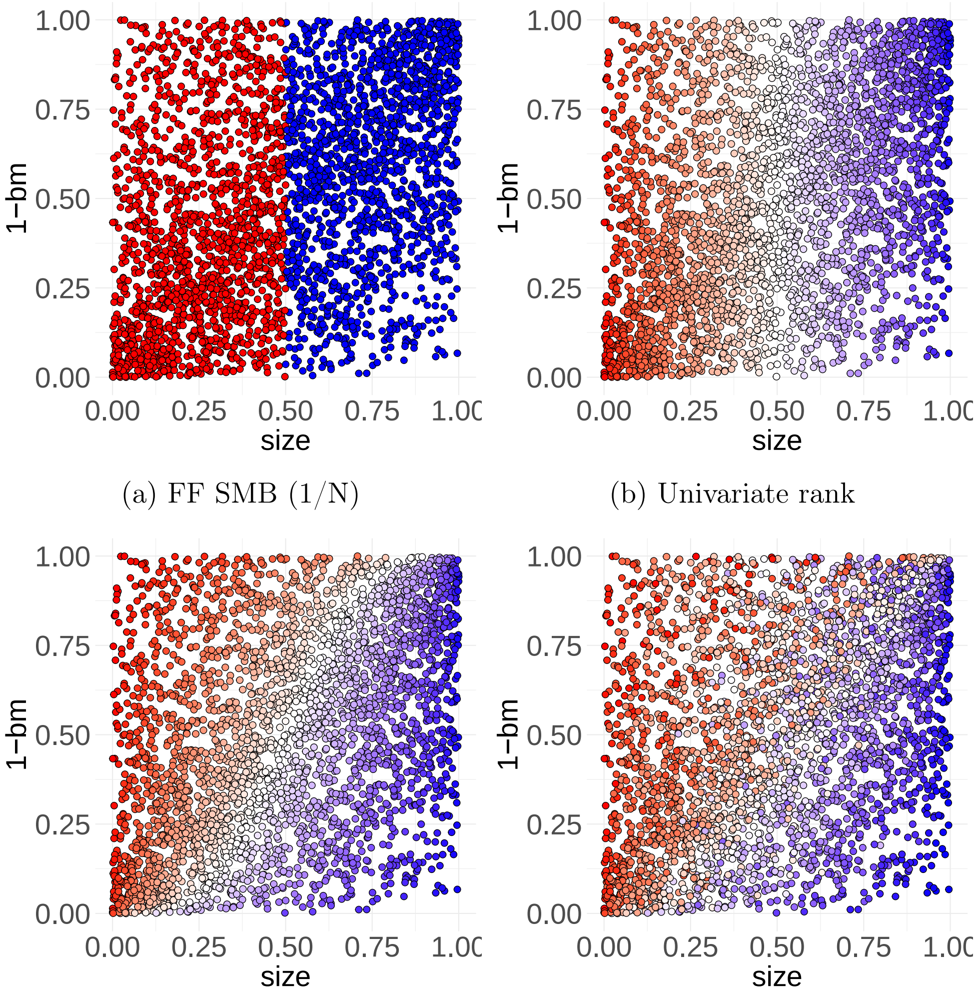
In order to calibrate the remaining error volatility, we make the restriction that $V(\eta) + \bar{\eta}^2 = \sigma_\epsilon^2$, and since the factor g is without a risk premium we can simply normalize its variance such that $\sigma_g^2 = 1$. This enables us to calibrate the variances such that:

$$\sigma_\epsilon^2 = \bar{V}(r_i)/2 - \sigma_f^2(V(\beta)+\mathrm{E}(x)^2)(\gamma/\lambda)^2. \tag{27}$$

The remaining term is the variance of the exposures to the unpriced factor, which with the restriction above is $V(\eta) = \sigma_\epsilon^2 - \bar{\eta}^2$.

B Figures and tables

Figure 1: Comparison of portfolio weights across sorting methods

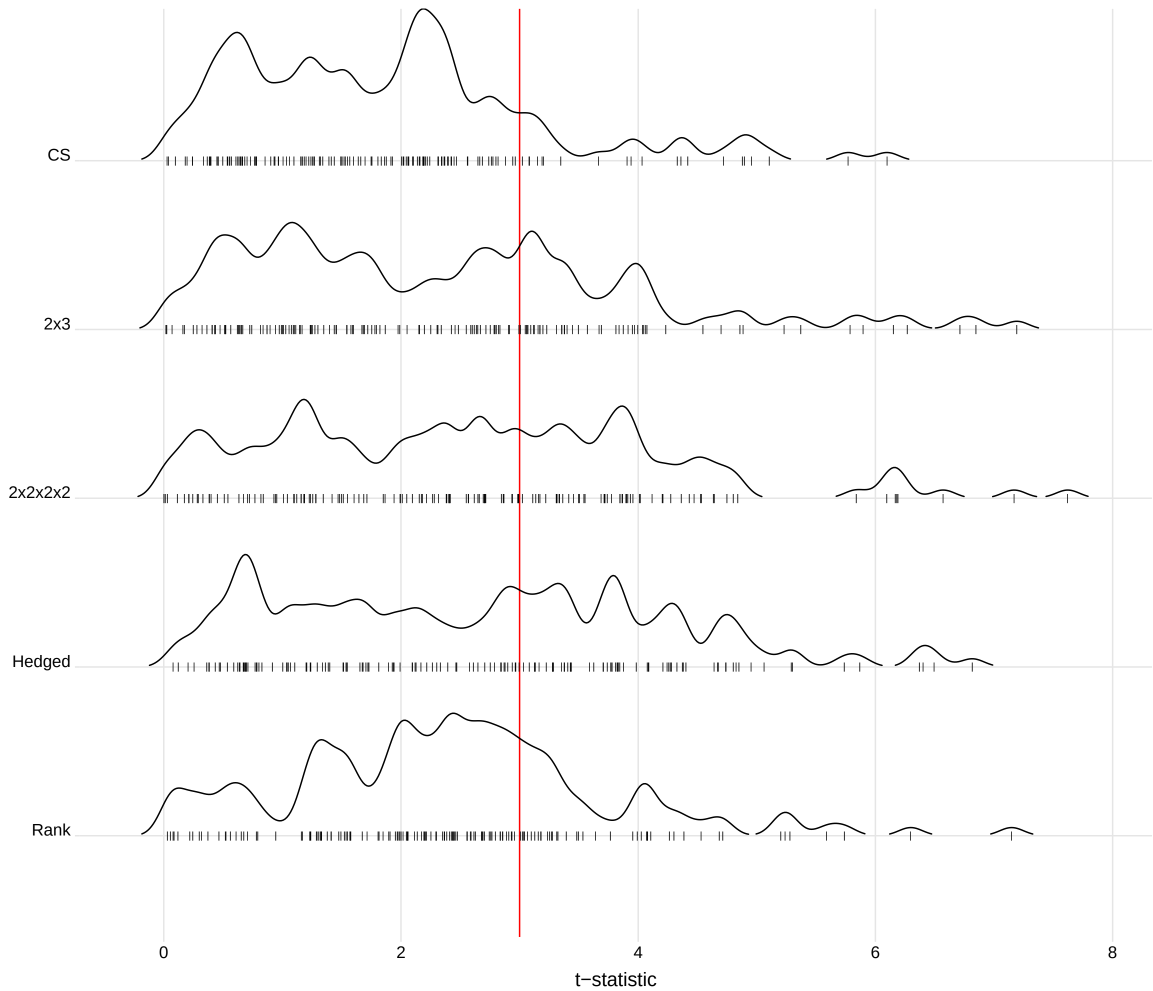


(a) FF SMB (1/N) (b) Univariate rank

(c) Copula: size conditional on value (d) Copula: size conditional on others

This figure plots the normalized rank of size and one minus normalized rank of value of all stocks in July 2020. The color of the dots corresponds to the stock portfolio weights obtained by the (a) equally-weighted FF3 SMB portfolio, (b) univariate rank size sort as in (10), (c) size copula sort obtained from estimating the two-dimensional Student-t copula based on size and value, and (d) size copula sort obtained from estimating the four-dimensional Student-t copula based on size, value, operational profitability and investment. Weights range from the most negative (dark blue) to the most positive (dark red).

Figure 2: Distribution of t-statistics of anomaly alphas across factor models



This figure plots the distribution of absolute t-statistics for the intercepts from regressing anomaly portfolio returns on different five-factor models. The red vertical line at 3.00 indicates the threshold for statistical significance accounting for multiple testing biases as in Harvey et al. (2016). The anomaly long-short portfolios are from Hou et al. (2020) and the sample ranges between January 1973 and March 2023.

Table 1: Summary of DGP parameters

		DGP - (6)	DGP - (14)
Assets	E(r)	$\beta_1\lambda_1 + \beta_2\lambda_2$	$\beta_1\lambda_1 + \beta_2\lambda_2$
	$\mathrm{V}(r)$	$\sigma_{f_1}^2\beta_1\beta_1' + \sigma_{f_2}^2\beta_2\beta_2' + \sigma_\epsilon^2 I$	$\sigma_{f_1}^2\beta_1\beta_1' + \sigma_{f_2}^2\beta_2\beta_2' + \sigma_g^2\eta\eta' + \sigma_\epsilon^2 I$
Empirical	SR	0.8	0.8
	$\mathrm{V}(\bar{r})$	0.20^2	0.20^2
	$\mathrm{V}(\mu_i)$	0.10^2	0.10^2
	$\bar{\mathrm{V}}(r)$	0.60^2	0.60^2
	R^2	-	0.5
Restrictions	γ_1	1	1
	γ_2	γ_1	γ_1
	$\rho_{(x1,x2)}$	0.5	0.5
	$\bar{x}_2$	$\bar{x}_1$	$\bar{x}_1$
	σ_{x2}^2	σ_{x1}^2	σ_{x1}^2
	σ_g^2	-	1
	$\sigma_{f1,f2}$	0	0
	σ_η^2	-	$\sigma_\epsilon^2 - \bar{\eta}^2$
Derived parameters	σ_{f1}^2	$SR^2/2$	$SR^2/2$
	σ_{f2}^2	$SR^2/2$	$SR^2/2$
	λ_1	$SR^2/2$	$SR^2/2$
	λ_2	$SR^2/2$	$SR^2/2$
	$\bar{\eta}$	-	$\sqrt{(1-R^2)\,\mathrm{V}(\bar{r})/\sigma_g^2}$
	$\bar{x}_1$	$\frac{\lambda}{\gamma}\sqrt{\mathrm{V}(\bar{r})/(2\sigma_f^2)}$	$\frac{\lambda}{\gamma}\sqrt{R^2\,\mathrm{V}(\bar{r})/(2\sigma_f^2)}$
	σ_{x1}^2	$\mathrm{V}(\mu)/(3\gamma^2)$	$\mathrm{V}(\mu)/(3\gamma^2)$
	σ_ϵ^2	$\bar{\mathrm{V}}(r) - \frac{2\gamma^2}{\lambda}(\sigma_x^2 + \bar{x}^2)$	$\bar{\mathrm{V}}(r)/2 - \frac{\gamma^2}{\lambda}(\sigma_x^2 + \bar{x}^2)$

The return-generating processes follow the models (6) and (14). The top panel *Assets* give the respective moments for the vector of returns in the respective DGPs. *Empirical* give the values that we impose based on empirical assumptions. The panel *Restrictions* provides the constraints that we impose on the three asset pricing models and *Derived parameters* summarise the internal restrictions that we find.

Table 2: Out-of-sample Sharpe ratio deviation ($Loss(w_k^{(m)})$)

	N		p(5)	p(20)	p(50)	mean	p(80)	p(95)
DGP 1	500	$Q \times P$	1.50	3.26	5.04	5.01	6.75	8.33
		CW	0.12	0.52	1.34	1.51	2.37	3.63
	1500	$Q \times P$	6.75	7.74	8.75	8.74	9.76	10.67
		CW	0.05	0.20	0.51	0.60	0.93	1.46
	3500	$Q \times P$	8.52	9.18	9.84	9.83	10.50	11.11
		CW	0.02	0.10	0.27	0.32	0.51	0.80
DGP 2	500	$Q \times P$	0.94	3.19	5.91	5.73	8.12	9.99
		CW	0.12	0.51	1.38	2.00	3.01	6.23
	1500	$Q \times P$	5.72	7.52	8.92	8.74	10.12	11.15
		CW	0.05	0.21	0.56	0.78	1.16	2.33
	3500	$Q \times P$	8.18	9.09	9.89	9.84	10.64	11.29
		CW	0.03	0.11	0.30	0.40	0.60	1.12

Distribution of the relative deviation from the optimal Sharpe ratio over $L = 10000$ simulations. Columns p(.) denote the corresponding percentile in the distribution. We consider independent 3×3 sorts for the $Q \times P$ factors.

Table 3: Summary statistics and correlations across different factors

Panel A: Summary statistics

	2 × 3 sorts				Copula sorts				Rank sorts			
	SMB	HML	RMW	CMA	SMB_{cs}	HML_{cs}	RMW_{cs}	CMA_{cs}	SMB_{rk}	HML_{rk}	RMW_{rk}	CMA_{rk}
Avg.	0.23	0.28	0.29	0.27	0.22	0.20	0.25	0.26	0.23	0.35	0.08	0.34
Vol.	3.05	3.01	2.24	2.09	1.76	1.44	1.48	0.76	2.13	1.51	1.89	1.07
t-stat.	1.99	2.48	3.46	3.49	3.32	3.72	4.45	9.21	2.90	6.25	1.09	8.54
	2 × 2 × 2 × 2 sorts				Hedged sorts							
Avg.	0.24	0.21	0.24	0.10	0.16	0.19	0.23	0.21				
Vol.	2.84	2.28	1.58	1.23	1.90	1.76	1.49	1.28				
t-stat.	2.29	2.50	4.08	2.24	2.21	2.81	4.14	4.28				

Panel B: Correlations

	2 × 3 sorts				Copula sorts				Rank sorts			
SMB	1.00				1.00				1.00			
HML	−0.01	1.00			−0.27	1.00			0.25	1.00		
RMW	−0.36	0.09	1.00		−0.37	0.79	1.00		−0.71	0.20	1.00	
CMA	−0.09	0.69	−0.01	1.00	−0.06	0.37	0.23	1.00	0.32	0.65	−0.13	1.00
	2 × 2 × 2 × 2 sorts				Hedged sorts							
SMB	1.00				1.00							
HML	0.25	1.00			0.19	1.00						
RMW	−0.71	0.20	1.00		−0.28	−0.45	1.00					
CMA	0.32	0.65	−0.13	1.00	0.15	0.68	−0.43	1.00				

Panel A reports average monthly returns, volatilities and t-statistics for the average returns of the factors across different sorts, while Panel B reports the correlation matrices of the factors for each sorting methodology. The sample ranges from July 1964 to December 2023, which gives 714 observations in the time series, with exception of the Hedged factors that are available from January 1963 to March 2023.

Table 4: Correlations across different sorts

	Copula sorts			
	SMB	HML	RMW	CMA
2×3	0.55	0.80	0.72	0.69
Rank	0.96	0.86	0.90	0.90
$2 \times 2 \times 2 \times 2$	0.56	0.85	0.77	0.70
Hedged	0.49	0.42	0.37	0.40

This table reports the correlation between each of the copula factors and the corresponding factor obtained from another sorting methodology. The sample ranges from July 1964 to December 2023, with exception of the sample used to compare with the Hedged factors, which spans from July 1964 to March 2023.

Table 5: Factor spanning regressions

	Coefficients						t-statistics						
2×3	α	Mkt	SMB	HML	RMW	CMA	α	Mkt	SMB	HML	RMW	CMA	R^2
SMB_{cs}	0.23	−0.11	0.35	−0.04	−0.04	−0.02	4.13	−8.19	13.54	−1.50	−1.40	−0.40	0.37
HML_{cs}	0.07	−0.03	−0.06	0.34	0.18	0.04	2.38	−3.69	−4.91	23.87	10.35	1.66	0.78
RMW_{cs}	0.08	−0.02	−0.03	0.19	0.43	0.02	2.50	−2.46	−2.20	9.51	21.06	0.72	0.71
CMA_{cs}	0.20	−0.00	−0.03	−0.04	0.01	0.29	9.10	−0.51	−3.20	−3.41	0.88	17.29	0.51
Rank	α	Mkt	SMB	HML	RMW	CMA	α	Mkt	SMB	HML	RMW	CMA	R^2
SMB_{cs}	0.07	−0.01	0.97	−0.19	0.22	−0.07	3.53	−2.75	54.98	−7.58	7.35	−2.02	0.96
HML_{cs}	−0.03	−0.01	−0.22	0.94	0.09	−0.16	−2.24	−2.21	−15.32	49.53	5.60	−6.28	0.95
RMW_{cs}	0.08	−0.02	0.13	0.27	0.75	−0.02	5.89	−8.07	8.66	12.81	37.48	−0.57	0.95
CMA_{cs}	0.04	−0.01	−0.04	−0.16	0.09	0.82	4.16	−2.22	−4.26	−11.80	6.72	38.11	0.92
$2{\times}2{\times}2{\times}2$	α	Mkt	SMB	HML	RMW	CMA	α	Mkt	SMB	HML	RMW	CMA	R^2
SMB_{cs}	0.20	−0.10	0.38	−0.02	−0.04	−0.07	3.82	−8.14	13.22	−0.45	−0.76	−1.23	0.37
HML_{cs}	0.09	−0.02	−0.06	0.43	0.14	0.11	3.26	−2.45	−4.72	24.83	6.37	4.02	0.77
RMW_{cs}	0.10	−0.03	−0.06	0.07	0.60	0.10	3.01	−2.76	−2.94	3.20	18.54	2.27	0.65
CMA_{cs}	0.23	−0.01	−0.03	0.03	−0.03	0.40	10.97	−0.96	−2.68	2.16	−1.30	20.24	0.50
Hedged	α	Mkt	SMB	HML	RMW	CMA	α	Mkt	SMB	HML	RMW	CMA	R^2
SMB_{cs}	0.19	−0.05	0.41	−0.07	−0.09	0.17	3.06	−2.07	9.30	−1.17	−1.43	2.33	0.27
HML_{cs}	0.06	0.02	−0.07	0.53	0.28	−0.12	1.27	0.74	−1.84	8.31	3.25	−2.34	0.29
RMW_{cs}	0.11	0.04	−0.05	0.34	0.44	−0.18	2.07	1.23	−1.25	5.45	5.38	−3.10	0.24
CMA_{cs}	0.20	−0.02	−0.03	0.01	0.07	0.27	7.94	−1.38	−1.86	0.20	2.07	8.55	0.19

This table reports, across four panels, the regression coefficients, t-statistics and R^2 associated with regressing each copula sorted factor onto the corresponding factor obtained from 2×3, Rank, $2 \times 2 \times 2 \times 2$ or Hedged sorts and a constant. The t-statistics associated with the respective coefficients are computed using heteroskedasticity robust standard errors. The sample ranges from July 1964 to December 2023, with exception of the sample for the Hedged factors, which spans from July 1964 to March 2023.

Table 6: Tangency spanning regressions and tangency weights

	Tangency spanning regressions					Tangency weights				
	Avg.	Vol.	SR	α	$t(\alpha)$	Mkt	SMB	HML	RMW	CMA
τ(Copula)	0.28	0.54	0.52			0.08	0.16	-0.09	0.26	0.59
$\tau(2 \times 3)$	0.32	1.05	0.30	0.17	10.39	0.16	0.13	-0.07	0.34	0.44
τ(Rank)	0.29	0.66	0.44	0.06	7.40	0.11	0.15	-0.03	0.34	0.37
$\tau(2\times2\times2\times2)$	0.23	0.86	0.27	0.19	11.44	0.12	0.12	-0.06	0.44	0.37
τ(Hedged)	0.24	0.62	0.38	0.19	9.81	0.11	0.15	0.03	0.34	0.37

This table reports, in the left panel, the monthly average return, volatility and Sharpe ratio of the tangency portfolio of each set of factors, together with the intercept and t-statistics from regressing the tangency allocation of the copula factors onto the tangency allocation of the other factors. The t-statistics are computed under heteroskedsticity robust standard errors. In the right panel, the fraction of wealth allocated to each of the factors in the tangency allocation of the respective sorts is reported. The sample ranges from July 1964 to December 2023, with exception of the sample for the Hedged factors, which spans from July 1964 to March 2023.

Table 7: Pricing anomaly portfolios

		Avg($\|\alpha_i\|$)	GRS	$\#\|t\| > 3.00$
All (150)	CS	0.31	2.41	21
	2×3	0.34	2.62	49
	Rank	0.39	2.48	42
	$2 \times 2 \times 2 \times 2$	0.38	2.82	57
	Hedged	0.43	2.51	63
Fric. (10)	CS	0.24	3.35	0
	2×3	0.20	2.61	1
	Rank	0.31	3.65	0
	$2 \times 2 \times 2 \times 2$	0.31	3.08	2
	Hedged	0.41	3.66	2
Intang. (20)	CS	0.38	4.90	3
	2×3	0.46	7.17	5
	Rank	0.45	5.86	4
	$2 \times 2 \times 2 \times 2$	0.48	6.53	7
	Hedged	0.56	7.12	9
Inv. (32)	CS	0.19	1.68	5
	2×3	0.19	1.93	7
	Rank	0.21	1.77	3
	$2 \times 2 \times 2 \times 2$	0.26	2.05	6
	Hedged	0.24	1.73	5
Mom. (32)	CS	0.41	2.20	5
	2×3	0.54	2.58	21
	Rank	0.48	2.45	12
	$2 \times 2 \times 2 \times 2$	0.62	2.70	26
	Hedged	0.70	2.82	29
Prof. (32)	CS	0.37	3.73	8
	2×3	0.40	3.60	15
	Rank	0.51	3.58	17
	$2 \times 2 \times 2 \times 2$	0.44	3.81	15
	Hedged	0.50	3.12	18
V-v-g (24)	CS	0.25	1.80	0
	2×3	0.15	1.93	0
	Rank	0.37	2.22	6
	$2 \times 2 \times 2 \times 2$	0.12	2.15	1
	Hedged	0.18	1.78	0

The table presents results across factor models for pricing anomaly long-short portfolios from Hou et al. (2020) between January 1973 and March 2023. The first column reports the average absolute intercepts from regressing the anomaly portfolios onto the five factors from the respective sorting methods. The second column reports the GRS statistic for the five-factor model with respect to the anomaly portfolios. The last column gives the number of obtained t-statistics in excess of 3.00, for the null hypothesis of zero pricing error.