Journal Of Investment Management, Vol. 21, No. 2, (2023), pp. 4–20

www.joim.com

INSIGHTS

"Insights" features the thoughts and views of the top authorities from academia and the profession. This section offers unique perspectives from the leading minds in investment management.

A PRACTITIONER'S GUIDE TO ADDRESS FAT TAILS AND DOWNSIDE RISK IN PORTFOLIO CONSTRUCTION

Eva A. Xu[a] *and Eric L. Tarkin*[b]

Standard models of risk and return are known to underestimate the frequency of extreme events and cannot account for the observed phenomena of increasing correlations in times of stress. This was most salient during the global financial crisis. Despite all of this, practitioners still rely heavily on the ubiquitous mean–variance optimization (MVO) for portfolio construction. This paper proposes a flexible framework, based on an explicit parametric model, that can address many of the shortcomings of MVO. We demonstrate that the proposed CVaR optimization is a superior descriptor of multi-asset returns and downside risk, and can lead to improvements in investment performance.

1 Modern Portfolio Theory in the 21st Century

The central problem of portfolio construction is determining the right combination of assets or securities in accordance with an investor's preferences and stated investment goal. Since Markowitz's (1952) original publication, modern portfolio theory (MPT) has set the foundation for portfolio construction and the widespread use of mean–variance optimization (MVO). MPT tells us that there is a tradeoff between risk and return, and that rational investors should choose a portfolio that lies along an efficient frontier which maximizes expected return for a given level of risk. Risk, in the context of MPT, is measured as the variance (or standard deviation) of returns.

Putting aside common statistical violations of independence and stationarity, MVO inherently relies on the simplifying assumption that investment returns can be approximated by a joint elliptical distribution, i.e., a multivariate normal distribution. This, in turn, implies that the variance of portfolio returns is a complete and adequate measure of investment risk. Said differently,

[a]Asset Allocation, Charles Schwab Investment Advisory, San Francisco, CA, USA. E-mail: Eva.Xu@Schwab.com
[b]Multi-Asset Research, Charles Schwab Investment Management (Formerly Charles Schwab Investment Advisory), Lone Tree, CO, USA. E-mail: Eric.Tarkin@schwab.com

the higher moments of the return distribution—such as skewness and kurtosis—can be ignored. These limiting assumptions of MVO make the portfolio construction problem tractable and analytically elegant.

Yet MPT is now almost 70 years old, and while the central tenets of MPT still hold true today, many have questioned its practical application and some of the simplifying assumptions inherent in the theory. Critics have pointed to MVO's inability, with variance and correlation as its risk measure, to model extreme events and account for increased correlations between asset classes during periods of market stress—essentially undermining diversification benefits when investors need them the most (see, for example, Lucas and Klaassen, 1998). Extreme events can have an outsized impact on investment performance which makes accurate modeling of them all the more critical. In addition, the presence of fat tails and autocorrelation in financial data can further result in estimates with a large degree of error (Campbell *et al.*, 2007). Empirical research has demonstrated that asset returns exhibit skewed and fat-tailed tendencies (for example, Sheikh and Qiao, 2010), and that elliptical and linear assumptions about risk and correlation may not adequately describe the dynamic and complex interdependency of global financial markets.

Portfolio construction has continued to evolve since Markowitz's trailblazing article, and academics and practitioners alike have tried to address some of the deficiencies of the MVO framework. Black and Litterman (1991, 1992) incorporate the practitioner's views as a prior distribution to reduce estimation error and force consistency between the expected return and covariance matrix inputs. Adler and Kritzman (2006) utilize historical returns and what they coined "full-scale optimization (FSO)" to explicitly model investors' asymmetric preference with respect to upside and downside risks outside the MVO framework.

We follow a long line of researchers attempting to bridge the gap between theory and practice and contribute to the growing body of literature known collectively as Post-Modern Portfolio Theory. In this paper, we advocate for an elegant and pragmatic framework that can address many of the shortcomings of MPT and MVO; we also provide the reader with a simple example and step-by-step instructions for implementation. We make note of ways this framework can be further refined and adapted to suit the needs, beliefs, and specific modeling choices of the practitioner.

2 A Simulation-Based, Fat-Tailed CVaR Optimization

In this paper, we propose an explicit fat-tailed, multivariate parametric model to describe the joint return distribution of a set of commonly found asset classes, numerically simulate return histories from the multivariate density, and then find optimal asset weights to maximize expected portfolio return subject to a conditional value-at-risk (CVaR) target. Others refer to this or similar approaches as expected shortfall optimization. The fat-tailed parametric assumptions make for a more realistic model of asset returns, the expected shortfall risk measure focuses on mitigating downside risk, and the stochastic approach allows us to reduce the degree of estimation error relative to the traditional MVO approach. Our proposed methodology does not require asset allocators to formulate their prior distribution with expressed views, a prerequisite for Black–Litterman and admittedly a difficult task for even the most experienced investment professionals; nor does it require explicit modeling of the investor's asymmetric utility and risk preference, as is necessary in utility-based portfolio construction methods such as FSO.

This approach extends MPT to accommodate more realistic and non-normal assumptions for asset returns, expands the optimization from MVO to Mean-CVaR, and can be shown to improve portfolio performance. It sets forth a more flexible and versatile foundational framework for portfolio construction that can also be enhanced in several additional directions. In the sections below, we review the main components of this approach, making note of alternative methods and opportunities for further extension. We first cover modeling choices for marginal and joint distributions of asset class returns, then focus on the choice of risk metric, combine these elements into a standard optimization problem and, finally, demonstrate that this methodology can improve portfolio performance over MVO using a simple illustrative example.

2.1 *Modeling of asset class returns* (*marginal distribution*)

The start of any portfolio construction process is typically some expectation of future asset class returns. MVO, for example, requires estimation of a vector of expected returns and a variance–covariance matrix describing an elliptical dispersion of returns around the mean. These inputs are typically estimated using historical data but can also be generated or augmented based on any process deemed suitable by the practitioner.

A common practice is to build asset class-specific capital market expectations (CMEs) to feed into MVO or other portfolio construction processes (Ilmanen, 2011). Many others in the industry have adopted Bayesian approaches such as Black–Litterman to incorporate forward-looking views, reduce the impact of estimation error, and force consistency between the expected mean return and covariance inputs.

Regardless of how asset class expected returns are generated, it is well documented that in many cases financial returns exhibit skew and leptokurtotic tendencies that may not be adequately described by the first two moments of a probability distribution, as is implicitly assumed by MVO. To address this directly, Tsai (2011) proposed the Johnson-SU distribution (henceforth, "Johnson") as an alternative to model asset class returns, as it is an extension and transformation of the normal distribution and includes additional parameters that can account for skewness and kurtosis, when detected (Johnson, 1949). The Johnson distribution is one of many models suitable in the presence of heavy tails; its advantages over other distributional options are its relatively simple density function and straightforward random number generation. Others have advocated for similar parametric or non-parametric distributional assumptions to address the non-Gaussian behavior of financial markets in portfolio construction (Bianchi *et al.*, 2019).

Each of these is an improvement over MVO and offers a more realistic assumption about asset class returns and the frequency of extreme (positive or negative) events. The proposed approach and example in this paper is similar in construction to Tsai (2011) and Sheikh and Qiao (2010), combining elements and attributes of both. The Johnson distribution was chosen for this paper due to its accessibility to practitioners. We provide details on how one can estimate necessary parameters using maximum likelihood in the Appendix.[1]

2.2 *Modeling dependence among asset class returns* (*joint distribution*)

After choosing a model to approximate marginal asset class returns, one needs to model the dependence structure among the asset classes. The conventional approach that asset class dependence is measured (or in the case of MVO, modeled) is to use the Pearson correlation coefficient. Pearson correlation is not without its limitations. First, putting aside the adage that "correlation doesn't imply causation," it is a measurement of the strength of a *linear* relationship.[2] Second, the observed correlation can be highly sensitive to the chosen representative estimation or sample period.

The stock–bond correlation is a favorite example to illustrate some of the nuances and limitations with standard correlation measures. Practitioners have noted that the observed Pearson correlation between stocks and bonds can be either positive or negative, depending on the chosen estimation period. It can also vary based on the contemporaneous magnitude of returns; i.e., correlations tend to rise in times of crisis. Page and Panariello (2018), for example, show that sample correlation can be materially different when one conditions on certain subsets or percentiles of historical data.

These empirical realities suggest that portfolio construction methods utilizing simple correlation coefficients may be incomplete and overly simplistic. To address this, many practitioners have adopted the use of copulas to model the dependence structure amongst asset returns. Copulas are increasingly important in the presence of fat tails, as they provide a means to incorporate tail dependence where other traditional methods may not (Kole *et al.*, 2007; Jaworski *et al.*, 2009).

Furthermore, through Sklar's theorem (Sklar, 1959), copulas allow for the separation of marginal and joint distributions. Theoretically, a practitioner could construct unique models for each marginal distribution and stitch them together in a conceptually sound manner via a chosen copula. This is particularly useful for univariate distributions that are not easily or directly extended to the multivariate case.

The multivariate-*t* (MVT) copula is a logical choice for modeling asset class dependence in portfolio construction, since it provides a means to incorporate both fat tails and tail dependence (Fischer *et al.*, 2009). We provide details on estimation of the MVT copula in the Appendix.

2.3 Choice of risk metric

Since Markowitz's seminal publication, volatility (calculated as the standard deviation of returns) has been the near universal metric used to measure and model investment risk. But despite its widespread adoption, it is well understood to be an incomplete and limited descriptor of investment risk. For one, standard deviation penalizes both positive and negative deviations from the mean return, and weights them equally. Markowitz himself noted this limitation and proposed downside risk ("semi-variance") as a preferred measure of investment risk. He ultimately used variance due to its analytical elegance and limited computational resources available at the time (Sharpe, 1964). Second, variance (or volatility) is a measurement of the *average* deviation from the mean, the typical risk one should expect to experience two-thirds of the time, based on an assumption of normality. But, in the presence of fat tails, and due to the outsized impact extreme events can have on investment performance, industry participants have favored more coherent risk metrics that focus on these extreme negative events, e.g., value-at-risk, conditional value-at-risk, and maximum drawdown. For example, Rom and Ferguson (1993) distinguish between bad and good variability and advanced the research on augmenting MPT with alternative and more coherent risk metrics, bolstering the literature and movement known as post-modern portfolio theory (PMPT). PMPT has continued to evolve and expand as academics worldwide have tested these theories and verified that they have merit.

We agree with the conclusions and advancements made by PMPT. Our preferred risk metric utilized in this framework is conditional value-at-risk (CVaR).[3] It has many desirable properties—notwithstanding it is part of the class of coherent risk measures—and is a direct measurement of the size of the tail of a distribution (Artzner *et al.*, 1999). A description of coherent risk measures and further detail on formulation of CVaR is provided in the Appendix.

2.4 Portfolio optimization

Putting all the pieces together, we have now explicitly defined a more realistic and flexible model for marginal and joint returns, and a risk metric. We can use standard statistical techniques to generate a sufficiently large sample of random returns from this multivariate distribution. We then input these returns into the augmented MVO framework with the objective to maximize the portfolio's expected return, subject to a CVaR inequality constraint. Uryasev and Rockafellar (2001) proved that the CVaR optimization problem can be well-approximated as a (convex) linear program, which makes the problem computationally tractable and guarantees a global optimum; the Appendix contains details on the derivation and construction of the optimization problem. This method can easily accommodate additional linear or quadratic constraints typical in industry practice (e.g., factor or sector exposure, tracking error) and is most appropriate for constructing multi-asset portfolios. One notable limitation of the approach is that estimation may be difficult for high-dimensional problems, e.g., constructing portfolios of individual equities or securities.

A large benefit of the approach is its reliance on simulated random returns. One of the innate challenges of empirical finance compared to other disciplines is the relatively short history of available data. Parametric simulation or other non-parametric approaches such as bootstrapping allow for estimation or, in this case, optimization to be performed on a larger sample, adding robustness to results. Practitioners with conviction in their views or capital market expectations (CMEs), or who want to construct a portfolio based on a particular macroeconomic outcome or environment (e.g., rising interest rates, above average inflation), can easily generate conditional samples of returns. Simulation adds robustness and allows for scenario analysis as part of the portfolio construction process.

3 An Example

We now present a simple example to demonstrate the application and potential benefits of the approach outlined above. We choose the Johnson distribution and MVT copula to model the marginal and joint asset class returns respectively; the resulting portfolio weights are benchmarked against those obtained from the traditional implementation of MVO and a conventional 60/40 portfolio.

We use monthly benchmark returns for 10 representative asset classes commonly included in a global portfolio and known to exhibit fat-tailed tendencies. Data used in the example are from the period March 1997 to December 2021. This period was the longest complete history of data available for the selected asset classes.[4] This 24-year period, which includes several market cycles and notable risk events such as the dot-com bubble, the Global Financial Crisis, E.U. referendum, and the COVID-19 pandemic, contains sufficient history to test the effectiveness of the different portfolio construction methodologies.

Univariate summary statistics are provided below in Table 1. From these statistics, it is evident that many of the asset classes display skew and leptokurtosis and may benefit from a more flexible parametric description than a Normal or Gaussian distribution.

Table 1 Descriptive statistics of monthly asset class returns: March 1997–December 2021.

Asset class	U.S. large cap equity	U.S. small cap equity	Intl large cap equity	Emerging market equity	U.S. REITs	Preferred stock	U.S. intermediate-term treasuries	U.S. investment grade corporate bonds	U.S. TIPS	U.S. high-yield bonds
Annualized return (arithmetic average)	8.4%	11.2%	6.4%	5.7%	13.6%	7.2%	6.6%	7.4%	7.8%	6.0%
Annualized return (geometric average)	7.2%	9.3%	5.3%	2.3%	13.3%	7.3%	6.7%	7.5%	7.9%	5.5%
Volatility	16.9%	21.6%	16.2%	25.6%	14.0%	4.1%	4.2%	4.8%	5.1%	11.5%
Minimum	−14.5%	−19.4%	−12.4%	−29.2%	−14.8%	−4.6%	−2.9%	−4.3%	−4.9%	−11.7%
Maximum	9.8%	16.5%	10.4%	13.8%	9.6%	3.3%	3.3%	3.7%	4.7%	9.4%
Median	1.0%	1.4%	0.9%	1.2%	1.5%	0.7%	0.6%	0.7%	0.6%	0.7%
Skewness	−0.48	−0.43	−0.48	−0.88	−0.61	−1.44	−0.34	−0.48	−0.67	−0.44
Excess kurtosis	2.95	3.45	3.00	4.69	4.52	7.76	3.35	4.16	6.10	5.12
VaR ($\alpha = 5\%$)	7.9%	8.1%	7.6%	12.4%	4.8%	1.3%	1.5%	1.6%	1.6%	5.7%
CVaR ($\alpha_g = 5\%$)	10.1%	13.3%	10.2%	17.5%	8.2%	2.8%	2.3%	2.7%	3.2%	7.9%
Maximum drawdown	44.7%	35.1%	48.0%	56.4%	20.3%	6.0%	3.3%	4.6%	5.6%	28.5%
Sharpe Ratio	0.50	0.52	0.40	0.22	0.97	1.77	1.56	1.52	1.53	0.53
Sortino Ratio	0.86	0.90	0.65	0.34	1.76	2.98	3.23	3.02	2.80	0.86

Historical data from Morningstar Direct. Benchmark indexes for the asset classes: S&P 500 Total Return Index (U.S. Large Cap Equity), Russell 2000 Total Return Index (U.S. Small Cap Equity), MSCI EAFE Net Total Return Index (International Large Cap Equity), MSCI Emerging Markets Net Total Return Index (Emerging Market Equity), S&P U.S. REIT Total Return Index (U.S. REITs), ICE BofAML Diversified Core U.S. Preferred Securities Total Return Index (Preferred Stock), Bloomberg U.S. Treasury 3–7 Year Total Return Index (U.S. Intermediate-Term Treasuries), Bloomberg U.S. Corporate Bond Total Return Index (U.S. Investment Grade Corporate Bonds), Bloomberg U.S. Treasury TIPS Total Return Index (U.S. TIPS) and Bloomberg VLI High Yield Total Return Index (U.S. High Yield Bonds).

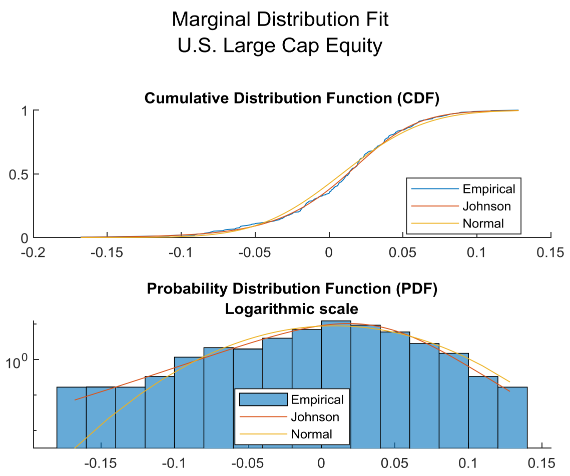


Figure 1 Comparison of marginal distribution fit: Johnson vs. Gaussian, U.S. large cap equity, March 1997–December 2021.

Historical data from Morningstar Direct.

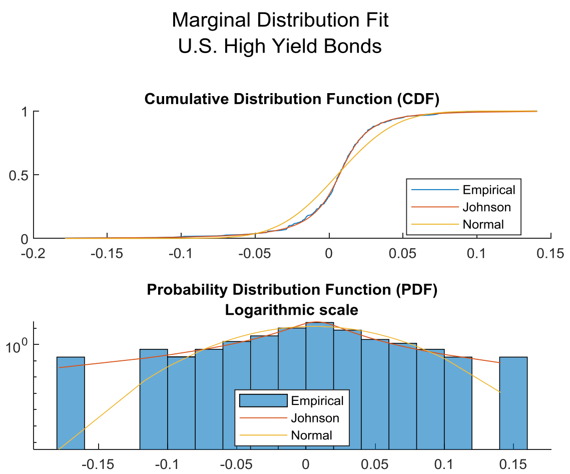


Figure 2 Comparison of marginal distribution fit: Johnson vs. Gaussian, U.S. High Yield Bonds, March 1997–December 2021.

Historical data from Morningstar Direct.

We fit marginal asset class distributions with a Johnson distribution following the method outlined in the Appendix. Figures 1 and 2 provide a visual comparison of the marginal fit of a Gaussian and Johnson distribution for U.S. Large Cap Equity and U.S. High Yield Bonds.

Not surprisingly, the differences between the fit of the Gaussian and Johnson distribution is most pronounced in the tails, where the Gaussian distribution underestimates the observed density for large negative returns. The steep rise in the middle of the cumulative density, particularly evident for example in high-yield bonds, demonstrates the degree of leptokurtosis inherent in the empirical data. Comparison charts for all 10 representative asset classes are included in the Appendix.

Once the Johnson distribution parameters are estimated for each asset class, we next fit the joint distribution assuming an MVT copula. Figure 3 is a visual representation of the bivariate fit of a multivariate normal distribution compared to an MVT copula, contrasting how our modelling choices represent the covariation between, in this case, U.S. Large Cap and U.S. Small Cap.

Even though the correlations in each model are the same, as is evident in Figure 3, simulated returns from a bivariate normal distribution fail to adequately represent and encapsulate the observed data; whereas the MVT copula produces substantially more extreme events and, in our opinion, is more representative of the downside tail risks that the portfolio we are constructing is likely to experience. Relative to a multivariate normal distribution, the larger dispersion of returns generated from an MVT copula also provides for a more robust estimate that is less sensitive to the choice of estimation period.

Once we have our marginal and joint distributions specified and estimated, we simulate 10,000 *ex-ante* histories of 12-month periods which are, in turn, used to determine optimal portfolio weights—maximizing portfolio expected return subject to not exceeding a specified risk target.

We use the period from March 1997 through December 2004 for estimation, and reserve the remaining data history from January 2005

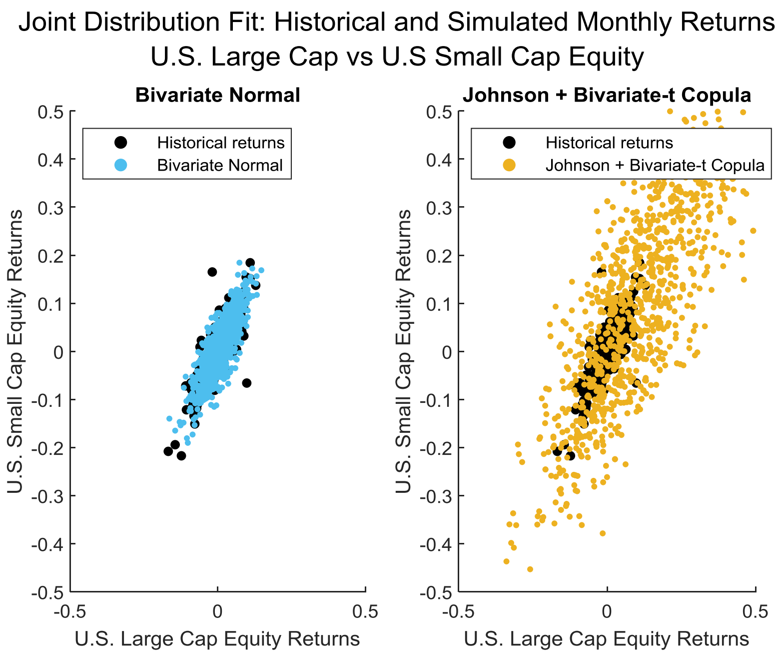


Figure 3 Comparison of joint distribution fit: Bivariate normal vs. bivariate t-copula, March 1997–December 2021.

Historical data from Morningstar Direct.

through December 2021 to evaluate out-of-sample performance. The estimation period includes notable market and political events like the dot-com bubble, September 11, 2001, and the Enron and Worldcom accounting scandals. The out-of-sample period includes the Great Financial Crisis, Greek Debt Crisis, EU referendum, and COVID-19 global pandemic. For this illustrative example, we used the simple empirical mean return over the sample estimation period for the expected return of each asset class. As mentioned earlier, one could easily supplement this framework with proprietary or survey-based return expectations.

The risk target used for the example is constructed based on the performance of a conventional 60/40 portfolio: the sample estimated variance in the case of MVO, and the estimated CVaR, based on simulated returns estimated from the in-sample data, for the Mean-CVaR algorithm. As is typical industry practice, we also applied a consistent set of business constraints to both optimizations to avoid overly concentrated portfolio weights and limit tracking error to conventional benchmarks.[5]

Table 2 compares the optimized weights from our Mean-CVaR optimization and MVO.

Comparing the Mean-CVaR optimal portfolio with the MVO optimal weights, it is notable that the Mean-CVaR portfolio holds more TIPS, U.S. Large Cap Equity, U.S. Small Cap Equity, and fewer International Developed Market Large Cap Equities, and U.S. Preferred Stocks. International Developed Market Large Cap Equities, for example, have relatively muted observed volatility but more pronounced downside risk, as estimated by CVaR and Maximum Drawdown metrics provided in Table 1. Given this observation, the resulting differences in weights are not surprising and consistent with intuition.

We finally test the performance of these portfolio weights using an out-of-sample period from

Table 2 Optimized portfolio weights, CVaR, MVO, and conventional 60/40, March 1997–December 2004 estimation period.

Asset class	CVaR	MVO	60/40
U.S. Large Cap Equity	25.2%	17.2%	60.0%
U.S. Small Cap Equity	25.2%	17.2%	0.0%
Intl Large Cap Equity	3.4%	15.0%	0.0%
Emerging Market Equity	3.4%	3.1%	0.0%
U.S. REITs	10.0%	10.0%	0.0%
Preferred Stock	1.6%	18.8%	0.0%
U.S. Intermediate-Term Treasuries	1.6%	1.9%	40.0%
U.S. Investment Grade Corporate Bonds	11.5%	13.1%	0.0%
U.S. TIPS	16.4%	1.9%	0.0%
U.S. High Yield Bonds	1.6%	1.9%	0.0%

Table 3 Out-of-sample performance: CVaR, MVO, and conventional 60/40, January 2005–December 2021.

	CVaR	MVO	60/40
Annualized return (arithmetic average)	8.7%	8.0%	8.1%
Annualized return (geometric average)	8.3%	7.5%	8.1%
Volatility	12.0%	12.3%	8.4%
Minimum	−16.6%	−13.4%	−9.6%
Maximum	10.9%	13.0%	7.8%
Median	1.0%	1.0%	0.9%
Skewness	−0.95	−0.83	−0.59
Excess kurtosis	7.18	6.64	4.95
VaR ($\alpha = 5\%$)	5.4%	5.4%	4.1%
CVaR ($\alpha = 5\%$)	8.5%	9.2%	5.5%
Maximum drawdown	41.3%	47.0%	30.1%
Sharpe Ratio	0.73	0.65	0.97
Sortino Ratio	1.15	1.02	1.67

January 2005 through December 2021. This most recent 17-year period includes the longest equity bull market in U.S. history, as well as two U.S. recessions: the Great Financial Crisis and the COVID-19 global pandemic. Table 3 compares portfolio performance statistics for the out-of-sample period and Figure 4 shows the growth of $100 over this same period.

The CVaR optimization outperforms the MVO framework for the 17-year out-of-sample period tested in this simplified example, delivering superior risk-adjusted returns and smaller downside risk and drawdowns. These results are illustrative of the benefits of strategic asset allocation frameworks that focus on downside risk metrics and incorporate more sophisticated and realistic assumptions of returns. One would be remiss not to mention the impressive risk-adjusted performance that a conventional 60/40 portfolio also delivered over this out-of-sample period. As it so happens, the 60/40 portfolio experienced meaningfully less volatility in the out-of-sample period than during the estimation period used for the MVO and CVaR algorithms, with annualized volatility of 8.4% vs. 9.9%, respectively. This period included the marked and prolonged outperformance of U.S. growth-oriented stocks,

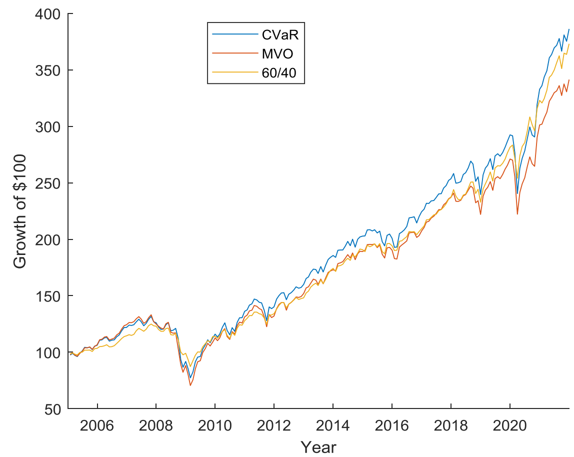


Figure 4 Out-of-sample performance: Growth of $100, January 2005–December 2021.

Historical data from Morningstar Direct.

accompanied by historically low interest rates. It is our and others' belief that the 60/40 portfolio may not continue to enjoy the favorable macroeconomic and financial environment that it has the past quarter century.

4 Conclusion

In this paper, we have outlined a portfolio construction framework that can explicitly estimate and account for downside risk and fat tails, and demonstrated improved performance relative to MVO, largely due to the outsized impact negative events can have on compounded returns. The main elements of the approach are to model marginal and joint return distributions using a Johnson-SU and multivariate-*t* copula, use these distributional assumptions to generate a sufficiently large random sample of returns, and then utilize CVaR to augment Markowitz's mean–variance framework, finding optimal portfolio weights that maximize expected portfolio return subject to a conditional value-at-risk constraint. We believe this framework makes substantial advancements over standard academic and practitioners' methods of portfolio construction, addressing empirical realities and challenges faced by industry participants and practitioners.

The framework utilizes randomly simulated returns rather than a single data history, which provides a degree of robustness and can reduce estimation error in the face of the inescapably large amounts of uncertainty inherent in quantitative finance. The framework can be easily augmented or supplemented with alternative parametric assumptions for marginal or joint distribution, non-parametric approaches, or proprietary views for return expectations. For example, based on our own research, incorporating *ex-ante* forward-looking return expectations can meaningfully enhance realized portfolio performance.

Appendix 1 Johnson-SU Distribution

The Johnson distribution's probability density function (PDF) is:

$$
\begin{aligned}
f(x; \delta, \lambda, \xi, \gamma) &= \frac{\delta}{\lambda\sqrt{2\pi}} \frac{1}{\sqrt{1+\left(\frac{x-\xi}{\lambda}\right)^2}} \\
&\times \exp\left\{-\frac{1}{2}\left[\gamma + \delta \sinh^{-1}\left(\frac{x-\xi}{\lambda}\right)\right]^2\right\}
\end{aligned}
$$

where the parameters $\delta, \lambda > 0$ and γ, ξ may take all real values. The first two moments are:

$$
\mathbb{E}(X) = \xi - \lambda \exp\left(\frac{\delta^{-2}}{2}\right) \sinh\left(\frac{\gamma}{\delta}\right)
$$

$$
\begin{aligned}
\mathrm{Var}(X) &= \frac{\lambda^2}{2}[\exp(\delta^{-2}) - 1] \\
&\times \left[\exp(\delta^{-2}) \cosh\left(\frac{2\gamma}{\delta}\right) + 1\right].
\end{aligned}
$$

As with Student's-*t* distribution, the Johnson distribution can model leptokurtic tails; additionally, the Johnson distribution is able to model distributions with both a negative and positive skew. Unlike the Gaussian and Student's-*t* distributions, built-in routines for parameter estimation and random variable simulation are usually not available for the Johnson distribution. Therefore, to estimate parameters, one can numerically maximize the log-likelihood function, which is given by:

$$
\begin{aligned}
\ell &= N\left(\log\delta - \log\lambda - \frac{1}{2}\log 2\pi\right) \\
&- \frac{1}{2}\sum_{i=1}^{N} \log\left[1 + \left(\frac{x_i - \xi}{\lambda}\right)^2\right] \\
&+ \left[\gamma + \delta \sinh^{-1}\left(\frac{x_i - \xi}{\lambda}\right)\right]^2.
\end{aligned}
$$

Sequential Quadratic Programming is a standard technique in nonlinear optimization that can be used to numerically maximize the log-likelihood, and is known to converge very quickly when the gradient and Hessian of the target function are well defined.

The corresponding cumulative distribution function (CDF) of the Johnson distribution is:

$$F = \Phi\left[\gamma + \delta \sinh^{-1}\left(\frac{x-\xi}{\lambda}\right)\right]$$

where $\Phi(\cdot)$ is the CDF of a standard Gaussian distribution. The associated quantile function $Q(p)$ is:

$$Q(p) = \xi + \lambda \sinh\left[\frac{\Phi^{-1}(p)-\gamma}{\delta}\right], \quad p \in [0, 1].$$

In addition to being a flexible descriptor of asset returns, a primary virtue of the Johnson distribution is its simplicity for Monte Carlo simulation. Using the quantile function of a distribution, a powerful technique called *inverse transform sampling* may be used; see Robert and Casella (2004, Chapter 2) for details.

Appendix 2 Multivariate-*t* Copula

Define the joint CDF for p variables as:

$$F(x_1, \ldots, x_p) = \mathbb{P}(X_1 \leq x_1, \ldots, X_p \leq x_p).$$

Based on Sklar's theorem (Sklar, 1959), which provides the theoretical foundation for the use of copulas to model dependency among random variables, every joint CDF can be expressed in terms of its marginal CDFs $F_i(x_i) = \mathbb{P}(X_i \leq x_i)$ and a copula $\mathcal{C}$:

$$F(x_1, \ldots, x_p) = \mathcal{C}[F_1(x_1), \ldots, F_p(x_p)].$$

An alternative representation for the copula, which is far more common, rests on the understanding that the CDF of a random variable is uniformly-distributed $U_i = F_i(X_i)$, which leads to the form:

$$\mathcal{C}(u_1, \ldots, u_p) = \mathbb{P}[X_1 \leq F_1^{-1}(u_1), \ldots, X_p \leq F_p^{-1}(u_p)].$$

This result comes out of the *probability integral transform* for each of the individual components. Unlike the standard Pearson correlation, a copula is invariant under monotone increasing transformations of the random variables. Once a copula has been chosen, the multivariate distribution for the set of random variables has the density:

$$f(x_1, \ldots, x_p) = c[F_1(x_1), \ldots, F_p(x_p)] \prod_{i=1}^{p} f_i(x_i)$$

where $c(\cdot)$ is the copula density and $f_i(x_i)$ is the PDF of the ith random variable. Besides providing a solid theoretical foundation for the use of copulas, Sklar's theorem shows that we may separate the modeling of the marginal distributions from the dependence structure induced by the copula.

The multivariate-t (MVT) copula is defined by taking $T \sim t_{p,\nu}(0, R)$ to be a set of random variables distributed according to an MVT distribution, with t_ν being the standard CDF for a Student's-t distribution and R as the correlation matrix:

$$\mathcal{C}_{R,\nu}(u) = \mathbb{P}[t_\nu(T_1) \leq u_1, \ldots, t_\nu(T_p) \leq u_p].$$

The MVT copula has fat-tailed attributes and the ability to model tail dependence; in the extreme case of no tail dependence, the MVT copula is simply the Gaussian copula.

There are several methods for estimating copulas that are commonly used. One option is to take the full maximum likelihood estimation route for the joint distribution, using the copula definition and the product of marginal distributions:

$$L = \prod_{j=1}^{n} \left\{ c[F_1(x_{1j}), \ldots, F_p(x_{pj})] \prod_{i=1}^{p} f_i(x_{ij}) \right\}.$$

While this is relatively simple to grasp, it is computationally difficult. A more commonly used option is the canonical maximum likelihood (CML) method that aims to maximize:

$$L = \prod_{j=1}^{n} c(\hat{u}_{1j}, \ldots, \hat{u}_{pj})$$

where $\hat{u}_{pj} = \text{rank}(x_{pj})/n$ is the nonparametric empirical CDF estimator.

Lower tail dependence, λ_l, is defined formally as:

$$\lambda_l = \lim_{u \to 0} \mathbb{P}[X_i \le F_i^{-1}(u) | X_j \le F_j^{-1}(u)]$$
$$= \lim_{u \to 0} \frac{1}{u} \mathcal{C}_{ij}(u, u).$$

The MVT copula exhibits positive lower tail dependence between asset returns, even when two assets are uncorrelated; this is seen from the explicit formula which may be derived for λ_ℓ:

$$\lambda_\ell = 2t_{\nu+1}\left[-\sqrt{(\nu+1)\left(\frac{1-R_{ij}}{1+R_{ij}}\right)}\right]$$

where $t_{\nu+1}(\cdot)$ is the CDF of Student's-t distribution with $\nu + 1$ degrees of freedom. The Gaussian copula, which is the limiting case of the MVT copula, has a tail dependence of 0—regardless of the correlation matrix—suggesting that it may not be appropriate for analyzing asset returns that exhibit tail dependence.

Appendix 3 Risk Measures

Any risk measure possessing the traits of *normalization*, *monotonicity*, *subadditivity*, *positive homogeneity*, and *translational invariance* is said to be a *coherent risk measure*. For a general risk measure $\rho(\lambda)$ of a portfolio λ, these attributes have the following properties:

- $\rho(0) = 0$ (Normalization)
- $\lambda_1 \le \lambda_2 \to \rho(\lambda_1) \ge \rho(\lambda_2)$ (Monotonicity)
- $\rho(\lambda_1 + \lambda_2) \le \rho(\lambda_1) + \rho(\lambda_2)$ (Subadditivity)
- $\alpha \ge 0 \to \rho(\alpha\lambda) = \alpha\rho(\lambda)$ (Positive Homogeneity)
- If R is a portfolio with deterministic return r, $\rho(\lambda + R) = \rho(\lambda) - r$ (Translation Invariance)

The value-at-risk (VaR) of a portfolio is defined as:

$$\text{VaR}_\alpha(X) = -\inf\{\ell \in \mathbb{R} : \mathbb{P}(X \le l) \ge \alpha\}$$

VaR is equivalent to the quantile function of a distribution evaluated at the α-percentile,[6] and is commonly used in quantitative finance; however, it does not satisfy the subadditivity property and therefore is not a coherent risk measure.[7] However, the CVaR of a portfolio, defined below, is a coherent risk measure:

$$\text{CVaR}_\alpha(X) = \mathbb{E}[-X \mid X \le -\text{VaR}_\alpha(X)].$$

Appendix 4 CVaR Portfolio Optimization

The computation of CVaR may at first look difficult, if not intractable; however, Uryasev and Rockafellar (2001) were able to show that it may be computed using linear programming. This section follows the notation conventions of Uryasev and Rockafellar (2001). For a loss function $f(w, x)$ and joint PDF $p(w, x)$ of the portfolio weights w and returns x, CVaR may be formally defined as:

$$\text{CVaR}_\alpha = \frac{1}{1-\alpha} \int_{f(w,x) \ge \text{VaR}_\alpha} \times f(w, x) p(w, x) dx.$$

To approximate the nonlinear problem, Uryasev and Rockafellar introduced an auxiliary function:

$$F_\alpha(w, \theta) = \theta + \frac{1}{1-\alpha} \int_{f(w,x) \ge \theta} \times [f(w, x) - \theta] p(w, x) dx$$
$$\equiv \theta + \frac{1}{1-\alpha} \int_R^d [f(w, x) - \theta]_+ \times p(w, x) dx.$$

The second line comes from the fact that the boundary only considers the positive parts of the integral. When F is minimized, the function is equivalent to CVaR. If we can generate J return samples using a Monte Carlo algorithm, then the integral may be approximated by:

$$\int_{\mathbb{R}^d} [f(w,x) - \theta]_+ p(w,x)dx$$

$$\approx \frac{1}{J}\sum_{j=1}^{J}[f(w,x_j) - \theta]_+.$$

Now by defining $z_j = f(w,x_j) - \theta$, with $\theta = \text{VaR}$, and choosing a linear loss function $f(w,x_j) = -w^T x_j$, the classic CVaR minimization problem may then be represented as a linear program and solved using standard techniques:

$$\min_{\theta,w,z} \theta + \frac{1}{(1-\alpha)J}\sum_{j=1}^{J} z_j$$

$$\text{s.t.} \quad z_j \geq -w^T x_j - \theta \geq 0$$

$$w^T \mathbf{1} = 1, \quad w \geq 0.$$

Equivalently, one can maximize the portfolio's expected return, subject to a CVaR constraint:

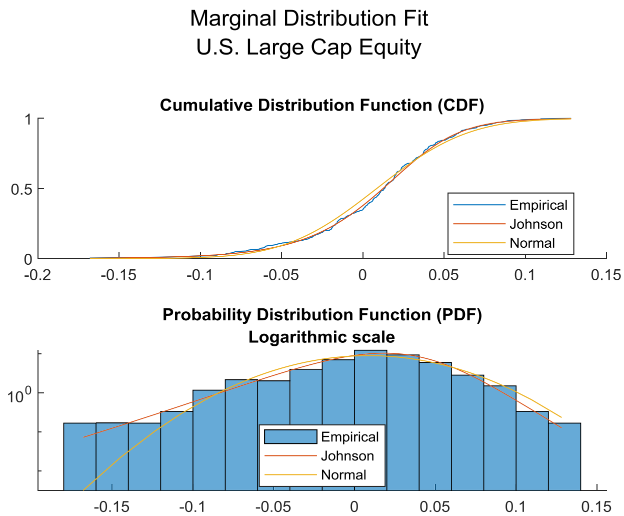
$$\max_{\theta,w,z} \quad w^T \mathbb{E}(x)$$

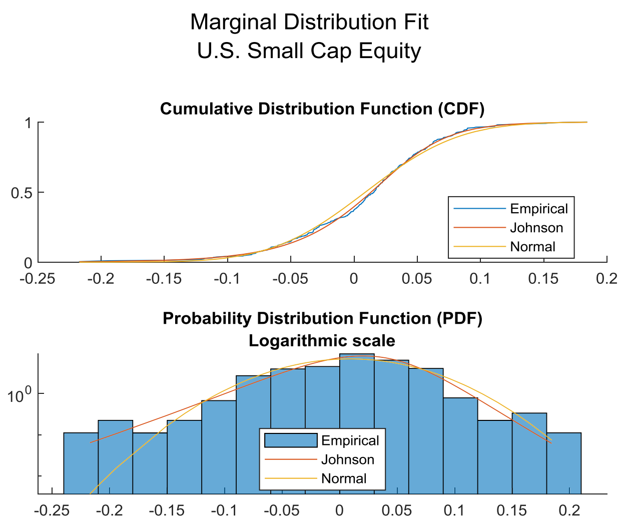
$$\text{s.t.} \quad \theta + \frac{1}{(1-\alpha)J}\sum_{j=1}^{J} z_j \leq T$$

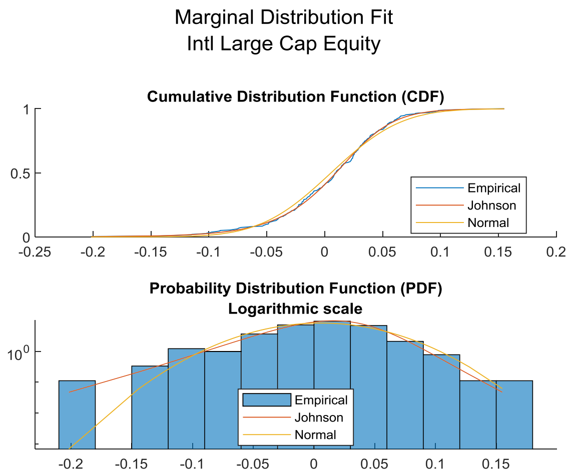
$$z_j \geq -w^T x_j - \theta \geq 0$$

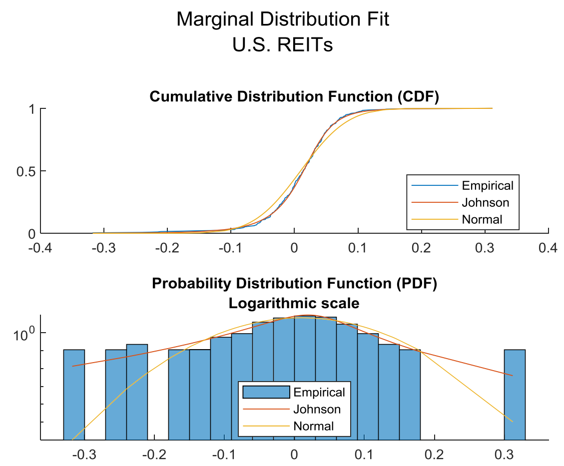
$$w^T \mathbf{1} = 1, \quad w \geq 0.$$

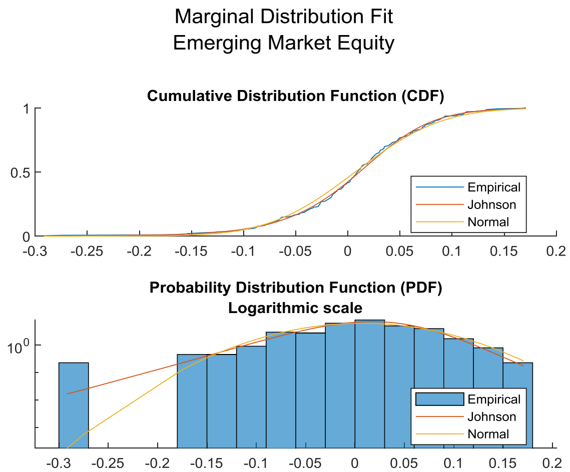
Appendix 5 Marginal Distribution Charts

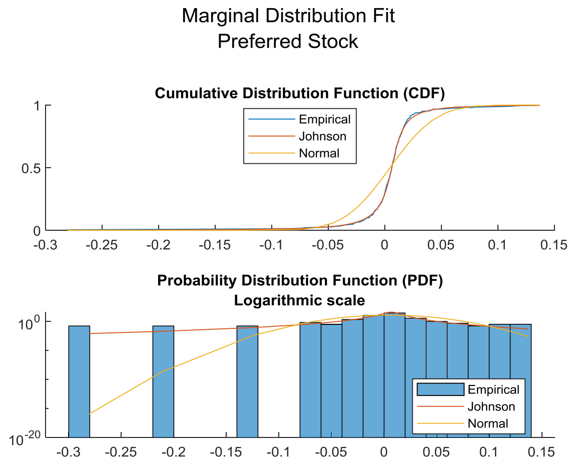


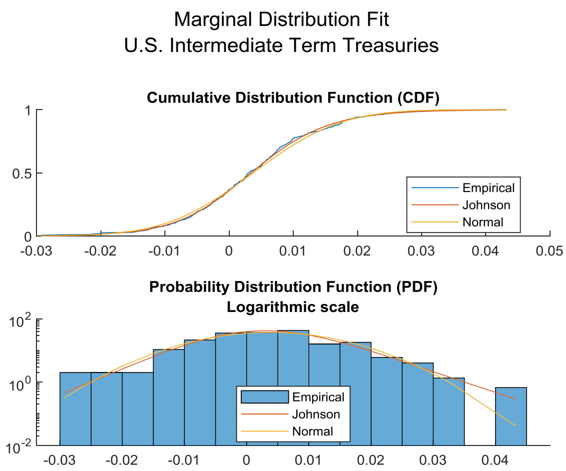


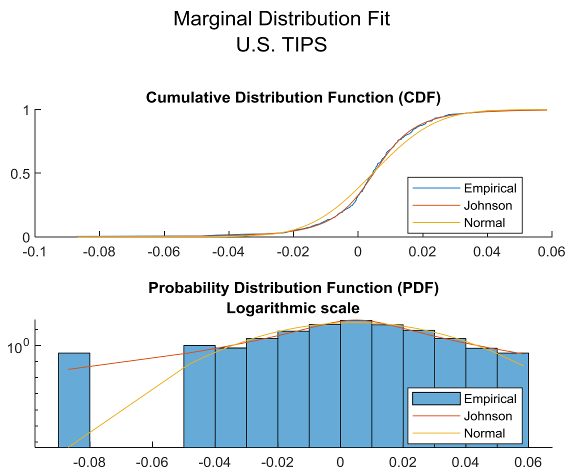

Marginal Distribution Fit
Intl Large Cap Equity
Cumulative Distribution Function (CDF)
Empirical
Johnson
Normal
Probability Distribution Function (PDF)
Logarithmic scale
Empirical
Johnson
Normal

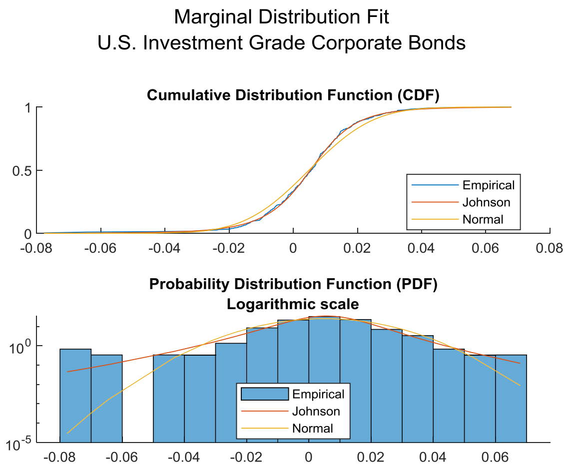

Marginal Distribution Fit
U.S. REITs
Cumulative Distribution Function (CDF)
Empirical
Johnson
Normal
Probability Distribution Function (PDF)
Logarithmic scale
Empirical
Johnson
Normal

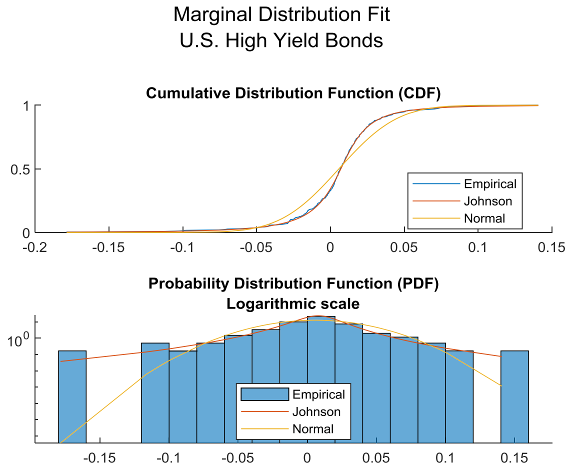

Marginal Distribution Fit
Emerging Market Equity
Cumulative Distribution Function (CDF)
Empirical
Johnson
Normal
Probability Distribution Function (PDF)
Logarithmic scale
Empirical
Johnson
Normal


Marginal Distribution Fit
Preferred Stock
Cumulative Distribution Function (CDF)
Empirical
Johnson
Normal
Probability Distribution Function (PDF)
Logarithmic scale
Empirical
Johnson
Normal


Marginal Distribution Fit
U.S. Intermediate Term Treasuries
Cumulative Distribution Function (CDF)
Empirical
Johnson
Normal
Probability Distribution Function (PDF)
Logarithmic scale
Empirical
Johnson
Normal


Marginal Distribution Fit
U.S. TIPS
Cumulative Distribution Function (CDF)
Empirical
Johnson
Normal
Probability Distribution Function (PDF)
Logarithmic scale
Empirical
Johnson
Normal


Marginal Distribution Fit
U.S. Investment Grade Corporate Bonds
Cumulative Distribution Function (CDF)
Empirical
Johnson
Normal
Probability Distribution Function (PDF)
Logarithmic scale
Empirical
Johnson
Normal


Marginal Distribution Fit
U.S. High Yield Bonds
Cumulative Distribution Function (CDF)
Empirical
Johnson
Normal
Probability Distribution Function (PDF)
Logarithmic scale
Empirical
Johnson
Normal

Endnotes

[1] Marginal distributional choices can be easily augmented to account for forward-looking return expectations (CMEs). Historical returns used for fitting parameters are simply recentered around the corresponding CME, which shifts the arithmetic average return but preserves higher moments inherent in the data.

[2] As a counterexample, consider the quadratic function $Y = X^2$. The Pearson correlation between Y and X is zero, despite the functional dependence between the variables.

[3] We suggest starting with an alpha value of $\alpha = 5\%$. For those who may want to put even more emphasis on extreme events, a smaller value of alpha could be chosen, i.e., $\alpha = 1\%$, 5% and 1% are standard choices in the literature and industry.

[4] Treasury Inflation-Protected Securities (TIPS) were first auctioned in January 1997.

[5] *Ex-ante* business constraints applied to both MVO and Mean-CVaR optimization problems were the following: U.S. Large Cap allocation $\geq$ U.S. Small Cap allocation, International Large Cap allocation $\leq$15%, Total REITs allocation $\leq$10%, Emerging Market Equity allocation $\leq$20% of total equity allocation, All equity (fixed income) allocations $\geq$5% of total equity (fixed income) allocation, All equity (fixed income) allocations $\leq$50% of total equity (fixed income) allocation.

[6] Conventional values for α are 1% and 5%.

[7] Interestingly, for elliptical distributions such as the normal distribution, using VaR as a risk measure in optimization is equivalent to MVO.

Acknowledgments

We thank Kyle Gschwend for his analytical support and contributions to this paper, Stoyan Stoyanov, Ph.D. for his peer review and useful discussions, and Jim Peterson, Ph.D. for his encouragement and leadership.

References

Adler, T. and Kritzman, M. (2006). "Mean-Variance versus Full-Scale Optimization: In and out of Sample," MIT Sloan Research Paper No. 4589–05.

Artzner, P., Delbaen, F., Eber, J. M., and Heath, D. (1999). "Coherent Measures of Risk," *Mathematical Finance* **9**(3), 203–228.

Bianchi, M. L., Stoyanov, S. V., Tassinari, G. L., Fabozzi, F. J., and Focardi, S. M. (2019). *Handbook of Heavy-Tailed Distributions in Asset Management and Risk Management*. World Scientific Handbook in Financial Economics Series **7**.

Black, F. and Litterman, R. (1991). "Asset Allocation Combining Investor Views with Market Equilibrium," *Journal of Fixed Income* **1**(2), 7–18.

Black, F. and Litterman, R. (1992). "Global Portfolio Optimization," *Financial Analysts Journal* **48**(5), 28–43.

Campbell, R. A. J., Forbes, C. S., Koedijk, K. G., and Kofman, P. (2007). "Increasing Correlations or Just Fat Tails?" *Journal of Empirical Finance* **15**(2), 287–309.

Fischer, M., Köck, C., Schlüter, S., and Weigert, F. (2009). "An Empirical Analysis of Multivariate Copula Models," *Quantitative Finance* **9**(7), 839–854.

Ilmanen, A. (2011). *Expected Returns: An Investor's Guide to Harvesting Market Rewards* (Chichester, UK: John Wiley & Sons).

Jaworski, P., Durante, F., Härdle, W. K., and Rychlik, T. (2009). "Copula Theory and its Applications," Lecture Notes in Statistics (Berlin, DE: Springer Verlag).

Johnson, N. L. (1949). "Systems of Frequency Curves Generated by Methods of Translation," *Biometrika* **36**(1/2), 149–176.

Kole, E., Koedijk, K., and Verbeek, M. (2007). "Selecting Copulas for Risk Management," *Journal of Banking & Finance* **31**(8), 2405–2423.

Lucas, A. and Klaassen, P. (1998). "Extreme Returns, Downside Risk, and Optimal Asset Allocation," *Journal of Portfolio Management* **25**(1), 71–79.

Markowitz, H. (1952). "Portfolio Selection," *The Journal of Finance* **7**(1), 77–91.

Page, S. and Panariello, R. A. (2018). "When Diversification Fails," *Financial Analysts Journal* **74**(3), 19–32.

Robert, C. P. and Casella, G. (2004). *Monte Carlo Statistical Methods*. Springer Texts in Statistics, 2nd edition (New York, NY: Springer Science + Business Media).

Rom, B. M. and Ferguson, K. W. (1993). "Post-modern Portfolio Theory Comes of Age," *The Journal of Investing* **2**(4), 27–33.

Sharpe, W. F. (1964). "Capital Asset Prices: A Theory of Market Equilibrium under Considerations of Risk," *The Journal of Finance* **19**(3), 425–442.

Sheikh, A. Z. and Qiao, H. (2010). "Non-normality of Market Returns: A Framework for Asset Allocation Decision Making," *The Journal of Alternative Investments* **12**(3), 8–35.

Sklar, A. (1959). “Fonctions de répartition à n dimensions et leurs marges,” *Publ. Inst. Statist. Univ. Paris* **8**, 229–231.

Tsai, C. S. (2011). “The Real World is Not Normal. Introducing the New Frontier: An Alternative to the Mean-Variance Optimizer,” *Morningstar Alternative Investments Observer*.

Uryasev, S. and Rockafellar, R. T. (2001). “Conditional Value-at-Risk: Optimization Approach,” In *Stochastic Optimization: Algorithms and Applications* (Berlin, DE: Springer Verlag), pp. 411–443.

Keywords: Portfolio construction; portfolio optimization; CVaR; expected shortfall; fat tails; downside risk