

Stenšin, Andrej; Bloznelis, Daumantas

## Article

# Copulas and portfolios in the electric vehicle sector

Journal of Risk and Financial Management

### Provided in Cooperation with:

MDPI – Multidisciplinary Digital Publishing Institute, Basel

*Suggested Citation:* Stenšin, Andrej; Bloznelis, Daumantas (2022) : Copulas and portfolios in the electric vehicle sector, Journal of Risk and Financial Management, ISSN 1911-8074, MDPI, Basel, Vol. 15, Iss. 3, pp. 1-20,  
<https://doi.org/10.3390/jrfm15030132>

This Version is available at:

<https://hdl.handle.net/10419/258855>

#### Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

#### Terms of use:

*Documents in EconStor may be saved and copied for your personal and scholarly purposes.*

*You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.*

*If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.*



<https://creativecommons.org/licenses/by/4.0/>

Article

# Copulas and Portfolios in the Electric Vehicle Sector

Andrej Stenšin \* and Daumantas Bloznelis 

School of Economics and Business, Norwegian University of Life Sciences, Christian Magnus Falsens Road 18, 1433 Aas, Norway; daumantas.bloznelis@nmbu.no

\* Correspondence: andrej.stensin@nmbu.no

**Abstract:** How can investors unlock the returns on the electric vehicle industry? Available investment choices range from individual stocks to exchange traded funds. We select six representative assets and characterize the time-varying joint distribution of their returns by copula-GARCH models. They facilitate portfolio optimization targeted at a chosen combination of risk and reward. With daily data from 2012 to 2020, we illustrate the models' applicability by building a minimum expected shortfall portfolio and comparing its performance to that of an equally weighted benchmark. Our results should be of interest to investors and risk managers seeking or facing exposure to the electric vehicle sector.

**Keywords:** portfolio optimization; electric vehicles; copula; vines; expected shortfall; value-at-risk

**JEL Classification:** G11; G17; C58; Q02



**Citation:** Stenšin, Andrej, and Daumantas Bloznelis. 2022. Copulas and Portfolios in the Electric Vehicle Sector. *Journal of Risk and Financial Management* 15: 132. <https://doi.org/10.3390/jrfm15030132>

Academic Editors: Aviral Kumar Tiwari and Wing-Keung Wong

Received: 21 January 2022

Accepted: 4 March 2022

Published: 10 March 2022

**Publisher's Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Copyright:** © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

The world's advanced economies have been experiencing slow growth and low interest rates ever since the Great Recession, and the recent COVID-19 pandemic has turned the economic outlook increasingly uncertain. In times of stalling economies and cheap money, investors are hard pressed to find attractive investment opportunities. Looking at growth cases and breakthrough industries, the electric vehicle (EV) sector is an emerging success story. Subsidized by national governments, electric vehicles are about to change the world's automotive landscape. The switch from internal combustion engines to electric motors is set to reduce carbon emissions and fuel imports and improve air quality. Global sales of new electric vehicles reached an estimated 2.3 million units in 2020, and the EV share has risen from 0.1% in 2011 to an estimated 3.2% in 2020 (see [IEA \(2020\)](#) and [Hertzke et al. \(2019\)](#)). While the total number of EVs is still small in comparison with new gasoline and diesel cars, the electric share of the total sales has been rapidly increasing and is projected to continue on an upward path ([Woodward et al. 2020](#)). The sector has produced the world's most valuable car manufacturer, Tesla ([FT 2020](#)) and the world's richest man, Elon Musk ([BBC News 2021](#)). However, while the returns on EV-related assets have been high, so has volatility. Tesla's stock had 5-year monthly *beta* of 2.15, which is considerably higher than the *betas* of other growing tech giants such as Alphabet (0.99), Apple (1.30) or Microsoft (0.83) and is something that not all investors will find acceptable.

How do we unlock the returns on the electric vehicle industry without taking on excessive risk? Should one invest in shares of companies that are producing electric cars or their components? Or should one purchase exchange-traded funds (ETFs) exposed to the industry? These questions about risk and reward relate to optimal portfolio construction. By virtue of diversification, portfolios tend to offer more attractive risk-reward characteristics than individual assets. A portfolio may be optimized to produce minimal risk for a given target return or maximum return for a given target risk. Building an optimal portfolio and characterizing the distribution of its returns requires a statistically adequate model of the

joint distribution of the component assets. The groundwork of model building is the focus of this paper.

We select three stocks and three exchange-traded funds exposed to the EV industry. We model the joint distribution of their daily returns using copula GARCH models estimated iteratively over a sequence of rolling windows. We first estimate marginal models for each asset's returns and then a model for the dependence between the assets. For the latter, we try out two different copula approaches. The first approach, called traditional copula, considers different six-variate copulas and selects the best-fitting one. The second approach, called vine copula, models the multivariate dependence using bivariate copulas as building blocks (Aas et al. 2009; Bedford and Cooke 2001, 2002). Vine copula allows for different bivariate copulas for each block, providing considerable flexibility in constructing the joint distribution. Using daily data from February 2012 to January 2020, we carry out a pseudo out-of-sample evaluation of the models' statistical adequacy. We assess the fit of the marginal distributions as well as the distributions of two portfolios, where one is equally weighted and another minimizes the expected shortfall subject to a desired expected return. We also present graphically how the dependence structure between the assets varies over time. Our findings suggest the vine copula GARCH model with Johnson's  $S_U$  marginal distributions is statistically adequate and superior to alternative modeling choices. The corresponding portfolios are well behaved, and the one minimizing the expected shortfall delivers a reduction in risk without giving up the expected return on an equally weighted portfolio (barring transaction costs).

The main contribution of the paper is a statistically sound joint distribution model of selected asset returns from the EV sector. The model can be used for obtaining return distributions of arbitrary portfolios, e.g., for the purpose of risk assessment. It also facilitates portfolio optimization with respect to different optimality criteria, as illustrated by the expected shortfall example. In addition, the paper documents the inadequacy of the Normal distribution and traditional elliptical and Archimedean copulas for modelling asset returns in the EV sector. This should be of interest to investors seeking attractive rewards from future-proof industries who are not comfortable taking on the risk of investing in a single asset. It may also be useful for risk assessors and managers in the EV sector, while the modeling framework is transferable to other sectors as well.

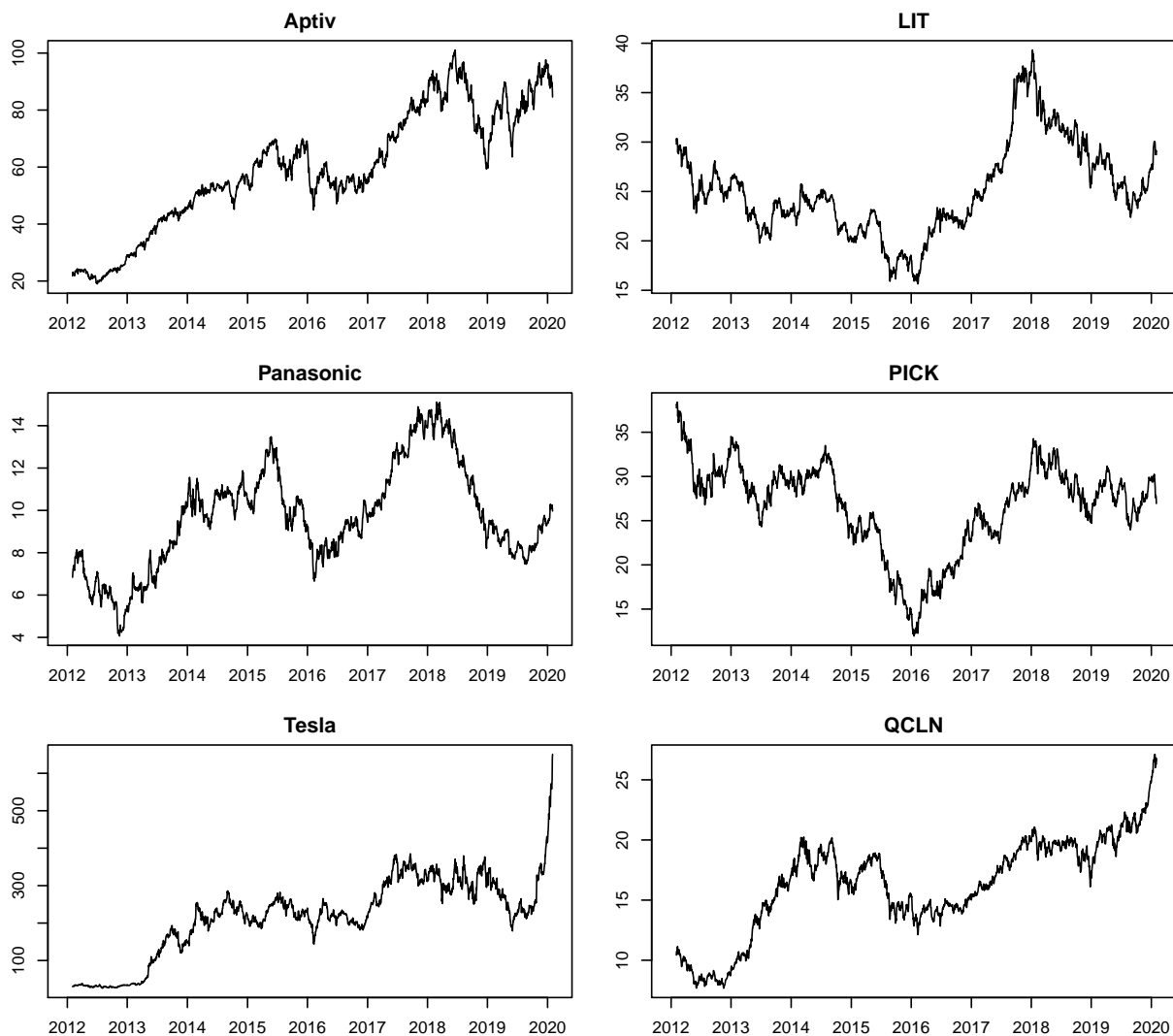
The remainder of the paper is organized as follows. Section 2 presents data and descriptive statistics. Section 3 specifies the models, their estimation and performance evaluation. Empirical results are presented in Section 4, and Section 5 summarizes and concludes.

## 2. Data

The EV industry is young, and there are hardly any companies in the sector that have been listed on a stock exchange for a nontrivial amount of time, with Tesla being a notable exception. The lack of historical data on the industry's asset returns poses a challenge for empirical analysis. However, there are a few exchange-listed firms, ETFs and commodities that are broadly related to the EV industry. Some firms manufacture cars while others provide their components and raw materials. Several ETFs facilitate investments in a variety of companies exposed to the EV industry. We explore three firms and three ETFs (with tickers of exchanges on which they are traded given in parentheses): Tesla (NASDAQ), Panasonic (OTCMKTS) and Aptiv (NYSE); and PICK (BATS), LIT (NYSEARCA) and QCLN (NASDAQ), respectively. Tesla, Inc., became the world's largest automaker by market capitalization (Forbes 2020) and was the world's best-selling plug-in electric passenger car manufacturer in 2020 (Statista 2021b). Tesla Model 3 was by far the world's best-selling plug-in electric car in 2020 (Statista 2021a). Panasonic is the main battery supplier for Tesla, with a dedicated Automotive and Industrial Systems division providing electrical components, lithium-ion and automotive batteries, dry batteries and electric motors. Aptiv (APTV) is a company supplying automakers with power solutions, mostly charging ports, high voltage connectors, shielding and sealing. iShares MSCI Global Select Metals & Mining Producers (PICK) is an ETF that consists of an international basket of companies

engaged in extraction and production of metals, including copper and cobalt, which are used for electric batteries. Global X Lithium (LIT) is an ETF that tracks companies engaged in the lithium industry and battery supply. First Trust NASDAQ Clean Edge Green Energy Index Fund (QCLN) is an ETF that invests in companies engaged in several green energy subsectors and also has Tesla among its holdings.<sup>1</sup>

Figure 1 depicts daily prices of Aptiv, Panasonic and Tesla as well as LIT, PICK and QCLN. The prices are obtained from Yahoo Finance and quoted in U.S. dollars.



**Figure 1.** Asset prices in U.S. dollars; daily data from 2 February 2012 to 31 January 2020.

Our sample contains daily data from 2 February 2012 to 31 January 2020, a total of 2011 observations. The start of the sample period is determined by data availability and the commercialization of mass production of electric cars. The end is selected to exclude the COVID-19 pandemic that was just starting at the time of data collection. Logarithmic price returns (abbreviated to *returns* and denoted by  $r_t$ ) are computed as  $r_t = \ln(P_t/P_{t-1})$ , where  $P_t$  and  $P_{t-1}$  are current and lagged daily prices, respectively.

Table 1 provides descriptive statistics for the six logarithmic return series.

**Table 1.** Descriptive statistics of logarithmic price returns.

|                 | Aptiv      | Panasonic  | Tesla      | LIT        | PICK       | QCLN       |
|-----------------|------------|------------|------------|------------|------------|------------|
| Mean            | 0.064      | 0.016      | 0.158      | −0.001     | −0.022     | 0.047      |
| Std. dev.       | 1.746      | 1.847      | 3.181      | 1.363      | 1.624      | 1.460      |
| Min             | −13.032    | −10.137    | −18.845    | −6.908     | −9.015     | −5.959     |
| Max             | 10.356     | 11.520     | 21.829     | 8.260      | 6.544      | 6.239      |
| Skewness        | 0.283      | 0.083      | 0.264      | −0.146     | −0.147     | −0.288     |
| Excess kurtosis | 3.569      | 3.080      | 5.705      | 2.146      | 1.596      | 0.919      |
| Jarque-Bera     | 1109.5 *** | 808.44 *** | 2787.6 *** | 399.14 *** | 224.29 *** | 100.37 *** |
| ARCH-LM(15)     | 42.129 *** | 118.69 *** | 132.14 *** | 156.79 *** | 141.95 *** | 151.00 *** |

Note: The sample spans 2011 daily observations from 2 February 2012 to 31 January 2020. Jarque–Bera is the statistic of [Jarque and Bera \(1980\)](#) test for normality. ARCH-LM is the statistic of [Engle \(1982\)](#) Lagrange multiplier test for autoregressive conditional heteroskedasticity computed using 15 lags. For the latter two tests, asterisks indicate significance at 1% (\*\*\*), 5% (\*\*) and 10% (\*) level.

Average returns are close to zero for all assets, and the corresponding standard deviations are larger by an order of one to three magnitudes, as is common for daily data. The skewness of returns is positive for stocks and negative for ETFs. All returns show high excess kurtosis, ranging from 0.919 for QCLN to 5.705 for Tesla, indicating fat tails of the returns’ distributions. Both skewness and excess kurtosis suggest that each return series is not normally distributed. Accordingly, the Jarque–Bera test ([Jarque and Bera 1980](#)) strongly rejects normality of returns for all assets. Finally, the autoregressive conditional heteroskedasticity Lagrange multiplier (ARCH-LM) test of [Engle \(1982\)](#) indicates the presence of ARCH effects in all assets’ returns.

### 3. Theoretical Framework and Methods

This section explains copula-based joint distribution models of returns on multiple assets. It elaborates on model selection, estimation and out-of-sample evaluation of statistical adequacy. It also outlines portfolio optimization targeted at minimizing the expected shortfall subject to a desired level of expected return.

#### 3.1. Marginal Models

We start by considering the anticipated characteristics of the marginal distributions of the return series. Daily asset returns are known to exhibit volatility clustering, skewness and both unconditionally and conditionally fat tails ([Cont 2001](#)). [Bollerslev \(1986\)](#) proposed the generalized autoregressive conditional heteroskedasticity (GARCH) model that accounts for unconditionally fat tails and volatility clustering. The GARCH(1,1) model for the logarithmic return  $r_t = \ln(P_t/P_{t-1})$  is specified as follows.

$$r_t = \mu_t + \varepsilon_t, \tag{1}$$

$$\varepsilon_t = \sigma_t z_t, \tag{2}$$

$$\sigma_t^2 = \omega + \alpha \varepsilon_{t-1}^2 + \beta \sigma_{t-1}^2, \tag{3}$$

$$z_t \stackrel{i.i.d.}{\sim} \text{JSU}(\gamma, \delta, \lambda, \xi) \text{ with } \mathbb{E}(z_t) = 0 \text{ and } \text{Var}(z_t) = 1. \tag{4}$$

Equation (1) additively decomposes the return at time  $t$  into a deterministic conditional mean,  $\mu_t$ , and a stochastic residual (error, shock and innovation),  $\varepsilon_t$ . In our case, the conditional mean is modeled as a constant,  $\mu_t = \mu$ . Equation (2) multiplicatively decomposes the stochastic residual into a conditional standard deviation,  $\sigma_t$ , and an independently and identically distributed (i.i.d.) standardized innovation,  $z_t$ . Equation (3) represents the structure of the conditional variance,  $\sigma_t^2$ , due to the GARCH specification. There,  $\omega$  is a constant,  $\varepsilon_{t-1}^2$  is the squared residual from the previous period and  $\sigma_{t-1}^2$  is the conditional variance from the previous period. To reflect the skewness and kurtosis of the standardized innovations,  $z_t$ , we employ Johnson’s S<sub>U</sub> (JSU) distribution ([Johnson 1949](#)) in Equation (4). The density function is given by the following.

$$f(z; \gamma, \delta, \lambda, \xi) = \frac{\delta}{\lambda \sqrt{2\pi} \sqrt{1 + \left(\frac{z - \xi}{\lambda}\right)^2}} \exp \left[ -\frac{1}{2} \left( \gamma + \delta \sinh^{-1} \left( \frac{z - \xi}{\lambda} \right) \right)^2 \right]. \quad (5)$$

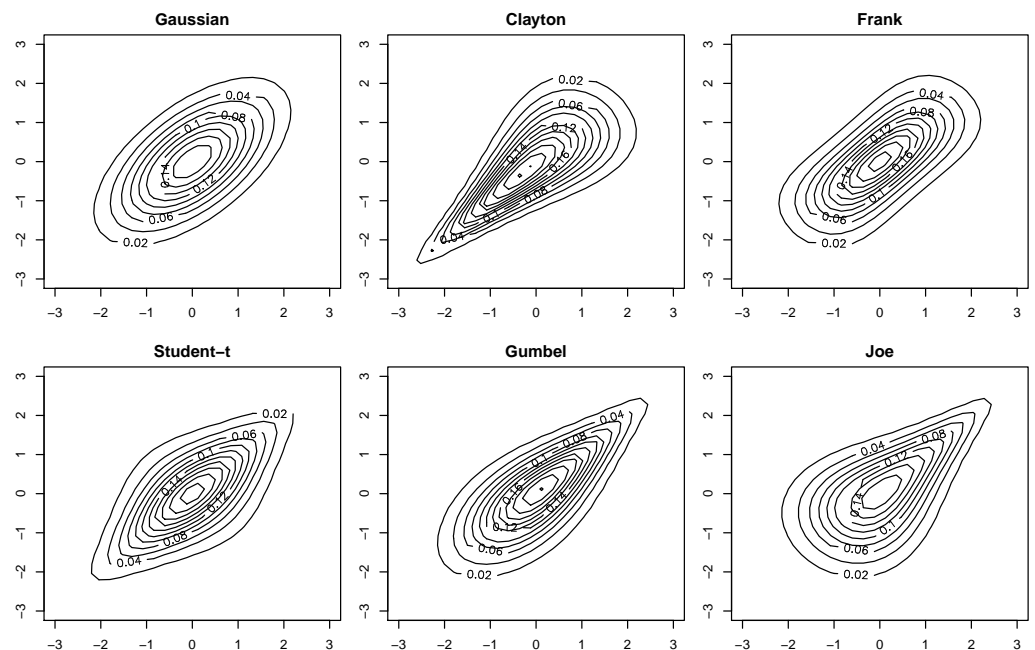
$\xi$  and  $\lambda > 0$  are the location and scale parameters, respectively. They are set to ensure zero mean and unit variance due to Equation (4), yielding a standardized version of Johnson’s  $S_U$  distribution.  $\gamma$  determines the skewness with  $\gamma > 0$  indicating positive and  $\gamma < 0$  negative skewness.  $\delta > 0$  determines the kurtosis.

In addition to this GARCH(1,1)-JSU model, we also consider the more basic and widespread GARCH(1,1)-Normal model that will serve as our benchmark. (Alternative model specifications were considered as well but did not yield statistically adequate results and will not be reported.)

### 3.2. Copula Models

According to Sklar’s theorem, any multivariate distribution can be decomposed into corresponding univariate marginal distributions and a dependence function called copula (Sklar 1959). Conversely, a copula function enables linking separate marginal distributions into a joint one. As such, copulas provide a large degree of flexibility in the specification of multivariate distributions.

Among other properties, copulas allow for a variety of dependence structures for the tails of the marginal distributions. They facilitate an explicit characterization of the lower and upper dependence in the orthants<sup>2</sup> of the joint distribution. Lower-tail dependence describes the relative frequency of co-occurrence of events from the left tails of the marginal distributions, while upper-tail dependence concerns the right tails. This property is important in finance, because simultaneous negative returns on multiple assets are of particular interest for risk and portfolio management. Contour plots in Figure 2 illustrate tail dependence in the bivariate case for selected copulas with standard normal marginal distributions.

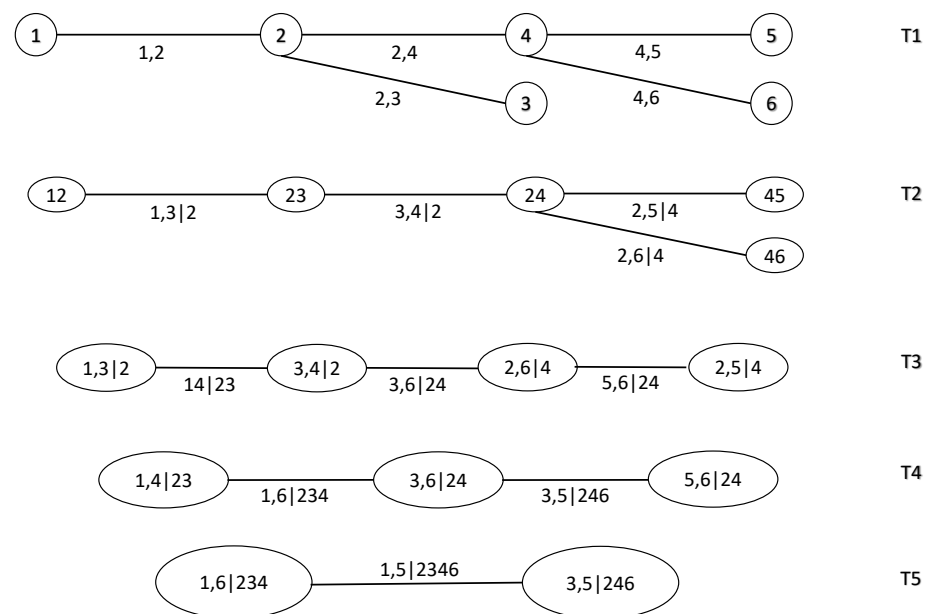


**Figure 2.** Contour plots of bivariate copula densities. Elliptical copulas in the left column: Gaussian ( $\rho = 0.6$ ) and Student- $t$  ( $\rho = 0.6, df = 3$ ). Bivariate Archimedean copulas in the middle and right columns: Clayton, Frank ( $\alpha = 5$ ), Gumbel and Joe ( $\alpha = 2$ ). All marginal distributions are standard normal.

Given the marginal distribution models from Section 3.1, we can employ two different copula approaches to allow for a variety of dependence structures among the return series. The traditional copula approach considers 6-variate copulas and selects the best one based on Akaike’s information criterion, AIC (Akaike 1973; Breymann et al. 2003; Rodriguez 2007), thus defining the multivariate distribution. The vine copula approach selects the best bivariate copula model for pairs of the marginal distributions and proceeds to connect the pairs until multivariate distribution has been fully determined. Traditional copulas are more restrictive and, thus, less prone to overfitting and easier to interpret. Vine copulas add more flexibility and may accommodate multivariate dependence structures more precisely. We present vine copulas next, while the traditional copulas are covered in Appendix B.

### 3.2.1. Vine Copulas

Building on the works of Joe (1996); Bedford and Cooke (2001, 2002); and Kurowicka and Cooke (2006); Aas et al. (2009) popularized a wide and flexible class of multivariate copula models called vine copulas; see Czado (2019) for an accessible introduction. The idea is to break a  $k$ -variate copula into  $K = k(k - 1)/2$  bivariate copulas or pair-copulas. It could be considered an adaptation of the well-known chain rule, or product rule, that allows factoring a multivariate probability density function into conditional and unconditional univariate densities. The vine copula breakdown replaces the conditional univariate densities in the chain rule by unconditional and conditional bivariate densities corresponding to pair-copulas.  $k - 1$  of these model pairs of random variables unconditionally, while the remaining ones characterize such pairs conditionally on one or more of the other random variables. The collection of pair-copulas is organized hierarchically so that it encodes the  $k$ -variate copula similarly to how the chain rule encodes the multivariate probability density function as a product of multiple conditional densities and an unconditional density. The organized collection is known as an R-vine and can be represented by a nested set of trees. The first tree,  $T_1$ , contains  $k$  original variables as nodes and  $k - 1$  unconditional pair-copulas as edges. The second tree  $T_2$  contains  $k - 1$  pairs of variables (corresponding to the edges of  $T_1$ ) as nodes and  $k - 2$  conditional pair-copulas as edges, where each pair-copula is conditioned on a single variable. The higher-order trees  $T_j$  contain  $j$  pairs of variables (corresponding to the edges of  $T_{j-1}$ ) as nodes and  $j - 1$  conditional pair-copulas, each conditioned on  $j - 1$  variables, as edges. For an illustration, see Figure 3; alternatively, consult Figure 1 in Dissmann et al. (2013) or Figure 5.3 in Czado (2019, p. 101).



**Figure 3.** An R-vine with 6 variables, producing 5 trees and 15 edges. Each node represents a variable or a pair of variables, and each edge represents a pair-copula.

Figure 3 shows a graphical model corresponding to a six-dimensional R-vine distribution. It consists of five trees:  $T_1$  to  $T_5$ . Tree  $T_j$  is made up of  $7 - j$  nodes and  $6 - j$  edges. The edges correspond to pair-copulas, while the edge labels indicate their subscripts, e.g., edge 1,4|23 corresponds to the copula density  $c_{1,4|23}$ . The entire structure is defined by  $6(6 - 1)/2 = 15$  pair copulas (edges) and the six marginal densities of the individual variables.

### 3.2.2. Model Building and Selection with Vine Copulas

The R-vine construction of multivariate copulas offers an enormous amount of flexibility. First, there is a large number of methods to break down any  $k$ -variate copula into pair-copulas; e.g., for  $k = 6$ , we have 23,040 possible breakdowns (Czado 2019, p. 159). Second, for a given R vine structure (a nested set of trees), each of the  $K$  pair-copulas can be selected independently. If there are  $m$  bivariate copulas to choose from, this produces  $m^K$  possible pair-copula combinations and, consequently,  $m^K$  corresponding multivariate copulas. Each of them can be specified further by setting parameter values of the pair-copulas. At this level of flexibility, it does not hurt much to impose a simplifying assumption for regular vine distributions, namely, that the conditional copula is independent of the conditioning variables (Dissmann et al. 2013; and Remark 5.12 in Czado 2019, p. 103).

Given the amount of ways to build a  $k$ -variate distribution with R-vines, selecting a vine copula for a set of variables is a formidable task. In the six-dimensional problem in our focus, the number of possibilities is too great for a brute-force full subset selection. Considering each possible combination, estimating its parameters, evaluating its fit and choosing the best among all is computationally infeasible. Instead, we employ Algorithm 3.1 of Dissmann et al. (2013), as implemented in the `rvinecopulib` package (Nagler and Vatter 2019) of the R statistical computing environment (R Core Team 2020).

This is a greedy algorithm containing  $k - 1$  stages corresponding to the  $k - 1$  trees of the R-vine. Each stage contains three steps: (1) selection of pairs to be connected by pair-copulas; (2) estimation of multiple candidate bivariate copulas for each pair; and (3) selection among them. In step (1), pairs are chosen by finding the strongest pairwise associations between variables, as measured by Kendall's rank correlation coefficient or Kendall's  $\tau$  (Kendall 1938). Step (2) proceeds in a brute-force manner, proceeding through a large number of candidate bivariate copulas. (The full list of candidates in our application can be found in the documentation of the `rvinecopulib` package (Nagler and Vatter 2019).) Step (3) selects the copulas that minimize the AIC (Akaike 1973) independently for each pair.

In stage  $s = 1$ , the algorithm connects individual variables into pairs by choosing the strongest pairwise associations, until a minimum spanning tree  $T_1$  is formed. For each pairwise connection, all possible bivariate copulas are estimated and the one maximizing AIC is chosen. In stage  $s = 2$ , the algorithm finds all pairs of variables (edges of  $T_1$ ) that share a variable (a node). For each such variable, there will be at least two pairs. If there only are two pairs, the algorithm conditions on the variable estimates all possible pair-copulas and selects the one minimizing AIC. If there are more than two pairs, it finds the pair of pairs that has the strongest relationship with respect to Kendall's  $\tau$  and then proceeds as in the case with only two pairs. All pairs of variables from  $T_1$  are, thus, connected by conditional pair-copulas, and a minimum spanning tree  $T_2$  is formed. The nodes of  $T_2$  are the edges of  $T_1$ , while the edges of  $T_2$  are the new conditional pair-copulas. In higher stages  $s > 2$ , the algorithm carries on in a similar fashion. It expands the conditioning set to an increasing number of variables ( $s - 1$  conditioning variables in stage  $s$ ) and joins the pairs (the edges of the tree  $T_{s-1}$ ) with either common conditioning sets or common sets of conditioning and regular variables, preferring the connections with the strongest pairwise association if there is a choice. The newly formed connections (conditional pair-copulas) become the edges of a minimum spanning tree  $T_s$ . The algorithm concludes at stage  $s = k - 1$ , where a final pair-copula joins the remaining two pairs of variables conditioned on all the other variables.



For a more detailed, formal description, see Algorithm 3.1 of [Dissmann et al. \(2013\)](#) or the Dissmann algorithm in Section 8.3 of [Czado \(2019\)](#), pp. 159–63).

While alternative methods of building and selecting an R-vine copula model have been proposed in the literature (e.g., [Kurowicka and Cooke 2006](#) as cited in [Dissmann et al. 2013](#)), the one described above has several statistical advantages, as elaborated upon in [Dissmann et al. \(2013\)](#). Moreover, it has a convenient and fast implementation in the `rvinecopulib` package ([Nagler and Vatter 2019](#)) of R, making it our preferred choice.

### 3.3. Joint Models

Bringing the marginal models and the copulas together, we consider three classes of joint distribution models of the six assets' returns. The first class uses GARCH(1,1)-Normal marginals and the Normal copula. This is equivalent to marginal GARCH(1,1) models combined with the assumption of multivariate normality of the standardized innovations. Being conceptually and computationally simple, models of this class will serve as our benchmark. The second class builds on GARCH(1,1)-JSU marginals using the Student- $t$  copula. The third class combines the GARCH(1,1)-JSU marginals with vine copulas and, thus, contains highly flexible joint distribution models. Taken together, the three classes span a range of complexity, starting from some of the simplest multivariate joint distributions and extending to ones of considerable flexibility. The parameters of the first two joint models are estimated using a procedure called inference for the margins (IFM) ([Joe and Xu 1996](#)). The last model is estimated using a sequential procedure implemented in the `rvinecopulib` package ([Nagler and Vatter 2019](#)).

### 3.4. Rolling-Window Portfolio Optimization with Expected Shortfall

The distributional characteristics of assets and portfolios that are of special interest to investors are the expected returns as well as measures of risk, e.g., variance, value at risk (VaR) and expected shortfall (ES). Consider an asset or a portfolio with return  $r$  modeled as a realization of a random variable  $R$ . The  $\alpha$ -level value at risk,  $\text{VaR}_\alpha(R)$ , is defined as the negative of the  $\alpha$ -level quantile of the return distribution,  $F_R$ :  $\text{VaR}_\alpha(R) = -F_R^{-1}(\alpha)$ . Similarly, the  $\alpha$ -level expected shortfall,  $\text{ES}_\alpha(R)$ , is defined as the negative of the expected value of all returns that are lower (thus less desirable) than the negative of  $\text{VaR}_\alpha(R)$ :  $\text{ES}_\alpha(R) = -\frac{1}{\alpha} \int_0^\alpha F_R^{-1}(\tilde{\alpha}) d\tilde{\alpha}$ .

Given the estimated joint distribution models, we can construct portfolios of assets targeting a desired balance between risk and return. There are at least three ways of formulating the investor's problem. First, it can be maximization of a linear combination of expected return (with a positive weight) and a risk measure (with a weight reflecting the undesirability of risk); a representative example is the mean-variance optimization based on the Modern Portfolio Theory of [Markowitz \(1952\)](#). Second, one may wish to maximize expected return subject to a constraint on risk, e.g., requiring that variance does not exceed a given number. Third, one may seek to minimize risk subject to achieving a desired level of expected return. It is the latter approach that fits an investor eyeing the expected return characteristic to the EV sector but wishing to minimize risk. We thus proceed with the third formulation using 2.5%-level expected shortfall as our risk measure, the choice being motivated by Basel Accords ([Basel Committee 2016, 2017](#)). Specifically, we target the expected return on an equally weighted portfolio of the six EV-related assets. We are, thus, not giving up any amount of the expected return (barring transaction costs) that an agent would obtain from naively investing in the assets in equal proportions. Subject to this constraint and avoiding short sales, we then minimize the 2.5%-level expected shortfall.

Mathematically, we have six assets with their return vector  $r_t = (r_{1,t}, \dots, r_{6,t})^\top$  on day  $t$  modeled as a realization of a random vector  $R_t = (R_{1,t}, \dots, R_{6,t})^\top$ . We consider portfolios that can be obtained by weighting the assets with a weight vector  $w_t = (w_{1,t}, \dots, w_{6,t})^\top$  that is non-negative (ruling out short sales), i.e.,  $w_i > 0$  for each  $i$ , and where the weights sum to unity,  $\sum_{i=1}^6 w_i = 1$ . The return on such a portfolio,  $\pi_t(w_t) = w_t^\top r_t$ , is thus modeled as a realization of a random variable  $\Pi_t(w_t) = w_t^\top R_t$ . An example is the equally weighted

portfolio with weights  $\bar{w} = (\frac{1}{6}, \dots, \frac{1}{6})^\top$  and a return  $\pi_t(\bar{w}) = \frac{1}{6} \sum_{i=1}^6 r_{i,t}$ . In each time period  $t$ , we solve the following:

$$ES_{\alpha,t}(\Pi_{t+1}(w_{t+1})) \rightarrow \min_{w_{t+1}}$$

subject to the following three conditions:

1. Target expected return:  $\mathbb{E}_t(\Pi_{t+1}(w_{t+1})) = \mathbb{E}_t(\Pi_{t+1}(\bar{w}))$ ;
2. No short sales:  $w_{i,t+1} \geq 0 \forall i = 1, \dots, 6$ ;
3. Normalization of weights:  $\sum_{i=1}^6 w_{i,t+1} = 1$ .

$\mathbb{E}_t$  above is an expectation conditional on the information available up to (and including) time  $t$ . Since the true conditional expectation is not observable, we replace it with its estimate from the copula-GARCH model. The true expected shortfall (conditional on the same information)  $ES_{\alpha,t}(\cdot)$  is also unknowable, and it is also replaced by its empirical counterpart; for details, see step 6 of the algorithm below.

We borrow the optimization algorithm for R from Yollin (2009). The routine relies on simulating a large number of multivariate returns from which the optimal portfolio weights are obtained by numerical optimization. We embed this in a loop of rolling windows in which we first select and estimate the models and then optimize the portfolio based on the one-day-ahead forecast return distribution. This results in the following algorithm (repeated in a rolling window fashion):

1. For each asset, estimate a marginal model of its returns;
2. Estimate a copula model and obtain a one-day-ahead density forecast of the multivariate joint distribution of the asset returns;
3. Simulate a large sample of multivariate returns from the forecast density;
4. Apply numerical optimization of the given criterion (i.e., minimize ES for a given expected value) on the simulated returns from step 3 to obtain optimal portfolio weights;
5. Multiply the simulated returns from step 3 by the weights from step 4 to obtain simulated portfolio returns.
6. Estimate the portfolio VaR and ES nonparametrically: (i)  $\widehat{VaR}_{\alpha,t}(\Pi_{t+1}(w_{t+1})) = -\hat{Q}_8(\alpha)$  where  $\hat{Q}_8(\alpha)$  is the empirical quantile defined in Hyndman and Fan (1996) and used as the default option in the `quantile` function in R; (ii)  $\widehat{ES}_{\alpha,t}(\Pi_{t+1}(w_{t+1})) = -\frac{1}{n_i} \sum_{i: \pi_i \leq \hat{Q}_8(\alpha)} \pi_i$  where  $\pi_i$  are the simulated portfolio returns and  $n_i$  is the number of them satisfying the condition  $\pi_i \leq \hat{Q}_8(\alpha)$ .

In each window, the algorithm produces out-of-sample estimates of one-day-ahead joint distributions and their derived features (estimated conditional expectations and VaR and ES values). The models' performance across the windows can then be assessed by comparing the out-of-sample estimates with the corresponding actual (realized) values.

There are 1012 rolling windows, with 1000 observations each. Thus, every training sample contains 1000 observations, while the evaluation sample is 1011 observations long. (One observation is lost as we are forecasting one step ahead and do not have the realized values corresponding to the last window's forecasts.) To emphasize, the models are evaluated entirely out of sample; we do not track any in-sample fit diagnostics but only out-of-sample forecast quality.

### 3.5. Performance Evaluation

#### 3.5.1. Statistical Adequacy

Adequate specification of marginal models is crucial for modeling the joint distribution with copulas, since a copula model is a function of the marginal distributions. As noted by Patton (2006), a misspecified marginal model leads to a misspecified copula model. To evaluate the marginal models' goodness of fit, we follow the logic of Diebold et al. (1998). If the marginal distributions are correctly specified, then the probability integral transform of the returns on each asset should be i.i.d. Uniform(0,1). We use the estimated marginal models to obtain one-step-ahead marginal distributions and employ them to produce

probability integral transforms of the realized one-step-ahead observations. We then perform the Kolmogorov–Smirnov (KS) test to assess whether they indeed are Uniform(0,1) and, thus, evaluate the one-step-ahead predictive adequacy of the marginal models across the rolling windows.

Assessing the general adequacy of joint distributions via nonparametric (e.g., Kolmogorov–Smirnov type) tests in dimensions as high as ours (six) ranges between impractical and impossible. Therefore, we resort to assessing distributions of weighted averages of the individual variables instead, and these distributions are implied by the joint distribution models. In practical terms, we construct portfolios of assets and assess the adequacy of the portfolios' distribution models in the same way as of the individual assets. This is relevant for practitioners, as only the distribution of portfolio returns matters for a holder of a given portfolio or for choosing among given alternative portfolios.

To test the adequacy of VaR estimates, we implement both the unconditional Kupiec (1995) and the conditional Christoffersen (1998) coverage tests as well as the Christoffersen and Pelletier (2004) duration test. The unconditional coverage test assesses whether the proportion of  $\text{VaR}_\alpha$  violations is indeed  $\alpha$ . The conditional coverage test additionally probes the independence of violations over time. Similarly, the duration test evaluates whether the durations of time between VaR violations are independent and do not cluster.

For the adequacy of ES estimates, we first use the test of McNeil and Frey (2000). Under the null hypothesis of correctly specified ES estimates, the returns on occasions when VaR is violated are i.i.d. and have the estimated ES values as their means. It is the latter observation that forms the basis of the test for unbiasedness of the ES estimates. We also implement three new tests of Bayer and Dimitriadis (2020). The tests are derived from joint linear models of VaR and ES, but the null hypotheses explicitly concern the ES model only. The auxiliary expected shortfall regression (A-ESR) and the strict expected shortfall regression (S-ESR) tests are two-sided and, thus, sensitive to both overestimation and underestimation of ES. The difference between them is that the A-ESR requires VaR estimates in addition to ES estimates as inputs, while S-ESR only requires ES estimates. In this regard, the S-ESR test is unique among currently available ES tests. The intercept-expected shortfall regression (I-ESR) test is a restricted version of the S-ESR test and allows for both two-sided and one-sided hypotheses. In the case of one-sided hypothesis, the alternative of interest is that the ES estimates are too liberal, i.e., the estimated losses are too small. Bayer and Dimitriadis (2020) show the three tests to have superior size and power properties compared to McNeil and Frey (2000) and other ES tests.

### 3.5.2. Financial Performance

For an investor, two implications of our models are of primary interest. First, it is the risk-reward characteristics of the portfolios. We expect the optimally weighted portfolio to have lower risk but about the same return as the equally weighted portfolio (barring transaction costs). Second, it is the portfolio weights and their development over time. Since implementing the optimally weighted portfolio requires knowledge of the weights, their average values as well as variation around the averages are instrumental to portfolio managers. We will analyze both aspects of the financial performance in the next section.

## 4. Empirical Results

### 4.1. Marginal Models

We start by reporting the performance of the univariate GARCH(1,1)-JSU models and the GARCH(1,1)-Normal benchmarks estimated in 1012 rolling windows of 1000 consecutive observations each. As explained in the previous section, their statistical adequacy is assessed entirely out of sample. The results are presented in Table 2.

**Table 2.** Statistical adequacy of marginal models.

| Normal                   | Aptiv | Panasonic | Tesla | LIT   | PICK  | QCLN  |
|--------------------------|-------|-----------|-------|-------|-------|-------|
| KS                       | 0.006 | 0.002     | 0.000 | 0.004 | 0.003 | 0.003 |
| VaR unc.                 | 0.010 | 0.884     | 0.355 | 0.193 | 0.265 | 0.137 |
| VaR cond.                | 0.008 | 0.390     | 0.120 | 0.428 | 0.208 | 0.235 |
| VaR dur.                 | 0.699 | 0.672     | 0.683 | 0.772 | 0.898 | 0.829 |
| ES unb.                  | 0.026 | 0.020     | 0.000 | 0.029 | 0.090 | 0.009 |
| S-ESR                    | 0.234 | 0.218     | 0.052 | 0.006 | 0.353 | 0.011 |
| A-ESR                    | 0.234 | 0.197     | 0.055 | 0.006 | 0.341 | 0.020 |
| I-ESR-1s                 | 0.050 | 0.023     | 0.008 | 0.002 | 0.053 | 0.006 |
| I-ESR-2s                 | 0.101 | 0.047     | 0.017 | 0.004 | 0.105 | 0.012 |
| Johnson’s S <sub>U</sub> | Aptiv | Panasonic | Tesla | LIT   | PICK  | QCLN  |
| KS                       | 0.174 | 0.194     | 0.417 | 0.315 | 0.142 | 0.704 |
| VaR unc.                 | 0.193 | 0.642     | 0.956 | 0.270 | 0.642 | 0.500 |
| VaR cond.                | 0.150 | 0.749     | 0.074 | 0.387 | 0.525 | 0.488 |
| VaR dur.                 | 0.263 | 0.703     | 0.609 | 0.895 | 0.692 | 0.211 |
| ES unb.                  | 0.900 | 0.194     | 0.100 | 0.171 | 0.398 | 0.544 |
| S-ESR                    | 0.648 | 0.307     | 0.045 | 0.336 | 0.677 | 0.000 |
| A-ESR                    | 0.650 | 0.343     | 0.044 | 0.347 | 0.645 | 0.001 |
| I-ESR-1s                 | 0.535 | 0.473     | 0.581 | 0.206 | 0.346 | 0.717 |
| I-ESR-2s                 | 0.931 | 0.946     | 0.838 | 0.413 | 0.692 | 0.565 |

Note: The table contains the *p*-values of statistical adequacy tests for univariate GARCH(1,1) marginal models with standardized residuals from Normal and Johnson’s S<sub>U</sub> distributions. VaR and ES are calculated for  $\alpha = 2.5\%$  level. KS stands for Kolmogorov–Smirnov test, VaR unc. for VaR unconditional coverage test of Kupiec (1995), VaR cond. for VaR conditional coverage test of Christoffersen (1998), VaR dur. for VaR duration test of Christoffersen and Pelletier (2004), ES unb. for ES unbiasedness test of McNeil and Frey (2000), S-ESR for strict expected shortfall regression test of Bayer and Dimitriadis (2020), A-ESR for auxiliary expected shortfall regression test (ibid.), and I-ESR-1s and I-ESR-2s for one-sided and two-sided intercept expected shortfall regression tests (ibid.), respectively.

First, consider the Kolmogorov–Smirnov test of the null hypothesis that the probability integral transform of the marginal distribution of an asset’s returns is Uniform(0,1). The GARCH(1,1)-JSU models pass the test at the conventional 5% significance level for all six assets, while GARCH(1,1)-Normal benchmarks fail for each asset even at the generous 1% significance level. Second, examine the Kupiec (1995), Christoffersen (1998) and Christoffersen and Pelletier (2004) tests for the adequacy of VaR estimates. The tests do not produce evidence against the GARCH(1,1)-JSU models for any asset, while GARCH(1,1)-Normal models fail for APTIV in two cases. Third, focus on the McNeil and Frey (2000) and the Bayer and Dimitriadis (2020) tests for evaluating ES estimates. The GARCH(1,1)-JSU models pass all the tests for each asset, with the exception of Tesla and QCLN, which fail the S-ESR and A-ESR tests.<sup>3</sup> Meanwhile, the GARCH(1,1)-Normal models fail in all cases but one (PICK, although it passes some of the tests only marginally). Overall, we find rather limited evidence against the statistical adequacy of the GARCH(1,1)-JSU marginal models and, therefore, feel reasonably confident in proceeding to building copula-based joint distribution models with them. Meanwhile, the GARCH(1,1)-Normal benchmarks suffer from multiple test violations for each asset, making them a poorly justified basis for subsequent copula modeling. Nevertheless, we do continue with them too, as we aim to illustrate the performance of the basic models alongside the complex ones in the context of modeling the joint distributions.

#### 4.2. Copula-Based Portfolios

The results of the statistical adequacy tests for the portfolios are presented in Table 3.

The benchmark class of joint distribution models based on GARCH(1,1)-Normal marginals and a Normal copula fails all of the adequacy tests (except for the VaR duration test) for both the optimally weighted and the equally weighted portfolios. This comes as no surprise, given that the marginal GARCH(1,1)-Normal models had failed a number of tests for each individual asset. Thus, the simple structure of GARCH(1,1) combined with multivariate normality of standardized innovations is too restrictive for our data.

**Table 3.** Statistical adequacy of portfolio models.

|                    | Normal–Normal | Student–Johnson’s S <sub>U</sub> | Vine–Johnson’s S <sub>U</sub> |
|--------------------|---------------|----------------------------------|-------------------------------|
| <b>Panel A: EW</b> |               |                                  |                               |
| KS                 | 0.043         | 0.111                            | 0.441                         |
| VaR unc.           | 0.027         | 0.193                            | 0.589                         |
| VaR cond.          | 0.075         | 0.286                            | 0.838                         |
| VaR dur.           | 0.351         | 0.313                            | 0.466                         |
| ES unb.            | 0.007         | 0.629                            | 0.733                         |
| S-ESR              | 0.017         | 0.767                            | 0.985                         |
| A-ESR              | 0.016         | 0.774                            | 0.969                         |
| I-ESR-1s           | 0.002         | 0.237                            | 0.499                         |
| I-ESR-2s           | 0.004         | 0.474                            | 0.999                         |
| <b>Panel B: OW</b> |               |                                  |                               |
| KS                 | 0.046         | 0.236                            | 0.469                         |
| VaR unc.           | 0.004         | 0.137                            | 0.095                         |
| VaR cond.          | 0.014         | 0.235                            | 0.245                         |
| VaR dur.           | 0.799         | 0.843                            | 0.656                         |
| ES unb.            | 0.005         | 0.186                            | 0.232                         |
| S-ESR              | 0.018         | 0.352                            | 0.358                         |
| A-ESR              | 0.021         | 0.400                            | 0.349                         |
| I-ESR-1s           | 0.002         | 0.051                            | 0.071                         |
| I-ESR-2s           | 0.003         | 0.102                            | 0.142                         |

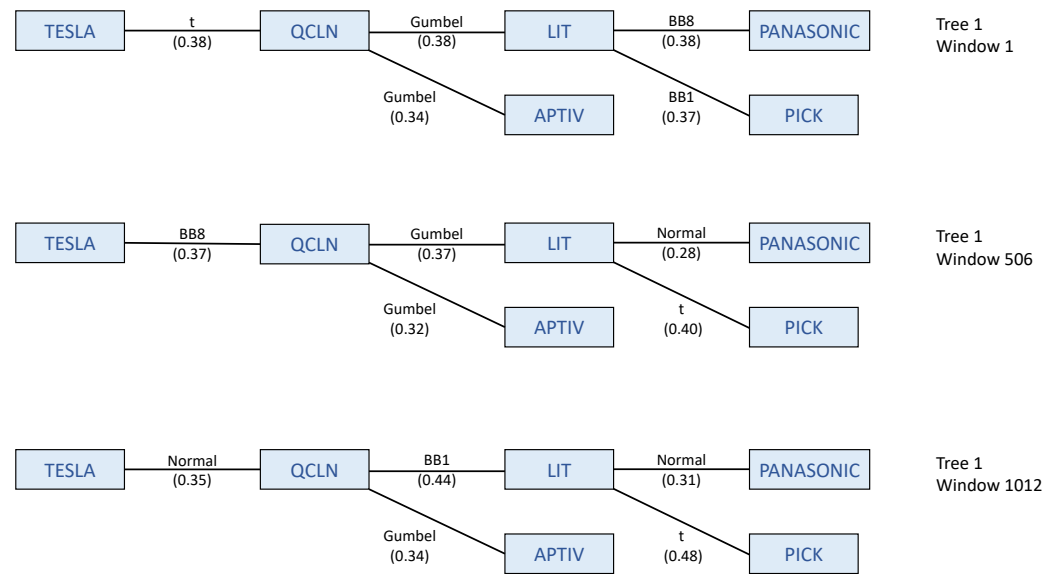
Note: The table contains the *p*-values of statistical adequacy tests for equally weighted (EW) and optimally weighted (OW) portfolios. VaR and ES are calculated for  $\alpha = 2.5\%$  level. KS stands for Kolmogorov–Smirnov test, VaR unc. for VaR unconditional coverage test of Kupiec (1995), VaR cond. for VaR conditional coverage test of Christoffersen (1998), VaR dur. for VaR duration test of Christoffersen and Pelletier (2004), ES unb. for ES unbiasedness test of McNeil and Frey (2000), S-ESR for strict expected shortfall regression test of Bayer and Dimitriadis (2020), A-ESR for auxiliary expected shortfall regression test (ibid.), and I-ESR-1s and I-ESR-2s for one-sided and two-sided intercept expected shortfall regression tests (ibid.), respectively.

The second class of joint distribution models that uses GARCH(1,1)-JSU marginals and Student-*t* copula passes all of the tests both for the optimally weighted and equally weighted portfolios. Hence, flexible marginals combined with a rather simple copula structure are sufficient for implying statistically adequate distributions of selected linear combinations of the individual random variables. In other words, we do not find evidence against a conjecture that the second class of joint distribution models offers a valid description of the equally as well as optimally weighted portfolios.

While the relatively parsimonious Student-*t* copula is sufficient, we obtain even better results from the flexible vine copulas combined with the same GARCH(1,1)-JSU marginals. The joint distribution models of the third class pass the majority of the diagnostic tests with higher *p*-values than do models of the second class. This indicates the superiority of vine copulas over the Student-*t* copula in predicting returns one step ahead, in terms of both the entire portfolio distribution (as evidenced by the Kolmogorov–Smirnov test) and the risk measures (as evidenced by the majority of VaR and ES tests). Therefore, GARCH(1,1)-JSU marginals combined with vine copulas can be considered the most statistically adequate attempt at modeling the joint distribution of the six assets’ returns.

#### 4.2.1. R-Vine Structure

To dig deeper into the superior performance of vine copulas, we look at their R-vines. While vine copulas are a rather intricate branch of copula modeling, their tree structures yield themselves easily to visual examination and interpretation. Figure 4 shows the first tree,  $T_1$ , for three windows (1st, 506th and 1012th) representing the beginning, the middle and the end of the sample, respectively. We focus on  $T_1$  alone, as the first tree is often the most important with regards to model fit and contains the strongest pairwise dependencies among all trees (Dissmann et al. 2013).



**Figure 4.** Trees  $T_1$  of R-vines from windows number 1 (first), 506 (middle) and 1012 (last). Note: Numbers in parentheses are Kendall’s rank correlation coefficients (Kendall’s  $\tau$ s).

One immediately notices the node-edge combination to be the same in all three windows. The dependence between the assets varies over time, but the variation is small. Most pairs are connected by elliptical copulas, while the pairs QCLN-LIT and QCLN-Aptiv are linked by the Gumbel copula, one that implies positive dependence between the upper tails of the marginal distributions. Kendall’s  $\tau$ s are low for all pairs, varying between 0.28 and 0.48, and indicate the dependencies are not particularly strong.

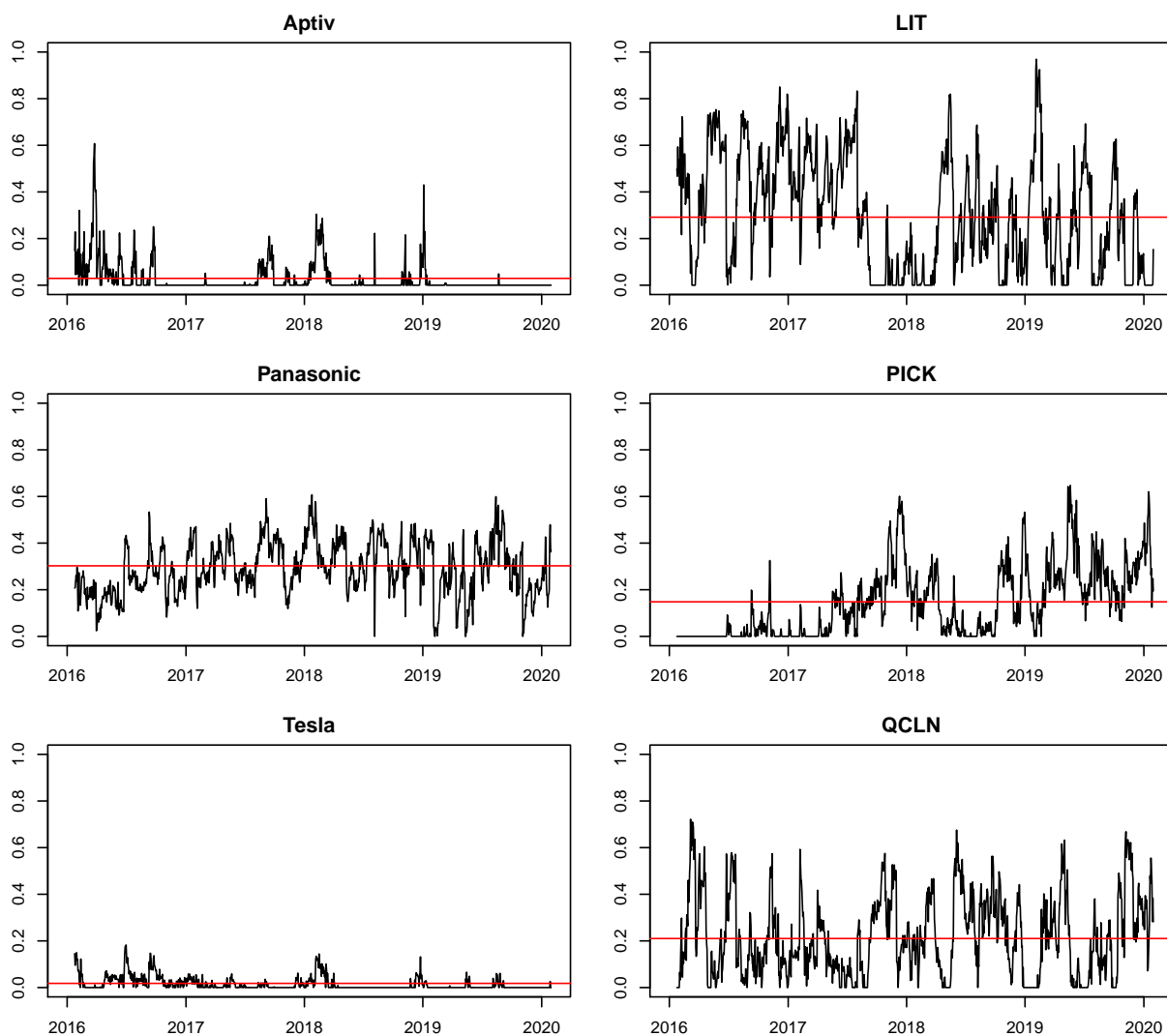
There is also a financial logic to the trees. First, the three ETFs (LIT, PICK and QCLN) are connected. Second, QCLN that invests in green energy and contains some of Tesla’s shares is linked to Tesla itself and to Aptiv, the power solution supplier. Third, LIT that tracks companies from lithium and battery industries shares an edge with Panasonic that manufactures lithium-ion batteries, among other products. Thus, the superior statistical adequacy of the vine copula models might be due to their capturing of genuine, explainable dependencies between the businesses represented by the assets.

#### 4.2.2. Optimally vs. Equally Weighted Portfolios

Having compared the three classes of joint distribution models and observing that two of them are statistically adequate, we proceed to a statistical comparison of equally vs. optimally weighted portfolios. For both the second and the third classes of models, the equally weighted portfolio delivers (much) higher  $p$ -values from the majority of diagnostic tests, indicating a better fit. The exceptions are the VaR duration test and the Kolmogorov–Smirnov test where the differences are not particularly large. Notably, models of the third class deliver excellent results for the equally weighted portfolio. We do not have a hypothesis for why the equally weighted portfolio should be easier to model than its optimally weighted counterpart; a deeper analysis would be required to cast light on the issue but is beyond the scope of the study.

#### 4.3. Financial Assessment of Portfolios

From the financial perspective, we are interested in portfolio weights and their development over time as well as the risk and return on equally vs. optimally weighted portfolios. Regarding the weights, Figure 5 presents them for the third class of the joint distribution models, i.e., GARCH(1,1)-JSU marginals with vine copulas. (Weights of the second class of models are not presented but are qualitatively similar.)

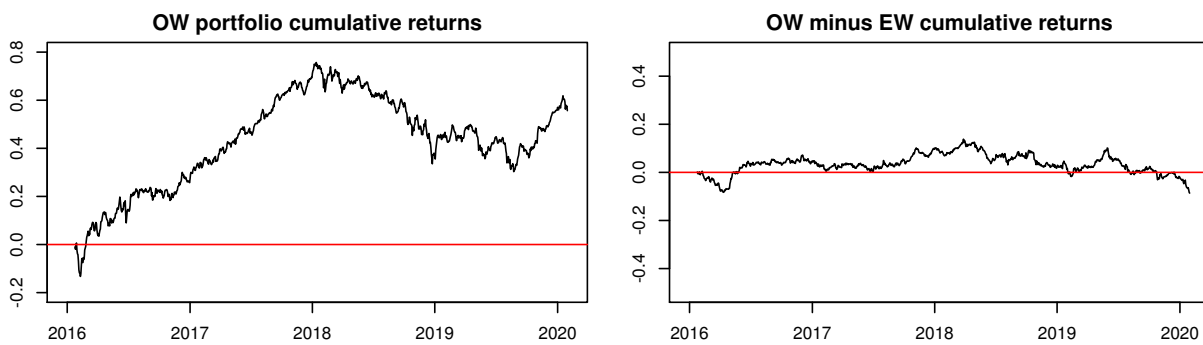


**Figure 5.** Optimal portfolio weights for GARCH(1,1)-JSU marginals and vine copulas. Note: The horizontal lines are the average weights over the test subsample.

The weights are rather volatile across the rolling windows for most assets.<sup>4</sup> Two exceptions are Tesla and, to a smaller extent, Aptiv. Tesla receives persistently low weights averaging around 0.02, which can be explained by the high risk-reward ratio, as it is a stock with exceptionally high volatility. Aptiv has somewhat similar characteristics to Tesla and receives an average weight of 0.03. The high volatility of both stocks is unsurprising, as these are relatively undiversified companies, unlike the Panasonic Corporation and the three ETFs. Offering lower volatility, the latter four are the main constituents of the optimally weighted portfolio for the majority of the time periods. The leaders with about 0.3 average weights each are Panasonic and LIT, followed by QCLN and PICK with average weights of about 0.2 and 0.15, respectively.

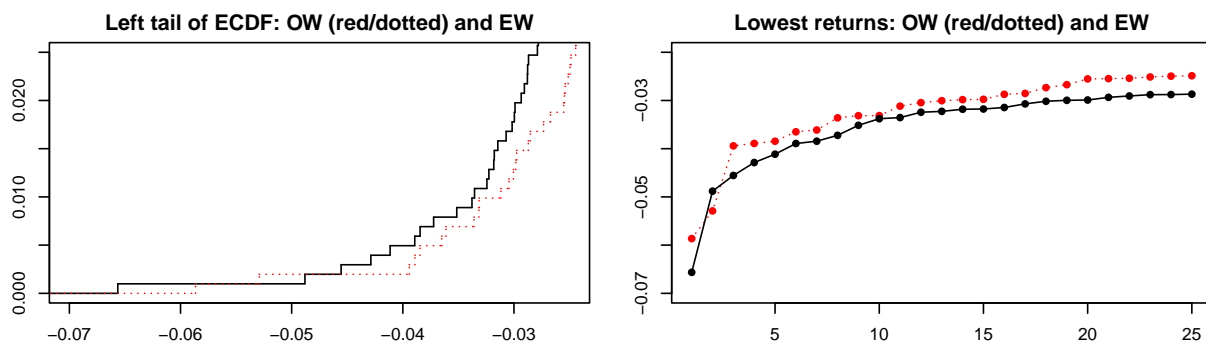
The cumulative returns on the optimally weighted portfolio and the difference between the cumulative returns on the two portfolios are shown in Figure 6.

The portfolios track each other closely. Barring transaction costs, the mean returns on the equally and the optimally weighted portfolios are about the same at 16.4% and 14.4% annually.<sup>5</sup> This is as expected given the optimization criterion of the optimally weighted portfolio; its expected return is targeted at the expected return on the equally weighted portfolio.



**Figure 6.** Cumulative returns on the optimally weighted (OW) portfolio on the left. Difference between the cumulative returns on the OW and the equally weighted (EW) portfolio on the right. (Transaction costs not accounted for).

Furthermore, one may wish to confirm that the optimally weighted portfolio has lower realized shortfall (less extreme left-tail returns) than the equally weighted one. This is indeed the case, as the means of their respective 2.5% lowest returns are  $-3.26\%$  and  $-3.54\%$ , the former being higher than the latter. Figure 7 shows the left tails of the two portfolios’ return distributions. The left panel depicts the two empirical cumulative distribution functions (ECDFs). With a single exception, the ECDF of the optimally weighted portfolio lies below its equally weighted counterpart, indicating lower risk. The same insight can be gained from inspecting the right panel that shows the lowest returns from each of the portfolios. Again, with a single exception, the returns on the optimally weighted portfolio lie above these on the equally weighted one.



**Figure 7. (Left panel):** Left tails of the empirical cumulative distribution functions of the optimally weighted (OW) and equally weighted (EW) portfolios; OW in red/dotted and EW in black/solid. **(Right panel):** Lowest 2.5% of returns from the two portfolios; OW in red/dotted and EW in black/solid; the horizontal axis marks ranks.

However, is the difference between the two realized shortfalls statistically significant? We are not aware of any formal tests to warrant a decision at a desired confidence level. However, should the optimally weighted portfolio not have smaller expected shortfall than its equally weighted counterpart, we would certainly not expect to see 24 out of the 25 lowest returns pointing in the present direction. We view this as highly suggestive evidence for genuine risk reduction by the optimally weighted portfolio.

### 5. Conclusions

The electric vehicle sector has gripped investor attention with its attractive future prospects and recent success stories such as Tesla. However, the sector is plagued by high uncertainty. The future pace of technological development is hard to predict, and governmental policies regarding green economy and climate change can be highly unstable,



as exemplified by the sharp changes of course in the U.S. over the past few years. A question relevant to many is as follows: how do we capture the sector’s attractive returns?

We have proposed a portfolio approach based on individual stocks and exchange traded funds all of which are related to the electric vehicle sector. We noted that, to build a portfolio with a desired balance between risk and return, one first needs to model the joint distribution of the asset returns. Employing a flexible approach of copulas, we found a statistically adequate class of joint distribution models. It combines vine copulas with GARCH(1,1) marginals and Johnson’s  $S_U$ -distributed standardized innovations. It passed a battery of out-of-sample goodness-of-fit tests for both the entire distribution in general and the left tails in particular, thereby indicating the model’s adequacy and its superiority to simpler multivariate models.

To exemplify the applicability of our results, we produced out-of-sample density, value at risk and expected shortfall estimates of equally and optimally weighted portfolios for the test period of 2016 to early 2020. The optimal weights of a mean-expected shortfall portfolio favored the well-diversified exchange traded funds and shares of a diversified corporation, with undiversified companies such as Tesla receiving minimal weights. (A more aggressive criterion of portfolio optimality would likely have increased the weights of the riskier, less diversified assets.) Our model, thus, provides a relevant foundation for investors and portfolio managers pondering upon entrance into, or already holding assets from, the electric vehicle sector.

The paper has a limitation that is typical of studies of young industries. Only a small number of representative assets are considered, as only they had a price history that is long enough for fitting sophisticated joint distribution models. To reflect the sector’s characteristics more comprehensively, a larger collection of publicly traded assets would be desirable. As they enter financial exchanges over time, one may take a new look and form a more representative view of the matter. This remains a research avenue for future studies.

**Author Contributions:** Conceptualization, A.S. and D.B.; methodology, A.S. and D.B.; data acquisition, A.S. and D.B.; computer programming, A.S. and D.B.; writing—original draft preparation, A.S. and D.B.; visualization, A.S. and D.B. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research received no external funding.

**Data Availability Statement:** The data presented in this study are openly available in OPEN ICPSR at <https://doi.org/10.3886/E139981V1>, reference number 139981.

**Conflicts of Interest:** The authors declare no conflict of interest.

## Appendix A. Exchange Traded Funds

**Table A1.** iShares MSCI Global Select Metals & Mining Producers (PICK) holdings.

| Holding                   | Ticker     | Weighting |
|---------------------------|------------|-----------|
| BHP Billiton Ltd          | BHP:AUX    | 10.84%    |
| Rio Tinto PLC             | RIO:LSE    | 8.78%     |
| Vale SA                   | VALE3:SAO  | 5.44%     |
| Glencore PLC              | GLEN:LSE   | 4.66%     |
| Anglo American PLC        | AAL:LSE    | 4.00%     |
| GMK Noril’skiy Nikel’ PAO | GMKN:MCX   | 3.22%     |
| Freeport-McMoRan Inc      | FCX        | 2.51%     |
| Nucor Corp                | NUE        | 2.09%     |
| Posco                     | 005490:KSC | 2.03%     |
| ArcelorMittal SA          | MT:AEX     | 1.59%     |

Note: Top 10 out of 207 holdings (45.17% of total weight) as of 19 February 2020. Source: <http://etfdb.com/etf/PICK/> (accessed on 20 February 2020).

**Table A2.** Global X Lithium (LIT) holdings.

| Holding                               | Ticker     | Weighting |
|---------------------------------------|------------|-----------|
| Albemarle Corp                        | ALB        | 19.78%    |
| Sociedad Quimica y Minera de Chile SA | SQM        | 12.74%    |
| Tesla Inc                             | TSLA       | 12.53%    |
| Livent Corp                           | LTHM       | 5.86%     |
| Samsung SDI Co Ltd                    | 006400:KSC | 5.41%     |
| LG Chem Ltd                           | 051910:KSC | 4.73%     |
| Panasonic Corp                        | 6752:TYO   | 4.56%     |
| Byd Co Ltd                            | 1211:HKG   | 4.51%     |
| Simplo Technology Co Ltd              | 6121:TWO   | 4.16%     |
| EnerSys                               | ENS        | 4.09%     |

Note: Top 10 out of 38 holdings (78.37% of total weight) as of 19 February 2020. Source: <http://etfdb.com/etf/LIT/> (accessed on 20 February 2020).

**Table A3.** First Trust NASDAQ Clean Edge Green Energy Index Fund (QCLN) holdings.

| Holding                          | Ticker | Weighting |
|----------------------------------|--------|-----------|
| Tesla Inc                        | TSLA   | 16.54%    |
| Brookfield Renewable Partners LP | BEP.UN | 5.97%     |
| Albemarle Corp                   | ALB    | 5.53%     |
| ON Semiconductor Corp            | ON     | 5.01%     |
| Universal Display Corp           | OLED   | 4.92%     |
| Enphase Energy Inc               | ENPH   | 4.24%     |
| Solaredge Technologies Inc       | SEDG   | 3.76%     |
| Cree Inc                         | CREE   | 3.14%     |
| First Solar Inc                  | FSLR   | 3.09%     |
| TerraForm Power Inc              | TERP   | 2.93%     |

Note: Top 10 out of 43 holdings (55.16% of total weight) as of 19 February 2020. Source: <http://etfdb.com/etf/QCLN/> (accessed on 20 February 2020).

## Appendix B. Traditional Copulas

For the traditional approach, we consider six copulas: Gaussian, Student-*t*, Clayton, Frank, Gumbel and Joe. They belong to two popular families: elliptical and Archimedean.

### Appendix B.1. Elliptical Copulas

An elliptical copula is a copula corresponding to a multivariate elliptical distribution such as the multivariate Gaussian or the Student-*t*. These copulas have become popular in finance and risk management because of their simple implementation and interpretation.

### Appendix B.2. Gaussian Copula

The Gaussian copula is defined as follows:

$$C(u_1, \dots, u_k; \Sigma) = \Phi_{\Sigma}(\Phi^{-1}(u_1), \dots, \Phi^{-1}(u_k)), \tag{A1}$$

where  $\Phi_{\Sigma}$  is the multivariate standard normal c.d.f. with a correlation matrix  $\Sigma$ , and  $\Phi^{-1}$  is the univariate standard normal quantile function. The Gaussian copula has zero dependence in the tails of the distribution.

### Appendix B.3. Student-*t* Copula

Let  $t_{\nu, \Sigma}$  be the standardized multivariate Student-*t* distribution with a correlation matrix  $\Sigma$  and  $\nu$  degrees of freedom. The multivariate Student-*t* copula can be defined as follows:

$$C(u_1, \dots, u_k; \Sigma, \nu) = t_{\nu, \Sigma}(t_{\nu_1}^{-1}(u_1), \dots, t_{\nu_k}^{-1}(u_k)), \tag{A2}$$

where  $t_{\nu_j}^{-1}$  is the quantile function of the univariate Student- $t$  distribution with  $\nu_j$  degrees of freedom. The Student- $t$  copula implies equal upper-tail and lower-tail dependence. The lower the degrees of freedom,  $\nu$ , the higher the tail dependence implied by the Student- $t$  copula. When  $\nu$  approaches infinity, the tail dependence tends to zero; in this case, we obtain the Gaussian copula.

*Appendix B.4. Archimedean Copulas*

An Archimedean copula is a copula constructed through a generator function  $\phi$ . If  $\phi$  is a strict generator with  $\phi^{-1}$  completely monotonic on  $[0, \infty]$ , then a  $k$ -variate Archimedean copula is defined as follows.

$$C(u_1, \dots, u_k) = \phi^{-1}(\phi(u_1) + \dots + \phi(u_k)). \tag{A3}$$

These copulas allow for asymmetric dependence in the tails of the distribution. Four Archimedean copulas, namely Clayton, Frank, Gumbel and Joe, are presented next. A further examination of copulas and measures of dependence can be found in Joe (1997); Cherubini et al. (2004); Nelsen (2006); Czado (2019).

*Appendix B.5. Clayton Copula*

The generator is given by  $\phi(u) = u^{-\alpha} - 1$ . Thenm the Clayton copula has the following form.

$$C(u_1, \dots, u_k) = (u_1^{-\alpha} + \dots + u_k^{-\alpha} - k + 1)^{-1/\alpha}, \quad \alpha > 0. \tag{A4}$$

The copula is asymmetric and allows for dependence in the lower orthant, while dependence in the upper orthant is restricted to zero.

*Appendix B.6. Frank Copula*

The generator is given by  $\phi(u) = \ln\left(\frac{\exp(-\alpha u) - 1}{\exp(-\alpha) - 1}\right)$ . Then, the Frank copula is defined by the following.

$$C(u_1, \dots, u_k) = -\frac{1}{\alpha} \ln \left[ 1 + \frac{(e^{-\alpha u_1} - 1) \cdot \dots \cdot (e^{-\alpha u_k} - 1)}{(e^{-\alpha} - 1)^{k-1}} \right], \quad \alpha > 0 \text{ and } k \geq 3. \tag{A5}$$

The copula does not have dependence in either upper or lower orthants of the distribution.

*Appendix B.7. Gumbel Copula*

The generator is given by  $\phi(u) = (-\ln(u))^\alpha$ . The Gumbel copula is as follows.

$$C(u_1, \dots, u_k) = \exp \left( - \left[ (-\ln(u_1))^\alpha + \dots + (-\ln(u_k))^\alpha \right]^{1/\alpha} \right), \quad \alpha > 1. \tag{A6}$$

The copula is an asymmetric copula with dependence in the upper orthant but no dependence in the lower one.

*Appendix B.8. Joe Copula*

The generator is given by  $\phi(u) = -\ln(1 - (1 - u)^\alpha)$ . Then, the Joe copula is as follows.

$$C(u_1, \dots, u_k) = 1 - \left( 1 - [1 - (1 - u_1)^\alpha] \cdot \dots \cdot [1 - (1 - u_k)^\alpha] \right)^{1/\alpha}, \quad \alpha \geq 1. \tag{A7}$$

The Joe copula resembles the Gumbel copula with dependence in the upper but not the lower orthant of the distribution.

## Notes

- <sup>1</sup> For more details on PICK, LIT and QCLN, see Tables A1–A3, respectively, in Appendix A.
- <sup>2</sup> Orthants are a multivariate generalization of the bivariate quadrants.
- <sup>3</sup> None of the alternative, unreported model specifications that have been tried delivered more satisfactory results for these cases, leaving GARCH(1,1)-JSU as our top choice.
- <sup>4</sup> To minimize transaction costs, a private investor would normally prefer weights that do not vary substantially over time and do not require daily rebalancing of the portfolio. However, institutional investors acting on behalf of a large number of clients receive capital inflows daily and need to make new investments every day. Whenever a new investment is to be made (a new portfolio formed), it may be natural to look for a portfolio that has been estimated to be optimal for that day. This does not mean that the already existing portfolios would or should be rebalanced as frequently.
- <sup>5</sup> A null hypothesis that the expected annual returns on the two portfolios are equal cannot be rejected at any reasonable significance level, as a two-sided (one-sided) paired *t*-test yields a *p*-value of 0.58 (0.29).

## References

- Aas, Kjersti, Claudia Czado, Arnoldo Frigessi, and Henrik Bakken. 2009. Pair-copula constructions of multiple dependence. *Insurance: Mathematics and Economics* 44: 182–98. [CrossRef]
- Akaike, Hirotogu. 1973. Information theory and an extension of the maximum likelihood principle. In *2nd International Symposium on Information Theory*. Budapest: Publishing House of the Hungarian Academy of Sciences, pp. 268–81.
- Basel Committee. 2016. Minimum capital requirements for market risk. In *Technical Report, Bank for International Settlements*. Basel: Basel Committee.
- Basel Committee. 2017. Pillar 3 disclosure requirements—consolidated and enhanced framework. In *Technical Report, Basel Committee on Banking Supervision*. Basel: Basel Committee.
- Bayer, Sebastian, and Timo Dimitriadis. 2020. Regression-based expected shortfall backtesting. *Journal of Financial Econometrics*. nbaa013. [CrossRef]
- BBC News. 2021. Elon Musk Becomes World's Richest Person as Wealth Tops \$185bn. BBC. Available online: <https://www.bbc.com/news/technology-55578403> (accessed on 7 January 2021).
- Bedford, Tim, and Roger M. Cooke. 2001. Probability density decomposition for conditionally dependent random variables modeled by vines. *Annals of Mathematics and Artificial Intelligence* 32: 245–68. [CrossRef]
- Bedford, Tim, and Roger M. Cooke. 2002. Vines: A new graphical model for dependent random variables. *The Annals of Statistics* 30: 1031–68. [CrossRef]
- Bollerslev, Tim. 1986. Generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics* 31: 307–27. [CrossRef]
- Breymann, Wolfgang, Alexandra Dias, and Paul Embrechts. 2003. Dependence structures for multivariate high-frequency data in finance. *Quantitative Finance* 3: 1–14. [CrossRef]
- Cherubini, Umberto, Elisa Luciano, and Walter Vecchiato. 2004. *Copula Methods in Finance*. Hoboken: John Wiley & Sons.
- Christoffersen, Peter F. 1998. Evaluating interval forecasts. *International Economic Review* 39: 841–62. [CrossRef]
- Christoffersen, Peter F., and Denis Pelletier. 2004. Backtesting value-at-risk: A duration-based approach. *Journal of Financial Econometrics* 2: 84–108. [CrossRef]
- Cont, Rama. 2001. Empirical properties of asset returns: Stylized facts and statistical issues. *Quantitative Finance* 1: 223–36. [CrossRef]
- Czado, Claudia. 2019. *Analyzing Dependent Data with Vine Copulas*. Berlin/Heidelberg: Springer.
- Diebold, Francis X., Todd A. Gunther, and Anthony Tay. 1998. Evaluating density forecasts with applications to financial risk management. *International Economic Review* 39: 863–83. [CrossRef]
- Dissmann, Jeffrey, Eike C. Brechmann, Claudia Czado, and Dorota Kurowicka. 2013. Selecting and estimating regular vine copulae and application to financial returns. *Computational Statistics & Data Analysis* 59: 52–69.
- Engle, Robert F. 1982. Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation. *Econometrica* 50: 987–1007. [CrossRef]
- Forbes. 2020. Tesla Is Now the World's Most Valuable Car Company with a Valuation of 208 Billion. Forbes. Available online: <https://www.forbes.com/sites/sergeiklebnikov/2020/07/01/tesla-is-now-the-worlds-most-valuable-car-company-with-a-valuation-of-208-billion/> (accessed on 3 March 2022).
- FT. 2020. Tesla Overtakes Toyota to Become World's Most Valuable Carmaker. Financial Times. Available online: <https://www.ft.com/content/1fd51f91-ff3c-4723-995c-f75f6629c66e> (accessed on 1 July 2020).
- Hertzke, Patrick, Nicolai Muller, Patrick Schaufuss, Stephanie Schenk, and Ting Wu. 2019. *Expanding Electric-Vehicle Adoption Despite Early Growing Pains*. Boston: McKinsey & Company.
- Hyndman, Rob J., and Yanan Fan. 1996. Sample quantiles in statistical packages. *The American Statistician* 50: 361–5.
- IEA. 2020. *Global Electric Car Sales by Key Markets, 2010–2020*. Paris: IEA.
- Jarque, Carlos M., and Anil K. Bera. 1980. Efficient tests for normality, homoscedasticity and serial independence of regression residuals. *Economics Letters* 6: 255–9. [CrossRef]
- Joe, Harry. 1996. Families of *m*-variate distributions with given margins and  $m(m - 1)/2$  bivariate dependence parameters. In *Lecture Notes-Monograph Series*. Hayward: Institute of Mathematical Statistics, pp. 120–41.

- Joe, Harry. 1997. *Multivariate Models and Multivariate Dependence Concepts*. Boca Raton: CRC Press.
- Joe, Harry, and James Jianmeng Xu. 1996. *The Estimation Method of Inference Functions for Margins for Multivariate Models*. Vancouver: UBC Faculty Research and Publications. [CrossRef]
- Johnson, Norman L. 1949. Systems of frequency curves generated by methods of translation. *Biometrika* 36: 149–76. [CrossRef] [PubMed]
- Kendall, Maurice G. 1938. A new measure of rank correlation. *Biometrika* 30: 81–93. [CrossRef]
- Kupiec, Paul. 1995. Techniques for verifying the accuracy of risk measurement models. *The Journal of Derivatives* 3: 73–84. [CrossRef]
- Kurowicka, Dorota, and Roger M. Cooke. 2006. *Uncertainty Analysis with High Dimensional Dependence Modelling*. Hoboken: John Wiley & Sons.
- Markowitz, Harry. 1952. The utility of wealth. *Journal of Political Economy* 60: 151–58. [CrossRef]
- McNeil, Alexander J., and Rüdiger Frey. 2000. Estimation of tail-related risk measures for heteroscedastic financial time series: An extreme value approach. *Journal of Empirical Finance* 7: 271–300. [CrossRef]
- Nagler, Thomas, and Thibault Vatter. 2019. *rvinecopulib: High Performance Algorithms for Vine Copula Modeling*. R Package Version 0.5.1.1.0.
- Nelsen, Roger B. 2006. *An Introduction to Copulas*. Berlin/Heidelberg: Springer.
- Patton, Andrew J. 2006. Modelling asymmetric exchange rate dependence. *International Economic Review* 47: 527–56. [CrossRef]
- R Core Team. 2020. *R: A Language and Environment for Statistical Computing*. Vienna: R Foundation for Statistical Computing.
- Rodriguez, Juan Carlos. 2007. Measuring financial contagion: A copula approach. *Journal of Empirical Finance* 14: 401–23. [CrossRef]
- Sklar, Abe. 1959. Fonctions de répartition à n dimensions et leurs marges. *Publications de l'Institut Statistique de l'Université de Paris* 8: 229–31.
- Statista. 2021a. *Best-Selling Plug-in Electric Vehicle Models Worldwide in 2020*. Hamburg: Statista.
- Statista. 2021b. *Estimated Plug-in Electric Vehicle Sales Worldwide in 2020, by Automaker*. Hamburg: Statista.
- Woodward, Michael, Jamie Hamilton, and Bryn Walton. 2020. Battery Electric Vehicles: New Markets, New Entrants, New Challenges. Deloitte. Available online: <https://www2.deloitte.com/uk/en/insights/industry/automotive/battery-electric-vehicles.html> (accessed on 20 February 2021).
- Yollin, Guy. 2009. R tools for portfolio optimization. In *R/Finance Conference*. Bellevue: Quantitative Research Analyst, Rotella Capital Management, vol. 2009.