Journal of Financial Market Infrastructures **11**(3), 1–22
DOI: 10.21314/JFMI.2024.006

Research Paper

Correlation breakdowns, spread positions and central counterparty margin models

David Li,[1] Fernando Cerezetti[2] and Roy Cheruvelil[1]

[1]Securities and Exchange Commission, 100 F Street, NE, Washington, DC 20549, USA;
emails: liyu@sec.gov, cheruvelilr@sec.gov
[2]EACH Risk Committee, Avenue des Arts 6, Brussels 1210, Belgium;
email: fernando.valvanocerezetti@ice.com

(Received March 21, 2024; revised June 7, 2024; accepted June 12, 2024)

ABSTRACT

The default of a member of the Nasdaq Clearing commodities market in 2018 and the Covid-19 events in 2020 brought the importance of appropriately measuring breakdowns in market correlation to the attention of risk managers at central counterparties. The sizable price dislocations registered on these occasions suggested that traditional risk models may not be fully equipped to capture such breakdowns. Because correlations are directly impacted by the statistical properties of each variable, any model that lacks the capacity to deal with nonstationarity may inappropriately represent correlations or their alterations. Using an approach that combines a generalized autoregressive conditional heteroscedasticity model with dynamic conditional correlation (GARCH-DCC) to accommodate such properties, we aim to study correlation behavior during adverse market conditions and the potential impact on central counterparty margins. We propose a case study on energy commodities, with a specific focus on spread positions for the electricity market. The analysis suggests that correlation breakdowns are more frequent than traditionally expected. When different types of shocks (ie, those of September 2018 and March–May 2020) are considered, it becomes evident that while the magnitudes of the breakdowns may differ, their cycles present a number of similarities. We also recognize the potentially increased

Corresponding author: F. Cerezetti

Print ISSN 2049-5404 | Online ISSN 2049-5412

margin procyclicality that may be entailed by model corrections to deal with correlation breakdowns, highlighting the challenges of balancing margin responsiveness and stability during adverse market conditions.

Keywords: correlation breakdown; multivariate t-copula; value-at-risk (VaR); dynamic conditional correlation (DCC); spread positions; Nasdaq default.

1 INTRODUCTION

Instances of correlation breakdowns are not new to financial markets. After almost every major crisis over the past 30 years, studies were performed aiming at understanding the reasons leading to, and the impacts of, a rupture in the dependency structure of different instrument prices and returns. Early reviews of the literature (see Bertero and Mayer 1990; King and Wadhwani 1990; Boyer *et al* 1999) showed that correlation breakdowns were typically associated with sudden moves to the extremes, ie, 1 or -1 in the case of the standard Pearson correlation. It appears, however, that as markets develop, altering the way demand and supply are established, so does the way in which correlations behave during times of adverse conditions.

The more recent crises demonstrated that prices may not all go in the same direction when shocks hit (Kazi and Salloy 2014; Bank for International Settlements 2017). The default of a member of the Nasdaq Clearing commodities market in September 2018 was caused by an extreme variation in the price spread between German and Nordic power futures; on September 10 the spread reached €5.56 per megawatt hour (MWh), compared with the previous high of €1.85 per MWh, cutting the typical (close to 1) correlation by almost half. Similar abrupt changes in the correlation structure of price returns were also observed in a number of markets during the initial phase of the Covid-19 pandemic. For instance, on April 20, 2020 the spread between the first and second expiries of West Texas Intermediate (WTI) futures jumped to almost 900% compared with the previous day's close.[1] Similarly, the spread between the spot gold price (ie, the buy/sell price of physical gold for immediate delivery) and the price of the gold futures contract reached almost US$80 in April 2020.[2] Note that this spread is typically near zero.

Correlation breakdowns represent a problem for investors and risk managers, as the protection endowed by portfolio diversification is reduced, if not lost. Under

[1] The main driver of this change was the price of the first expiry of the WTI contract, which decreased from almost US$18 per barrel on April 17 to approximately minus US$38 per barrel on April 20.

[2] The dislocation was associated with transportation issues in moving physical gold from London, the global hub for the physical contracts, for delivery against futures contracts in New York, where most of the derivatives were traded.

the classical asset pricing theories (eg, minimum-variance portfolios, the capital asset pricing model or arbitrage price theory), correlations play a key role, as they determine the market or beta-zero portfolio (Markowitz 1959; Sharpe 1964; Lintner 1965; Ross 1976). When intertemporal or consumption models are considered, the stochastic discount factor cannot be characterized without accounting for the dependency between assets (Hansen and Jagannathan 1991; Duffie 2001; Delbaen and Scharchermayer 2006). Similarly, traditional risk metrics such as parametric value-at-risk (VaR), Monte Carlo VaR and filtered historical simulation depend on correlations to define the expected losses of an asset or portfolio. Any change in the estimated values of correlations will, therefore, impact the behavior of risk managers (higher margins, reduction in stop-loss metrics and change in hedging strategies are some examples) (see, for example, Jorion 2006; Alexander 2009; Gregory 2012).

In central counterparty (CCP) risk management the importance of such correlation dislocations is no different. As shown by Vicente *et al* (2015), most risk methodologies currently employed by CCPs can trace their origins to a common forebear: the VaR approach. The concept of a single risk value strongly influenced the development of traditional CCP risk management models, as it could be easily translated into a margin requirement. However, the standard implementations of the VaR approach face challenges in fully reflecting the reality of multi-asset risk management at CCPs. When historical or parametric VaR are considered, the ability to deal with correlation breakdowns may be limited, exposing CCPs to the risks of undermargining more complex portfolios.

This paper aims to contribute to the literature on correlation behavior during adverse market conditions. To address the nonstationary statistical properties of price returns, we use an approach that combines a generalized autoregressive conditional heteroscedasticity (GARCH) model with dynamic conditional correlation (DCC). Boyer *et al* (1999) proposed that changes in correlation may simply be a reflection of alterations in the volatility of variables under analysis. Therefore, any model that lacks the capacity to deal with nonstationarity may represent correlation or its alterations inappropriately. Similarly, multivariate fat-tail VaR models are reviewed in the paper to assess how different configurations of CCP margin models can influence the way correlation breakdowns are considered. Li and Cheruvelil (2019) discussed some of these configurations, highlighting that the introduction of correlation ceilings could reduce the impact of breakdowns.[3]

---

[3] Our research does not advocate for or promote a specific margin model or methodology (eg, the GARCH–DCC framework) for use by any CCP; such research would involve a different analysis, inclusive of other considerations such as efficiency and procyclicality effects. However, it should help practitioners gain insights into the nature of correlation behavior for risk management purposes.

Our analysis of correlation breakdowns is performed through a case study on energy commodities, and the electricity market in particular. As electricity cannot be stored in a flexible and cost-effective manner, preventing rapid responses to sudden changes in supply and demand, it is a good candidate for the study of shocks to correlation. Moreover, the events of September 2018 and March–May 2020 allow us to contrast the correlation behavior in response to two distinct types of shocks. While the former is mainly an endogenous shock and largely related to climate conditions that affect only the electricity market, the latter event can be viewed as an exogenous macroeconomic shock spreading across several asset classes.

Using spread positions on these electricity futures contracts as a mechanism to capture the dependency between different assets, we are able to confirm some of the well-known features of the energy markets, such as excess kurtosis (ie, fat tails), jumpiness and nonstationarity (Carmona and Durrelman 2003). The implementation of an enhanced multivariate GARCH model enables the identification of correlation breakdowns, suggesting that these dislocations are more frequent than traditionally expected. When the different types of shocks are considered, it becomes evident that while the magnitude of the breakdowns may differ, their shock cycles present a number of similarities. This enhanced methodology also suggests that traditional VaR models may face challenges in capturing the dynamics of correlation dislocations.

The rest of the paper is organized as follows. Section 2 presents a literature review. Section 3 introduces our case study and a potential enhancement to the standard VaR approach that could more appropriately filter the statistical properties of spread positions. Section 4 shows the results for the empirical implementation of the enhanced VaR. Section 5 presents a "crisis replay", backtesting and procyclicality exercises. Section 6 states our conclusions.

2 LITERATURE REVIEW

2.1 Correlation and correlation breakdowns

It is not a novel fact that correlation presents limitations as a metric of dependence. Correlation is a linear measure, and any relationship structure that is nonlinear might not be appropriately captured (eg, the correlation between a normally distributed variable X and a χ^2 (or chi-squared) variable is zero). Moreover, the Fréchet–Hoeffding boundaries establish that correlation does not necessarily vary between 1 and -1, with its range depending on the statistical distribution of the variables being measured (Nelsen 1990). It is only when distributions are close to the elliptical family (eg, a normal or Student t distribution) that the metric starts to be more appropriate. Nonetheless, correlation is still widely used in the financial economics literature, with the understanding of breakdowns being an important topic for study, due to the distressing effects on markets arising from correlation dislocations.

Different approaches have been pursued to explain the origins of correlation breakdowns. The important work of Loretan and English (2000) suggested that periods with high correlations tend to be accompanied by higher-than-average volatilities. They emphasized that correlation breakdowns reflect the time-varying volatility pattern of financial markets rather than changes in the relationship between asset returns (see also Boyer *et al* 1999). When there exists a linear dependency between variables, the conditional correlation becomes a function of the volatility of the underlying. If the volatility for a particular moment in time increases above its unconditional value, so will the correlation. Chang and Cheng (2016) added to this perspective and, considering a vector autoregression model, suggested that linkages between assets tend to increase following market shocks.

Using a slightly different approach, Eydeland and Wolyniec (2003) proposed that there are four general reasons why correlations may have unexpected behavior: a change in conditional correlation, for which the correlation coefficient might be time-dependent or stochastic; the nonexistence of an unconditional correlation; a nonlinear structure of dependency; and estimation noise. Moreover, when confidence intervals given by the Fisher transformation are considered, any appropriate estimation of correlations would typically require over 100 points before it starts to stabilize, although this number may increase for lower correlation levels (Fisher 1915).

Besides the purely statistical perspective, the literature has also tried to explain correlation breakdowns based on the behavior of investors. Falbo and Grassi (2015) proposed that correlation breakdowns are related to corrections rational agents make to their estimation of asset dependency in periods of high volatility. Specifically, using an investor's decision model, Falbo and Grassi suggested that the herding behavior of speculators and the procyclical attempt of rational investors to protect against dislocations are the root causes of correlation breakdowns. From a similar perspective, Dornbusch *et al* (2000) suggested that restrictions on the relocation of institutional investors' portfolios may also intensify breakdowns, as these investors have to recalibrate their portfolios to comply with laws or contractual clauses.

2.2 Spread positions

In finance, spreads broadly relate to one or a set of contracts whose final value is based on the difference between prices of two or more distinct risk factors (in some cases, spreads are based on a finite linear combination of three or more of these risk factors). Typically, this type of contract, or portfolio of contracts, is used to hedge against adverse movements in several risk factors at the same time. Defined as such, spread positions are key mechanisms for understanding the behavior of correlation across time, as their price formation in particular reveals important information about the future expectations of dependency.

The modeling of spread contracts, however, it not a simple task, and a number of different approaches have been proposed. Jump-diffusion models gained popularity in the 1980s due to their ability to replicate, at least qualitatively, the spiky nature of spread prices. Use of these models, initially applied to single-asset time series, soon expanded to other types of data such as spreads (see Bates (1988) for a review of the literature). Nonetheless, the inefficient and unstable estimates of jump-diffusion models made such attempts short-lived. The instability would typically stem from a combination of problem misspecification and a lack of extreme events in the time series. Moreover, when estimating the model, the parameters of the jump distribution could be biased by the presence of stochastic volatility.

Borrowing from the option pricing literature, an alternative modeling approach is based on the proposition that spreads behave as a normal distribution, and the instantaneous change as an arithmetic Brownian motion. The supporting argument for this approach is based on empirical evidence, as plots of spread histograms typically resemble a normal distribution. The framework led to the well-known Bachelier option pricer. A challenge to this modeling approach lies in the fact that an assumption of arithmetic Brownian motion is irreconcilable with the standard view that prices behave as lognormal distributions, with their dynamics given by geometric Brownian motions (Goldenberg 1991).

A complementary approach is to model each underlying of the spread as a geometric Brownian motion and to use a Wiener process to represent the relationship between them. The use of geometric Brownian motion is appealing when modeling spread positions, as it also allows for a straightforward derivation of prices for options. The key challenge of the geometric Brownian motion approach, however, is its limited capacity to replicate the statistical properties of contracts that exhibit excess kurtosis, skewness and an autoregressive pattern in the variance. In addition, the implied volatility extracted from these option pricers seems to show low resemblance with the realized volatility observed in the market (Eydeland and Wolyniec 2003).

Due to the above limitations, alternative approaches based on higher-dimensional models have become more popular for modeling spread positions, with the GARCH model gaining considerable popularity (Engle 1982; Bollerslev 1986). Using the GARCH specification with t-distributed (fat-tailed) innovations improves the in-sample fit, helping to remove unconditional skewness and excess kurtosis from the data. For energy markets in particular, a number of papers (see, for example, Zanottia *et al* 2010; Du and Laib 2017) have emphasized these and other positive features of the GARCH approach for single contracts as well as spread positions, as introduced in the next section.

3 THEORY AND CASE STUDY

Energy markets saw rapid changes in the decades following the so-called deregulation phase. In the United States, after the opening up of competition in fuel markets in the 1980s, electricity markets in the 1990s and weather/emissions markets in the late 1990s, the whole energy sector was marked by an expansion in trading volumes, contract types and market participants (Eydeland and Wolyniec 2003). Together with these developments, new price-discovery mechanisms were established, substantially changing the dynamics of price formation and behavior through time. Demand and supply emerged to replace the previous target-based approach to the valuation of energy commodities. In this new world, the dynamics of energy prices were characterized by seasonality, mean reversion, excess kurtosis and jumpiness.

3.1 Data

The proposed case study is based on the electricity energy market, with a specific focus on spread positions. According to Carmona and Durrelman (2003), spreads are probably the most useful, prevalent and important structure in the world of energy. Spreads are used to describe the costs associated with power plants, refineries, storage facilities and transmission lines. In some cases, spreads can even be used as a way to quantify the cost of production for refined products made from a combinations of raw materials. Practically every aspect of energy production can potentially be explained using spreads. For our case study, we assess the historical series of the spread between two energy futures contracts: Intercontinental Exchange (ICE) Endex German Power Financial Base Futures (GAB), and ICE Endex Nordic Power Financial Base Futures (NRB). These are financially settled futures contracts based on monthly futures, with a valuation dependent on the hourly prices of electricity arising from controlled areas.

The data used in the assessment are daily prices, and the first 12 expiries are considered for the period from June 6, 2016 to June 1, 2020. This period includes September 2018, when prices for the spread between GAB and NRB spiked to levels not previously observed, as well as the first few months of the Covid-19 pandemic (ie, the severe shocks observed during March 2020 and May 2020). The dislocation in 2018 was due to a combination of excessive rainfall in the Nordic regions, which sharply increased hydropower supply, and the concurrent price surge of the European carbon allowance, which drove the price of German power high. While the 2018 shock was endogenous and affected only the electricity market, the Covid-19 event was a global and systemic market stress, exogenously spreading across several asset classes.

3.2 Methodology: multivariate conditional correlation

3.2.1 GARCH with constant conditional correlation

Bollerslev (1990) proposed a multivariate GARCH model that uses time-varying conditional variances and covariance but constant conditional correlation (CCC). The conditional covariance matrix is given by

$$\boldsymbol{H}_t = \boldsymbol{D}_t \bar{\boldsymbol{R}} \boldsymbol{D}_t, \tag{3.1}$$

where $\boldsymbol{D}_t$ is an $n \times n$ stochastic diagonal matrix with elements $\sigma_{i,t}$, which follows a univariate GARCH process, and $\bar{\boldsymbol{R}}$ is an $n \times n$ time-invariant unconditional correlation matrix of the standardized error $\boldsymbol{\epsilon}_t$. Specifically,

$$\begin{aligned}
\boldsymbol{D}_t &= \operatorname{diag}(\sigma_{1,t}, \sigma_{2,t}, \dots, \sigma_{n,t}), &(3.2)\\
\boldsymbol{\epsilon}_t &= \boldsymbol{D}_t^{-1} \boldsymbol{\varepsilon}_t, &(3.3)\\
\bar{\boldsymbol{Q}} &= \operatorname{Cov}(\boldsymbol{\epsilon}_t \boldsymbol{\epsilon}_t^{\mathrm{T}}) = E[\boldsymbol{\epsilon}_t \boldsymbol{\epsilon}_t^{\mathrm{T}}], &(3.4)\\
\bar{\boldsymbol{R}} &= \operatorname{diag}(\bar{\boldsymbol{Q}})^{-1/2} \bar{\boldsymbol{Q}} \operatorname{diag}(\bar{\boldsymbol{Q}})^{-1/2}, &(3.5)
\end{aligned}$$

where $\sigma_{i,t}^2$ follows a GARCH process as defined in (3.2) and $\boldsymbol{\varepsilon}_t$ is the de-autocorrelated residual as defined in (3.3).

Although the estimation of GARCH–CCC is computationally attractive because the correlation matrix is constant, the empirical evidence suggests it may be too restrictive. Therefore, to appropriately account for the heteroscedasticity of the series, the model is generalized in this paper by assuming the correlation matrix varies with time, as discussed in the following.

3.2.2 GARCH–DCC and the DCC-copula

The DCC model was introduced by Engle and Sheppard (2001). The key design idea is that the dynamic covariance matrix $\boldsymbol{H}_t$ can be decomposed into the matrix of conditional standard deviations $\boldsymbol{D}_t$ and a correlation matrix $\boldsymbol{R}_t$. Both $\boldsymbol{D}_t$ and $\boldsymbol{R}_t$ are time-varying.

The conditional correlation estimator under the multivariate DCC representation is

$$\begin{aligned}
\boldsymbol{R}_t &= \operatorname{diag}(\boldsymbol{Q}_t)^{-1/2} \boldsymbol{Q}_t \operatorname{diag}(\boldsymbol{Q}_t)^{-1/2}, &(3.6)\\
\boldsymbol{Q}_t &= (1 - \alpha - \beta)\bar{\boldsymbol{Q}} + \alpha \boldsymbol{\epsilon}_{t-1} \boldsymbol{\epsilon}_{t-1}^{\mathrm{T}} + \beta \boldsymbol{Q}_{t-1}, &(3.7)
\end{aligned}$$

where $\boldsymbol{R}_t$ is the DCC at time t, $\bar{\boldsymbol{Q}}$ is the unconditional covariance matrix as defined in (3.4), α represents the dynamic term introduced by the interaction between the two innovations and β represents the persistence term. To ensure the matrix $\boldsymbol{R}_t$ is positive-definite, the scalars α and β must satisfy

$$\alpha \geqslant 0, \quad \beta \geqslant 0, \quad \alpha + \beta < 1.$$

The DCC-copula can be represented as

$$\boldsymbol{F}(\boldsymbol{\epsilon}_{1t}, \boldsymbol{\epsilon}_{2t}, \dots, \boldsymbol{\epsilon}_{nt}) = \boldsymbol{C}(\boldsymbol{F}_1(\boldsymbol{\epsilon}_{1t}), \boldsymbol{F}_2(\boldsymbol{\epsilon}_{2t}), \dots, \boldsymbol{F}_n(\boldsymbol{\epsilon}_{nt}); \boldsymbol{\psi}_t), \tag{3.8}$$

where $\boldsymbol{\psi}_t$ is the copula parameter including the dependence-structure parameter $\boldsymbol{R}_t$ and the multivariate degree of freedom (DoF) or copula DoF $\boldsymbol{\nu}_\mathrm{c}$.

To estimate the DCC-copula, there are generally two key steps, each using the maximum loglikelihood. The loglikelihood is obtained from the following formula:

$$\begin{aligned}\ln(L(\theta)) = \sum_{t=1}^{T} \bigg(\ln\bigg[\Gamma\bigg(\frac{\boldsymbol{\nu}_\mathrm{c}+n}{2}\bigg)\bigg] - \ln\bigg[\Gamma\bigg(\frac{\boldsymbol{\nu}_\mathrm{c}}{2}\bigg)\bigg] - \frac{n}{2}\ln[\pi(\boldsymbol{\nu}_\mathrm{c}-2)] \\ - \tfrac{1}{2}\ln[|\boldsymbol{D}_t\boldsymbol{R}_t\boldsymbol{D}_t|] - \frac{\boldsymbol{\nu}_\mathrm{c}+n}{2}\ln\bigg[1 + \frac{\epsilon_t^\mathrm{T} R_t^{-1}\boldsymbol{\epsilon}_t}{\boldsymbol{\nu}_\mathrm{c}-2}\bigg]\bigg).\end{aligned} \tag{3.9}$$

The parameter set θ is divided into two groups:

$$(\phi, \psi) = (\phi_1, \phi_2, \dots, \phi_n, \psi),$$

where $\phi_i = (\alpha_{1i}, \dots, \alpha_{ni}, \beta_{1i}, \dots, \beta_{ni})$, $i = 1, \dots, n$, are the parameters of the univariate GARCH model for the risk factor and $\psi = (\alpha, \beta, \boldsymbol{\nu}_\mathrm{c})$. The DCC-copula is then estimated as follows.

STEP 1 Estimate the univariate GARCH model to get ϕ for calculating $\boldsymbol{D}_t$.[4]

STEP 2 Estimate ψ to simultaneously obtain the time-varying dependence structure $\boldsymbol{R}_t$ (ie, α and β) and $\boldsymbol{\nu}_\mathrm{c}$ using the standardized residuals from the first step.

The advantage of the DCC-copula is that the loglikelihood of the volatility and of the correlation can be maximized independently, as long as consistency is ensured within these two steps (Engle 2002).[5]

4 EMPIRICAL ASSESSMENT AND MODEL FITTING

4.1 Univariate statistical analysis and model fitting

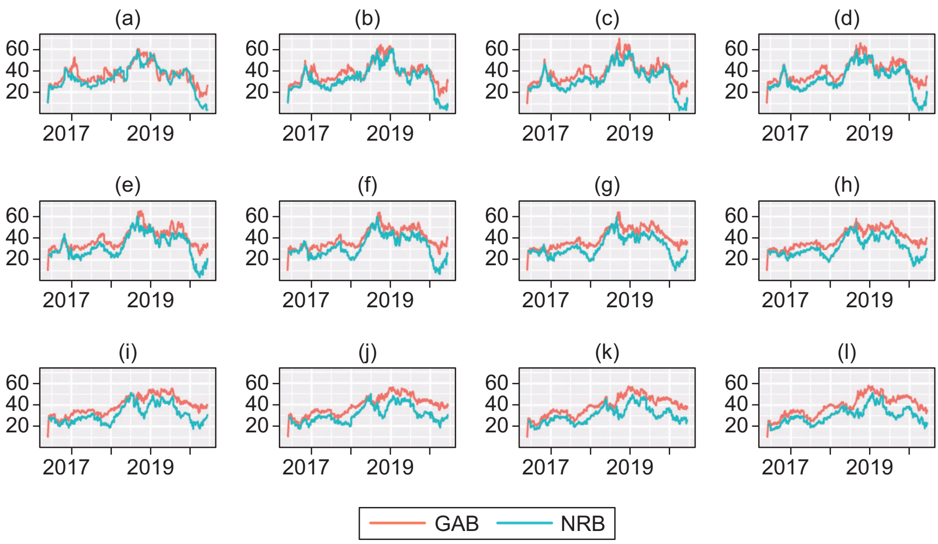
The historical price series of the first 12 monthly expiries for GAB and NRB are displayed in Figure 1 (where 1M denotes the first monthly expiry, 2M the second

---

[4] The autocorrelation coefficient and individual DoF are also estimated in the first step.

[5] The technical limitations of DCC have been widely discussed in the academic literature (see, for example, Caporin and McAleer 2013). One of the limitations is that two-step estimators for DCC may not be consistent, because in (3.7) the matrix $\boldsymbol{Q}_t$ is not the expectation of the standardized residuals' cross products. The various attempts at enhancement are beyond the scope of this paper.

FIGURE 1 Historical price series of GAB and NRB, in euros per MWh.



(a) 1M. (b) 2M. (c) 3M. (d) 4M. (e) 5M. (f) 6M. (g) 7M. (h) 8M. (i) 9M. (j) 10M. (k) 11M. (l) 12M.

monthly expiry, and so on). Studying the univariate statistical properties of these series reveals the excess kurtosis (ie, fat tails), jumpiness and heteroscedasticity of the data (for details see Tables A1 and A2 and Figures A1 and A2 in the online appendix). In particular, for 12M, discussed here for illustration, we observe that the series exhibit a DoF between 2 and 3. Both GAB and NBR price return time series show little skewness but sizable kurtosis. The quantile–quantile test also confirms significant deviation from the normal distribution.

To accommodate the nonstationarity of the series, several symmetric and asymmetric volatility and innovation models are tested for goodness-of-fit. The Glosten–Jagannathan–Runkle (GJR) GARCH is tested as the asymmetric volatility model, and Hansen's skewed Student t is tested as the asymmetric innovation model. For the time series up to September 7, 2018 (the business day before the September 10 dislocation), the symmetric GARCH(1, 1) and symmetric Student t exhibit the best-fitting results. The combinations including either asymmetric GJR-GARCH or Hansen's asymmetric Student t produce slightly lower but comparable likelihood. However, neither the GJR term nor the skew term from Hansen's Student t is statistically significant, which is consistent with the relatively small skewness seen in the statistical analysis of the empirical time series. Therefore, in the analysis below, GARCH(1, 1) and Student t are selected for univariate volatility and innovation modeling for GAB and NRB, with both time series displaying small autocorrelations. For details see Tables A3–A6 in the online appendix.[6]

4.2 Multivariate model calibration

For the multivariate analysis, joint maximum likelihood estimation is used to calibrate the model as described in Section 3.2. The multivariate dynamic copula model parameters (α, β and $\boldsymbol{\nu}_{\mathrm{c}}$) are all statistically significant.[7] The multivariate DoFs also reflect their respective univariate DoFs reasonably well. All three parameters are stable, whether we are comparing before and after the September 2018 dislocation or looking at the Covid-19 period. We do not account for the Samuelsson effect as it is not strong for long-term contracts, such as most of those considered in the analysis. Tables 1–3 illustrate these parameters for 12M and specific days preceding sizable shocks in the spread contracts.

4.3 Correlation analysis

Appropriately calibrating the univariate and multivariate models enables us to deal with the nonstationarity of the return time series. Correlation can then be studied without the undesirable effects of these features. Under this approach, we observe that the correlation between GAB and NRB fluctuates around the long-term historical average, as indicated by the unconditional correlations. This effect is prominent for the longer expiries (namely, those equal to or greater than four months), as detailed in Figure A3 of the online appendix.

If correlation breakdowns are defined as dislocations beyond a 95% confidence interval around the long-term average, then several breakdowns can be seen throughout the analyzed historical period. This evidence suggests that these dislocations are more frequent than traditionally expected, and that the events of September 2018 and March–May 2020 are not exceptions. Moreover, in contrast to what is commonly postulated, the breakdowns occur in both directions and are not necessarily associated with the correlation values moving closer to the unit.

Another important observation is that correlation breakdowns are not homogeneous across different parts of the futures curve (ie, different expiries). This is observable for both the September 2018 and March–May 2020 shocks and, even

[6] To study the stability of model parameters, and how they evolve over time during different time periods, we segmented the entire lookback period into smaller subsets. These display the overall stability of the parameters before or after certain major market events (see Tables A7 and A8 in the online appendix).

[7] The parameters β and $\boldsymbol{\nu}_{\mathrm{c}}$ are statistically significant above the 99% level, while the α term is significant above the 90% level.

TABLE 1 GARCH–DCC dynamic copula model parameters on September 7, 2018, for 12M.

	α	β	ν_{C}
Parameter value	0.0220972	0.8870032	2.70
Error	(0.01422)	(0.081214)	(0.07)

TABLE 2 GARCH–DCC dynamic copula model parameters on January 24, 2019, for 12M.

	α	β	ν_{C}
Parameter value	0.0154947	0.9006684	2.90
Error	(0.010254)	(0.0941639)	(0.07)

TABLE 3 GARCH–DCC dynamic copula model parameters on May 15, 2020, for 12M.

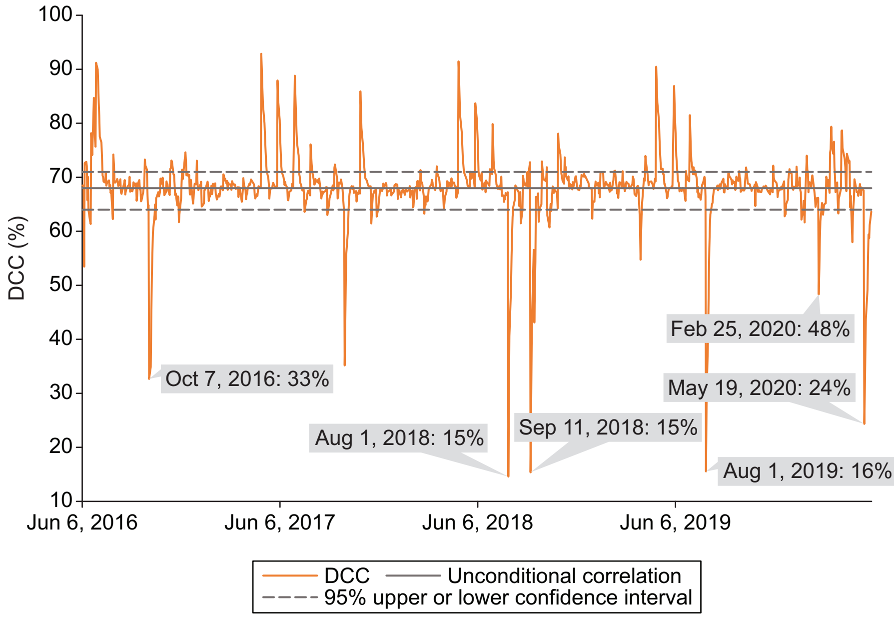
	α	β	ν_{C}
Parameter value	0.0188728	0.9014135	2.82
Error	(0.011134)	(0.0941639)	(0.06)

when the systemic stress of the Covid-19 period is accounted for, no consistent pattern is seen for contracts with different maturity dates. Similarly, a comparison of the effects of the shocks in the September 2018 and March–May 2020 periods suggests that while the magnitude of the breakdowns may differ, their cycles present a number of similarities. More specifically, after the initial dislocation is observed, in general the correlation tends to return quickly to its long-term average.

To illustrate the above aspects, we consider 12M, as displayed in Figure 2. The 95% confidence interval for the long-term average has a lower bound of 64% and an upper bound of 71%, with an average value around 68%. The DCC figure goes as high as 93% and as low as 15%, and there are several episodes of major correlation breakdowns in this window (ie, from June 2016 to June 2020). The correlation on September 7, 2018 is around 72%. At close of business on September 10, the DCC drops sharply, getting close to 15%.[8] However, the correlation figure does not fully reflect the breakdown until a day later, September 11. After this episode, there are three more major correlation breakdowns: one on August 1, 2019; another on February 25, 2020, which is at the onset of the Covid-19 period; and finally one on May 19,

[8] The DCC at close of business on September 10 is 53%, based on the calibration using the time series up to September 10.

FIGURE 2 GARCH–DCC dynamic correlation from June 2016 to June 2020.



2020, during the continued Covid-19 stress. The patterns of the breakdowns are all to some extent similar; that is, they have mild warning signs prior to the sudden dislocation and converge quickly following the initial shock.

5 CRISIS REPLAY AND MODEL COMPARISON ANALYSIS

Li and Cheruvelil (2019) observed that risk metrics for balanced portfolios are quite sensitive to correlation changes under high-correlation regimes (eg, above 75%).[9] A "crisis replay" therefore represents an efficient way of illustrating how different statistical treatments of correlation can impact risk metrics. We calibrate five distinct models using the time series up to the day before the crisis event or the crisis period. Following that, we calibrate models every day for the selected window of the crisis and use them to calculate a risk metric (ie, a theoretical margin) for the next day. In addition to the multivariate t-copula, other models selected for comparison are

[9] One approach to mitigate this effect could be the introduction of "correlation ceilings".

historical simulation (HS) VaR and filtered historical simulation (FHS) VaR.[10] For the FHS VaR models, we consider decay factors ranging from 0.95 to 0.99 for the "devolatilizing" and "revolatilizing" volatilities.[11] The above models are also tested under the ES configuration. All risk models consider a 99% quantile with a two-day liquidation period, and a portfolio composed of a short position on 1 million MWh for GAB and a long position on 1 million MWh for NRB, for 12M. The two-day liquidation period, or margin period of risk, is consistent with the way these contracts are currently margined by CCPs and compliant with the applicable European Union regulation (European Union 2012). The portfolio size follows the minimum contract specifications and is used on an illustrative basis.

5.1 September 2018 event

On September 7, 2018 the GAB contract price was US$54.27 per MWh, while the NRB contract was valued at US$44.59 per MWh. At the end of the first business day of the crisis, September 10, the one-day loss of the theoretical portfolio is US$5.79 million. The cumulative two-day loss amounts to US$7.91 million on September 11, the second day. Table 4 summarizes the different risk metrics derived from the models discussed above, including the measurements of how much they fall short of the realized losses (ie, their shortfall). In particular, the table shows that the FHS VaR is the worst performer when the shortfall is considered. The unfiltered HS VaR performs better than the FHS VaR, but still with a sizable shortfall.[12] The multivariate t-copula VaR with either CCC or DCC performs much better than either HS or FHS. This is not only related to the commonly known fat-tail distribution and robust tail interdependence of the t-copula, but also attributable to the GARCH volatility forecast and, for DCC only, the correlation forecast.

The above results are in line with the observation of Gurrola-Perez and Murphy (2015) that HS or FHS VaR models cannot handle correlation breakdowns very well, and that a correlation-updating mechanism for these nonparametric models would be desirable. Comparing the DCC and CCC models, we find that the former does not generate higher margins than the CCC model. The main reason is that the September shock is a two-day stress event with no (or a very mild) sign of a correlation breakdown leading up to September 10, 2018. The DCC (along with the corresponding

---

[10] Gurrola-Perez and Murphy (2015) provided in-depth discussions on these models and also specifically discussed their performance on spread positions (eg, long natural gas and gasoil futures and short WTI).

[11] The devolatilizing and revolatilizing exponentially weighted moving-average volatilities start from the same date: June 6, 2016.

[12] The reason that the unfiltered HS VaR performs better than the FHS VaR in this case is that there was a scenario of stressed returns and correlation breakdown on August 1, 2018, just one month before the Nasdaq event.

TABLE 4 Model comparisons and portfolio loss for 12M on September 10 and September 11, 2018

Model or P/L	VaR/ES (US$ m) Sep 10	VaR/ES (US$ m) Sep 11	One-day shortfall (US$ m)	Two-day shortfall (US$ m)
PortfolioP/L			(5.79)	(7.91)
HS VaR	2.62	3.21	(3.17)	(5.29)
FHS VaR ($\lambda = 0.99$)	2.12	2.54	(3.67)	(5.79)
FHS VaR ($\lambda = 0.97$)	2.13	2.63	(3.66)	(5.78)
Multivariate t-copula VaR CCC	3.04	4.14	(2.75)	(4.87)
Multivariate t-copula VaR DCC	3.03	5.27	(2.76)	(4.88)
HS ES	4.15	4.86	(1.64)	(3.76)
FHS ES ($\lambda = 0.99$)	3.84	4.65	(1.95)	(4.07)
FHS ES ($\lambda = 0.97$)	3.96	4.79	(1.83)	(3.95)
Multivariate t-copula ES CCC	5.75	7.62	(0.04)	(2.16)
Multivariate t-copula ES DCC	5.39	8.99	(0.40)	(2.52)

Both the GARCH-CCC and DCC in this table have a GARCH volatility forecast. The GARCH-DCC also uses a correlation forecast.

forecasted correlation) is actually higher (73%) than the CCC (71%) on September 7, 2018.[13] The 99% ES with DCC or CCC is more robust than the 99% VaR with the HS or FHS model. This shows that the DCC and CCC models are even more conservative at higher quantiles (eg, 99.5%) than these nonparametric models.

5.2 Covid-19 pandemic event (May 2020)

The Covid-19 period was marked by more than one correlation breakdown. Although it does not exhibit an extreme price scenario like the September 2018 event, on May 19, 2020 the correlation drops 25 percentage points, gradually returning to its long-term value a week later. The GAB contract recorded a 1% return on May 19, while NRB experienced a 20% decrease. The one-day portfolio loss is US$4.78 million, and the two-day cumulative loss is much smaller (US$1.41 million) due to the NRB price bouncing back up on May 19.

The GARCH–DCC model has limited power to forecast the new correlation regime in this case. However, because the dynamic correlation can reflect the new correlation regime after the first day of the breakdown, the GARCH–DCC model is still able to catch up to the lower correlation after the first day, driving margins responsively higher for the rest of the low-correlation regime. In general, the

[13] The unconditional correlation based on the information up to September 7, 2018 is 71%. It drops to 68% after the Covid-19 period in May and June 2020.

TABLE 5 Model comparisons21 and portfolio loss for 12M on May 15 and May 18, 2020

Model or P/L	VaR/ES (US$ m) May 15	VaR/ES (US$ m) May 18	One-day shortfall (US$ m)	Two-day shortfall (US$ m)
Portfolio P/L			(4.78)	(1.41)
HS VaR	2.17	2.38	(2.61)	0.76
FHS VaR ($\lambda = 0.99$)	2.22	1.88	(2.56)	0.81
FHS VaR ($\lambda = 0.97$)	2.46	2.08	(2.32)	1.05
Multivariate t-copula VaR CCC	2.42	2.35	(2.36)	1.01
Multivariate t-copula VaR DCC	2.46	3.22	(2.32)	1.05
HS ES	3.51	3.60	(1.27)	2.10
FHS ES ($\lambda = 0.99$)	3.66	3.51	(1.12)	2.25
FHS ES ($\lambda = 0.97$)	4.22	3.94	(0.56)	2.81
Multivariate t-copula ES CCC	3.66	3.60	(1.12)	2.25
Multivariate t-copula ES DCC	3.81	4.91	(0.97)	2.40

Both the GARCH-CCC and DCC in this table have a GARCH volatility forecast. THe GARCH-DCC also uses a correlation forecast.

multivariate t-copula VaR or ES are more responsive to correlation stress than the nonparametric models (see details in Table 5).

5.3 Backtesting and procyclicality analysis

To further assess the impact of correlation on risk metrics, we perform two additional tests: backtesting and procyclicality analysis. In the backtesting exercises, we calculate margins in the form of 99% VaR and compare them with the two-day profit and loss of the spread position. If on a particular day the margin is smaller than the profit and loss, then a breach (also known as an exceedance) is recorded. The widely adopted Kupiec likelihood ratio test is used to statistically evaluate the frequency of exceedances. The Kupiec test is defined by the null hypothesis that the expected proportion of violations is equal to 0.01.

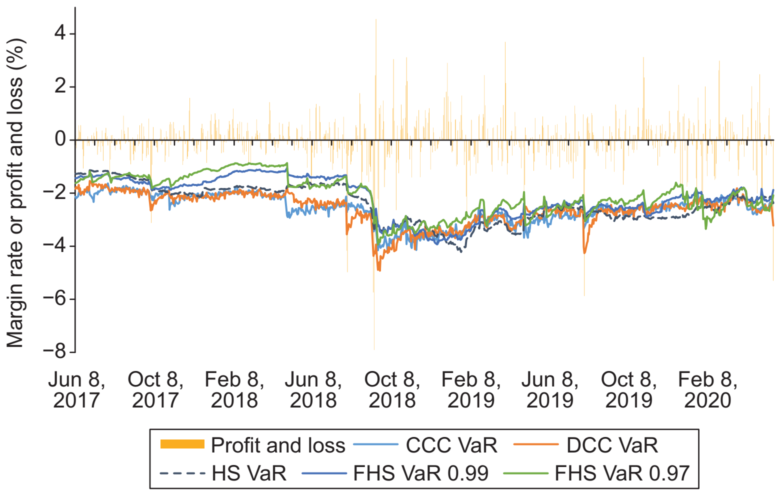
The backtesting exercise is performed for the same theoretical portfolio as the crisis replay and considers the following three periods.

(1) *June 8, 2017 to September 7, 2018.* This is the period right before the September 2018 event, and it is categorized by a relatively calm market.

(2) *June 8, 2018 to June 1, 2020.* This period includes both the September 2018 event and the Covid-19 pandemic. It is categorized as a stressed market period.

(3) *June 8, 2017 to June 1, 2020.* This is the whole period for the backtesting study.

FIGURE 3 Model backtesting for the five selected models from June 2017 to June 2020.



The margin rate is calculated using the 99% VaR compared with the two-day profit and loss.

In all three periods the multivariate t-copula DCC and CCC perform the most conservatively out of all models, with DCC exhibiting the lowest number of breaches. The Kupiec value of the HS VaR is close to, but below, the critical value, while the two FHS VaR models fail the test in the three periods. For the relatively calm period, all three nonparametric models pass the Kupiec test, with FHS VaR ($\lambda = 0.99$) close to the critical threshold. However, for the stressed period, HS VaR comes very close to the borderline for failure, while the two FHS VaR models fail the test. All test results are displayed in Tables A9–A11 in the online appendix, and a visual illustration is given by Figure 3.

To assess whether the conservatism of the multivariate t-copula DCC and CCC comes from reactive models, we assess procyclicality. Procyclicality is measured by the one-day and three-day margin rate changes over the entire lookback period. The results show that the FHS VaR ($\lambda = 0.97$) and the multivariate t-copula DCC and CCC are the most procyclical models, with daily one-day margin jumps of over 30%. For the three-day margin rate change, the FHS VaR ($\lambda = 0.97$) and the multivariate t-copula DCC are similarly the most reactive models, displaying changes of over 40% in a few instances (for more details on the one-day and three-day margin jump analysis, see Section 8.5 of the online appendix). We also assess procyclicality from a cost perspective, calculating different statistics of the margin level (ie, average, median, minimum, maximum and different quartiles). The overall conclusion is

TABLE 6 Model comparisons for margin levels from June 2017 to June 2020.

	FHS VaR				
Margin rate	$\lambda = 0.97$	$\lambda = 0.99$	**HS VaR**	**DCC VaR**	**CCC VaR**
Average	−2.097	−2.174	−2.415	−2.599	−2.604
Median	−2.130	−2.176	−2.442	−2.500	−2.528
Maximum	−3.942	−3.930	−4.218	−4.924	−4.252
75% quartile	−2.607	−2.647	−2.938	−2.869	−2.894
50% quartile	−2.130	−2.176	−2.442	−2.500	−2.528
25% quartile	−1.476	−1.450	−1.831	−2.135	−2.127
Minimum	−0.865	−1.050	−1.136	−1.537	−1.632

that the multivariate t-copula DCC and CCC, together with the HS VaR, are the most expensive models (see Table 6).

6 CONCLUSIONS AND FURTHER WORK

In this paper we studied correlation behavior during adverse market conditions using the GARCH–DCC model to filter out the nonstationary statistical properties of price returns. In particular, we considered a case study based on German and Nordic electricity power futures and the spread positions comprising them. These electricity power futures exhibit certain salient statistical characteristics, such as fat tails, volatility clustering, low autocorrelation, low skewness and sudden correlation breakdowns (and convergence). The study focused on the dynamic correlations and correlation breakdowns of these spread positions. The use of fat-tails and volatility clustering via the Student-t copula (ie, the GARCH volatility framework and DCC) allowed us to single out the correlation behavior. The shocks observed during the September 2018 event and the Covid-19 pandemic in early 2020 were the focus of our analysis.

Among other aspects discussed previously, we conclude that correlation breakdowns in the electricity power market seem to be more frequent than expected. Even though the nature of the shocks differs between the September 2018 event and the Covid-19 period, the pattern of the correlation breakdowns is relatively similar. It is also important to highlight that the correlation between GAB and NRB varies through time, exhibiting different structures for distinct expiry dates. To illustrate the importance of the modeling techniques used to capture correlation, we performed a crisis replay and backtesting exercises. The multivariate t-copula DCC or CCC with volatility and correlation forecasts demonstrated more conservative metrics in terms of coverage rates as well as breach amounts during the September 2018 event and the Covid-19 period than several nonparametric models, although they also displayed more procyclical behavior.

Several lines of further work could be explored to further improve model performance for spread positions from different asset classes. Typical spread positions such as those on Treasury cash and futures, Standard & Poor's 500 index options and Chicago Board Options Exchange Volatility Index futures are candidates to be included in a potential future study. Certain unique features of an asset class (eg, seasonal effects in certain futures products) could also be added to the framework. From the standpoint of procyclicality, since the GARCH–DCC copula framework would inevitably increase the margin responsiveness during correlation breakdowns, an effective margin floor may need to be explored to make margins less procyclical. A volatility floor (eg, long-term volatility) is commonly used for this purpose. For spread positions with a dynamic correlation assumption, the concept of a correlation ceiling could be explored to give them a reasonable level of offset and yet maintain a sufficient level of margin.[14] These features could be integrated into a multi-asset portfolio VaR measure and used to benchmark existing offsets provided by CCPs and their margin models.

DECLARATION OF INTEREST

The authors report no conflicts of interest. The authors alone are responsible for the content and writing of the paper. The Securities and Exchange Commission (SEC) disclaims responsibility for any private publication or statement of any SEC employee or Commissioner. This paper expresses the authors' views and does not necessarily reflect those of the Commission, the Commissioners or members of the staff. Fernando Cerezetti has written this paper strictly in his capacity as a member of the European Association of CCP Clearing Houses (EACH) Risk Committee. The views expressed in this paper are the author's own and do not necessarily reflect the views, position or policy of the author's employer, any affiliated company or EACH. The author's employer and EACH disclaim all responsibility for the private publications or statements of their employees or members.

ACKNOWLEDGEMENTS

The authors thank the anonymous referees and David Murphy, Gerardo Ferrara and Pedro Gurrola-Perez for their comments and suggestions. Any remaining inconsistencies are solely the responsibility of the authors.

[14] For the spread position studied in this paper, the correlation is in the range $[0.15, 0.93]$. If the correlation level is assumed to follow a gamma distribution, then, very intuitively, a certain high quantile (eg, 95%) could be used as a correlation ceiling.

REFERENCES

Alexander, C. (2009). *Market Risk Analysis, IV: Value-at-Risk Models*. Wiley.

Bank for International Settlements (2017). Beyond swings in risk appetite. *BIS Quarterly Review*, March. BIS, Basel. URL: www.bis.org/publ/qtrpdf/r_qt1703.htm.

Bates, D. S. (1988). Pricing options under jump-diffusion processes. Working Paper 37-88, The Wharton School, University of Pennsylvania, Philadelphia, PA. URL: https://bit.ly/4cvajVe.

Bertero, E., and Mayer, R. C. (1990). Structure and performance: global interdependence of stock markets around the crash of October 1987. *European Economic Review* **34**(6), 1150–1180 (https://doi.org/10.1016/0014-2921(90)90073-8).

Bollerslev, T. (1986). Generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics* **31**(3), 307–327 (https://doi.org/10.1016/0304-4076(86)90063-1).

Bollerslev, T. (1990). Modelling the coherence in short-run nominal exchange rates: a multivariate generalized ARCH model. *Review of Economics and Statistics* **72**(3), 498–505 (https://doi.org/10.2307/2109358).

Boyer, B. H., Gibson, M. S., and Loretan, M. (1999). Pitfalls in tests for changes in correlations. International Finance Discussion Paper 597 (revised), March, Board of Governors of the Federal Reserve System, Washington, DC.

Caporin, M., and McAleer, M. (2013). Ten things you should know about the dynamic conditional correlation representation. *Econometrics* **1**(1), 115–126 (https://doi.org/10.3390/econometrics1010115).

Carmona, R., and Durrelman, V. (2003). Pricing and hedging spread options. Working Paper, February, Princeton University. URL: https://carmona.princeton.edu/download/fe/sirev.pdf.

Chang, G., and Cheng, P. (2016). Evidence of cross-asset contagion in US markets. *Economic Modelling* **58**, 219–226 (https://doi.org/10.1016/j.econmod.2016.05.014).

Delbaen, F., and Scharchermayer, W. (2006). *The Mathematics of Arbitrage*. Springer (https://doi.org/10.1007/978-3-540-31299-4).

Dornbusch, R., Park, Y. C., and Claessens, S. (2000). Contagion: understanding how it spreads. *World Bank Research Observer* **15**(2), 177–197 (https://doi.org/10.1093/wbro/15.2.177).

Du, J., and Laib, K. K. (2017). Modeling dependence between European electricity markets with constant and time-varying copulas. *Procedia Computer Science* **122**, 94–101 (https://doi.org/10.1016/j.procs.2017.11.346).

Duffie, D. (2001). *Dynamic Asset Pricing Theory*, 3rd edn. Princeton University Press.

Engle, R. F. (1982). Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation. *Econometrica* **50**(4), 987–1007 (https://doi.org/10.2307/1912773).

Engle, R. F. (2002). Dynamic conditional correlation: a simple class of multivariate generalized autoregressive conditional heteroscedasticity models. *Journal of Business and Economic Statistics* **20**(3), 339–350 (https://doi.org/10.1198/073500102288618487).

Engle, R. F., and Sheppard, K. (2001). Theoretical and empirical properties of dynamic conditional correlation multivariate GARCH. Working Paper 8554, National Bureau of Economic Research (https://doi.org/10.3386/w8554).

European Union (2012). Regulation (EU) No 648/2012 of the European Parliament and of the Council of 4 July 2012 on OTC derivatives, central counterparties and trade repositories. *Official Journal of the European Union* **55**(L201), 1–59. URL: https://data.europa.eu/eli/reg/2012/648/oj.

Eydeland, A., and Wolyniec, K. (2003). *Energy and Power Risk Management: New Developments in Modeling, Pricing, and Hedging*. Wiley.

Falbo, P., and Grassi, R. (2015). Does expectation of correlation breakdown in financial market fulfil itself? *Discrete Dynamics in Nature and Society* **2015**, Paper 263908 (https://doi.org/10.1155/2015/263908).

Fisher, R. A. (1915). Frequency distribution of the values of the correlation coefficient in samples of an indefinitely large population. *Biometrika* **10**(4), 507–521 (https://doi.org/10.2307/2331838).

Goldenberg, D. H. (1991). A unified method for pricing options on diffusion processes. *Journal of Financial Economics* **29**(1), 3–34 (https://doi.org/10.1016/0304-405X(91)90011-8).

Gregory, J. (2012). *Counterparty Credit Risk and Credit Value Adjustment: A Continuing Challenge for Global Financial Markets*. Wiley (https://doi.org/bf7c).

Gurrola-Perez, P., and Murphy, D. (2015). Filtered historical simulation value-at-risk models and their competitors. Working Paper 525, Bank of England (https://doi.org/10.2139/ssrn.2574769).

Hansen, I., and Jagannathan, R. (1991). The implications of security market data for models of dynamic economies. *Journal of Political Economy* **99**(2), 225–262 (https://doi.org/10.1086/261749).

Jorion, P. (2006). *Value at Risk: The New Benchmark for Managing Financial Risk*, 3rd edn. McGraw-Hill, New York.

Kazi, I. K., and Salloy, S. (2014). Dynamics in the correlations of the credit default swaps' G14 dealers: are there any contagion effects due to Lehman Brothers bankruptcy and the global financial crisis? Working Paper 2014-237, IPAG Business School, Paris.

King, M. A., and Wadhwani, S. (1990). Transmission of volatility between stock markets. *Review of Financial Studies* **3**(1), 5–33 (https://doi.org/10.1093/rfs/3.1.5).

Li, D., and Cheruvelil, R. (2019). Study of correlation impact on credit default swap margin using a GARCH–DCC-copula framework. *The Journal of Financial Market Infrastructures* **8**(1), 51–92 (https://doi.org/10.21314/JFMI.2018.113).

Lintner, J. (1965). The valuation of risky assets and the selection of risky investments in stock portfolios and capital budgets. *Review of Economics and Statistics* **47**(1), 13–37 (https://doi.org/10.2307/1924119).

Loretan, M., and English, W. B. (2000). Evaluating "correlation breakdowns" during periods of market volatility. In *International Financial Markets and the Implications for Monetary and Financial Stability*, pp. 214–231. BIS Conference Papers, Volume 8. Bank for International Settlements, Basel (https://doi.org/10.2139/ssrn.231857).

Markowitz, H. (1959). *Portfolio Selection: Efficient Diversification of Investments*. Wiley.

Nelsen, R. (1990). *An Introduction to Copulas*. Springer Series in Statistics. Springer.

Ross, S. (1976). The arbitrage theory of capital asset pricing. *Journal of Economic Theory* **13**(3), 341–360 (https://doi.org/10.1016/0022-0531(76)90046-6).

Sharpe, W. (1964). Capital asset prices: a theory of market equilibrium under conditions of risk. *Journal of Finance* **19**(3), 425–442 (https://doi.org/10.1111/j.1540-6261.1964.tb02865.x).

Vicente, L. A. B. G., Cerezetti, F. V., De Faria, S. R., Iwashita, T., and Pereira, O. R. (2015). Managing risk in multi-asset class, multimarket central counterparties: the CORE approach. *Journal of Banking and Finance* **51**, 119–130 (https://doi.org/10.1016/j.jbankfin.2014.08.016).

Zanottia, G., Gabbib, G., and Geranioc, M. (2010). Hedging with futures: efficacy of GARCH correlation models to European electricity markets. *Journal of International Financial Markets, Institutions and Money* **20**(2), 135–148 (https://doi.org/10.1016/j.intfin.2009.12.001).