Quantitative Finance

ISSN: (Print) (Online) Journal homepage: https://www.tandfonline.com/loi/rquf20

Portfolios of value and momentum: disappointment aversion and non-normalities

Simon Lalancette & Jean-Guy Simonato

Quantitative Finance, 2022
Vol. 22, No. 7, 1247–1263, https://doi.org/10.1080/14697688.2022.2040742

Portfolios of value and momentum: disappointment aversion and non-normalities

SIMON LALANCETTE and JEAN-GUY SIMONATO*

Department of Finance, HEC Montréal, 3000 Cote-Sainte-Catherine, Montréal, Québec, H3T 2A7, Canada

(*Received 5 November 2019; accepted 28 January 2022; published online 11 March 2022*)

Professional money managers often combine value and momentum to benefit from diversification gains resulting from their negative return correlation. We go beyond the mean-variance and constant relative risk aversion environments to investigate this fundamental issue by assuming that investors exhibit disappointment aversion, and empirically account for the non-normalities and non-linear dependencies of the returns associated with these strategies. We perform an exploration exercise on the interplay of these two general sets of assumptions. In accordance with the literature, we find that disappointment preferences lead to optimal portfolio choices that are not attainable under the traditional constant relative risk aversion case. The findings reveal that higher disappointment aversion utility levels can be reached when non-normalities are introduced in the optimization process. The intensity of this phenomenon rises as disappointment aversion increases. However, taking advantage of these non-linearities and non-normalities requires higher portfolio turnovers. Therefore, when introducing transaction costs in the analysis, we find significant improvement in certainty equivalent returns only for investors having high disappointment aversion levels and facing low transaction costs.

Keywords: Value; Momentum; Disappointment aversion; Copula; Portfolio choices

JEL classifications: G11, G17, G41, C32, C53

1. Introduction

Value (VA) and momentum (MO) are two of the most prominent equity strategies used by quantitative money managers. These strategies have received considerable attention in the literature as they directly challenge the asset pricing paradigm. The work of Asness *et al.* (2013) appears particularly relevant for investors given their critical finding that the returns of VA and MO investment strategies show a strong negative correlation. When VA and MO strategies are considered as individual assets, combining them into a two-asset portfolio with a mean-variance analysis should improve the risk-return profile. However, further empirical investigations of this issue seem warranted for at least two reasons.

First, as Kinlaw *et al.* (2017) point out, mean-variance analysis is valid for determining optimal portfolios when the underlying joint return distribution of assets belongs to the elliptical family (for example, the multivariate normal distribution). However, the recent literature recognizes that VA and MO investment strategies display return distributions with skewness and kurtosis incompatible with those of an elliptical distribution. For example, Barroso and Santa Clara (2015) and Daniel and Moskowitz (2016) document the presence of tail risk in MO return distribution. Furthermore, Daniel *et al.* (2017) demonstrate that the correlation between VA and MO returns is state-conditional, which hints that return comovements may also display non-normalities. Investors who use an elliptical distribution in their analysis are thus likely to obtain suboptimal results.

Second, Levy and Markowitz (1979) mention that if the investor's utility is well approximated by a quadratic function, the mean-variance analysis can provide a good approximation of the optimal portfolio, even when the return distribution is not elliptical. However, experimental evidence suggests that investors have utility functions that are incompatible with quadratic approximations. Barberis (2018) points out that the traditional expected utility framework might not accurately describe the investment decision-making process. Robust empirical support for this contention is offered by Kahneman and Tversky (1979) who document that investors maintain an asymmetric attitude towards market losses as opposed to otherwise symmetrical gains.

Given the above concerns, our paper presents an exploration exercise in the formation of VA and MO portfolios in a non-expected utility framework that might better describe the investor's behavior. More specifically, considering VA and

*Corresponding author. Email: jean-guy.simonato@hec.ca

MO as individual assets, we study the performance of portfolios that simultaneously consider (i) disappointment aversion preferences as proposed in Gul (1991) and (ii) non-normalities and non-linearities of the joint return distribution in a dynamic setting.

To achieve our objectives, we build on the work of Ang *et al.* (2005), who rely on the disappointment aversion (DA) preferences of Gul (1991), which provides a setup that is solidly grounded in decision theory. The endogenous asymmetric weighting of wealth outcomes of this setup is an attractive feature implying that managers become preoccupied not only by the mean and variance of the joint return distribution, but also by its higher moments in a fashion consistent with the predictions made by Scott and Horvath (1980). The DA framework is also appealing from an empirical point of view. For example, experimental evidence reported in Choi *et al.* (2007) finds that half of the subjects have risk preferences that are best described by disappointment aversion. To the best of our knowledge, there is no empirical evidence that portfolio managers exhibit these asymmetries in their preferences. However, since they act on behalf of their clients, we consider that such asymmetric preferences are relevant to portfolio managers. DA preferences implemented in the study by Ang *et al.* (2005) offer two advantages. First, they nest the standard one-period CRRA-expected utility framework that allows for a direct comparison with the standard case found in the literature. Second, unlike prospect theory, which is silent about the reference point from which the investor imposes asymmetric weights on losses, the DA framework provides an endogenous threshold. This is an important advantage since our results will not depend on an ad-hoc reference point.

In the literature, papers that examine non-normality and portfolio allocations in a dynamic setting include those of Patton (2004), Jondeau and Rockinger (2012) and Christoffersen and Langlois (2013). Among these papers, Patton (2004), and Christoffersen and Langlois (2013) are closest to ours in that both studies use copula approaches to characterize non-normal returns.† They find that accounting for non-normalities is beneficial mainly for highly levered portfolios under high risk aversion. We contribute to the literature by extending these studies to the non-expected utility case.

In order to capture the dynamics of the data, we estimate NGARCH models for the VA and MO returns together with the dynamic conditional correlation (DCC) model of Engle (2002). To capture the non-normalities, we first specify the marginal distributions of VA and MO with a Johnson S_u distribution, which is easier to manipulate than the commonly used skewed t distribution, while allowing for a similar locus of attainable skewness and kurtosis. Because the Johnson S_u distribution is a simple transformation of a Gaussian distribution, its implementation facilitates the use of quadrature methods to compute optimal portfolios. This is especially relevant in a DA preference setting because the identification of the endogenous reference point requires a double layer of numerical work. Given the marginal Johnson S_u distributions, we then use a copula approach to characterize the joint VA and MO bivariate distribution that is required to compute the optimal portfolios. We examine three copulas: the bivariate normal, the bivariate symmetric t, and the bivariate skewed t copula of Demarta and McNeil (2005), which is advocated by Christoffersen and Langlois (2013) to capture non-linear dependencies among similar return series.

Our contributions to the literature are as follows. First, we complement the results of Asness *et al.* (2013) for optimal portfolios combining VA and MO investment strategies by explaining how different preference structures impact portfolio choices. We examine the benefits of building such portfolios when the investor exhibits disappointment aversion, a more realistic utility framework for many investors. Second, we also contribute to the literature by examining different VA and MO return generating processes. In particular, we model the joint distribution of VA and MO monthly return series via a copula approach that captures the non-normalities and non-linearities of these series. This methodology provides an empirical platform that goes beyond the unconditional sample return correlation framework of Asness *et al.* (2013). Finally, we extend the work of Ang *et al.* (2005), who examine the impact of DA preferences on wealth allocation, assuming normally distributed returns.

Our findings are as follows. The marginal distributions of VA and MO return series are well captured by the NGARCH model based on Johnson S_u innovations. For the joint return distribution, the symmetric t copula performs the best while the DCC conditional correlation on which it is based is positive in several periods, therefore challenging the diversification benefits entailed by combining the VA and MO investments. Consistent with Ang *et al.* (2005), we find that DA preferences provide wealth allocation cases that are not attainable with CRRA preferences i.e. portfolio choices for which little wealth is allocated to equity. For reasonable risk aversion levels, richer statistical approaches cannot significantly outperform the mean-variance benchmark under CRRA preferences. These statistical approaches also tend to entail higher turnovers as they are more responsive to changing market conditions. Without transaction costs, when investors exhibit DA preferences, we find that the symmetric t copula approach, as opposed to the normal copula approach, leads to significantly higher certainty equivalents in two circumstances: (i) for investors who have a low DA sensitivity to negative outcomes but at unrealistically high levels of risk aversion and (ii) for investors who show a very high DA sensitivity to negative outcomes for all levels of risk aversion. However, similarly to the CRRA case, the symmetric t copula approach leads to portfolio choices that are more sensitive to market conditions, resulting in higher portfolio turnovers. Therefore, the introduction of transaction costs in this context takes away almost all of the overperformance. More specifically, out of 40 calibrated cases with DA preferences, the symmetric t copula approach outperforms the benchmark only in 15 cases when transaction costs are set at 0.2% and only in five cases when

† The series examined here are different from the value and momentum factors used in Christoffersen and Langlois (2013). First, whereas those authors examine weekly returns, the series used here and in Asness *et al.* (2013) are monthly returns, a horizon that is more relevant for portfolio management. Second, the VA and MO series we examine are the results of pure value and momentum investment strategies, unlike the typical factor series which have been transformed to obtain returns that are neutral to market capitalization in order to ease asset pricing testing procedures.

transaction costs are set at 0.33%. These findings indicate that DA investors should use the symmetric t copula approach to generate optimal portfolio choices only if they possess a high DA sensitivity to negative outcomes and low transaction costs.

The remainder of the paper is organized as follows. The empirical approach is presented in Section 2; the estimation results are presented in Section 3; the portfolio analysis is presented in Section 4; the empirical results are examined in Section 5; concluding remarks are offered in Section 6.

2. Empirical approach

As mentioned in the introduction, we explore VA and MO portfolio choices under two general sets of assumptions: investors showing non-expected utility and return strategies exhibiting non-normalities and non-linear dependencies. In this section, we introduce our statistical approach, which accounts for non-normalities and non-linear dependencies. To compute an optimal portfolio in a non-gaussian context, we must first characterize the joint VA and MO return density for the next period. For this purpose, we use a dynamic conditional copula approach similar to Patton (2004) and Christoffersen and Langlois (2013). This approach is based on Sklar's (1959) theorem. It essentially states that the next period joint conditional density is equivalent to a combination of the conditional marginal densities and a conditional copula linking these marginals.

2.1. Marginal densities

We assume that the continuously compounded returns are well described by the following NGARCH process

$$\begin{aligned} R_{i,t} &= \varphi_i + \sigma_{i,t}\varepsilon_{i,t}, \\ \sigma_{i,t}^2 &= \beta_{i,0} + \beta_{i,1}\sigma_{i,t-1}^2 + \beta_{i,2}\sigma_{i,t-1}^2\left(\varepsilon_{i,t-1} - \theta_i\right)^2, \end{aligned} \qquad (1)$$

for $i = \{v, m\}$ where φ_i and $\sigma_{i,t}^2$ are the mean and the conditional variance of $R_{i,t}$. The random innovations $\varepsilon_{v,t}$ and $\varepsilon_{m,t}$ are zero-mean, unit variance random variables distributed according to a non-normal Johnson S_u distribution (Johnson 1949). A Johnson S_u innovation is defined by a continuous and monotonic transformation of a standard Gaussian innovation corresponding to

$$\varepsilon_{i,t} = c_i + d_i \times \sinh\left(\frac{z_{i,t} - a_i}{b_i}\right), \qquad (2)$$

for $i = \{v, m\}$ where $\sinh(u) = [\mathrm{e}^u - \mathrm{e}^{-u}]/2$, and $z_{i,t}$ is a zero-mean, unit variance Gaussian innovation. The parameters c_i and d_i consist of continuous functions of a_i and b_i. They are defined in Appendix 1, and ensure that the innovations $\varepsilon_{i,t}$ have a mean of zero and a unit variance. Here, $-\infty < a_i < \infty$, and $b_i > 0$ are the shape parameters of the Johnson S_u distribution. The log likelihood function from which we compute the parameter estimates for φ_i, $\beta_{i,0}$, $\beta_{i,1}$, $\beta_{i,2}$, θ_i, a_i, and b_i for $i = \{v, m\}$ is obtained from the Johnson S_u density. This log likelihood function is presented in Appendix 1.

Several reasons motivate our choice of the Johnson S_u distribution over the well known skewed t distribution of Hansen (1994). First, because the Johnson S_u distribution is rooted in its Gaussian counterpart, analytical expressions are easily obtained. Second, the constraints imposed on the shape parameters are less binding than those of the skewed t, although the Johnson S_u distribution generates a similar locus of attainable skewness and kurtosis pairs (see, for example, Lalancette and Simonato 2017). Finally, for some of the copula choices introduced below, quadrature approaches are conveniently implemented to compute the expected values required in the optimization procedure.†

2.2. Copula density functions

Based on the previous marginal Johnson distributions, we consider three different copula functions: (i) the bivariate standard normal distribution, (ii) the bivariate symmetric t copula, and (iii) the bivariate skewed t copula of Demarta and McNeil (2005). We also allow the copula correlation matrix associated with the bivariate copula density to vary over time for the different copula functions using the Dynamic Conditional Correlation (DCC) model of Engle (2002). The work of Christoffersen and Langlois (2013) motivates these choices.

Apart from the unconditional correlation matrix estimate which can be estimated from a sample covariance matrix of the standardized residuals, the other parameters involved with the different copula choices are as follows:

When the copula density consists of the bivariate standard normal distribution, we need to estimate two parameters capturing the dynamics of the correlation matrix in the DCC model, namely π_1 and π_2 in table 3.

When the copula density consists of the bivariate symmetric t, besides the parameters π_1 and π_2, one additional parameter needs to be estimated. This parameter is denoted as ν in table 3 and captures the degree of freedom of the joint standardized t distribution. Since this density has thicker tails than the normal distribution, it can capture more extreme tail behavior displayed by the VA and MO return series in comparison to the normal copula function.

When the copula density consists in the bivariate skewed t, besides the three previous parameters, two additional asymmetry parameters, denoted as λ_v and λ_m in table 3, are estimated. Under this distribution thicker tails are allowed to display asymmetric features, thereby providing more flexibility to reflect extreme events driving VA and MO returns.

† In theory, when working with unbounded fat-tailed distributions, infinite expected values might arise when computing exponentials of continuously compounded returns as a result of converting the latter to periodic returns in the utility function. In a technical appendix available from the authors, we have empirically verified that all expected values are finite in our study. To go about this task, we have compared the expected values and characteristics of a truncated Johnson S_u distribution (which guarantees finite expected values) with those of its unbounded counterpart used herein. We find that the first four moments of the returns of both distributions are virtually identical. We also have replicated some of the results appearing in later sections using a truncated Johnson S_u distribution. The results are, up to a Monte Carlo error, identical to those produced by the unbounded distribution with no impact on the conclusions.

Table 1. Summary statistics.

	Mean	Median	Std	Skew	Kurt	Q(1%)	Q(5%)	Q(95%)	Q(99%)
VA	0.005	0.000	0.064	2.582	24.813	− 0.139	− 0.077	0.098	0.200
MO	0.012	0.015	0.079	− 2.334	20.293	− 0.240	− 0.102	0.115	0.204
ρ	− 0.449								

Note: This table presents descriptive statistics for the monthly discretely compounded returns for the value and momentum strategies for the period January 1928–December 2016. The table shows the mean, median, standard deviation, skewness, kurtosis, and the 1%, 5%, 95%, and 99% quantiles. ρ denotes the unconditional correlation.

Table 2. Estimation results for the NGARCH.

	Value			Momentum		
	Gaussian	Johnson	Skewed t	Gaussian	Johnson	Skewed t
φ_0	0.002	0.002	0.002	0.012	0.013	0.013
	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)	(0.002)
β_1	0.865	0.858	0.859	0.631	0.591	0.588
	(0.022)	(0.026)	(0.025)	(0.022)	(0.042)	(0.043)
β_2	0.105	0.109	0.108	0.196	0.205	0.204
	(0.012)	(0.020)	(0.020)	(0.012)	(0.024)	(0.025)
θ	0.371	0.353	0.357	− 0.880	− 0.912	− 0.920
	(0.062)	(0.126)	(0.124)	(0.062)	(0.121)	(0.121)
a		− 0.276			0.431	
		(0.136)			(0.111)	
b		1.832			1.533	
		(0.190)			(0.131)	
ν			6.316			4.618
			(1.105)			(0.612)
λ			0.088			− 0.178
			(0.044)			(0.041)
JB	159.685	0.021		717.871	0.178	
	[0.000]	[0.988]		[0.000]	[0.918]	
$Q(12)$	27.952	28.015	27.884	16.573	16.656	16.640
	[0.006]	[0.006]	[0.006]	[0.166]	[0.163]	[0.164]
$Q^2(12)$	13.088	13.793	13.868	18.929	18.167	18.445
	[0.363]	[0.314]	[0.309]	[0.090]	[0.111]	[0.103]
loglik	1689.7	1717.7	1717.5	1470.1	1543.0	1542.2

Note: This table reports the estimated parameters for the NGARCH model based on the VA and MO return time series for the period January 1928–December 2016. Standard errors are reported in parentheses below the parameter estimates. Gaussian refers to NGARCH estimations when return innovations are driven by a Gaussian distribution. Johnson refers to NGARCH estimations when return innovations are driven by the Johnson S_u distribution. Skewed t refers to NGARCH estimations when return innovations are driven by a Skewed t distribution. JB is the Jarque-Bera test statistic for testing whether return innovations from the NGARCH estimations are normally distributed while $Q(12)$ and $Q^2(12)$ are the Ljung-Box statistics based on 12 lags for the standardized innovations and squared standardized innovations. The p-values for these tests are reported below in square brackets. Loglik corresponds to the likelihood value at the optimum.

Fortin *et al.* (2022) provide the exact formulas for the copula functions, as well as all the details required for estimation and implementation.†

3. Estimation results

Our monthly data set covers the period starting in January 1928 and ending in December 2016 for a total of 1,068 time series observations. The data are from the Kenneth French Library. As in Daniel *et al.* (2017), the returns of the VA and MO strategies are the monthly differences between the top and bottom decile portfolio returns based on the book/(market value of equity) ratio and momentum criterion.

† Our procedures and codes have been validated by reproducing the results obtained by Christoffersen and Langlois (2013) with their original data.

Table 1 presents descriptive statistics for the VA and MO return series. The MO strategy displays significantly higher mean and median, is more volatile and contains more tail risk than the VA series when looking at the 1% and 5% quantiles of the empirical distribution. This characteristic is the manifestation of the well-known crash risk attribute of this strategy. The sample correlation between the two series is −0.45, a value similar to the one reported by Asness *et al.* (2013).

Table 2 displays the maximum likelihood parameter estimates of the NGARCH models for the VA and MO investment strategies. For each strategy, we report two different specifications. The first (second) specification involves estimates obtained under the assumption of normal (Johnson S_u) innovations. The specifications based on the Johnson S_u innovations provide a better fit to the data since the likelihoods increase by large amounts compared with the normal case. Further indications of the better fit are also provided by the Jarque-Bera statistics whose underlying null hypothesis

Table 3. Estimation results for the Copulas.

	Normal	Symmetric t	Skewed t
π_1	0.119	0.123	0.131
	(0.018)	(0.024)	(0.029)
π_2	0.770	0.772	0.784
	(0.040)	(0.048)	(0.048)
ν		5.623	4.086
		(1.009)	(0.170)
λ_1			− 0.045
			(0.052)
λ_2			0.047
			(0.056)
loglik	63.090	83.046	84.413

Note: This table reports the estimated parameters for the normal, symmetric t and skewed t copulas for the period January 1928–December 2016. Standard errors are reported in parentheses below the parameter estimates. Loglik corresponds to the likelihood value at the optimum.

is (not) rejected under the normal (Johnson S_u) distribution. We observe that the a parameter is significantly negative (positive) for the VA (MO) strategy in conformity with the skewness coefficients reported in table 2.† We also report in table 2, estimates involving the skewed t distribution. We present these results because this distribution is more familiar and more broadly applied. The results show almost identical estimates for the parameters, standard errors and likelihood values. Our choice of a Johnson S_u probability distribution is thus comparable to the skewed t, with the added benefit of computational convenience.

Table 3 displays the maximum likelihood estimates for the three copula models, which use as inputs the innovations of the univariate NGARCH model with Johnson S_u random shocks. For the three specifications, parameters π_1 and π_2 are significant, of similar magnitude, and sum to values lower than one. This behavior leads us to believe that grouping the VA and MO strategies into a portfolio which looks beyond the unconditional correlation is warranted. Looking at the likelihood values for the three specifications, we observe that the symmetric t and skewed t copulas have much higher likelihoods than the normal counterpart, showing that these models provide a better fit with only with a few additional parameters. Interestingly, parameters λ_v and λ_m, which cause the asymmetry in the skewed t copula, are not significant. The likelihood values indicate that the skewed t copula does not bring much improvement over the symmetric t copula. We conclude from these findings that the symmetric t copula, which includes a dynamic correlation of copula innovations, provides an adequate specification of the conditional joint distribution of the VO and MA returns.

To complement the analysis we show in the upper graph of figure 1 the conditional volatilities for both investment strategies obtained under the NGARCH-Johnson specifications. We also present in the lower graph of figure 1 the time series of the correlation estimates between VA and MO under the symmetric t copula model. Not only does the conditional correlation displays a time-varying behavior, but it also becomes positive in several periods. The graph also displays bands showing the US recessions based on the NBER definition. There does not seem to be any persistent link between the correlations and the economy. The graph suggests that positive correlations do not necessarily coincide with expansion or recession periods. While some of the periods showing positive (negative) correlations fall within the vicinity of the bands, many other peaks (troughs) fall outside.

We also examine the links between financial crises and the conditional correlation time series. For this purpose, as in Bekaert *et al.* (2013), we attempt to match the correlations with recent financial crises and events (such as the Gulf war, Mexican crisis, Asian crisis, LTCM crisis, 09/11 crisis, corporate scandals, Lehman aftermath, and the Euro debt crisis). Again, we find no systematic link between positive or negative correlation peaks and financial crises and events. In another attempt to understand the dynamic behavior of the conditional correlation, we investigate the connection between financial uncertainty and the conditional correlation time series. For this purpose, we examine three measures of financial uncertainty: the VXO index,‡ the investor sentiment index of Baker and Wurgler's (2006), and the Chicago Fed National Activity Index (Basistha and Kurov 2008). These series' relationships with the estimated correlations are weak with sample correlations of − 12%, − 19%, and 6%, implying that financial uncertainty does not seem to be linked to the conditional correlation. Finally, we attempt to match the conditional correlation with the stance of the monetary policy proxied by the Fed Funds rate. Again, the empirical findings do not support the existence of a connection between these components.

4. Portfolio analysis

This section describes the type of investors' preferences that we use to explore VA and MO portfolio choices.

4.1. Portfolio returns and constraints

Here, we closely follow the previous literature (Christoffersen and Langlois 2013) as we describe the constraints and returns for a portfolio of long-short investment strategies.

Building the VA and MO return series requires a long portfolio position and the simultaneous short sell of another portfolio in an equal amount. In the absence of market frictions, these assets are zero-investment strategies that do not require any cash inflow since the amount from the short sell is used to buy the long portfolio. The investable amount is thus unlimited. In practice, however, investors are restricted in the amount they can allocate to such strategies. Regulation T, which applies to customers of U.S. broker/dealers, requires that 50% of the invested amount be deposited in a margin account. The rule implies that the ratio consisting of the size of the total position over invested capital must not exceed two.

† The parametrization we use for the Johnson S_u distribution implies a negative value of the parameter a for positively skewed distributions and vice versa.

‡ The VXO is available for a more extended period than the VIX. The correlation between the two indexes stands around 98%.

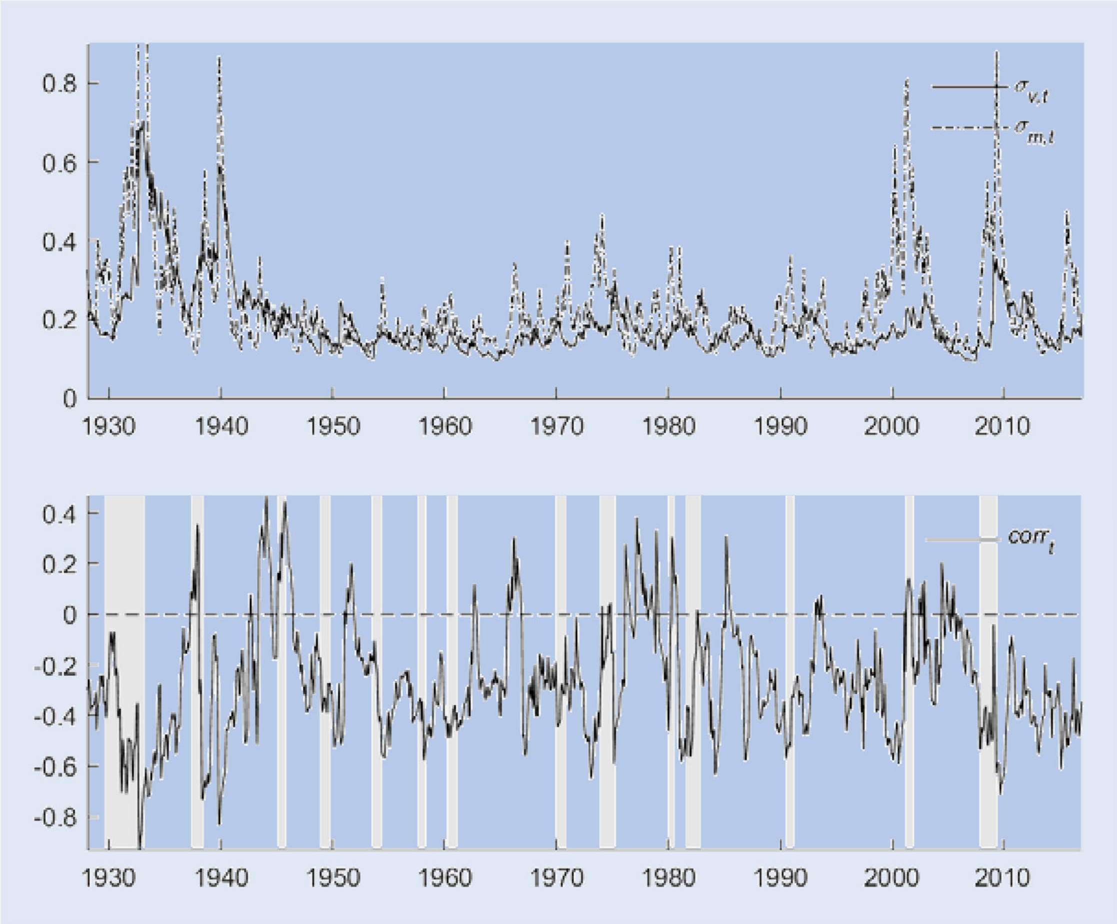

Figure 1. Estimated conditional volatilities and correlation.

Using the analysis of Pastor and Stambaugh (2003) and the assumption that the amounts placed in the margin account earn the risk-free rate, we define the return on the VA-MO portfolio as

$$r_{p,t+1} = r_{f,t+1} + \alpha_v r_{v,t+1} + \alpha_m r_{m,t+1}, \tag{3}$$

where $r_{p,t+1}$ is the $t+1$ portfolio return, $r_{f,t+1}$ is the $t+1$ risk-free rate, $r_{v,t+1}$ and $r_{m,t+1}$ respectively denote the $t+1$ monthly holding period returns on the VA and MO strategies, and where α_v and α_m denote the corresponding proportions of wealth allocated at time t. These weights must satisfy the constraint

$$2\alpha_v + 2\alpha_m \leq 1/M, \tag{4}$$

where M represents the fraction of the invested dollar amount placed in the margin account. The weights are multiplied by 2 since they simultaneously apply to long and short investment positions, while M is set to 50% under Regulation T. In this particular case, the portfolio weight constraint simplifies to

$$\alpha_v + \alpha_m \leq 1.$$

Pastor and Stambaugh (2003) mention that investors can exceed this limit by dealing with non-U.S. brokers or through joint back-office arrangements. Therefore, we also examine the impact of imposing lower margin requirements such as 25%, which implies

$$\alpha_v + \alpha_m \leq 2.$$

In addition to a margin requirement, we restrict the analysis to positive weights because Jagannathan and Ma (2003) explain that this constraint results in more stable portfolio returns.

4.2. Portfolio optimization and disappointment aversion

We base our analysis on a setup where the portfolio manager exhibits an aversion for losses that manifests itself as a DA behavior (Gul 1991) in a one-period working horizon. We also study the nested case of CRRA preferences that takes place whenever the DA mechanism by which bad outcomes are overweighted is neutralized. We follow Ang *et al.* (2005) to establish our empirical DA framework to which we add the portfolio constraints discussed previously. The DA utility problem the portfolio manager attempts to maximize is

$$\max_{\alpha_v,\alpha_m} U(\mu_W(\boldsymbol{\alpha})), \tag{5}$$

under the constraints given in the previous section, with

$$U(\mu_W(\boldsymbol{\alpha})) = \frac{1}{K}\left(\int_{-\infty}^{\mu_W(\boldsymbol{\alpha})} U(W_{t+1})\,\mathrm{d}F(W_{t+1})\right.$$

$$+ A \int_{\mu_W(\boldsymbol{\alpha})}^{\infty} U(W_{t+1})\,\mathrm{d}F(W_{t+1})\Bigg), \qquad (6)$$

where $W_{t+1} = W_t(1 + r_{p,t+1})$, $A \leqslant 1$ is the disappointment aversion parameter, $\boldsymbol{\alpha} = [\alpha_v \ \ \alpha_m]$ is the vector of portfolio weights, $F(\cdot)$ is the distribution function of next period's wealth, and K is a scalar quantity given by

$$K = \Pr(W_{t+1} \leq \mu_W(\boldsymbol{\alpha})) + A \Pr(W_{t+1} > \mu_W(\boldsymbol{\alpha})).$$

It is important to note that this utility function is written as a function of an *endogenous* certainty equivalent, $\mu_W(\boldsymbol{\alpha})$. Equation (6) reveals that the portfolio manager exhibits DA whenever $A < 1$ where states materializing above the certainty equivalent are scaled down. This mechanism intensifies as $A \to 0$. Maximizing this utility function amounts to finding the set of weights (under the stated constraints) that entails the highest certainty equivalent. As did Ang *et al.* (2005), we assume that $U(W_{t+1})$ is a power utility function (the CRRA utility function). Under this choice, because utility is homogenous in wealth, we can set $W_t = 1$ to obtain

$$U(W_{t+1}) = \begin{cases} \dfrac{\left(1 + r_{p,t+1}\right)^{1-\eta} - 1}{1-\eta} & \text{if } \eta \neq 1 \\ \ln\left(1 + r_{p,t+1}\right) & \text{if } \eta = 1, \end{cases} \qquad (7)$$

where η denotes the relative risk aversion coefficient. Furthermore, this utility function collapses to the standard CRRA case when $A = 1$. The investor becomes disappointed-averse to losses when $A < 1$.

As explained by Ang *et al.* (2005) and shown in equation (6), the implicit nature of $\mu_W(\boldsymbol{\alpha})$ makes the optimization procedure a challenging problem. Each time the optimization algorithm inputs a new set of portfolio weights to reach the optimum, the value of $\mu_W(\boldsymbol{\alpha})$ that solves equation (6) must be found with a numerical procedure. The optimization process thus requires a double layer of numerical work. Appendix 2 gives the details of our numerical strategy, which uses quadrature approaches whenever possible and Monte Carlo simulations otherwise.

5. Empirical results

We examine the impact of our different specifications using an expanding window approach. We use a first window covering January 1928–December 1959 to estimate the model parameters, and, in turn, compute the one-period-ahead forecasts of the densities for January 1960. A set of optimal portfolio weights is then obtained for that month. We then use the returns observed in January 1960 to compute the realized portfolio return. We repeat this procedure by expanding the sample window by one month until November 2016 is reached. This sequence delivers 684 out-of-sample portfolio returns. In this exercise, forecasting the expected return is a challenging task and a controversial topic. We implement, accordingly, the approach of Fleming *et al.* (2001) where the expected return is assumed constant and estimated from the in-sample mean return for all of the competing specifications.

We examine four different approaches to construct the optimal portfolios. These approaches are progressively enriched in an effort to assess the marginal impact of each of the main features characterizing the methodological framework.

- The first approach uses a mean-variance utility function with a bivariate normal distribution. The expected utility function for this case is obtained from a second-order Taylor series of the expected CRRA utility evaluated with an initial wealth set to \$1. The resulting expected utility function is then

$$E[U(W_{t+1})] \approx E\left[r_{p,t+1}\right] - \frac{\eta}{2} E\left[r_{p,t+1}^2\right].$$

  The inputs for this approach are obtained from the dynamic NGARCH model estimated with normal innovations, combined with a dynamic normal copula (normal marginals combined with a normal copula generate a multivariate normal distribution). We label this approach MV (mean-variance).
- The second approach uses a DA utility function (CRRA when $A = 1$) and a bivariate normal distribution. As in the previous case, the inputs for this alternative are obtained from the dynamic NGARCH model estimated with normal innovations, combined with a dynamic normal copula. We label this approach NmNc (normal marginal normal copula).
- The third approach uses a DA utility function (CRRA when $A = 1$) and a non-normal bivariate distribution. The inputs for this approach are obtained from the dynamic NGARCH model estimated with Johnson S_u innovations, combined with a dynamic normal copula. We label this approach JmNc (Johnson marginal normal copula).
- The fourth approach uses a DA utility function (CRRA when $A = 1$) and a non-normal bivariate distribution. The inputs for this approach are obtained from the dynamic NGARCH model estimated with Johnson S_u innovations, combined with the dynamic symmetric t copula. We label this approach JmSc (Johnson marginal Student copula).

In the next two subsections, we explore how DA preferences interact with non-normal joint distributions to determine optimal portfolio choices using these four approaches. More specifically, we analyze the marginal impact of different factors on the respective performances of optimal portfolios of VA and MO investment strategies. These factors include non-linearities and non-normalities of the VA and MO return series, the intensity of disappointment aversion, the magnitude of the relative risk aversion, and the leverage imposed through maximum margin requirements.

5.1. CRRA ($A = 1$)

When $A = 1$, investors do not exhibit disappointment aversion. In this framework, we use the MV portfolio as a benchmark. The results obtained by Levy and Markowitz (1979) provide a motivation for this choice. They advocate that for

Table 4. Portfolio choices for $A = 1$ (CRRA case).

	$\eta = 1$				$\eta = 5$				$\eta = 9$			
	MV	NmNc	JmNc	JmSc	MV	NmNc	JmNc	JmSc	MV	NmNc	JmNc	JmSc
	Required margin of 50%											
Mean	20.06	20.07	20.01	20.28	15.84	16.04	15.71	16.21	12.94	13.43	12.90	13.87
Std	20.57	20.66	20.77	20.96	11.45	11.65	11.53	11.72	8.44	8.90	8.54	8.91
Skew	− 0.97	− 0.98	− 1.04	− 0.93	− 0.43	− 0.43	− 0.44	− 0.48	− 0.30	− 0.29	− 0.23	− 0.13
Kurt	7.49	7.62	8.26	7.34	4.69	4.66	4.87	4.68	5.01	4.92	5.33	4.94
α_v	0.05	0.05	0.05	0.04	0.37	0.37	0.37	0.36	0.33	0.34	0.34	0.34
α_m	0.95	0.95	0.95	0.96	0.58	0.59	0.57	0.59	0.42	0.45	0.42	0.45
CEQ	19.41	19.39	19.29	19.57	13.22	13.30	13.02	13.44	10.08	10.22	9.97	10.75
Sharpe	0.75	0.75	0.74	0.75	0.98	0.98	0.97	0.99	0.99	1.00	0.97	1.04
Turn	2.20	2.13	2.74	2.58	5.98	6.02	7.24	7.71	5.14	5.38	6.35	7.92
p-value		0.79	0.88	0.27		0.08	0.88	0.30		0.07	0.75	0.03
	Required margin of 25 %											
Mean	32.65	32.78	32.39	32.68	20.32	21.13	19.89	21.87	13.81	14.78	13.65	15.26
Std	31.34	31.60	31.53	32.25	15.76	16.49	15.66	16.40	9.52	10.47	9.47	10.26
Skew	− 0.62	− 0.63	− 0.67	− 0.70	− 0.26	− 0.26	− 0.20	− 0.10	− 0.22	− 0.14	− 0.13	− 0.02
Kurt	5.17	5.21	5.49	5.61	5.14	5.03	5.51	5.02	5.47	5.62	5.52	5.16
α_v	0.41	0.40	0.42	0.39	0.63	0.65	0.64	0.65	0.39	0.43	0.39	0.43
α_m	1.59	1.60	1.58	1.61	0.80	0.84	0.78	0.83	0.49	0.54	0.47	0.52
CEQ	31.36	31.41	30.92	30.95	14.85	15.07	14.48	16.13	10.05	10.15	9.96	10.99
Sharpe	0.90	0.89	0.88	0.87	1.00	1.00	0.98	1.05	0.97	0.97	0.96	1.04
Turn	10.50	10.37	13.25	13.45	9.89	10.33	12.04	15.58	6.63	7.37	8.15	12.26
p-value		0.24	0.87	0.79		0.07	0.87	0.03		0.26	0.65	0.01

Note: This table reports the mean, standard deviation, skewness and kurtosis for the out-of-sample annualized returns (in %) obtained with the MV, NmNc, JmNc, and JmSc portfolio strategies. α_v and α_m are the average weights allocated to the VA and MO strategies. CEQ is the annualized certainty equivalent return in %; Sharpe is the Sharpe ratio; Turn is the average turnover in %; the p-value is obtained from of a block bootstrap approach based on the difference (CEQ *port*—CEQ MV) where *port* indicates the portfolio alternatives NmNc, JmNc or JmSc.

an investor whose utility is well approximated by a quadratic function, a mean-variance analysis provides a good approximation of the optimal portfolio, even if the return distribution is not elliptical.

Table 4 reports the findings for $A = 1$ for two margin requirements, $M = 50\%$ (regulation T) and $M = 25\%$, and three values of the risk-aversion parameter, η, set to 1, 5 and 9. This table shows summary statistics computed from the resulting out-of-sample portfolio annualized return series. In addition to the mean portfolio weights, the table also reports the certainty-equivalent return (CEQ),† the Sharpe ratio, and the average portfolio turnover. The CEQ, defined as the empirical equivalent of the inverse of the CRRA utility function, is computed as

$$\mathrm{CEQ} = \begin{cases} \left(\dfrac{1}{T}\displaystyle\sum_{t=1}^{T}(1+r_{p,t})^{1-\eta}\right)^{\frac{1}{1-\eta}} - 1 & \text{if } \eta \neq 1 \\ \exp\left(\dfrac{1}{T}\displaystyle\sum_{t=1}^{T}\ln(1+r_{p,t})\right) - 1 & \text{if } \eta = 1, \end{cases} \tag{8}$$

where $r_{p,t}$ is a realized portfolio return based on the strategies defined above. We also report bootstrapped p-values for the significance of a given specification. The p-values are based on the differences in the CEQs between a given portfolio construction approach and the MV benchmark. We compute them with the block-bootstrap approach of Politis and Romano (1994), based on 10 000 samples of the optimal portfolio returns and an average block size of 20 observations. The p-value is computed as the proportion of bootstrapped samples with CEQ differences smaller than zero. A p-value near zero indicates that the difference between the CEQ measures is statistically different from zero.

Turning our attention to the asset allocations in table 4, we see that increasing risk aversion lowers the MO weights at the benefit of the VA strategy. For example, the case $\{\eta = 1, M = 50\%\}$ is almost entirely invested in momentum, while the proportions become approximately 45% for MO and 34% for VA when $\eta = 9$. This migration towards the more conservative VA investment strategy is intuitive. Also, such a shift brings the skewness of the realized portfolios to higher levels (less negative) when risk aversion is increasing.

Table 4 suggests that accounting for departures from normality through the symmetric t copula brings, in general, the CEQ measures and Sharpe ratios to higher levels than the other approaches. The p-values indicate significant results in three cases: $\{\eta = 9, M = 50\%\}$, $\{\eta = 5, M = 25\%\}$, and $\{\eta = 9, M = 25\%\}$. Such results are qualitatively similar to those obtained in Patton (2004) and Christoffersen and Langlois (2013). For the NmNc and JmNc approaches, despite higher CEQ measures, the differences are not statistically significant. As mentioned before, Levy and Markowitz (1979) emphasize that if a quadratic function offers a good approximation of the investor's utility, as in the CRRA case, a mean-variance analysis can provide a reasonable estimate of

† The certainty equivalent measure corresponds to the annualized risk-free return making the investor indifferent between holding the optimal portfolio and earning the certainty equivalent rate over the investment horizon.

the optimal portfolio. In this context, it seems more difficult for richer statistical approaches to produce superior performances. Also, for small values of η, which reflect a higher tolerance for risk, we observe much larger standard deviations of portfolio returns relative to their means. In such an environment, obtaining statistically significant results is more difficult. For larger values of η, this phenomenon disappears, and differences among the CEQ measures also become larger. We observe that for the highest level of relative risk aversion, $\eta = 9$, the JmSc approach as opposed to the NmNc and JmNc approaches, brings significant economic gains through higher average returns and higher (less negative) skewness. However, this relative risk aversion level is unrealistically high. As pointed out by Mehra and Prescott (2007), the vast consensus has settled on values close to 3. Anything larger than these values is traditionally considered 'too high' and should be viewed with a certain degree of skepticism.†

As in Christoffersen and Langlois (2013), we also report an average turnover (in %) based on the portfolio turnover at time t defined as

$$\text{turnover}_t = \frac{1}{2} \sum_{i=v,m} \left| \alpha_{i,t} - \alpha_{i,t-1} \right|, \tag{9}$$

where $\alpha_{i,t}$ and $\alpha_{i,t-1}$ are respectively the optimal proportions at time t and $t-1$. The average turnovers reported in table 4 show that the JmNc and JmSc approaches require more trading than the others. These differences can be substantial. When $\{\eta = 9, M = 25\%\}$, for example, the turnover for the JmSc approach is almost twice that of the benchmark. These differences in turnovers can be traced back to the behavior of the conditional covariance estimates between the VA and MO series. For the NmNc and JmNc approaches, we find that the conditional covariances are slightly more stable through time than those of the JmSc approach, which translates into higher autocorrelations (unreported). These more persistent conditional covariance series produce more stable portfolio weights, resulting in lower portfolio turnovers for the NmNc and JmNc approaches when $\eta = 5$ or 9. We note that for a required margin of 50%, a different dynamic seems at work for the case of $\eta = 1$ because the average turnovers seem less sensitive to the dynamics of the conditional covariances. In this case, the low risk aversion of the investor results in portfolios almost entirely concentrated in the asset with the highest mean return investment strategy (MO), regardless of risk. In such a situation, the covariance only plays a small role in the allocation.

These higher turnovers raise questions about the economic significance of the JmSc strategy when accounting for transaction costs. We discuss this issue in the context of table 9.

5.2. Disappointment aversion (*A* < 1)

We next analyze in table 5 the findings when the DA mechanism is activated with a disappointment coefficient A set to 0.8 and 0.6. In this subsection, we add an extra layer to the analysis by allowing portfolio managers to put more weight on bad outcomes. A portfolio manager endowed with these preferences scales up bad return outcomes located below the endogenous certainty equivalent by a factor of magnitude $1/A$ when building portfolios. Since we no longer operate in the CRRA environment, we adopt the NmNc strategy as the benchmark because the MV portfolio construction does not leave any room for asymmetric treatment of future outcomes. To save space, table 5 only reports findings for the more realistic case of a required margin of 50%.

We first examine portfolio holdings. Table 5 reports that under the lowest level of risk aversion, $\eta = 1$, DA portfolio holdings are substantially less polarized than those observed under the CRRA case in table 4. In the latter case, larger amounts of wealth are allocated to the MO strategy. With $A = 0.8$, table 5 shows that the portfolio allocations are more balanced, with weights hovering around 30% (68%) for the VA (MO) strategy. Interestingly, we point out that when moving from low ($\eta = 1$) to higher ($\eta = 5$) levels of risk aversion, the adjustment in portfolio shares is borne entirely by a reduction in the investment in the MO portfolio while the VA shares remain essentially the same. We also observe that regardless of the level of η and A, the JmSc approach delivers slightly higher portfolio shares for both VA and MO portfolios than the JmNc. This phenomenon may reflect the empirical fact that the mean of the out-of-sample conditional copula covariances is lower (more negative) under the JmSc than under the JmNc. The gaps between CRRA and DA portfolio choices become smaller when risk aversion increases and reaches $\eta = 9$. Nevertheless, we find that irrespective of the magnitude of the risk aversion level, a comparison between tables 4 and 5 clearly shows that DA preferences lead to more conservative portfolios since less wealth is allocated to risky strategies in this case. This phenomenon is amplified when $A = 0.6$. The combination $\{A = 0.6, \eta = 9\}$ for example, leads to portfolios where approximately 20% to 25% of the wealth is allocated to the equity strategies. In fact, substantially more wealth is devoted to equity for the most conservative portfolio choices obtained with CRRA than with DA preferences. This result is in line with Ang *et al.* (2005) who argue that the state-dependence of the wealth allocation process allowed under DA preferences, entails portfolio positions that are not replicable under CRRA preferences, even at higher levels of risk aversion. We will further investigate this issue in table 6.

The mean turnovers reported in table 5 display a similar pattern across the NmNc, JmNc, and JmSc approaches when compared to those of the CRRA case in table 4. This is not surprising because the optimal DA portfolio choices are based on the same series of conditional covariances. This perspective is also consistent with the results (unreported) obtained by computing the autocorrelations of the VA and MO portfolio holding series, where a slightly increasing persistence is observed going from the JmSc to the NmNc approaches.

As in the CRRA case, we investigate whether the JmNc and JmSc portfolios significantly outperform the NmNc benchmark. To do so, we compute an empirical certainty equivalent with the realized portfolio returns and a bisection algorithm

† In the same spirit, Gandelman and Hernandez-Murillo (2015) state that '…the most commonly accepted measures of the coefficient of relative risk aversion lie between 1 and 3'. Looking at the data from 80 countries, they estimated that individual country estimates vary between 0 and 3.

Table 5. Portfolio choices for $A = 0.8$ and 0.6 (required margin of 50%).

	$\eta = 1$			$\eta = 5$			$\eta = 9$		
	NmNc	JmNc	JmSc	NmNc	JmNc	JmSc	NmNc	JmNc	JmSc
				$A = 0.80$, required margin of 50 %					
Mean	16.77	17.05	17.09	13.42	13.10	14.16	10.61	10.26	11.39
Std	13.66	14.22	14.27	9.10	8.98	9.32	6.45	6.13	6.67
Skew	− 0.60	− 0.70	− 0.69	− 0.28	− 0.29	− 0.15	− 0.07	0.04	0.12
Kurt	5.57	6.10	5.94	5.52	5.57	5.12	6.44	6.42	5.95
α_v	0.33	0.31	0.31	0.34	0.33	0.35	0.25	0.23	0.26
α_m	0.66	0.68	0.69	0.45	0.44	0.46	0.32	0.30	0.33
CEQ	12.63	12.72	12.72	9.08	8.83	9.71	7.14	7.08	7.80
Sharpe	0.89	0.88	0.88	0.97	0.95	1.03	0.94	0.93	1.02
Turn	5.25	6.57	7.09	5.73	7.16	8.69	4.73	5.53	7.91
p-value		0.31	0.36		0.88	0.10		0.65	0.01
				$A = 0.60$, required margin of 50 %					
Mean	10.55	11.06	12.67	7.72	7.87	9.11	6.59	6.50	7.37
Std	8.08	8.51	8.80	4.68	4.51	5.22	3.20	2.91	3.50
Skew	− 0.09	− 0.25	0.10	0.58	0.72	0.64	0.36	0.40	0.45
Kurt	8.75	8.03	6.47	11.45	10.51	9.06	10.41	10.21	8.93
α_v	0.23	0.24	0.28	0.13	0.13	0.17	0.09	0.08	0.11
α_m	0.33	0.36	0.39	0.19	0.19	0.22	0.12	0.11	0.14
CEQ	6.02	6.12	7.35	4.89	5.12	5.78	4.54	4.65	5.04
Sharpe	0.74	0.76	0.92	0.67	0.73	0.87	0.63	0.66	0.80
Turn	6.41	8.86	11.36	4.48	5.19	8.41	3.18	3.48	6.23
p-value		0.37	0.04		0.07	0.00		0.16	0.00

Note: This table reports the mean, standard deviation, skewness and kurtosis for the out-of-sample annualized returns (in %) obtained with the NmNc, JmNc, and JmSc portfolio strategies. α_v and α_m are the average weights allocated to the VA and MO strategies. CEQ is the annualized certainty equivalent return in %; Sharpe is the Sharpe ratio; Turn is the average turnover in %; the p-value is obtained from of a block bootstrap approach based on the difference (CEQ *port*—CEQ NmNc) where *port* indicates the portfolio alternatives JmNc or JmSc.

solving the following equations for the CEQ

$$\mathrm{CEQ} = \begin{cases} \dfrac{1}{K}\displaystyle\sum_{t=1}^{T} \dfrac{(1+r_{p,t})^{1-\eta}-1}{1-\eta} \\ \quad \left(\mathbf{1}_{]-\infty,\mathrm{CEQ}]} + A\mathbf{1}_{[\mathrm{CEQ},\infty[}\right), & \text{if } \eta \neq 1 \\ \dfrac{1}{K}\displaystyle\sum_{t=1}^{T} \ln(1+r_{p,t}) \\ \quad \left(\mathbf{1}_{]-\infty,\mathrm{CEQ}]} + A\mathbf{1}_{[\mathrm{CEQ},\infty[}\right) & \text{if } \eta = 1 \end{cases} \tag{10}$$

with

$$K = \sum_{t=1}^{T} \left(\mathbf{1}_{]-\infty,\mathrm{CEQ}]} + A\mathbf{1}_{[\mathrm{CEQ},\infty[}\right),$$

where $1_{[\cdot]}$ is an indicator variable equal to 1 if the wealth level falls within the interval and 0 otherwise. The upper panel in table 5 where $A = 0.8$ shows that even though the JmSc approach achieves larger CEQ estimates in comparison to the NmNc benchmark, there is only one case where this dominance is significant and that is when $\eta = 9$. This is similar to that observed in the upper panel of table 4 under CRRA preferences. In this set up, the JmNc approach does not deliver additional significant economic gains. However, table 5 shows that when adverse outcomes are overweighted by a factor of $(A = 0.6)^{-1} = 1.66$, the JmSc portfolios significantly outperform the benchmark for all levels of risk aversion and produce portfolios with higher skewness than the alternative methods. Such is not the case for the JmNc approach. These findings may not be surprising however. When investors progressively overweigh market downturns, a statistical approach that better captures the tails of the distributions becomes more relevant and potentially allows the DA utility function to reach higher CEQs. Finally, increasing the risk aversion level has an impact similar to that observed in table 4. As risk aversion is increasing, the wealth is progressively shifted from the MO to the VA strategy.

In table 6, we further investigate the interplay between the DA intensity and the relative risk aversion levels, and their respective impacts on asset allocation, performance and turnover for the JmSc data-generating process. Table 6 reports different statistics where A varies from 1 to 0.5, while η varies from 1 to 15. To make the analysis manageable, we only focus on the regulation-T case where $M = 50\%$.

Before examining the DA framework, the upper panel of table 6 reports the findings for the special case of CRRA preferences, i.e. $A = 1$. We notice that for low risk aversion levels, most of the wealth is allocated to MO. As risk aversion is increasing, the allocations gradually shift from MO to VA. The allocations reach average proportions of 25% and 31% in VA and MO respectively when $\eta = 15$, a risk aversion level which is perhaps unrealistic. Nevertheless, we can use this case to assess the extent to which DA wealth allocations can be replicated in a CRRA framework.

The lower panels of table 6 report the results for DA preferences. We first generally observe that, for all values of A, the reduction in portfolio shares is much larger for MO than VA as η increases. This result maybe not surprising since MO portfolios display more volatility and tail risks. On the other hand, there are cases leading to similar allocations to the above CRRA example, but at levels of relative risk aversion that are

Table 6. JmSc portfolios (required margin of 50%).

	$\eta = 1$	$\eta = 3$	$\eta = 5$	$\eta = 7$	$\eta = 9$	$\eta = 11$	$\eta = 13$	$\eta = 15$
				CRRA (A = 1)				
CEQ	19.57	15.58	13.44	11.98	10.75*	9.75*	9.01*	8.42*
mean	20.28	17.61	16.21	15.07	13.87	12.73	11.80	10.98
std	20.96	14.06	11.72	10.17	8.91	7.81	6.93	6.17
skew	− 0.93	− 0.61	− 0.48	− 0.29	− 0.13	− 0.06	0.01	0.03
α_v	0.04	0.29	0.36	0.37	0.34	0.31	0.28	0.25
α_m	0.96	0.71	0.59	0.51	0.45	0.39	0.35	0.31
turn	2.58	7.48	7.71	7.79	7.92	7.65	7.26	6.83
				$A = 0.9$				
CEQ	15.84	13.15	11.63	10.33	9.27*	8.49*	7.90*	7.40*
mean	18.60	16.63	15.42	14.08	12.78	11.71	10.82	10.04
std	17.31	12.71	10.68	9.15	7.88	6.88	6.04	5.35
skew	− 0.81	− 0.57	− 0.37	− 0.14	− 0.03	0.04	0.08	0.06
α_v	0.17	0.35	0.37	0.35	0.31	0.27	0.24	0.21
α_m	0.83	0.63	0.53	0.46	0.39	0.34	0.30	0.27
turn	6.35	7.63	7.92	8.18	8.01	7.54	7.04	6.55
				$A = 0.8$				
CEQ	12.72	11.00	9.71	8.60*	7.80*	7.25*	6.77*	6.42*
mean	17.09	15.67	14.16	12.62	11.39	10.42	9.59	8.93
std	14.27	11.28	9.32	7.81	6.67	5.74	5.01	4.41
skew	− 0.69	− 0.46	− 0.15	0.02	0.12	0.17	0.14	0.11
α_v	0.31	0.37	0.35	0.30	0.26	0.22	0.20	0.17
α_m	0.69	0.55	0.46	0.39	0.33	0.28	0.25	0.21
turn	7.09	8.19	8.69	8.55	7.91	7.29	6.70	6.08
				$A = 0.7$				
CEQ	9.84	8.76	7.62*	6.89*	6.41*	5.98*	5.69*	5.51*
mean	15.55	13.86	11.97	10.62	9.59	8.74	8.11	7.65
std	12.03	9.27	7.45	6.17	5.17	4.41	3.82	3.36
skew	− 0.52	− 0.12	0.19	0.30	0.34	0.28	0.23	0.24
α_v	0.37	0.33	0.27	0.23	0.19	0.16	0.14	0.12
α_m	0.57	0.45	0.36	0.29	0.24	0.21	0.18	0.15
turn	9.28	9.94	9.50	8.45	7.60	6.80	6.05	5.39
				$A = 0.6$				
CEQ	7.35*	6.29*	5.78*	5.38*	5.04*	4.88*	4.75*	4.67*
mean	12.67	10.39	9.11	8.13	7.37	6.88	6.52	6.27
std	8.80	6.59	5.22	4.22	3.50	2.96	2.57	2.29
skew	0.10	0.61	0.64	0.63	0.45	0.41	0.42	0.43
α_v	0.28	0.21	0.17	0.13	0.11	0.09	0.08	0.07
α_m	0.39	0.29	0.22	0.18	0.14	0.12	0.10	0.09
turn	11.36	10.02	8.41	7.20	6.23	5.37	4.69	4.15
				$A = 0.5$				
CEQ	5.05*	4.61*	4.24*	4.10	4.03	4.02	3.96	3.94
mean	8.18	7.07	6.26	5.80	5.54	5.37	5.23	5.15
std	5.08	3.78	2.92	2.30	1.92	1.67	1.50	1.38
skew	1.19	0.97	0.64	0.33	0.37	0.39	0.38	0.38
α_v	0.12	0.09	0.07	0.05	0.04	0.03	0.03	0.03
α_m	0.18	0.13	0.09	0.07	0.06	0.05	0.04	0.04
turn	7.84	6.34	5.28	4.29	3.60	3.03	2.66	2.33

Note: This table reports the mean, standard deviation, and skewness for the out-of-sample annualized returns (in %) obtained with the JmSc portfolio strategies. α_v and α_m are the average weights allocated to the VA and MO strategies. CEQ is the annualized certainty equivalent return in %; A star indicates a significant difference between the certainty equivalent at 5% level based on a zero proportional transaction cost and a block bootstrap approach.

more reasonable. For example, the case $A = 0.7$ and $\eta = 7$ shows mean portfolio holdings of 23% (29%) in VA (MO). These proportions are roughly the same as those of the CRRA case where $\eta = 15$. At this point however, increasing the DA intensity levels generates average allocations that cannot be reached under the CRRA preferences when a reasonable level of relative risk aversion is assumed. Furthermore, at the minimum DA level, $A = 0.5$, our results are consistent with Ang *et al.* (2005) who observe non-participation in the stock market under DA preferences. Again, such results are not attainable for reasonable risk aversion levels under CRRA preferences.

Next, we discuss the results for the performance analysis. As observed earlier, increasing either the DA intensity level A or η decreases the magnitude of the CEQs, mean returns, and return standard deviations. In a consistent fashion, increasing the DA intensity level or raising the level of the relative risk aversion causes the portfolio skewness to rise. Of course, these tendencies are a manifestation of the portfolios becoming increasingly more conservative by moving wealth from MO to VA and, in turn, lowering the proportion allocated to equity. In terms of the CEQs based on the JmSc approach, significant overperformance occurs only at relatively higher levels of risk aversion when A is high to moderate and at lower levels of risk aversion when A is low. For example, when $A = 0.6$, the JmSc approach displays significant overperformance at low levels of risk aversion. At such low DA coefficients, the DA investor imposes a more important rescaling of tail outcomes. Since the JmSc is better able to capture these low tail events, this portfolio approach produces better performance. We also point out that the JmSc approach produces superior portfolio choices when $\eta \leqslant 5$ and $A = 0.5$. These are cases where almost no wealth is invested in the risky assets.

Table 6 shows that for a fixed level of risk aversion, the average turnover estimates first increase and then decrease when A becomes smaller. This phenomenon is rooted in the increasing sensitivity of DA investors to adverse outcomes as A becomes smaller which, in turn, results in a higher degree of responsiveness to variations in the portfolio return distributions. This phenomenon entails intensive portfolio rebalancings. However, past a certain point, as a lower proportion of wealth is allocated to risky assets, the turnovers mechanically decrease to lower levels because the portfolio weights decrease in magnitude. We observe a similar phenomenon when η becomes larger and the level of DA intensity is fixed.

A more complex portfolio approach, such as the JmSc, produces optimal portfolio holdings which are the results of different features designed to capture non-linearities and non-normalities. One such feature is the dynamic correlation captured by the symmetric t copula. To obtain a better intuition about how this feature and the portfolio holdings interact, we report in table 7 average portfolio holdings for periods of negative and positive correlations. These periods are determined according to the estimated correlations obtained from the maximum likelihood estimate of the symmetric t copula using the full sample. The table shows that in most cases, periods of negative correlations lead to invested proportions that are larger than in periods of positive correlations. VA and MO are both scaled down in roughly similar proportions in the latter state. These differences are robust to the different values of A, except when $\eta = 1$. At the lowest level of risk aversion, the sign of the conditional correlation has little impact unless A is very small.

A complementary research question arising with the more sophisticated JmSc approach is whether it outperforms the benchmark when VA and MO perform poorly. We investigate

Table 7. Differences in average proportions between periods of negative and positive correlation (required margin of 50%).

	$\eta = 1$	$\eta = 5$	$\eta = 9$	$\eta = 13$
	CRRA ($A = 1$)			
$\alpha_{v,rho-}$	0.04	0.37	0.36	0.30
$\alpha_{m,rho-}$	0.96	0.60	0.46	0.36
$\alpha_{v,rho-} - \alpha_{v,rho+}$	− 0.01	0.03	0.11	0.12
$\alpha_{m,rho-} - \alpha_{m,rho+}$	0.01	0.09	0.13	0.13
	$A = 0.9$			
$\alpha_{v,rho-}$	0.17	0.38	0.33	0.26
$\alpha_{m,rho-}$	0.83	0.55	0.41	0.32
$\alpha_{v,rho-} - \alpha_{v,rho+}$	− 0.01	0.08	0.13	0.12
$\alpha_{m,rho-} - \alpha_{m,rho+}$	0.01	0.11	0.14	0.13
	$A = 0.8$			
$\alpha_{v,rho-}$	0.31	0.36	0.28	0.21
$\alpha_{m,rho-}$	0.69	0.48	0.35	0.26
$\alpha_{v,rho-} - \alpha_{v,rho+}$	− 0.01	0.13	0.13	0.11
$\alpha_{m,rho-} - \alpha_{m,rho+}$	0.02	0.14	0.15	0.12
	$A = 0.7$			
$\alpha_{v,rho-}$	0.38	0.29	0.21	0.15
$\alpha_{m,rho-}$	0.58	0.38	0.26	0.19
$\alpha_{v,rho-} - \alpha_{v,rho+}$	0.09	0.16	0.13	0.10
$\alpha_{m,rho-} - \alpha_{m,rho+}$	0.10	0.17	0.14	0.11
	$A = 0.6$			
$\alpha_{v,rho-}$	0.30	0.18	0.12	0.09
$\alpha_{m,rho-}$	0.42	0.24	0.16	0.11
$\alpha_{v,rho-} - \alpha_{v,rho+}$	0.18	0.13	0.09	0.07
$\alpha_{m,rho-} - \alpha_{m,rho+}$	0.20	0.17	0.11	0.08
	$A = 0.5$			
$\alpha_{v,rho-}$	0.14	0.08	0.05	0.03
$\alpha_{m,rho-}$	0.20	0.11	0.06	0.05
$\alpha_{v,rho-} - \alpha_{v,rho+}$	0.11	0.07	0.04	0.03
$\alpha_{m,rho-} - \alpha_{m,rho+}$	0.16	0.09	0.06	0.04

Note: This table reports the differences in average proportions between periods of negative and positive correlation. $\alpha_{i,rho-}$ is the average proportion of asset i for negative correlation periods, while $\alpha_{i,rho+}$ is the average proportion of asset i for positive correlation periods.

Table 8. Portfolio Sharpe ratios when value and momentum perform poorly (required margin of 50%).

	$\eta = 1$	$\eta = 5$	$\eta = 9$	$\eta = 13$
	CRRA ($A = 1$)			
$Sh_{p,NmNc}$	− 1.2373	− 1.5925	− 1.5396	− 1.4673
$Sh_{p,JmSc}$	− 1.2215	− 1.6481	− 1.5715	− 1.4590
$Sh_{p,JmSc} - Sh_{p,NmNc}$	0.0158	− 0.0556	− 0.0319	0.0083
	$A = 0.9$			
$Sh_{p,NmNc}$	− 1.4457	− 1.5004	− 1.4575	− 1.3788
$Sh_{p,JmSc}$	− 1.4168	− 1.6186	− 1.4687	− 1.3594
$Sh_{p,JmSc} - Sh_{p,NmNc}$	0.0289	− 0.1182	− 0.0111	0.0194
	$A = 0.8$			
$Sh_{p,NmNc}$	− 1.5591	− 1.4031	− 1.2719	− 1.2095
$Sh_{p,JmSc}$	− 1.5396	− 1.5052	− 1.2583	− 1.1822
$Sh_{p,JmSc} - Sh_{p,NmNc}$	0.0195	− 0.1022	0.0136	0.0273
	$A = 0.7$			
$Sh_{p,NmNc}$	− 1.4033	− 1.1138	− 0.9942	− 0.8938
$Sh_{p,JmSc}$	− 1.3759	− 1.1585	− 0.9798	− 0.8764
$Sh_{p,JmSc} - Sh_{p,NmNc}$	0.0274	− 0.0447	0.0145	0.0175
	$A = 0.6$			
$Sh_{p,NmNc}$	− 0.7806	− 0.6578	− 0.5655	− 0.4724
$Sh_{p,JmSc}$	− 0.9167	− 0.6214	− 0.5347	− 0.4610
$Sh_{p,JmSc} - Sh_{p,NmNc}$	− 0.1362	0.0364	0.0308	0.0114
	$A = 0.5$			
$Sh_{p,NmNc}$	− 0.3542	− 0.2434	− 0.1365	− 0.0546
$Sh_{p,JmSc}$	− 0.3488	− 0.2773	− 0.1830	− 0.1095
$Sh_{p,JmSc} - Sh_{p,NmNc}$	0.0054	− 0.0339	− 0.0465	− 0.0548

Note: This table reports the Sharpe ratios of portfolio returns when the returns of both VA and MO are below their respective 30*th* percentile value. $Sh_{p,NmNc}$ are the Sharpe ratios for the NmNc strategy; $Sh_{p,JmSc}$ are the Sharpe ratios for the JmSc strategy.

Table 9. *p*-values for different transaction cost levels (required margin of 50%).

	$\eta = 1$	$\eta = 3$	$\eta = 5$	$\eta = 7$	$\eta = 9$	$\eta = 11$	$\eta = 13$	$\eta = 15$
				CRRA ($A = 1$)				
CEQ	19.57	15.58	13.44	11.98	10.75	9.75	9.01	8.42
pv 0.00%	0.27	0.29	0.30	0.11	0.03*	0.01*	0.01*	0.00*
pv 0.20%	0.30	0.44	0.39	0.17	0.08	0.04*	0.03*	0.02*
pv 0.33%	0.32	0.55	0.45	0.22	0.13	0.08	0.06	0.06
pv 0.50%	0.35	0.67	0.53	0.29	0.22	0.17	0.14	0.15
				$A = 0.9$				
CEQ	15.84	13.15	11.63	10.33	9.27	8.49	7.90	7.40
pv 0.00%	0.58	0.55	0.19	0.09	0.05*	0.02*	0.02*	0.02*
pv 0.20%	0.70	0.66	0.26	0.15	0.11	0.07	0.06	0.08
pv 0.33%	0.77	0.73	0.32	0.21	0.18	0.12	0.11	0.15
pv 0.50%	0.84	0.81	0.39	0.30	0.29	0.23	0.22	0.30
				$A = 0.8$				
CEQ	12.72	11.00	9.71	8.60	7.80	7.25	6.77	6.42
pv 0.00%	0.36	0.27	0.10	0.03*	0.01*	0.01*	0.01*	0.02*
pv 0.20%	0.51	0.37	0.17	0.09	0.05*	0.04*	0.06	0.07
pv 0.33%	0.61	0.44	0.23	0.14	0.09	0.08	0.12	0.15
pv 0.50%	0.73	0.53	0.32	0.25	0.19	0.19	0.27	0.33
				$A = 0.7$				
CEQ	9.84	8.76	7.62	6.89	6.41	5.98	5.69	5.51
pv 0.00%	0.30	0.07	0.02*	0.00*	0.00*	0.01*	0.01*	0.01*
pv 0.20%	0.41	0.13	0.05	0.02*	0.02*	0.04*	0.04*	0.05*
pv 0.33%	0.48	0.17	0.09	0.05	0.05*	0.10	0.11	0.12
pv 0.50%	0.59	0.25	0.19	0.12	0.13	0.24	0.29	0.29
				$A = 0.6$				
CEQ	7.35	6.29	5.78	5.38	5.04	4.88	4.75	4.67
pv 0.00%	0.04*	0.01*	0.00*	0.00*	0.00*	0.00*	0.01*	0.01*
pv 0.20%	0.09	0.03*	0.01*	0.01*	0.03*	0.03*	0.05*	0.05
pv 0.33%	0.13	0.07	0.03*	0.03*	0.09	0.09	0.14	0.15
pv 0.50%	0.19	0.16	0.08	0.12	0.26	0.28	0.34	0.38
				$A = 0.5$				
CEQ	5.05	4.61	4.24	4.10	4.03	4.02	3.96	3.94
pv 0.00%	0.00*	0.00*	0.02*	0.05	0.08	0.07	0.12	0.12
pv 0.20%	0.01*	0.01*	0.12	0.25	0.33	0.30	0.43	0.43
pv 0.33%	0.03*	0.03*	0.27	0.48	0.58	0.53	0.66	0.66
pv 0.50%	0.10	0.14	0.58	0.77	0.84	0.80	0.87	0.87

Note: CEQ is the annualized certainty equivalent return in %; pv *n*% corresponds to the *p*-value of a block bootstrap based on the difference (CEQ JmSc—CEQ NmNc); the percentage indicates the level of a proportional transaction cost. A star indicates a significant result at the 5% level.

this issue in table 8 by computing the Sharpe ratios of portfolio returns when VA and MO returns are jointly below their respective 30th percentile values for a portfolio margin of 50%. The table shows that the JmSc strategy produces higher Sharpe ratios than those of the benchmark when A is equal to 0.6 to 0.8. These DA intensity levels also correspond to those where we generally observe more significant results for the JmSc approach (see table 6). At higher values of A, mixed results are obtained, wherein half of the cases show higher ratios. Finally, when $A = 0.5$, a higher Sharpe ratio for the JmSc is obtained only for the lowest risk aversion level, $\eta = 1$. Again, this result is consistent with the findings of table 6 showing significant overperformance of the JmSc approach when $A = 0.5$ and $\eta \leqslant 5$.

The analysis in tables 4 and 6 reveals that the JmSc approach imposes larger turnovers. Therefore, it seems warranted to introduce transaction costs into the analysis to assess the robustness of the above results. We conduct our analysis as in DeMiguel *et al.* (2009) with the introduction of proportional transaction costs. More specifically, using the turnover measure provided by equation (9), we compute the wealth resulting from the portfolio holdings with

$$W_{t+1} = W_t \times \left(1 + r_{p,t+1}\right) \times \left(1 - \kappa \sum_{i=v,m} \left|\alpha_{i,t+1} - \alpha_{i,t}\right|\right),$$

where $W_t = 1$, κ is a proportional transaction cost parameter and $\alpha_{i,t+1}$ is the optimal portfolio weight. The portfolio return after transaction costs is then $W_{t+1}/W_t - 1$. We investigate three different levels of transaction costs: 0.20%, 0.33%, and 0.50%. The first two levels are borrowed from Lynch and Tan (2010) (based on an earlier analysis in Hasbrouck 2003). They advocate for these transaction costs for trading low and high book-to-market portfolios. The third level comes from DeMiguel *et al.* (2009). Table 9 reports the significance of the CEQ differences between the JmSc portfolio and the NmNc portfolio benchmark when A varies from 1 to 0.5 and η from 1 to 15.

We first examine the special case of CRRA preferences. The limitations of the JmSc approach when accounting for transaction costs are apparent when examining table 9. This is expected given the previous results in table 4. For $\eta \leq 9$, there is no significant result after accounting for transaction costs. We find significant results only for $\eta \geq 11$ and when transaction costs are set to 0.2%. Increasing transaction costs to 0.33% or higher eliminates any overperformance.

For a DA intensity level of $A = 0.9$, the findings reported in table 9 reveal an absence of significant overperformance when transactions costs are introduced regardless of the level of relative risk aversion. We observe different results, however, for lower values of A. For example, when $A = 0.6$, we find evidence that the JmSc approach shows signifiant overperformance when $1 < \eta < 15$ and when transaction costs are set to 0.20%. This suggests that seeking optimal portfolios with the JmSc approach is justified only for investors who substantially overweigh adverse outcomes. Furthermore, setting $\eta = 5$ or 7 when $A = 0.6$ leads the JmSc approach exhibiting significant overperformance that is robust to a 0.33% level of transaction costs in spite of the relatively high portfolio turnovers. At the highest DA intensity level, $A = 0.5$, the JmSc approach generates significant overperformance for transaction costs of 0.3% when $\eta \leq 3$. At higher levels of relative risk aversion (i.e. $\eta \geq 4$), the wealth is mostly allocated to the risk-free asset and no significant overperformance is observed after transaction costs.

Overall, accounting for transaction costs underlines the limitations of the JmSc portfolio approach, except when DA investors considerably overweigh disappointment outcomes. These results are perhaps not surprising since the JmSc approach typically displays a high portfolio turnover. In addition, we do not find any significant results when transaction costs reach the 0.5% level.

6. Concluding remarks

An important body of literature building on the studies by Barberis (2018), Ang *et al.* (2005), Gul (1991), Kahneman and Tversky (1979) challenges the use of standard expected utility models as the appropriate tool for portfolio choices. This literature suggests that portfolio decisions accounting for loss aversion and/or disappointment aversion can generate more realistic portfolio holdings. We believe that this phenomenon is important to consider for value and momentum investment strategies, whose joint return distribution displays significant departures from the Gaussian environment.

Overall, the evidence we document shows that for investors exhibiting disappointment aversion, higher levels of utility are difficult to reach by accounting for the non-linear dependencies between the value and momentum return series with non-normal multivariate densities. We observe that the most economically attractive portfolios are those showing the highest skewness coefficient, but a comparison with the Gaussian benchmarks demonstrates that taking advantage of the non-linearities and non-normalities entails higher turnovers in optimal portfolios. As a result, the statistical significance of most of the gains brought by these approaches disappears when considering transaction costs.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Funding

Lalancette acknowledges the financial support from HEC Montreal. Simonato acknowledges the financial support from the Social Sciences and Humanities Research Council of Canada (SSHRC).

ORCID

Simon Lalancette http://orcid.org/0000-0002-8319-7042

References

Ang, A., Bekaert, G. and Liu, J., Why stocks may disappoint. *J. Financ. Econ.*, 2005, **76**, 471–508.

Asness, C.S., Moskovitz, T.J. and Pedersen, L.H., Value and momentum everywhere. *J. Finance*, 2013, **68**, 929–985.

Baker, M. and Wurgler, J., Investor sentiment and the cross-section of stock returns. *J. Finance*, 2006, **61**, 1645–1680.

Barberis, N., Psychology-based models of asset prices and trading volume. *Handbook Behav. Econ.: Appl. Found.*, 2018, **1**, 79–175.

Barroso, P. and Santa Clara, P., Momentum has its moments. *J. Financ. Econ.*, 2015, **116**, 111–120.

Basistha, A. and Kurov, A., Macroeconomic cycles and the stock market's reaction to monetary policy. *J. Bank. Finance*, 2008, **32**, 2606–2616.

Bekaert, G., Hoerova, M. and Lo Duca, M., Risk, uncertainty and monetary policy. *J. Monet. Econ.*, 2013, **60**, 771–788.

Choi, S., Gale, D., Kariv, S. and Fisman, R, Consistency and heterogeneity of individual behavior under uncertainty. *Am. Econ. Rev.*, 2007, **97**, 1921–1938.

Christoffersen, P. and Langlois, H., The joint dynamics of equity market factors. *J. Financ. Quant. Anal.*, 2013, **48**, 1371–1404.

Daniel, K. and Moskowitz, T., Momentum crashes. *J. Financ. Econ.*, 2016, **122**, 221–247.

Daniel, K., Jagannathan, R. and Kim, S., Tail risk in momentum strategy returns. NBER Working Paper No. 18169, 2017.

Demarta, S. and McNeil, A., The *t* Copula and related Copulas. *Int. Statist. Rev.*, 2005, **73**, 111–129.

DeMiguel, V., Garlappi, L. and Uppal, R., Optimal versus diversification: How inefficient is the 1/*N* portfolio strategy? *Rev. Financ. Stud.*, 2009, **22**, 1915–1953.

Engle, R., Dynamic conditional correlation: A simple class of multivariate generalied autogressive conditional heteroskedasticity models. *J. Bus. Econ. Stat.*, 2002, **20**, 339–350.

Fleming, J., Kirby, C. and Ostdiek, B., The economic value of volatility timing. *J. Finance*, 2001, **56**, 329–352.

Fortin, A.-P., Simonato, J.-G. and Dionne, G., Forecasting expected shortfall: Should we use a multivariate model for stock market factors? *Int. J. Forecast.*, 2022. https://doi.org/10.1016/j.ijforecast.2021.11.010

Gandelman, N. and Hernandez-Murillo, R., Risk aversion at the country level. *Fed. Reserve Bank St-Louis Rev.*, 2015, **97**, 53–66.

Gul, F., A theory of disappointment aversion. *Econometrica*, 1991, **59**, 667–686.

Hansen, B., Autoregressive conditional density estimation. *Int. Econ. Rev.*, 1994, **35**, 705–730.

Hasbrouck, J., Intraday price formation in U.S. equity index markets. *J. Finance*, 2003, **58**, 2375–2399.
Kahneman, D. and Tversky, A., Prospect theory: An analysis of decision under risk. *Econometrica*, 1979, **47**, 263–291.
Jagannathan, R. and Ma, T., Risk reduction in large portfolios: Why imposing the wrong constraints helps. *J. Finance*, 2003, **58**, 1651–1684.
Jondeau, E. and Rockinger, M., On the importance of time variability in higher moments for asset allocation. *J. Financ. Econom.*, 2012, **10**, 84–123.
Johnson, N., Systems of frequency curves generated by methods of translation. *Biometrika*, 1949, **36**, 149–176.
Judd, K., *Numerical Methods in Economics*, 1998 (MIT Press: Cambridge, MA).
Kinlaw, W., Kritzman, M.P. and Turkington, D., *A Practitioner's Guide to Asset Allocation*, 2017 (Wiley: NJ).
Lalancette, S. and Simonato, J.G., The role of the conditional skewness and kurtosis in VIX index valuation. *Eur. Financ. Manag.*, 2017, **23**, 325–354.
Levy, H. and Markowitz, H., Approximating expected utility by a function of mean and variance. *Am. Econ. Rev.*, 1979, **69**, 308–317.
Lynch, A. and Tan, S., Multiple risky assets, transaction costs, and return predictability: Allocation rules and implications for U.S. investors. *J. Financ. Quant. Anal.*, 2010, **45**, 1015–1053.
Mehra, R. and Prescott, E.C., The equity premium: ABCs. In *Handbook of The Equity Risk Premium*, 2007 (North-Holland: Amsterdam).
Pastor, L. and Stambaugh, R., Comparing asset pricing models: An investment perspective. *J. Polit. Econ.*, 2003, **111**, 642–685.
Patton, A., On the out-of-sample importance of skewness and asymmetric dependence for asset allocation. *J. Financ. Econom.*, 2004, **2**, 130–168.
Press, W., Flannery, B., Teukolsky, S. and Vetterling, K., *Numerical Recipes in C: The Art of Scientific Computing*, 2012 (Cambridge University Press: Cambridge).
Politis, D.N. and Romano, J.P., The stationnary bootstrap. *J. Am. Stat. Assoc.*, 1994, **89**, 1303–1313.
Scott, R.C. and Horvath, P.A., On the direction of preference for moments of higher order than the variance. *J. Finance*, 1980, **35**, 915–919.
Sklar, A., Fonctions de Répartition à *n* Dimensions et Leurs Marges. *Institut Publique de Statistiques, Université Paris*, 1959, **8**, 229–231.

Appendices

Appendix 1. Johnson S_u random variables

Let z denote a standard normal random variable and consider x, a two-parameter Johnson S_u random variable defined as $x = \sinh((z - a)/b)$ where $-\infty < a < \infty$, $b > 0$, and $\sinh(u) = (\mathrm{e}^u - \mathrm{e}^{-u})/2$. The mean and variance of x are given by $M_x = -\omega^{\frac{1}{2}} \sinh(\Omega)$ and $V_x = \frac{1}{2}(\omega - 1) \times (\omega \cosh(2\Omega) + 1)$ with $\omega = \mathrm{e}^{\frac{1}{b^2}}$ and $\Omega = \frac{a}{b}$. Using the mean and variance, a standardized Johnson S_u random variable is defined as

$$\varepsilon = \frac{x - M_x}{\sqrt{V_x}}.$$

The above expression can be rewritten as

$$\varepsilon = c + d \times \sinh\left(\frac{z - a}{b}\right),$$

where $c = -M_x/\sqrt{V_x}$ and $d = 1/\sqrt{V_x}$, with $E(\varepsilon) = 0$ and $\mathrm{Var}(\varepsilon) = 1$.

The likelihood function of the time series model described in Section 2 is obtained from a Gaussian density with a change of variable. For a time-series of T observations, this log likelihood function is

$$\begin{aligned}\ln l = & -\frac{1}{2}\sum_{t=1}^{T} \ln \sigma_t^2 - \frac{T}{2} \ln 2\pi \\ & - \frac{1}{2}\sum_{t=1}^{T} \left(a + b \cdot \sinh^{-1}\left(M_x + \varepsilon_t\sqrt{V_x}\right)\right)^2 \\ & + \frac{T}{2} \ln V_x + T \ln b - \frac{1}{2}\sum_{t=1}^{T} \ln\left(\left(M_x + \varepsilon_t\sqrt{V_x}\right)^2 + 1\right).\end{aligned}$$

The standard normal residuals at the source of the Johnson errors can be recovered with

$$z = a + b \cdot \sinh^{-1}\left(M_x + \varepsilon\sqrt{V_x}\right),$$

where $\sinh^{-1}(u) = \ln(u + \sqrt{u^2 + 1})$. The validity of the Johnson S_u distribution assumption is assessed by testing whether the standard normal innovations obtained from the maximum likelihood estimation are indeed normally distributed.

Appendix 2. Numerical approaches

When we compute the integrals involved in the numerical optimization of the utility function, two different sets of contingencies appear depending on whether $A = 1$ or $A < 1$.

A.1. The CRRA case (A = 1)

For $A = 1$, the utility function is based on the standard CRRA framework, and we can rewrite the expected utility as

$$E_t\left[U(W_{t+1})\right] = \int_{\mathcal{R}_2} U(W_{t+1}) f(\mathbf{r}_{t+1})\, \mathrm{d}\mathbf{r}_{t+1}, \tag{A1}$$

where $f(\mathbf{r}_{t+1})$ is the bivariate density function associated with the next period VA and MO returns.

When the normal copula is used in conjunction with the Johnson S_u marginals, the resulting bivariate density function is known in closed form, and corresponds to the Johnson S_u distribution. Because this bivariate density is written as a transformed vector of normal random variates, the Gauss-Hermite (GH) quadrature technique can be used to compute the integrals. In short, if we denote by $h(\mathbf{r})$ the CRRA utility function (where the time subscript is dropped for notational convenience), the GH quadrature implies the following computations

$$E\left[h(\mathbf{r})\right] \simeq \frac{1}{\pi}\sum_{j=1}^{N_q}\sum_{k=1}^{N_q} h\left(\widehat{\mathbf{r}}_{j,k}\right) \times \omega_{v,j} \times \omega_{m,k},$$

where $\widehat{\mathbf{r}}_{j,k} = [\widehat{r}_{v,j}\ \widehat{r}_{m,k}]^\intercal$ are the monthly returns (discretely compounded) computed with the GH quadrature integration nodes, $\omega_{v,j}$ and $\omega_{m,k}$ are the quadrature weights associated with these nodes, and N_q corresponds to the number of quadrature nodes. Appendix 3 describes how these returns, nodes and weights are computed. Since the GH approach is specifically designed to yield efficient computations, setting $N_q = 6$ typically delivers very precise approximations.

Under the symmetric t copula combined with the Johnson S_u marginals, the resulting multivariate density does not lead to an analytical form. We use in this case a Monte Carlo simulation to compute the integrals. The symmetric t distribution has the following stochastic representation

$$\boldsymbol{u}_{t+1} = \sqrt{w}\boldsymbol{y}_{t+1}, \tag{A2}$$

where $\boldsymbol{u}_{t+1} = [u_v\ u_m]$ is a vector of probabilities of the marginal cumulative densities, w is an inverse gamma random variable with

$w \sim IG(\nu/2, \nu/2)$, $\boldsymbol{y}_{t+1}$ is a 2×1 vector of normal variables, $\boldsymbol{y}_{t+1} \sim \mathcal{N}(\mathbf{0}, \boldsymbol{\Sigma}_{t+1})$, with $\boldsymbol{y}$ and w being independent, while $\boldsymbol{\Sigma}_{t+1}$ is the estimate of the copula correlation matrix. By sequentially simulating the probabilities $\boldsymbol{u}_{t+1}$ using equation (A2), the corresponding $t+1$ VA and MO simulated returns are obtained from

$$z_{i,t+1} = N^{-1}(u_i),$$

$$\varepsilon_{i,t+1} = c_i + d_i \times \sinh\left(\frac{z_{i,t} - a_i}{b_i}\right),$$

and

$$R_{i,t+1} = \mu_i + \sigma_{i,t+1}\varepsilon_{i,t+1},$$

for $i = \{v, m\}$. Using 10,000 independent sample paths, we approximate $E[h(\mathbf{r})]$ with a sample average using the returns $r_{i,t+1} = \exp(R_{i,t+1}) - 1$ for $i = \{v, m\}$.

A.2. *The disappointment aversion case* ($A < 1$)

For the case where $A < 1$, it is convenient to rewrite the DA utility function (6) with indicator variables as

$$U(\mu_W(\boldsymbol{\alpha})) = \frac{1}{K}\int_{\mathcal{R}_2} h(\mathbf{r}_{t+1}) f(\mathbf{r}_{t+1})\, \mathrm{d}\mathbf{r}_{t+1}, \tag{A3}$$

with

$$h(\mathbf{r}_{t+1}) = U(r_{p,t+1})\left(\mathbf{1}_{]-\infty,\mu_W]} + A\mathbf{1}_{[\mu_W,\infty[}\right), \tag{A4}$$

and

$$K = \Pr(W_{t+1} \le \mu_W(\boldsymbol{\alpha})) + A\Pr(W_{t+1} > \mu_W(\boldsymbol{\alpha})), \tag{A5}$$

and where $\mathbf{1}_{[\cdot]}$ is an indicator variable equal to 1 if the $t+1$ wealth is in the interval, and 0 otherwise.

When the normal copula is used in conjunction with the Johnson S_u marginals, we use a quadrature approach to solve the integral in equation (A4) based on Simpson nodes and weights. As mentioned in Ang *et al.* (2005), the Gauss-Hermite approach produces a grid of nodes with uneven spacing between the points, which makes it ill-suited for identifying the certainty equivalent parameter $\mu_W(\boldsymbol{\alpha})$ under a numerical search. Simpson nodes and weights provide instead a uniform grid that allows the use of a bisection search to obtain $\mu_W(\boldsymbol{\alpha})$. In this setting, the double integral can be approximated with

$$U(\mu_W(\boldsymbol{\alpha})) \simeq \frac{1}{\pi}\sum_{j=1}^{N_q}\sum_{k=1}^{N_q} h\left(\widehat{\mathbf{r}}_{j,k}\right) \times \omega_{v,j} \times \omega_{m,k},$$

and

$$K \simeq \frac{1}{\pi}\sum_{j=1}^{N_q}\sum_{k=1}^{N_q}\left(\mathbf{1}_{]-\infty,\mu_W]} + A\mathbf{1}_{[\mu_W,\infty[}\right) \times \omega_{v,j} \times \omega_{m,k},$$

Appendix 3 provides the details on how the Simpson nodes, weights and returns are computed. We set herein $N_q = 51$, a bisection search with a convergence criterion of 1×10^{-12}, and a bracketed search between 0.8 and 1.2.

Under the symmetric t copula combined with the Johnson S_u marginals, we use a simulation procedure similar to the one described in the CRRA case to generate simulated portfolio returns. Equations (A3) to (A5) are then solved using their sample counterparts. We perform 10,000 sample paths, a bisection search with a convergence criterion of 1×10^{-12}, and a bracketed search between 0.8 and 1.2.

Appendix 3. Quadrature approaches

We provide in this section a description of the bivariate quadratures based on Gauss-Hermite and Simpson. Equations (1) and (2) define a continuous one-to-one mapping between the returns r_i and z_i, the Gaussian innovations at the source of the Johnson S_u innovations. This invertible mapping can be expressed generically as

$$\mathbf{r} = g(\mathbf{z}) \leftrightarrow \mathbf{z} = g^{-1}(\mathbf{r}),$$

where $\mathbf{r} = [r_v\ r_m]^\intercal$ and $\mathbf{z} = [z_v\ z_m]^\intercal$ are the 2×1 vectors of returns and Gaussian innovations. The portfolio allocation procedure in the CRRA context requires one to compute the expected value of a utility function that depends on the random vector $\mathbf{r}$. This computation, for a generic and continuous function $h(\mathbf{r})$, is written as

$$E\{h(\mathbf{r})\} = \int_{\mathcal{R}_2} h(\mathbf{r}) f(\mathbf{r})\, \mathrm{d}\mathbf{r},$$

where $f(\mathbf{r})$ is the multivariate density function of the returns. Using the explicit form of the Johnson S_u returns and the multivariate change of variable $\mathbf{r} = g(\mathbf{z})$, the expected value is written as

$$E\{h(\mathbf{r})\} = \frac{\det(\mathbf{C}_\mathbf{z})^{-1/2}}{2\pi}\int_{\mathcal{R}_2} h(g(\mathbf{z})) \times \mathrm{e}^{-\frac{1}{2}\mathbf{z}'\mathbf{C}_\mathbf{z}^{-1}\mathbf{z}}\, \mathrm{d}\mathbf{z},$$

with $\mathbf{C}_\mathbf{z}$ the 2×2 correlation matrix associated with the z's. As outlined by Judd (1998), to facilitate the use of quadrature methods, the above expression is transformed with the change of variable $\mathbf{y} = \boldsymbol{\Psi}^{-1}\mathbf{z}/\sqrt{2}$ where $\mathbf{C}_\mathbf{z} = \boldsymbol{\Psi}\boldsymbol{\Psi}'$ and $\mathbf{C}_\mathbf{z}^{-1} = (\boldsymbol{\Psi}^{-1})'\boldsymbol{\Psi}^{-1}$, which yields after some simplifications

$$E\{h(\mathbf{r})\} = \frac{1}{\pi}\int_{\mathcal{R}_2} h\left(g\left(\sqrt{2}\boldsymbol{\Psi}\mathbf{y}\right)\right)\mathrm{e}^{-\sum_{i=1}^{2} y_i^2}\, \mathrm{d}\mathbf{y}.$$

The Gauss-Hermite quadrature approach uses discrete sets of nodes and weights $\{y_{v,j}, \omega_{v,j}\}_{j=1}^{N_q}$ and $\{y_{m,k}, \omega_{m,k}\}_{k=1}^{N_q}$ which are designed to yield precise approximations of the above integral with a finite sum. In our applications, we use the algorithm described by Press *et al.* (2012) to compute the nodes and weights. Defining $\widehat{\mathbf{z}}_{j,k} = \sqrt{2}\boldsymbol{\Psi}\mathbf{y}_{j,k}$ with $\mathbf{y}_{j,k} = [y_{v,j}\ y_{m,k}]^\intercal$, the expected value is approximated with†

$$E\{h(\mathbf{r})\} \simeq \frac{1}{\pi}\sum_{j=1}^{N_q}\sum_{k=1}^{N_q} h\left(g\left(\widehat{\mathbf{z}}_{j,k}\right)\right) \times \omega_{v,j} \times \omega_{m,k},$$

where $g(\cdot)$ is the mapping between the Johnson S_u noise and the returns. Expressed in terms of returns, the approximation is

$$E\{h(\mathbf{r})\} \simeq \frac{1}{\pi}\sum_{j=1}^{N_q}\sum_{k=1}^{N_q} h\left(\widehat{\mathbf{r}}_{j,k}\right) \times \omega_{v,j} \times \omega_{m,k},$$

where the ith element of the 2×1 vector $\widehat{\mathbf{r}}_{j,k}$ is computed as $\widehat{\mathbf{r}}_{j,k}(i) = \exp\left(\widehat{R}_i\right) - 1$ with

$$\widehat{R}_i = \varphi_i + \sigma_i\left(c_i + d_i \times \sinh\left(\frac{\widehat{\mathbf{z}}_{j,k}(i) - a_i}{b_i}\right)\right) \tag{A6}$$

for $i = \{v, m\}$. If normal innovations are considered instead of Johnson S_u innovations, the term in front of σ_i in equation (A6) is replaced by $\widehat{\mathbf{z}}_{j,k}(i)$.

† We use a hat to distinguish between the continuous random variable and the discrete set of values used by the quadrature computing scheme.

As explained earlier, to solve the integrals underlying the DA ($A < 1$) environment, we use Simpson nodes and weights instead of their GH counterparts (see for example Press *et al.* 2012). They are computed as follows. We define a uniformly spaced grid over an interval that should be sufficient to cover the realizations of a normal standard variable. For this purpose, we use the interval $[-5, +5]$ with N_q discrete values. Denote the elements of this grid by y_j for $j = 1$ to Nq where Δy represents the difference between the equidistant points. For each point on the grid, we compute the corresponding weights as

$$\omega_{i,j} = \begin{cases} \Delta y/3 \times \exp(-y_j^2) & \text{if } j = 1 \text{ or } N_q, \\ 4\Delta y/3 \times \exp(-y_j^2) & \text{if } j \text{ is even}, \\ 2\Delta y/3 \times \exp(-y_j^2) & \text{if } j \text{ is odd}, \end{cases}$$

for $i = \{v, m\}$. Using these nodes and weights, the returns are computed as above.