Option-Implied Dependence and Correlation Risk Premium

Oleg Bondarenko* and Carole Bernard,†‡

This version: June 1, 2021

Abstract

We propose a novel model-free approach to obtain the joint risk-neutral distribution among several assets that is consistent with options on these assets and their weighted index. We implement this approach for the nine industry sectors comprising the S&P 500 index and find that their option-implied dependence is highly asymmetric and time-varying. We then study two conditional correlations: when the market moves down or up. The risk premium is strongly negative for the down correlation but positive for the up correlation. Intuitively, investors dislike the loss of diversification when markets fall, but they actually prefer high correlation when markets rally.

Keywords: Risk-neutral density, option-implied dependence, asymmetric dependence, down and up correlation, correlation risk premium, model-free dependence recovery.

JEL codes: G11, G12, G13, G17.

1 Introduction

Option markets provide rich information about assets future returns. There exists a no-arbitrage relationship that links prices of options to the risk-neutral density of an asset future return. First discovered by Ross (1976), Breeden and Litzenberger (1978), and Banz and Miller (1978),

*Corresponding author: Oleg Bondarenko, Department of Finance, University of Illinois at Chicago (email: olegb@uic.edu).

†Carole Bernard, Department of Accounting, Law and Finance, Grenoble Ecole de Management, and Department of Economics and Political Sciences at Vrije Universiteit Brussel (VUB) (email: carole.bernard@grenoble-em.com).

‡For valuable comments, we thank Torben G. Andersen (discussant), Ian Dew-Becker, Nicole Branger, Adrian Buss (discussant), Hamed Ghanbari, Jens Jackwerth, Philippe Mueller (discussant), Paul Schneider (discussant), Fabio Trojani, Steven Vanduffel, Grigoriy Vilkov, Alex Weissensteiner, and seminar participants at CDI 2018 Conference in Montreal, Research in Options 2018 Conference in Honor of Bruno Dupire, 2019 Conference on Risk Management and Financial Innovation Conference in Memory of Peter Christoffersen, 2019 ITAM Finance Conference, 2019 EFMA Meeting, 2019 Risk Day organized by ETH Zurich, CUNEF Madrid, 2020 WFA meeting, Northwestern University, NYU, University of Lugano, and University of Illinois at Chicago. An earlier version of this paper was titled "Option-Implied Dependence" (with Steven Vanduffel). We also gratefully acknowledge funding from the Canadian Derivatives Institute (formerly IFSID) and the Global Risk Institute (GRI).

this fundamental relationship is one of the most useful in financial economics. Building on the relationship, researchers have developed effective techniques to estimate the risk-neutral distributions in a model-free way; see, for example, Jackwerth and Rubinstein (1996), Aït-Sahalia and Lo (2000), and Bondarenko (2003). Option-implied distributions have since then been used in numerous applications. However, the fundamental relationship only works in one dimension. It allows researchers to obtain the *individual* risk-neutral distributions for future returns of stock A and stock B, but not their *joint* distribution. In this paper, a novel extension to higher dimensions is proposed. Specifically, the options written on individual assets *and* on their index are used to fully describe the forward-looking risk-neutral dependence among the assets.

We refer to our approach as MFDR, or Model-Free Dependence Recovery. The approach consists of three main steps. First, we estimate risk-neutral marginal distributions of individual assets and also of their weighted sum (the index). As inputs, this step requires traded options on the individual assets and the index. Second, we frame the problem of finding a joint distribution among the assets as an integer optimization problem, in which a matrix of asset returns must be arranged in a suitable manner. Third, to solve the resulting large-scale optimization problem, we rely on a combinatorial technique, termed the block rearrangement algorithm (BRA) by Bernard and McLeish (2016).The BRA is the key to our approach, as alternative solution techniques are simply not feasible for the problem at hand.

Our MFDR method derives a complete description of the implied dependence and, as such, compares favorably with existing methods that primarily focus on the average pairwise correlation. As Buss, Schönleber, and Vilkov (2019a) point out,"[c]omputing the historical pairwise correlation among any two stocks is rather easy; however, computing an expected pairwise correlation from option data is, in practice, not possible." Thus, additional simplifying assumptions must be imposed to obtain an estimate. A common assumption is that of a constant pairwise correlation, which makes the recovery of the single parameter possible by comparing the variances of the index and individual components. This approach has been employed by the Chicago Board Options Exchange (CBOE) that has been disseminating its S&P 500 Implied Correlation Indices since July 2009. These correlation indices are now widely accepted dependence measures. In academic literature, related approaches are followed among others by Driessen, Maenhout, and Vilkov (2009, 2013), Buraschi, Kosowski, and Trojani (2013), and Faria, Kosowski, and Wang (2018), who also assume constant pairwise correlations.

Compared to the existing methods, the MFDR approach offers two critical advantages. First, as the name suggests, the approach is completely model-free and requires no parametric assumptions. Second and most importantly, it yields a full dependence structure, not just a partial dependence measure. While the existing methods equate the index variance with that of the weighted sum of its components (one moment condition), MFDR matches their whole distributions (theoretically, infinitely many conditions). This results in an essentially perfect fit of the index implied volatility curve, instead of only matching one of its statistics. Since MFDR yields a feasible dependence structure, a proper correlation matrix is assured by construction, without the need for any additional assumptions. Finally, while there might be many compatible dependence structures, MFDR is shown to maximize entropy and thus it yields the "most likely" implied dependence among asset returns given the information contained in available options.

We implement MFDR using ETF options on S&P 500 index and its nine industry sectors. The industry sectors provide an ideal setting for our methodology. On the one hand, the joint distribution of the nine sectors is rich and economically interesting. On the other hand, the

dimensionality of the problem is not excessively large and, computationally, it can be readily handled by our methodology.[1] Importantly, options on the industry sectors have become sufficiently liquid in recent years, making it feasible to accurately estimate their marginal distributions. Starting from January 2007, we are able to estimate option-implied dependence daily. To assess accuracy and stability of MFDR, we conduct an extensive Monte-Carlo experiment, see Appendix D. This experiment confirms the viability the new methodology in empirically relevant applications. Limitations are discusssed in Section 3.3.

Empirically, we make several contributions. Our first contribution is to document empirical properties of option-implied dependence. For the first time, our methodology allows a direct examination of the dependence under the risk-neutral measure. We find that the dependence is time-varying and highly nonnormal. In particular, the dependence between the nine sectors was much stronger in the later part of the financial crisis than in the earlier part. For most trading days, the option-implied dependence is grossly inconsistent with the assumption of multinormality. The formal tests of Mardia (1970) demonstrate that violations of multinormality are predominantly due to nonzero skewness, while violations due to excess kurtosis are less prevalent.[2] Overall, we find that for most days the option-implied dependence is highly asymmetric, with large negative returns being much more correlated than large positive returns.

Our second contribution is to present novel evidence regarding the correlation risk premium (CRP), defined as the difference between the correlation under the real-world and risk-neutral probability measures. The realized correlations are computed from historical returns, while the risk-neutral ones are obtained from the option-implied joint distribution. We focus on three types of average correlations computed for the nine sectors. The first type is the standard unconditional (or *global*) correlation. The other two are *down* and *up* correlations, i.e., correlations conditional on the S&P 500 return being below or above its median value. While the global correlation has been studied extensively in the literature, without our methodology, it is impossible to investigate the other two types.

The early studies of the global correlation risk premium include Driessen, Maenhout, and Vilkov (2009, 2013). These papers document a strong negative CRP for stocks in the S&P 100, S&P 500 and the DJIA and link it to diversification. During turbulent times, correlations tend to increase, making diversification less effective. Therefore, the index options are expensive because they allow investors to hedge against the risk of reduced diversification. In our work, we also find that the global correlation risk premium is negative. However, our most intriguing empirical results pertain to the down and up correlations. Here we document that the risk premium is significantly negative for the down correlation but positive for the up correlation. These findings are consistent with the economic intuition that investors are mainly concerned with the loss of diversification when the market falls. As a result, they are willing to pay a considerable premium to hedge against increases in the down correlation. On the other hand, investors actually *prefer* high correlation when the market rallies. That is, investors view the down correlation as "bad" and the up correlation as "good." The net effect of the negative risk premium for the down correlation and the positive risk premium for the up correlation is a

[1] It is worth mentioning that even for this problem of "moderate" dimensionality, the computational demands are considerable. We estimate option-implied dependence for over 2,500 trading days. For each day, the first step of MFDR calls for 10 RND estimations (or multivariate quadratic optimizations). The third step of MFDR involves finding the optimal perturbation of an 1000×10 matrix. The latter is an NP-complete problem with $(1000!)^9$ possible permutations. However, the BRA finds an approximate solution in a reasonable amount of time.

[2] The latter observation is somewhat surprising, given the fact that the risk-neutral margins are highly leptokurtic, for both the index and the sectors. However, for the *dependence*, excess kurtosis is less of an issue.

negative risk premium for the global correlation.

The signs of the down and up correlation risk premium mirror those for the down and up variance risk premium, as recently reported in Kilic and Shaliastovich (2019), and Feunou, Jahan-Parvar, and Okou (2018). However, our results neither follow from nor imply the latter findings. The above papers study the market variance risk premium and are based exclusively on index options. Intuitively, these papers establish that index OTM puts are expensive and index OTM calls are cheap compared to the historical distribution of the index returns. Our results, on the other hand, make a statement on the *relative* pricing of individual options compared to the index options. Intuitively, we find that sector OTM puts are cheap compared to index OTM puts (thus, implying too high down correlations between the sectors), while sector OTM calls are expensive compared to index OTM calls (implying too low up correlations).

Perhaps the sign of the down correlation is not completely unexpected. If investors only care about volatility when it leads to losses, they will dislike the down correlation and its risk premium will be negative. As argued by Ang, Chen, and Xing (2006), similar conclusions are also obtained in an equilibrium with disappointment aversion preferences (see Gul (1991), Routledge and Zin (2010)). However, it might be more challenging to rationalize with standard preferences the positive risk premium for the up correlation. Furthermore, the *magnitudes* of the risk premia for the down and up correlations also appear quite remarkable. To put things in perspective, the magnitude of the risk premium for the down (respectively, up) correlation is approximately 3 (respectively, 2) times larger than for the global correlation. Motivated by these findings, we introduce the down minus up correlation (DUC) swap, which at maturity pays the difference between the realized down and up correlations. Historically, the strategy that sells the DUC swap would have been very profitable, as it takes advantage of both the "expensive" down correlation and the "cheap" up correlation. Thus, selling the swap earns the risk premium which is about 5 times larger than selling the global correlation.

As is well-known, the regular (Pearson) correlation is jointly affected by the margins and the dependence. Thus, an important but difficult question is whether the CRP is driven by the priced components of the margins (e.g., volatility, skewness, heavy tails) or from the priced dependence. Therefore, as our third contribution, we disentangle the respective roles on the CRP of the changes in the two components. To remove the effects of the margins in the computation of correlation, we consider Spearman (instead of Pearson) correlations. Spearman's is a rank correlation and it has the advantage of being unaffected by the margins. To compute Spearman correlation under the risk-neutral measure, however, the complete joint distribution is required, even for the case of the global correlation. Thus, our MFDR methodology is crucial. By contrasting the results for Pearson and Spearman correlations, we conclude that the CRP is mainly driven by the dependence and not by the margins. Along the same lines, when we investigate what causes the enormous spread between the risk-neutral down and up correlations, we find that only about 11% of the spread can be attributed to nonnormality of the margins, while the rest is due to the nonnormality (skewness) of the dependence.

Our empirical results highlight the importance of proper modeling of the dependence, especially under the risk-neutral measure. To match the salient features of the option data, it is critical to allow for highly asymmetric dependencies. Standard models in the literature do not always have this property. Therefore, as our fourth contribution, we develop an alternative approach to model the multivariate joint distribution. We refer to it as the *hybrid* model because it combines (i) fully nonparametric margins extracted from the individual options and (ii) a parsimonious parametric copula. The proposed copula is based on the homogeneous multivariate

skewed normal distribution driven by two parameters only. Despite its simplicity, the model captures reasonably well the most salient features of the option-implied dependence. Of course, in terms of fitting option prices the parsimonious hybrid model cannot compete with MFDR, because the latter produces (essentially) a perfect fit. Instead, the hybrid model offers different advantages: it is transparent, intuitive, and easy to implement. Our primary motivation for developing this model is two-fold. First, because the hybrid model does not rely on the somewhat opaque MFDR methodology, it provides an alternative confirmation of our key empirical results. Second, we believe that the hybrid model could prove useful in other applications where MFDR cannot be implemented due to data limitations.

Our paper is related to several important strands of the literature. A number of academic papers use traded options to infer various measures of dependence. Driessen, Maenhout, and Vilkov (2009, 2013), Buraschi, Kosowski, and Trojani (2013), and Faria, Kosowski, and Wang (2018) estimate the average implied correlation assuming that pairwise correlations are all equal. Buss, Schönleber, and Vilkov (2017) estimate a block-diagonal correlation matrix with two possible values for pairwise correlations (that is, pairwise correlations are constant for any two stocks within the same economic sector of the S&P 500 but take another value when two stocks belong to different sectors). Buss and Vilkov (2012) are able to relax the assumption of equal pairwise correlations by imposing a special structure on the correlation risk premium. Specifically, they assume a linear transformation between the correlations under the real-world and risk-neutral measures. Kelly, Lustig, and Van Nieuwerburgh (2016) assess the dependence among several assets by computing the spread between a portfolio of individual puts and the put on the index. This approach is also related to the comonotonicity index developed by Linders, Dhaene, and Schoutens (2015).

The down and up correlation risk premiums (CRP) introduced in this paper extend the literature on the global CRP. The pioneering work on the global CRP includes Driessen, Maenhout, and Vilkov (2009, 2013), who document a strong negative CRP for stocks in the S&P 100, S&P 500 and the DJIA. Further investigation of the CRP and its link to macroeconomic variables can be found in Buraschi, Kosowski, and Trojani (2013), Faria, Kosowski, and Wang (2018), Engle and Figlewski (2014), Mueller, Stathopoulos, and Vedolin (2017), Pollet and Wilson (2010) and Harvey, Liu, and Zhu (2016). The correlation risk is also closely related to the variance risk (Bollerslev and Todorov (2011), Bondarenko (2014), Carr and Wu (2009), Schneider and Trojani (2015)) and to the disagreement risk (Buraschi, Trojani, and Vedolin (2014)). The down and up variance risk premium is studied by Kilic and Shaliastovich (2019) and Feunou, Jahan-Parvar, and Okou (2018).

The asymmetric behaviour of the left and right tail of the asset returns is widely accepted among practitioners and academics building on a pioneering paper of Longin and Solnik (2001). Ang and Chen (2002), Hong, Tu, and Zhou (2006), Alcock and Hatherley (2017), Jiang, Wu, and Zhou (2018) and Alcock and Sinagl (2020) provide evidence of asymmetric correlation and dependence between stocks and the market. Furthermore, Hong, Tu, and Zhou (2006) assess the economic importance of asymmetric returns in the context of a portfolio choice problem. They find that investors can achieve over 2% annual certainty-equivalent gains when they account for asymmetric correlation. Longin and Solnik (2001) study the asymmetric tail dependence by computing the exceedance correlation, or the correlation between returns that are jointly above or below a given threshold. For international equity markets, they demonstrate that the correlation between large *negative* returns does not converge to zero but instead tends to increase deeper in the left tail. On the other hand, the correlation between large *positive* returns

decreases to zero. More recently, Chabi-Yo, Ruenzi, and Weigert (2018) study the dependence of stock returns. Like us, they are motivated by the limitations of Pearson correlation. After correcting for the effect of margins, they find evidence of asymmetry between the left and right tails. One of their key findings is that "correlation is mainly affected by the market trend" (and not "volatility" per se, which is an information about the marginal distribution). It is intuitive to believe that such observation pertains also when looking at option implied information (asset returns under the risk neutral distribution). However, without the use of our MFDR method, this question could not be answered without making model assumptions. In our study of the nine sectors of the S&P 500 index, we find that the dependence is much more asymmetric under the risk-neutral measure (using option implied information) than under the real-world one (using assets returns). Our work is also related to the literature on skewness (see Amaya, Christoffersen, Jacobs, and Vasquez (2015), Jondeau, Zhang, and Zhu (2019), DeMiguel, Plyakha, Uppal, and Vilkov (2013), and Patton (2004)) and on downside risk (see Bollerslev and Todorov (2011), Kelly and Jiang (2014), Kelly, Lustig, and Van Nieuwerburgh (2016), Farago and Tédongap (2018), and Orłowski, Schneider, and Trojani (2020)).

The block rearrangement algorithm (BRA) used in the third step of MFDR generalizes the standard rearrangement algorithm (RA) introduced by Puccetti and Rüschendorf (2012). The RA has found important applications in various disciplines. For example, Embrechts, Puccetti, and Rüschendorf (2013) use the RA in quantitative risk management to assess the impact of model uncertainty on Value-at-Risk estimates for portfolios. Other important applications of the algorithm include operations research (fair allocation of goods, optimization) and engineering (image reconstruction). Bernard, Bondarenko, and Vanduffel (2018) study the theoretical properties of the BRA for the problem of finding a joint distribution when the marginal distribution of several random variables and of their sum are specified. This paper provides for the first time an empirical application using option data. The first step of MFDR assumes that risk-neutral marginal distributions can be estimated from traded options in a model-free way, see, for example, Jackwerth and Rubinstein (1996), Aït-Sahalia and Lo (2000), and Bondarenko (2003). Figlewski (2018) reviews pros and cons of the various methods that have been proposed for extracting risk-neutral densities.

The rest of the paper is organized as follows. In Section 2, we formulate the problem of inferring the dependence from option prices and discuss the existing approaches. We present our MFDR methodology in Section 3. In Section 4, we implement the approach using options on the S&P 500 index and its sectors and document the properties of the option-implied dependence. Section 5 is dedicated to a detailed investigation of the CRP. Section 6 concludes and discusses further research directions and applications of MFDR.

2 Extracting Information from Options

In this section, we recall how to extract the option-implied probability distribution for a single asset and discuss how index options can provide information about the dependence among several assets. We then review existing approaches to constructing option-implied measures of dependence, namely implied correlations and discuss their underlying assumptions and limitations.

2.1 Univariate Case

For a given underlying asset, let X denote the return over a fixed time period $[t, T]$. Let $C(K)$ and $P(K)$ denote the time-t price of the European-style call and put options with moneyness K and maturity T written on the asset's *return* X. For simplicity, we assume that the asset pays no dividends and that the risk-free rate is zero.[3]

Under the standard assumptions, the option prices are equal to the expected value of their payoffs under a suitably chosen *risk-neutral* probability measure $\mathbb{Q}$:

$$C(K) = E^{\mathbb{Q}}\left[(X-K)^+\right] = \int_0^\infty (x-K)^+ f(x)\,dx,$$

$$P(K) = E^{\mathbb{Q}}\left[(K-X)^+\right] = \int_0^\infty (K-x)^+ f(x)\,dx,$$

where $f(x)$ denotes the *risk-neutral density* (RND). The RND satisfies the relationship first established by Ross (1976), Breeden and Litzenberger (1978), and Banz and Miller (1978):

$$f(x) = \left.\frac{\partial^2 C(K)}{\partial K^2}\right|_{K=x} = \left.\frac{\partial^2 P(K)}{\partial K^2}\right|_{K=x} \tag{1}$$

Similarly, the *risk-neutral cumulative distribution* (RNCD) satisfies

$$F(x) = \left.\frac{\partial C(K)}{\partial K}\right|_{K=x} = \left.\frac{\partial P(K)}{\partial K}\right|_{K=x} \tag{2}$$

Although not directly observable, the RNCD can be recovered using the relationship in (2), provided that options with a continuum of strikes K are available. In practice, options are only available for a finite number of strikes. Nevertheless, a number of efficient nonparametric approaches have been proposed in the literature that make it possible to circumvent this shortcoming; see, for example, Jackwerth and Rubinstein (1996), Aït-Sahalia and Lo (2000), and Bondarenko (2003). Note that for our approach, we only need the information about the RNCD and not the RND. The former can be estimated considerably more accurately – this is because the so-called curse of differentiation is not as severe when estimating the first, rather than the second, derivative of a function.

2.2 Implied Dependence of Several Assets

To obtain the joint distribution of several assets, the knowledge of individual marginal distributions is not enough. We also need their dependence. Dependence is implicit in the prices of multivariate options, such as index options. Consider an index comprising d assets. Let $X_{j,t+\tau}$

[3] In the empirical application, we use ETF options on the S&P 500 index and its industry sectors. These options are American-style and are written on an asset's price, not the return. Moreover, the underlying ETFs do pay quarterly dividends, and of course, the risk-free rate is not really zero. We address these real-world complications as follows. First, we convert *spot* prices of the options and the underlying asset into *forward* prices (for delivery at time-T). The forward prices account for dividends and the non-zero risk-free rate. Second, we adjust option prices for the early exercise feature by following the approach of Barone-Adesi and Whaley (1987). This gives us the prices of equivalent European-style options. Third, we rescale option prices and their strikes by the asset's forward price. This gives us options on the asset's *return*.

denote the return of the jth asset for the period $[t, t+\tau]$ for some fixed horizon τ (e.g., three months). The return of the index can be written as

$$S_{t+\tau} = \sum_{j=1}^{d} \omega_j X_{j,t+\tau}, \qquad \sum_{j=1}^{d} \omega_j = 1,$$

where $\omega_1, \ldots, \omega_d$ are the weights. When there is no confusion, we simplify the notation and drop the time index. For example, we write X_j instead of $X_{j,t+\tau}$ and S instead of $S_{t+\tau}$. We assume that the options market offers a sufficient range of strikes so that the risk-neutral distributions $F_S, F_1, \ldots, F_d$ for the returns $X_S, X_1, \ldots, X_d$ can be estimated accurately.

Our goal is to find a dependence structure (or copula) C that can explain the distribution of the index. Specifically, for a given copula C, we can define a new random variable:

$$Z := Z(C) = Z(C; F_1, \ldots, F_d, \omega_1, \ldots, \omega_d) := \sum_j \omega_j X_j, \tag{3}$$

where $(X_1, X_2, \ldots, X_d)$ has a joint distribution fully described by the copula C and the respective margins F_j. We can think about the random variable Z as the weighted return of the d components, or replicated index return. Ideally, we would like to find a copula C such that the weighted return Z is equal in distribution to the observed index return S:

$$Z \stackrel{d}{=} S. \tag{4}$$

This is a difficult problem, as an equality in distribution imposes infinitely many constraints on the choice of the dependence structure C. We could start by matching various moments of Z and S. We note that the risk-neutral densities have a mean of 1 (when X_j and S are defined as gross returns); thus, the first central moment of Z and of S are automatically matched:

$$E[Z] = \sum_j \omega_j E[X_j] = \sum_j \omega_j = 1 = E[S].$$

Matching the second central moment leads to the standard identity between the variance of the index and the variances of its components:

$$\mathrm{var}(S) = \sum_{j=1}^{d} \omega_j^2 \mathrm{var}(X_j) + 2 \sum_{j=1}^{d-1} \sum_{j<k} \omega_j \omega_k \sqrt{\mathrm{var}(X_j)} \sqrt{\mathrm{var}(X_k)} \rho_{jk}, \tag{5}$$

in which ρ_{jk} is the correlation between X_j and X_k. Most existing approaches rely on the identity in (5) and additional auxiliary assumptions, e.g., the assumption of equal pairwise correlations, $\rho_{jk} = \rho$. We review them in the next subsection and then present our approach in Section 3.

2.3 Existing Approaches

The CBOE Implied Correlation Index is an attempt to estimate the average pairwise correlation among the stocks in the S&P 500 index.[4] The basic idea of the CBOE index is to use the condition in (5) but to approximate the variances by the squares of the ATM Black-Scholes

[4]The underlying methodology is detailed in the white paper, Chicago Board Options Exchange (2009).

implied volatilities (BSIV, or simply IV) and to replace the different ρ_{jk} with a single correlation parameter that we denote as ρ_{cboe}. Then

$$\rho_{cboe} = \frac{\sigma_S^2 - \sum_{j=1}^{d} \omega_j^2 \sigma_j^2}{2 \sum_{j=1}^{d-1} \sum_{k>j} \omega_j \omega_k \sigma_j \sigma_k}. \tag{6}$$

Assuming that this approximation is exact, the expression of ρ_{cboe} in (6) can be rewritten as the weighted average pairwise correlation:

$$\rho_{cboe} = \frac{\sum_{j<k} \omega_j \omega_k \sigma_j \sigma_k \rho_{jk}}{\sum_{j<k} \omega_j \omega_k \sigma_j \sigma_k}. \tag{7}$$

However, since this approximation is in fact *not* exact, this interpretation no longer holds. The CBOE index cannot be viewed as a genuine correlation, it may be very different from the true average correlation, and it can potentially take values that are greater than 1.[5]

The academic literature has proposed several modifications to the CBOE methodology. Most notably, Driessen, Maenhout, and Vilkov (2009, 2013) improve the CBOE approach in two ways. First, they apply the condition in (5) to all components and not a small subset. Second, for the standard deviations of X_j and S they use the model-free implied volatility (MFIV).[6] The variances are now driven not by the ATM options only, but are based on the whole cross-section of options. For MFIVs, the condition in (5) holds exactly under the assumption of diffusion, but only approximately if prices could jump (which is likely in practice).

More recently, Buss, Schönleber, and Vilkov (2017, 2019a) use standard deviations for RNDs, $\sigma^{\mathbb{Q}}$, for which the condition in (5) now holds identically.[7] Furthermore, Buss, Schönleber, and Vilkov (2017) relax the equal pairwise correlation constraint by estimating a block diagonal correlation matrix with two correlation parameters. Buss and Vilkov (2012) replace the assumption of constant pairwise correlations $\rho_{jk} = \rho$ with a linear specification, which ensures a negative CRP ($\rho^{\mathbb{Q}} > \rho^{\mathbb{P}}$) that is is higher in magnitude for stocks with low or negative correlation consistently with the empirical evidence in Mueller, Stathopoulos, and Vedolin (2017) for the FX market.

A rather different approach to assess the risk-neutral dependence among several assets is proposed by Kelly, Lustig, and Van Nieuwerburgh (2016). It is based on the idea that a portfolio of individual options is always more expensive than an option on the portfolio, where the strikes are chosen appropriately. The spread between the two is larger when the correlation among the assets is lower (see also Kelly and Jiang (2014)). Kelly, Lustig, and Van Nieuwerburgh (2016) find that during the 2007–2009 financial crisis, the OTM put options for financial firms were extraordinarily expensive relative to the matched OTM put options for the financial sector

[5]For example, the KCJ index was 100.8 on November 6, 2008; 105.93 on November 13, 2008; and 103.4 on November 20, 2008. We observe that, strictly speaking, the condition in (5) does not hold for ATM IVs. Implied volatilities are standard deviations of log-returns under the assumption that these are normally distributed. Therefore, using implied volatilities in (5) would be justified when $\log(S) = \sum_j \omega_j \log(X_j)$. In reality, the index is an arithmetic, not a geometric, average. Hence, some bias is introduced and we can only state an approximate relation: $\sigma_S^2 \approx \sum_{j=1}^{d} \omega_j^2 \sigma_j^2 + 2 \sum_{j=1}^{d-1} \sum_{k>j} \omega_j \omega_k \sigma_j \sigma_k \widetilde{\rho}_{jk}$ where σ_S and σ_j are the ATM IVs for the index and its components, and $\widetilde{\rho}_{jk}$ is the pairwise correlation of $\log(X_j)$ and $\log(X_k)$. In addition, CBOE uses a subset of $d = 50$ largest components of the index, which introduces another bias: since the selected components are generally less volatile, the average implied correlation tends to be overstated.

[6]See Britten-Jones and Neuberger (2000), Carr and Madan (2001), and Bakshi, Kapadia, and Madan (2003).

[7]Note that $\sigma^{\mathbb{Q}}$ is also related to the concept of the simple model-free implied volatility (SMFIV) of Martin (2017) and can be computed from the *simple variance swap*.

index. Therefore, they conclude that a large amount of aggregate tail risk was missing from the cost of the financial sector crash insurance, likely due to a perceived sector-wide government bailout guarantee. Overall, their approach uses options to infer a particular indicator of tail dependence, but it does not provide the entire joint distribution.

2.4 A More General Approach

One common limitation of the existing approaches is that only partial dependence information is obtained. Furthermore, strong assumptions are typically imposed on the correlation matrix under the risk-neutral measure. The identifying restriction in (5), which equates the index variance to the variance of the portfolio of the components, is clearly insufficient to recover the entire correlation structure – there are many possible ways to satisfy this single restriction.

As stated previously, our approach is more ambitious, as it attempts to find a full dependence (copula) and to do so in a completely model-free fashion. How is this even possible? The key to our approach is that it matches not just one moment in (5) but a *continuum* of moments. Specifically, our method attempts to construct the random variable Z such that it is equal to S almost surely, which is even stronger than the condition in (4) and which implies that for *any* function $g(z, s)$,

$$E[g(Z, S)] = E[g(S, S)]. \tag{8}$$

In this respect, our approach generalizes the existing approaches. Our solution satisfies the second moment condition in (5) as a special case, meaning that our solution will yield exactly the same average global pairwise correlation ρ as the existing approaches. The existing approaches use just one summary statistic from each option-implied distribution F_j (either BSIV, or MFIV, or $\sigma^{\mathbb{Q}}$) and match just one restriction in (5).[8] In contrast, our approach uses the complete information contained in each distribution F_j and attempts to satisfy a continuum of restrictions implied by (8). We discuss in more detail various implications of (8) in Appendix A.

3 Model-Free Dependence Recovery

Recall that $S = \sum_{j=1}^{d} \omega_j X_j$ for some weights ω_j that sum to 1. For a moment, we assume that we observe the $(d + 1)$ risk-neutral distributions (F_j for the return X_j and F_S for the index return).

We want to construct a *discrete* multivariate distribution among the d returns $X_1, \ldots, X_d$. To do so, we first approximate F_j for each asset $j \in \{1, \ldots, d\}$ with a discrete distribution (which can be done to any degree of accuracy) as follows. There are n equiprobable states in which X_j takes the values x_{ij}, $i = 1, \ldots, n$, where the elements x_{ij} are defined as realizations $x_{ij} := F_j^{-1}(\frac{i-0.5}{n})$ $(i = 1, \ldots, n)$. Using this discretization, we represent the multivariate vector

[8] BSIV depends on the price of the ATM option only. MFIV and $\sigma^{\mathbb{Q}}$ use information from the entire cross-section of options with different strikes, but that information is still reduced to a single statistic. Clearly, there are many distributions, otherwise very different, but with identical MFIV (or $\sigma^{\mathbb{Q}}$). The existing approaches are unable to distinguish among them, yielding the same measure of dependence.

In Appendix E, we provide numerical examples, which illustrate that satisfying only condition (5) is not enough to reproduce the observed implied volatility curve of the index, see Table 8 and Table 9

of asset returns $(X_1, X_2, \ldots, X_d)$ by an $n \times d$ matrix:

$$\begin{bmatrix} x_{11} & x_{12} & \ldots & x_{1d} \\ x_{21} & x_{22} & \ldots & x_{2d} \\ \vdots & \vdots & \ddots & \vdots \\ x_{n1} & x_{n2} & \ldots & x_{nd} \end{bmatrix},$$

where the j-th column corresponds to the j-th asset return X_j and the i-th row represents a state of the world in which a joint outcome $(X_1 = x_{i1}, \ldots, X_d = x_{id})$ occurs with probability $1/n$. If one permutes the elements in the j-th column, the marginal distribution of X_j remains unchanged because all realizations are equally likely. In contrast, the dependence of X_j with the other variables X_k is affected, because permutations result in different states.

To simplify exposition, it is convenient to define transformations

$$Y_j := \omega_j X_j, \qquad \text{so that} \qquad y_{ij} := \omega_j x_{ij}.$$

Let G_j denote the distribution of Y_j. We then introduce the following $n \times (d+1)$ matrix $\mathbf{M}$:

$$\mathbf{M} = \begin{bmatrix} y_{11} & y_{12} & \ldots & y_{1d} & -s_1 \\ y_{21} & y_{22} & \ldots & y_{2d} & -s_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ y_{n1} & y_{n2} & \ldots & y_{nd} & -s_n \end{bmatrix}, \tag{9}$$

where the last column consists of the elements $-s_i$, which are all possible discrete realizations of the negative sum $-S$; i.e., $-s_i := -F_S^{-1}(\frac{i-0.5}{n})$.

3.1 A Toy Example

To explain the method for constructing a joint dependence, we begin with an oversimplified example. There are $d = 3$ assets and $n = 5$ states of the world. Therefore, Y_1, Y_2, Y_3 and $-S$ all take five values with probability $1/5$, which are collected in a matrix:

$$\mathbf{M} = \begin{bmatrix} 1 & 1 & 0 & -19 \\ 2 & 2 & 3 & -13 \\ 3 & 3 & 4 & -10 \\ 5 & 5 & 5 & -8 \\ 6 & 7 & 9 & -6 \end{bmatrix}. \tag{10}$$

The first three columns of matrix $\mathbf{M}$ depict the random vector (Y_1, Y_2, Y_3). Its joint outcomes are displayed in the five rows, each row reflecting one of the five states of the world. The random vector (Y_1, Y_2, Y_3) does not yet describe a compatible dependence, because the five row sums (taken over all four columns) do not equal to zero; i.e., we do not yet meet the constraint that $Y_1 + Y_2 + Y_3 - S = 0$. However, permuting elements within a column is allowed, as it does not affect the marginal distributions. We thus aim at permuting elements within columns to satisfy the condition that $Y_1 + Y_2 + Y_3 - S = 0$ at each of the five states.

With only five rows, it is feasible to try every possible permutation. However, in a realistic situation, this would be impractical, as the number of distinct configurations, $(n!)^d$, would be extremely large. In our empirical application, we employ our method using at least $n = 1,000$

states and, therefore, we need an efficient way to find a candidate solution. To achieve such efficiency, we observe that the condition $Y_1 + Y_2 + Y_3 - S = 0$ is equivalent to the condition that the random variable $Y_1 + Y_2 + Y_3 - S$ has zero variance. Clearly, to minimize the variance of $Y_1 + Y_2 + Y_3 - S$, it must hold that Y_1 is as negatively correlated as possible with $Y_2 + Y_3 - S$, which means that the elements of the first column of the matrix $\mathbf{M}$ in (10) should appear in opposite order (be antimonotonic) to those that correspond to $Y_2 + Y_3 - S$. Since permuting (rearranging) values within columns does not affect the marginal distributions, we rearrange the values in the first column to achieve this situation.

Let us illustrate this fundamental principle using the first column of the matrix in (10). We rearrange this column such that its realizations are placed in opposite order to the realizations of $Y_2 + Y_3 - S$. After this step, we obtain the matrix $\mathbf{M}^{(1)}$:

$$Y_2 + Y_3 - S = \begin{bmatrix} -18 \\ -8 \\ -3 \\ 2 \\ 10 \end{bmatrix} \qquad \mathbf{M}^{(1)} = \begin{bmatrix} 6 & 1 & 0 & -19 \\ 5 & 2 & 3 & -13 \\ 3 & 3 & 4 & -10 \\ 2 & 5 & 5 & -8 \\ 1 & 7 & 9 & -6 \end{bmatrix}. \tag{11}$$

For the starting configuration $\mathbf{M}$, the variance of row sums is $\text{var}(Y_1 + Y_2 + Y_3 - S) = 126$. After rearranging the first column, we obtain $\mathbf{M}^{(1)}$ and $\text{var}(Y_1 + Y_2 + Y_3 - S) = 58$ has been strictly decreased. We have not found a solution yet, but we are getting one step closer. We now repeat the same process for each of the subsequent columns of the matrix. We can further improve this procedure by noting that in order to yield zero row sums, we actually need $Y_1 + Y_2$ to be antimonotonic to $Y_3 - S$ and, likewise, $Y_1 + Y_3$ to be antimonotonic to $Y_2 - S$, and $Y_2 + Y_3$ to $Y_1 - S$. The application of this procedure to the toy example is fully described in Appendix B.1. After 12 steps, the procedure converges and yields the following output:

$$\widetilde{\mathbf{M}} = \mathbf{M}^{(12)} = \begin{bmatrix} 5 & 5 & 9 & -19 \\ 3 & 3 & 0 & -6 \\ 1 & 7 & 5 & -13 \\ 6 & 1 & 3 & -10 \\ 2 & 2 & 4 & -8 \end{bmatrix}. \tag{12}$$

We thus have arrived to the ideal situation in which the row sums of the rearranged matrix are all equal to zero; i.e., we have found an admissible multivariate model for the assets (Y_1, Y_2, Y_3) which is consistent with the distribution of their sum S and which can now be used to compute various statistics of interest. For instance, we can find the conditional probability $P(Y_2 > 3 | Y_1 > 3) = 0.5$. To specifically study the dependence among the three assets, we can remove the effect of the marginal distributions by applying the transformation $Y_j \to G_j(Y_j)$, $j = 1, 2, 3$. Then we obtain the following discrete dependence structure:

$$\widetilde{\mathbf{U}} = \begin{bmatrix} 0.8 & 0.8 & 1 \\ 0.6 & 0.6 & 0.2 \\ 0.2 & 1 & 0.8 \\ 1 & 0.2 & 0.4 \\ 0.4 & 0.4 & 0.6 \end{bmatrix}. \tag{13}$$

In practice, due to discretization errors and the fact that the algorithm is a heuristic, the variance of the row sums might not be exactly equal to zero. However, empirically the procedure usually performs extremely well. Although the final row sums do deviate from zero, deviations are trivial for all practical purposes and could be safely ignored. A formal exposition of the BRA method is relegated to Appendix B.2.

3.2 Illustration of the BRA for $d = 2$

The previous subsection explains the BRA methodology using a very simple example with only five states. This toy example allows us to trace every step of the BRA procedure. However, the example does not resemble anything close to what we would like to do in real applications. Therefore, to further develop intuition we provide another illustration of the algorithm, now using continuous margins which are discretized into a large number of states. This is still an unrealistic, but pedagogical example. It uses only two assets and allows us to easily visualize the impact on the dependence of altering the distribution of their weighted sum. The complete details of the example are delegated to Appendix B.3.

Specifically, we assume that the two returns X_1 and X_2 are normally distributed with standard deviations of 0.2 and 0.4, respectively. Their margins are thus fixed. We then consider three cases for the distribution of their weighted sum, $S = \frac{1}{2}X_1 + \frac{1}{2}X_2$. In the first two cases, S is also normally distributed with a variance chosen such that the implied correlation is equal to either zero (no dependence, the first and second panels in Figure 13) or 0.97 (strong dependence, the first and second panels in Figure 14), respectively. In the last case, S has a skewed distribution with a heavy left tail (asymmetric dependence, the first and second panels in Figure 15). To run the BRA, we discretize the marginal distributions of X_1, X_2 and S into $n = 1,000$ equiprobable states. The resulting joint distribution is represented by $1,000$ dots corresponding to pairs (X_1, X_2) in the fourth panel of each figure (which in turn correspond to the $1,000$ rows of the output matrix from the BRA). In Figure 13, the two variables X_1 and X_2 appear independent, which is expected when an implied correlation is equal to zero. In Figure 14, the implied dependence is strong and symmetric. This is also intuitive because the distribution of the sum is also symmetric. Finally, when the sum has a skewed distribution, the joint distribution is highly asymmetric, with a very pronounced left tail dependence, as can be seen in the third and fourth panels of Figure 15.

3.3 Discussion and Limitations of the New Approach

Intuitively, our approach can be compared to GMM. Asset returns are permutated in such a way as to keep the margins fixed and to enforce the "moment" condition $S = \sum_{j=1}^{d} Y_j$ state-by-state. Effectively, there are as many restrictions as states n, which could be chosen as large as desired. The approach finds a compatible joint distribution, which matches every available option on the assets and the index. Although there are many restrictions being enforced, the solution is, typically, not unique. This situation is not uncommon – when the market is incomplete, there are many risk-neutral measures which correctly price available options. However, out of many possible ones, the BRA finds an economically most sensible solution. As shown in Bernard, Bondarenko, and Vanduffel (2018), the BRA procedure has an important property: the obtained multivariate model for $(Y_1, Y_2, \ldots, Y_d)$ exhibits maximum entropy. This means that the procedure yields the "most likely" configuration given the information available and given

that no additional information is used.[9]

To make clear what is meant by "most likely," let us take a step back and assume for a moment that we only have information about the marginal distributions of the assets, and not the sum. That is, we only agree on the values that appear in the first d columns but not on the order in which they appear. Consequently, all permutations within columns are equally plausible, and there is no reason to prefer one permutation over another. Hence, randomizing the assignment of realizations to the different states leads to marginal distributions (reflected by the columns) that are most likely to be independent, which corresponds precisely to the maximum-entropy case. Suppose now that the additional information is known, namely, the marginal distribution of the sum. In this case, the set of admissible permutations is simply reduced to those that yield row sums that are zero. The BRA method implements the idea of randomizing the assignment of realizations to the different states but now under the additional constraint provided by the knowledge of the distribution for the index.[10]

In sum, our approach amounts to constructing a numerical dependence and we refer to it as MFDR, or Model-Free Dependence Recovery. The approach consists of three main steps. First, from traded options on the individual assets and the index, we estimate the corresponding risk-neutral marginal distributions. Second, we discretize the estimated marginal distributions into n states and reduce the problem of finding a joint distribution among d assets to finding a suitable permutation of an $n \times (d+1)$ matrix.[11] Finally, we solve the resulting integer optimization problem with the BRA.

A natural question arises about the maximum possible value for d that can be handled. In fact, the practical limitation does not really arise from the methodology itself but rather from the problem at hand that we want to solve. Observe that the larger d is, the less informative the distribution of a weighted sum of the d components is. Intuitively, the larger d is, the more flexible the dependence among the components is to fit the distribution of the index. When it exists (absence of arbitrage in the options market), the BRA makes it possible to provide one possible model but this model is not unique, as the knowledge of the distribution of the index is not enough to specify the joint model among the components. Practically speaking, there are two further limitations to consider d very large. The first limitation is the difficulty arising from the model-free estimation of the marginal distributions. The second limitation is of course the time for BRA to converge when d is large. A possible step to reduce it is for instance to consider a reduced number n of states. An alternative way to significantly reduce the time to convergence of the BRA is to choose a random subet of all partitions into 2 sets of columns instead of all of them. Considering the standard RA for instance will lead to a loss of accuracy but will guarantee a linear complexity of the algorithm. The latter can then handle very large d, such as 500 stocks. The loss of accuracy is significant when studying a specific element of the copula such as a given pairwise correlation, but is negligible when studying aggregate quantities, such as the down and up correlation that we estimate later. Finally, remember that

[9] In this regard, it is also worth citing Jaynes (2003, p. 370), who developed the principle of maximum entropy in its modern form and who stated the following: *"In summary, the principle of maximum entropy is not an oracle telling which predictions must be right; it is a rule for inductive reasoning that tells us which predictions are most strongly indicated by our present information."*

[10] In the context of extracting RNDs, the principle of maximum entropy has been explored in Rubinstein (1994), Jackwerth and Rubinstein (1996), and Stutzer (1996).

[11] Bernard, Bondarenko, and Vanduffel (2021) discuss an alternative formulation for finding a joint distribution based on the optimal transport (OT) problem. However, they conclude that for the problem at hand the OT approach would not be computationally feasible.

using marginal distributions information, the BRA returns a model that fits index options prices and the components options prices without making additional assumptions (such as constant pairwise correlation). Thus it is in any case better than any approaches in the literature.

4 Empirical Application

In this section, we implement our method to study option-implied dependence. We use daily closing prices of options on the nine SPDR Select Sector funds and on the SPDR S&P 500 Trust. These nine sector ETFs are capitalization-weighted portfolios comprising all stocks in the S&P 500 index. Option data is obtained directly from CBOE. Table 1 lists abbreviated names for the ETFs.

Description	Ticker	Abbreviation
Consumer Discretionary Sector SPDR Fund	XLY	cdi
Consumer Staples Sector SPDR Fund	XLP	cst
Energy Sector SPDR Fund	XLE	ene
Financial Sector SPDR Fund	XLF	fin
Health Care Sector SPDR Fund	XLV	hea
Industrial Sector SPDR Fund	XLI	ind
Materials Sector SPDR Fund	XLB	mat
Technology Sector SPDR Fund	XLK	tec
Utilities Sector SPDR Fund	XLU	uti
SPDR S&P 500 ETF Trust	SPY	spx

Table 1: **S&P 500 Sectors.** This table lists the underlying assets used in this study. The sample period is from January 1, 2007, to June 12, 2017.

Our sample covers the period from January 1, 2007 to June 12, 2017. Although options on sector ETFs were also traded in prior years, the availability of strikes and maturities was limited. By 2007, the liquidity of sector options has improved considerably, making it feasible to accurately estimate their RNDs on a day-to-day basis. We focus on horizon $\tau = 3$ months and infer the option-implied joint distribution of the nine sectors.[12] The details of the estimation procedure are provided in Appendix C. The realized volatilities and correlations are computed using daily returns of the ETFs, which are obtained from CRSP. The daily sector weights ω_j are obtained from Bloomberg. The risk-free rate is approximated by the rate of Treasury bills.

We first briefly discuss the main characteristics of the sectors and of the aggregate market over the studied period. Figure 1 plots the cumulative returns and the 3-month ATM implied volatilities for the S&P 500 and for three specific sectors (fin, tec, and ene). The performance of individual sectors often differs considerably. During the great financial crisis (defined in this paper as the period from August 1, 2007 to April 1, 2009), the financial sector starts to decline much earlier than other sectors, but from the last quarter of 2008 onwards the crisis spreads to the rest of the economy. During the 2015 energy crisis, it is the energy sector that experiences a considerable decline, while the other sectors are mostly unaffected. As expected, the ATM

[12]The choice of $\tau = 3$ months represents a practical compromise. The horizon is long enough to allow accurate estimation of the realized correlations. At the same time, it is short enough so that sector options are still liquid.

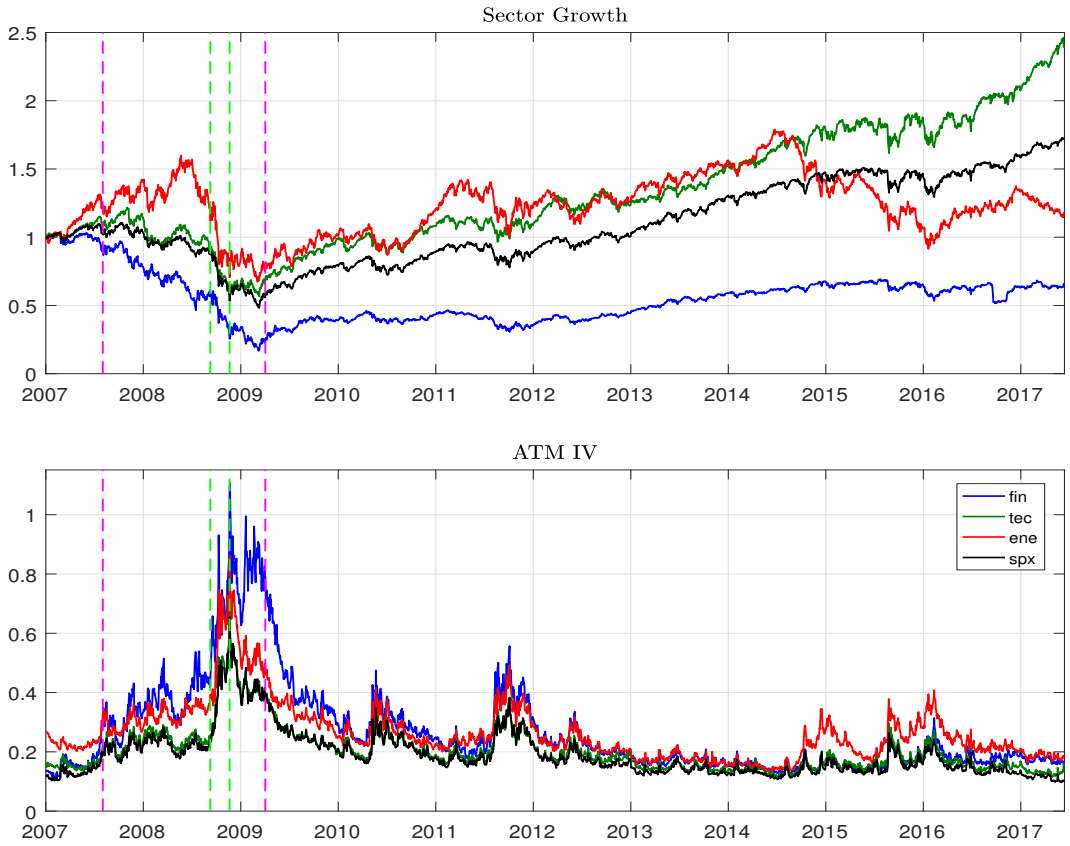

Figure 1: **Cumulative Returns and ATM Implied Volatilities.** Implied volatilities correspond to maturity of 3 months. The pink vertical lines indicate the financial crisis (August 1, 2007, to April 1, 2009). The green vertical lines show two selected days: September 8, 2008, and November 20, 2008. Shown are three sectors (fin, tec, ene) as well as the S&P 500 index (the black line).

implied volatilities (IV) are highly correlated over time. The IV of the financial sector is much larger than of the other sectors in 2008–2010, whereas the IV of the energy sector stands out in 2014–2015.

Table 2 provides descriptive statistics for the nine sectors and the S&P 500 index. In particular, it shows that the largest sectors of the S&P 500 index are tec, fin, and hea, with the median portfolio weights of 21.6%, 16.0%, and 13.0%, respectively. The smallest sectors are mat and uti, with the median weights of 3.3% and 3.4%, respectively. Some sectors experience considerable variation in the weight ω_j. For example, the weight for fin ranges from 8.6% to 22.4% over the sample period, while the weight for ene ranges from 5.9% to 16.2%. Table 2 also reveals that the ATM IVs are consistently smaller than the standard deviations of the estimated risk-neutral densities ($\sigma^{\mathbb{Q}}$). This pattern reflects nonnormality of risk-neutral densities and manifests itself in the pronounced volatility skews. Consistent with the literature on the variance risk premium (VRP) (see Bollerslev, Tauchen, and Zhou (2009), Todorov (2010), Carr and Wu (2009), Bondarenko (2014)), we observe evidence of a negative VRP, as the standard deviations under the risk-neutral measure are consistently larger than under the real-world measure (this holds true for the index and most sectors, with the only exceptions being fin and ene, due to extreme realized volatilities during the financial and energy crises).

	Mean	$\sigma^{\mathbb{P}}$	Skew	Kurt	Corr	Beta	SR	Weight Min	Weight Med	Weight Max	IV	$\sigma^{\mathbb{Q}}$
mat	0.109	0.246	-0.00	8.28	0.87	1.10	0.42	0.026	0.033	0.039	0.242	0.255
ene	0.092	0.287	-0.11	12.20	0.82	1.22	0.30	0.059	0.106	0.162	0.263	0.273
fin	0.083	0.341	0.41	16.19	0.86	1.50	0.22	0.086	0.160	0.224	0.268	0.285
ind	0.121	0.213	-0.02	7.89	0.92	1.01	0.54	0.095	0.106	0.122	0.213	0.226
tec	0.130	0.199	-0.04	9.73	0.92	0.94	0.62	0.184	0.216	0.256	0.202	0.215
cst	0.105	0.135	-0.28	7.04	0.80	0.55	0.73	0.091	0.106	0.141	0.148	0.165
uti	0.078	0.179	0.49	14.48	0.70	0.64	0.40	0.028	0.034	0.047	0.180	0.192
hea	0.120	0.165	0.27	12.74	0.81	0.69	0.68	0.106	0.130	0.170	0.174	0.189
cdi	0.131	0.214	-0.05	7.80	0.89	0.98	0.58	0.076	0.109	0.132	0.213	0.229
spx	0.104	0.194	0.25	16.22	1.00	1.00	0.50				0.189	0.203

Table 2: **Summary Statistics.** This table reports time series averages over our sample for the nine sectors and for the aggregate market (S&P 500). The first seven columns are computed from daily returns and include the mean, standard deviation $\sigma^{\mathbb{P}}$, skewness, kurtosis, correlation with the market (Corr), beta with the market, and the Sharpe ratio (SR). The next three columns are the minimum, median, and maximum for sector weight ω_j. The last two columns report $\mathbb{Q}$ statistics computed from 91-day options. They include the option ATM implied volatility (IV) and standard deviation of the RND. When applicable, statistics are reported in annualized form and as decimals.

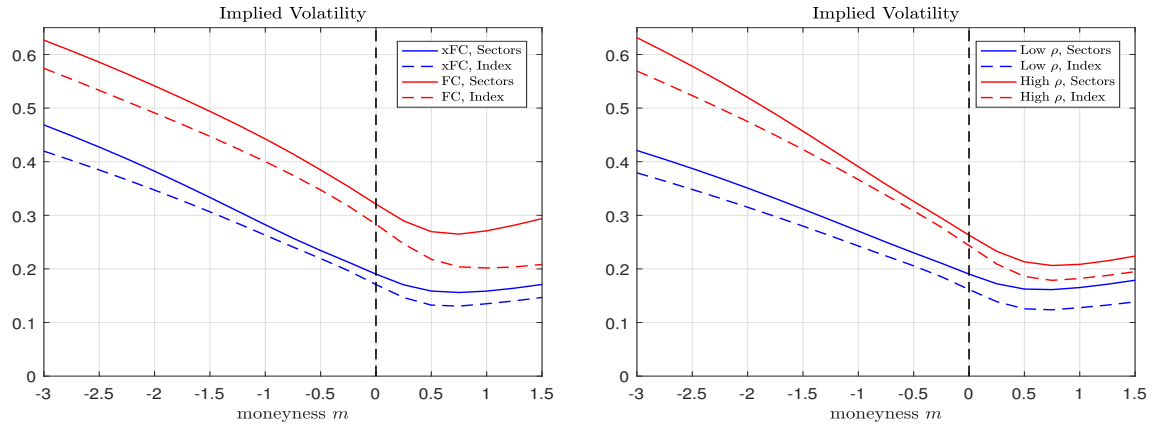

Figure 2: **Implied volatility.** The average implied volatilities of the nine sectors (solid lines) and the aggregate market (dashed lines) are shown for the nomalized moneyness m. Time to maturity $\tau = 3$ months. For the sectors, implied volatility smiles are value-weighted for each day. The left panel shows the averages computed separately for the financial crisis (FC, August 1, 2007 to April 1, 2009, red) and for the period excluding the crisis (xFC, blue). The right panel shows the averages computed for trading days when the realized correlation is low (blue) and high (red). Specifically, we first compute the weighted average pairwise correlation for the nine sectors over the 3-month trailing window. We then select the bottom and top quartiles and compute average implied volatilities for both groups.

To provide initial intuition, Figure 2 plots the average implied volatility smile for the nine sectors and for the index. The implied volatility smile corresponds to $\tau = 3$ months and is shown as a function of the normalized moneyness

$$m := \frac{\log \frac{K}{Z}}{\sigma^{ATM}\sqrt{\tau}}, \tag{14}$$

where σ^{ATM} is the ATM IV and Z is the forward price. To construct the figure, we compute the daily volatility smiles for the nine sectors and value-weight them. The average sector volatility smile and the index volatility smile are then averaged across various subsamples. In the left panel, we isolate the financial crisis (defined as the period from August 1, 2007, to April 1, 2009). Generally, the volatility smile for the sectors is higher than for the index, although the gap between the two is relatively narrow, indicating a high correlation between the sectors. During the financial crisis, as expected, the implied volatilities are much higher. It is also noteworthy that the gap between the two smiles is wider for positive m, implying a potentially lower correlation on the upside. In the right panel, we distinguish between days when the realized correlation is low or high. Specifically, we first compute the average pairwise realized correlation for the nine sectors over the trailing 3-month window. We then select the bottom and top quartiles and compute average implied volatilities for each group. For the high correlation quartile, the two volatility smiles are much steeper and the gap between the two smiles is more narrow, especially, for the middle and high values of m.

4.1 Three Types of Correlations

To describe the option-implied dependence of the nine sectors, we compute three types of pairwise correlations. The first type is the standard (or global) correlation

$$\rho_{j,k}^{\mathbb{Q}} = \operatorname{corr}^{\mathbb{Q}}\left(X_j, X_k\right), \tag{15}$$

which is the Pearson correlation coefficient between the returns for sectors X_j and X_k over the period $[0, \tau]$. The other two types are *down* and *up* correlations, or correlations conditional on the S&P 500 having a low or high return, respectively. Specifically,

$$\rho_{j,k}^{d,\mathbb{Q}} = \operatorname{corr}^{\mathbb{Q}}(X_j, X_k \,|\, S \leqslant S^m), \tag{16}$$

$$\rho_{j,k}^{u,\mathbb{Q}} = \operatorname{corr}^{\mathbb{Q}}(X_j, X_k \,|\, S > S^m), \tag{17}$$

where S^m denotes the median return of the S&P 500 index. In Section 5, we use the down and up correlations to better understand the nature of the correlation risk premium.

Since there are many sector pairs ($\frac{1}{2}d(d-1) = 36$), it is often convenient to work with weighted average correlations. Specifically, for positive weights π_j, we define

$$\rho^{a,\mathbb{Q}} = \frac{\sum_{j<k} \pi_j \pi_k \rho_{j,k}^{a,\mathbb{Q}}}{\sum_{j<k} \pi_j \pi_k}, \tag{18}$$

where the superscript $a \in \{g, d, u\}$ indicates the type of the correlation (global, down, or up). However, if there is no confusion, we often drop the superscript g when discussing the usual, global correlation. There are several sensible choices for the weights π_j, including $\pi_j = 1/d$ (equally-weighted) or $\pi_j = \omega_j$ (value-weighted). Here, we focus on the "risk-weighted" averaging

$$\pi_j = \omega_j \sigma_j^{\mathbb{Q}}, \tag{19}$$

where $\sigma_j^{\mathbb{Q}}$ is the standard deviation of the risk-neutral density. This case has been used in most existing approaches. In particular, assuming constant pairwise correlations, the average global correlation $\rho^{\mathbb{Q}}$ with the weights in (19) can be computed from the volatilities of the components

and the index without our MFDR but using (6) instead. Thus, focusing on this case allows for direct comparison with the existing literature.[13] However, the findings on the correlation risk premia (CRP) that we present in Section 5.2 are robust to the choice of the weights. When studying the CRP we need not only the risk-neutral correlations, but also their real-world counterparts, $\rho^{g,\mathbb{P}}$, $\rho^{d,\mathbb{P}}$, and $\rho^{u,\mathbb{P}}$. The latter are computed from the same equations (15)-(19) but using the realized correlations instead of the option-implied correlations.[14] The realized correlations are obtained from the daily ex-dividend returns of the underlying ETFs over the trailing 3-month window.[15] More precisely, the window is equal to 63 trading days, which approximately corresponds to 3 calendar months, or 91 calendar days.

4.2 Implied Dependence for Selected Days

We now illustrate our approach on two selected dates in the midst of the financial crisis: September 8, 2008 and November 20, 2008. The first date represents a relatively calm period, while the second represents an extremely turbulent one.

For each day, we proceed as follows. We use 91-day options to obtain the risk-neutral marginal distributions F_j and F_S for the nine sectors X_j and the index S. We discretize the $(d+1)$=10 distributions into $n = 1,000$ states, collect them in an $n \times 10$ matrix, and apply the method described in Section 3. The output of the BRA is another $n \times 10$ matrix that describes a joint model compatible with all marginal distributions. Since the full joint distribution is a 10-dimensional object, we need to make some choices about how to display it. Hence, from the output matrix, we extract triplets (x_i, y_i, z_i) for $i = 1, \ldots, n$ to examine the dependence among the S&P 500 index (x_i), the Financial sector (y_i), and the Utilities sector (z_i). (Of the nine sectors, we arbitrarily pick two, the most and least dramatic ones.)

We present the results in Figures 3 and 4, where the top panels display dependence for the pair (S,fin) and the bottom panels for the pair (S,uti). We remove the effect of the marginal distributions on the joint distribution and display in the first column the scatterplots $(F_S(x_i), F_{fin}(y_i))$ and $(F_S(x_i), F_{uti}(z_i))$. By doing so, we bring all returns to the same (uniform) scale and obtain a visualization of the true copula. However, it is typically easier to interpret the dependence between *normally* distributed variables instead of *uniformly* distributed ones. Therefore, in the second column we show the scatterplots of transformed variables that are now standard normal (normalized dependence). Specifically, we use the quantile function Φ^{-1} of the standard normal distribution to define transformation $G_S(x) := \Phi^{-1}(F_S(x))$ for the S&P 500 index; $G_{fin}(x)$ and $G_{uti}(x)$ are defined similarly for fin and uti. In the second column, we then show scatterplots for the couples $(G_S(x_i), G_{fin}(y_i))$ and for the couples $(G_S(x_i), G_{uti}(z_i))$. In the third column, we display the corresponding contour plots. When the dependence is normal,

[13] To obtain the correlation risk premium, Driessen, Maenhout, and Vilkov (2009), Buss, Schönleber, and Vilkov (2017, 2019a) have to rely on the relationship in (6) as their approach cannot produce pairwise correlations under $\mathbb{Q}$. A notable exception is Buss and Vilkov (2012) who impose a linear relationship between pairwise implied and realized correlations. They are then free to use an approach based on either (7) or (18).

[14] We also considered alternative definitions for the down and up correlations, where the mean return is used instead of the median S^m as the cutoff. The empirical results were very similar. However, there are important theoretical and practical advantages to defining the conditional correlations with respect to a specific quantile of the index return. In particular, using the median guarantees that the calculation of down and up correlations under $\mathbb{P}$ is always based on an equal number of realizations. Under Q, the down and up correlations can be readily computed because the full joint distribution is known for each trading day.

[15] See Jackwerth and Vilkov (2019) for a discussion of the impact of frequency on the estimation of real-world correlations.

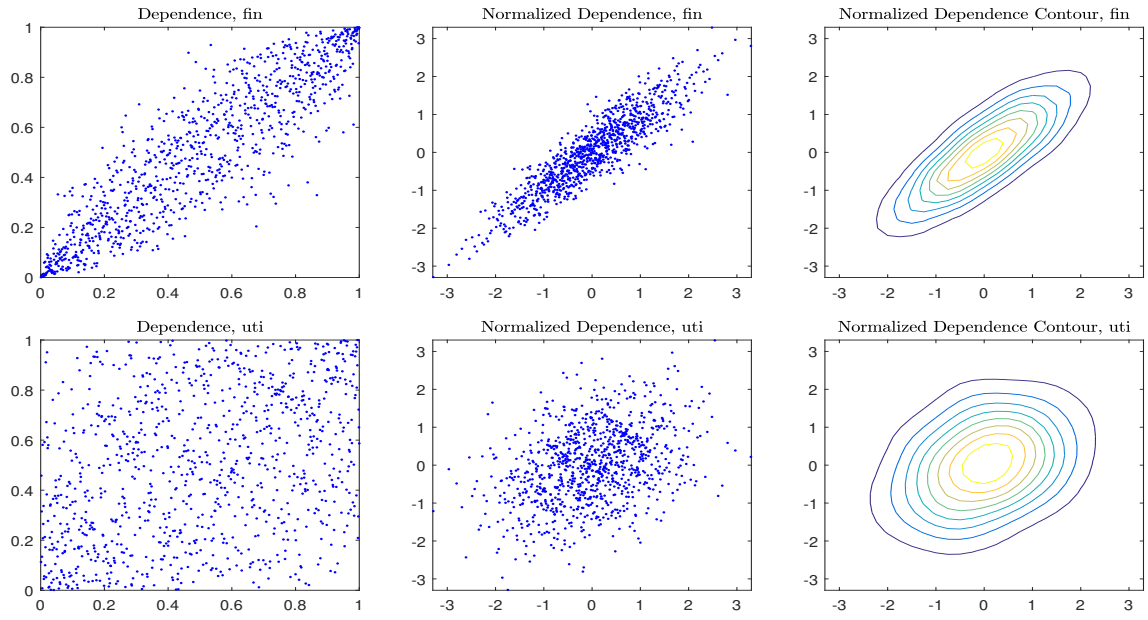

Figure 3: **Implied dependence on September 8, 2008.** The first column shows the dependence of the Financial sector (top panels) and the Utilities sector (bottom panels) relative to the S&P 500 index. The middle column shows the same dependence but after transformation to normally distributed variables. The third column displays the corresponding contour plots.

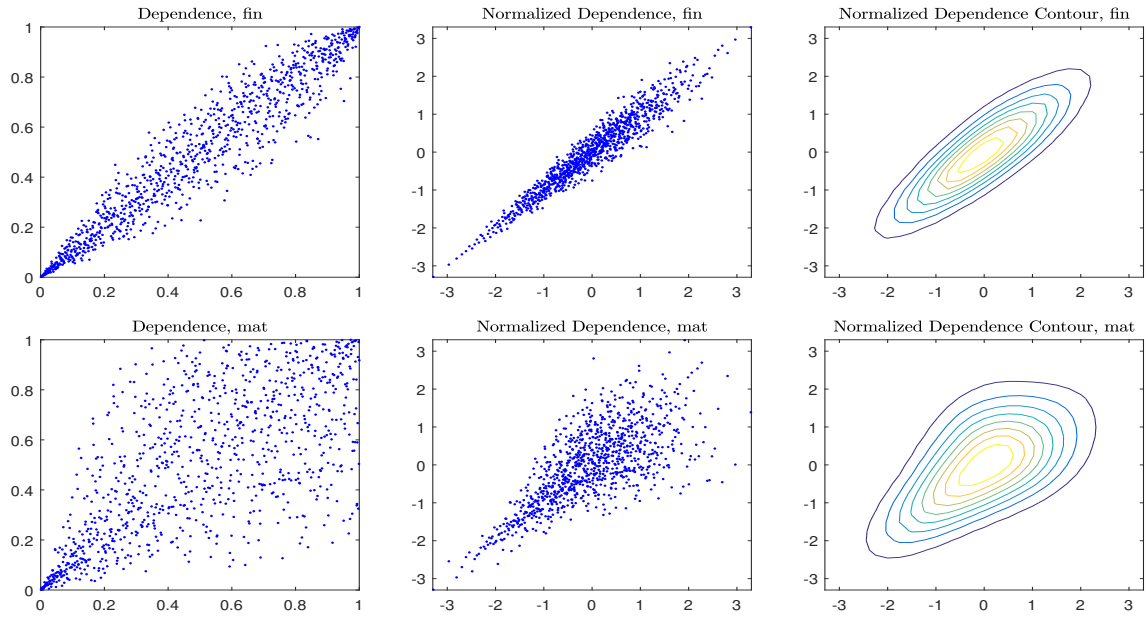

Figure 4: **Implied dependence on November 20, 2008.** The first column shows the dependence of the Financial sector (top panels) and the Utilities sector (bottom panels) relative to the S&P 500 index. The middle column shows the same dependence but for normally distributed variables. The third column displays the corresponding contour plots.

these contours must be perfect ellipsoids. On September 8, 2008, we find positive dependences for both sectors, but the one for the financial sector is much stronger. Both dependencies appear slightly asymmetric, with the left tail being stronger than the right tail. On November 20, 2008, the same trends become much stronger for both sectors: the dependency is now noticeably more pronounced, the asymmetry is very obvious, and the left tail dependence is extreme, even for the "calm" Utility sector.

For the two dates, we show all pairwise correlations $\rho^{\mathbb{Q}}_{j,k}$ on the left panels of Figures 5 and 6. We observe that Financial, Energy, and Technology sectors are highly correlated with themselves and the other sectors. The best diversifiers are Materials and Utilities. The pairwise correlations are higher across the board for the second date compared to the first date. The average global correlation $\rho^{\mathbb{Q}}$ is 0.58 for September 8, 2008, and 0.83 for November 20, 2008.

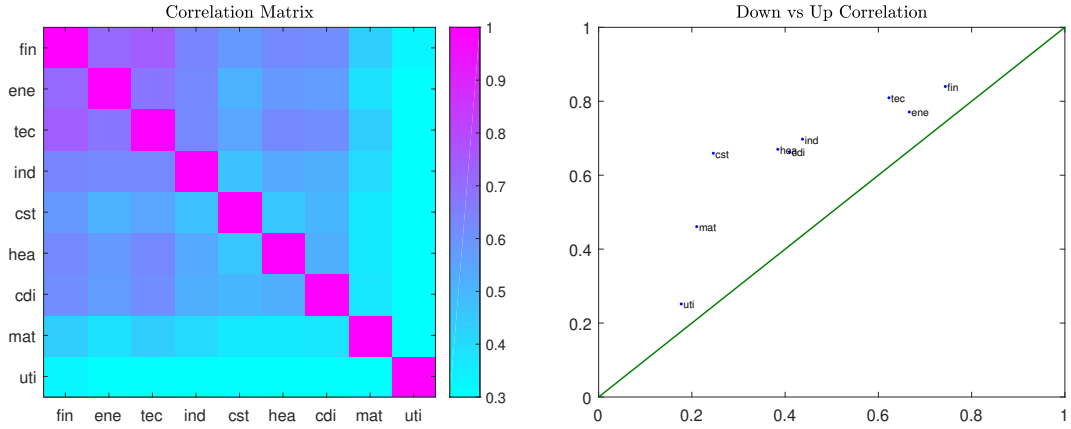

Figure 5: **Implied Correlations for the Nine Sectors on September 8, 2008.** The left panel shows the correlation matrix. The right panel shows the implied down correlation $\rho^{d,\mathbb{Q}}_{j,S}$ (y-axis) versus the implied up correlation $\rho^{u,\mathbb{Q}}_{j,S}$ (x-axis). Also shown is the 45-degree line.

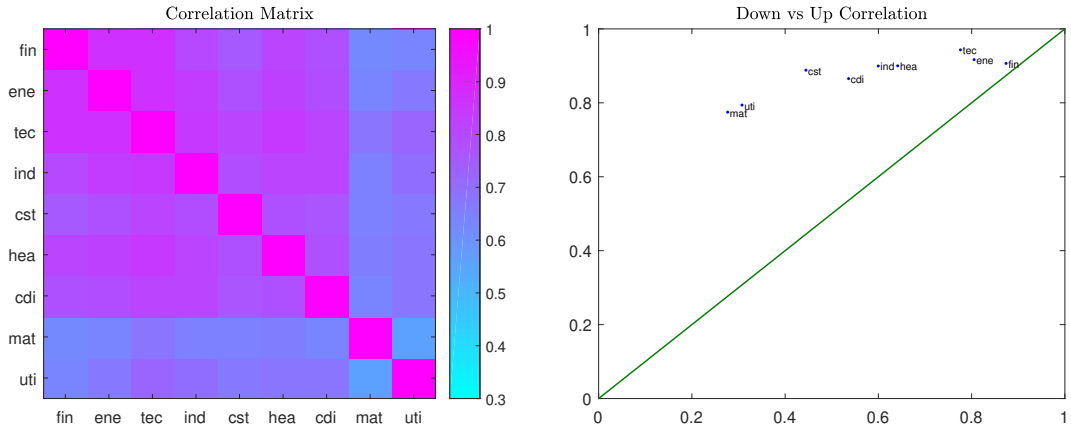

Figure 6: **Implied Correlations for the Nine Sectors on November 20, 2008.** The left panel shows the correlation matrix. The right panel shows the implied down correlation $\rho^{d,\mathbb{Q}}_{j,S}$ (y-axis) versus the implied up correlation $\rho^{u,\mathbb{Q}}_{j,S}$ (x-axis). Also shown is the 45-degree line.

If X_j is replaced with the index S in definitions (16)-(17), we obtain the global, down, and

up correlations $\rho_{j,S}^{\mathbb{Q}}$, $\rho_{j,S}^{d,\mathbb{Q}}$ and $\rho_{j,S}^{u,\mathbb{Q}}$ of sector j with the market. We display the scatter plot of the last two correlations in the right panels of Figures 5 and 6. There are nine points corresponding to the nine sectors. Also shown is the first bisectrix $\rho_{j,S}^{d,\mathbb{Q}} = \rho_{j,S}^{u,\mathbb{Q}}$. From Figures 5 and 6, it is clear that the down correlations tend to be much higher than the up correlations, i.e., $\rho_{j,S}^{d,\mathbb{Q}} > \rho_{j,S}^{u,\mathbb{Q}}$. In fact, the average down and up correlations are 0.44 and 0.17 for September 8, 2008, but 0.83 and 0.46 for November 20, 2008.[16] On both days, the correlation conditional on the market going down is thus considerably stronger than the correlation conditional on the market going up. This feature is not unique for the two selected days and provides a strong indication that the implied dependence for the nine sectors is asymmetric and thus nonnormal. In Section 4.3, we formally assess the extent to which the dependence deviates from normality.

4.3 Asymmetry of the Joint Risk-Neutral Distribution

In this subsection, we first formally test whether the option-implied dependence is consistent with the dependence of a multivariate normal distribution. The classical tests in this regard are Mardia's tests of multinormality and variants thereof. For general multivariate data, Mardia (1970) constructed two statistics for measuring multivariate skewness and kurtosis, which can be used to test the hypothesis of normality (Mardia (1974), Mardia (1975) and Mardia, Kent, and Bibby (1980)). Usually, the test for whether skewness (MS) and kurtosis (MK) are consistent with a normal model are performed separately; however, so-called omnibus tests can assess them simultaneously. In this paper, we perform these two tests on the normalized dependence defined as in the previous subsection. That is, we use option-implied dependence with attached normal margins and assess whether multivariate normality holds. The two tests are performed separately on each trading day in the sample, with the results reported in Table 3. Both tests provide strong evidence that the risk-neutral dependence among assets is not of a normal nature, although the evidence for nonzero skewness is more pronounced. Moreover, given the p-values and the proportion of rejection for the respective tests MS and MK, it is clear that the asymmetry (skewness) is a key feature to reject the normal dependence hypothesis.

	N_{obs}	Mean	$p > 0.01$	$p > 0.05$	$p > 0.10$
MS	2584	0.0010	0.0054	0.0031	0.0015
MK	2584	0.0791	0.2562	0.1935	0.1633

Table 3: **Test for Normal Dependence.** The table reports the results of two multivariate normal tests, MS and MK, which are run for all trading days. Shown are the average p-values and the fraction of the days when the p-value exceeds a threshold of 0.01, 0.05, or 0.10 (that is, when the null hypothesis of normal dependence is not rejected).

This asymmetry of the risk-neutral joint distribution echoes the evidence of asymmetry for the real-world distribution; see Longin and Solnik (2001), Ang and Chen (2002) and Hong, Tu, and Zhou (2006). These authors focus on asymmetric correlations between an individual stock and the market and find evidence of a stronger correlation when the market goes down. Their "down" and "up" states are defined based on the coexceedance criterion, i.e., realizations

[16] The reported average down and up correlations may seem too low when compared to correlations $\rho_{j,S}^{d,\mathbb{Q}}$ and $\rho_{j,S}^{u,\mathbb{Q}}$ shown in Figures 5 and 6. Recall, however, that the average correlations are computed from *pairwise* correlations $\rho_{j,k}^{d,\mathbb{Q}}$ and $\rho_{j,k}^{u,\mathbb{Q}}$, which are lower than $\rho_{j,S}^{d,\mathbb{Q}}$ and $\rho_{j,S}^{u,\mathbb{Q}}$.

that are *jointly* below or above a given threshold. Ang and Chen (2002) is the first paper to implement the formal test for correlation asymmetry. Hong, Tu, and Zhou (2006) subsequently propose a more general test, which is model-free and checks whether a pair is consistent with *any* symmetric distribution. Their test is best suited to a two-dimensional distribution, while we have to deal with 36 sector pairs. Fortunately, a very simple model-free test can be implemented in our setting. For day t, let b_t denote a binary random variable that is equal to one if the average down correlation across the nine sectors exceeds the up correlation (i.e., $\rho^{d,\mathbb{Q}} > \rho^{u,\mathbb{Q}}$) and zero otherwise. Under the null hypothesis of a symmetric distribution, b_t is a Bernoulli random variable with the parameter 0.5, a hypothesis that can be readily tested. The p-value for this test is 0, as the average down correlation exceeds the average up correlation on *every* trading day; see also Figure 7.

5 Correlation Risk

5.1 Implied and Realized Correlations over Time

Figure 7 plots the average implied correlations for each trading day in our sample, where the top panel displays the time series of $\rho^{\mathbb{Q}}$, while the bottom panel displays both $\rho^{d,\mathbb{Q}}$ and $\rho^{u,\mathbb{Q}}$. In particular, this figure makes it clear that the gap between the down and up correlations is always positive and consistently very wide. We observe that the sample average for the global correlation (0.770) is larger than for the down correlation (0.731) and is much larger than for the up correlation (0.332). It is important to stress that the "truncated" correlations are not directly comparable to the global correlation. Due to the conditioning bias, the former are generally closer to zero than the latter. In the multivariate normal setting, this conditioning bias can be shown analytically.[17] Assuming normality and accounting for the conditioning bias as in (20), the above average down and up correlations translate into much higher equivalent global correlations of 0.872 and 0.504, respectively.

Figure 7 reveals that all three correlations drop sharply during the second half of 2008, which might seem counterintuitive, as this period is in the midst of the financial crisis. To explain this feature and to put it in proper perspective, we also produce Figure 8, which compares the average global correlations under both $\mathbb{P}$ and $\mathbb{Q}$ measures. The former is plotted over a longer period since it does not require option prices. First, we observe that over the common period, the two types of correlations behave very similarly. This is reassuring: the forward-looking implied correlation extracted from options using our BRA methodology appears to be closely related to the real-world correlation computed from asset historical returns. Second, we observe that the real-world correlation, which is available over the longer time horizon, displays considerable variation over time and that the magnitude of the fall in the correlations that we observe during 2008 is not that unusual. For example, between January 1999 and October 2000, the realized correlation first dropped from 60% to 8% and then rebounded to nearly 80% two years later.

[17] Under multivariate normality, the following relation between the down (up) correlation $\rho_{j,S}^{d,\mathbb{Q}}$ $\left(\rho_{j,S}^{u,\mathbb{Q}}\right)$ and the global correlation $\rho_{j,S}^{\mathbb{Q}}$ holds (for convenience, we suppress the reference to the measure $\mathbb{Q}$ and the assets j and S):

$$\rho^d = \rho^u = \rho\sqrt{\frac{1-\frac{2}{\pi}}{1-\frac{2\rho^2}{\pi}}}, \tag{20}$$

which immediately implies that $|\rho^d| = |\rho^u| < |\rho|$. This formula is consistent with derivation in Appendix B of Ang and Chen (2002). See also equation (3) in Campbell, Forbes, Koedijk, and Kofman (2008).

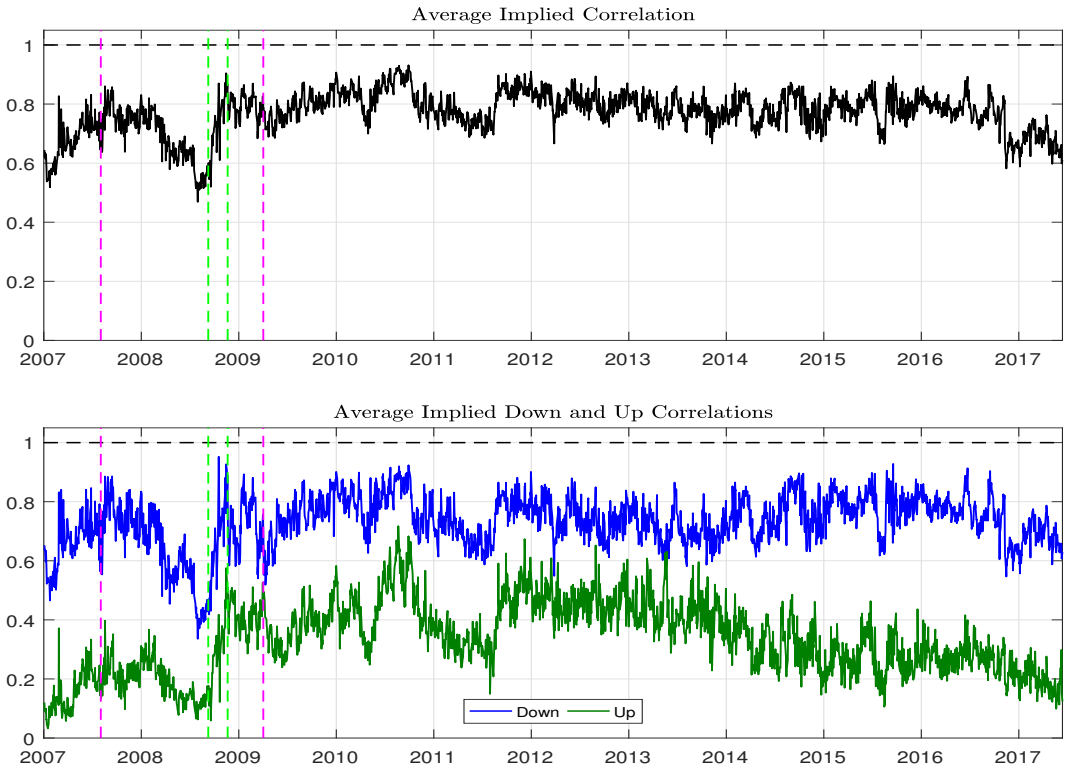

Figure 7: **Average Implied Correlations**. The top panel shows the average implied global correlation $\rho^{\mathbb{Q}}$. The bottom panel shows the average implied down (blue) and up (green) correlations $\rho^{d,\mathbb{Q}}$ and $\rho^{u,\mathbb{Q}}$. The pink vertical lines indicate the period corresponding to the financial crisis. The green vertical lines show two selected days: September 8, 2008, and November 20, 2008.

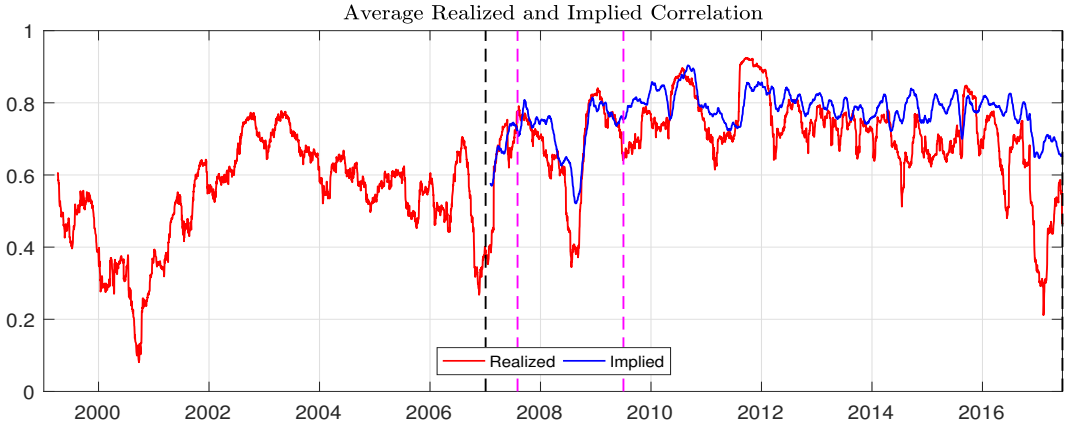

Figure 8: **Average Realized and Implied Correlation.** The realized correlation (red line) is shown for a longer period (starting from January 1999), while the implied correlation (blue line) is shown for a shorter period (starting from January 2007, indicated with the black vertical line) and as a 3-month moving average. The pink vertical lines indicate the period corresponding to the financial crisis.

Next, we aim at explaining the considerable drop in correlations between March 2008 and September 2008. To do so, we split the time period August 2007 - July 2009 into three subperiods

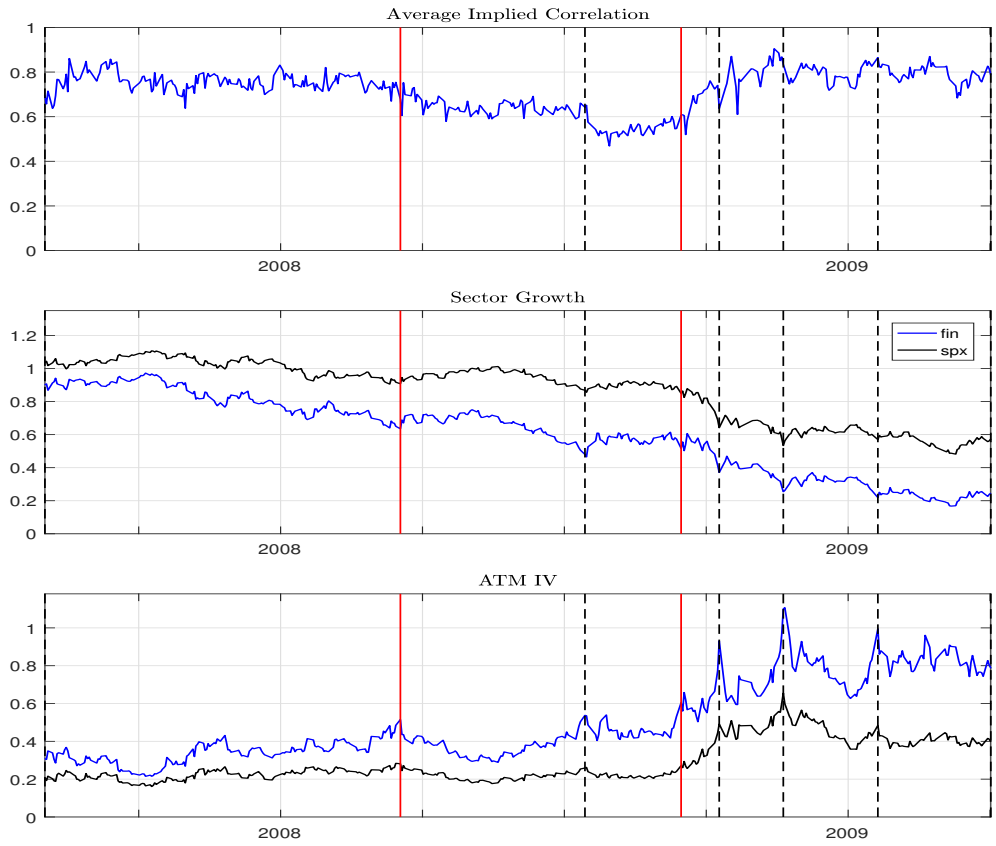

Figure 9: **Average Implied Correlation, Cumulative Returns, and ATM Implied Volatilities.** This figure focuses on the period of the financial crisis (August 1, 2007, to April 1, 2009). The red solid lines show the dates that separate Period I, Period II, and Period III of the financial crisis (March 17, 2008, and September 15, 2008). The black dashed lines indicate several extreme trading days (July 14, 2008, October 9, 2008, November 20, 2008, and January 20, 2009). The second and third panels show the financial sector (blue line) and the S&P 500 index (black line).

as shown in Figure 9: Period I runs until March 2008 (the bailout of Bear Sterns), Period II is from March 2008 until September 2008 (the failure of Lehman Brothers), and Period III is from September 2008 until July 2009.

In 2006, the US financial industry was highly exposed to subprime mortgages. When house prices started to fall in July 2006, this had an immediate impact on banks. The financial sector declined, and its implied volatility increased, as confirmed by the blue lines in Figure 1. By August 2007, correlations had also increased considerably, as confirmed by the red line in Figure 8. However, the crisis had yet not spread to other sectors, and the market implied volatility remained fairly stable, as evidenced by the black line in the third panel of Figure 9. At that point, there was a widespread belief among market participants that by lowering interest rates, the Federal Reserve could boost market liquidity and restore confidence. In March 2008, Bear Stearns became the first of several financial institutions to be bailed out by the government. This event created an expectation that if needed, the government would rescue any other banks and thus that the crisis would not spill over to other sectors. Correlations and volatilities temporarily decreased during Period II, returning to the pre-crisis levels as evidenced by Figure 8 and by the first and third panels of Figure 9. However, the situation changed drastically

after the bankruptcy of Fannie Mae and Freddie Mac on September 7, followed by that of Lehman Brothers a week later. As the government decided to let Lehman fail, the market stress increased sharply and investors started to panic. A new crisis period started as investors rushed to the safest investments, such as cash or government securities. During the last quarter of 2008, markets fell worldwide, and volatilities and correlation peaked and remained at high levels afterward, as confirmed by all three panels of Figure 9 during Period III.

5.2 Risk Premia

We use the average correlations defined in Section 4.1 to study down and up correlation risk premia. Formally, we define the global, down, and up correlation premia as

$$\theta^a = \rho^{a,\mathbb{P}} - \rho^{a,\mathbb{Q}}, \tag{21}$$

where, as before, $a \in \{g, d, u\}$ corresponds to the global, down, and up correlations. Recall that the correlations under $\mathbb{P}$ are computed over the 3-month trailing window, so that the risk premia are defined in the *ex ante* fashion. The same approach has been used, among others, in Driessen, Maenhout, and Vilkov (2009), Buss, Schönleber, and Vilkov (2017, 2019a), Bollerslev, Tauchen, and Zhou (2009).[18]

For each type, Figure 10 plots across time the average implied correlation (blue line) and realized correlation (red line). The first panel displays the global correlations $\rho^{\mathbb{Q}}$ and $\rho^{\mathbb{P}}$. The corresponding correlation risk premium (CRP) θ, which appears as the difference between the blue and the red lines, is mostly negative. Across the sample period, it has an average of -0.066 and is statistically highly significant with a t-statistic of -6.0, see Table 4. Buss, Schönleber, and Vilkov (2017) also estimate the CRP for the nine sectors of the S&P 500 and 3-month maturity, but over a different sample period from 1998 to 2015. Their estimate (-0.059) is of similar magnitude as ours.[19]

The two bottom panels of Figure 10 illustrate our most intriguing contribution to the CRP literature. The second panel visually demonstrates that the average realized down correlations are systematically lower than their implied counterparts (the red line is consistently below the blue line). The opposite is true for the third panel: the average realized up correlations are systematically higher than their implied counterparts.

Table 4 reveals that on average, the up correlations are lower than the down correlations under $\mathbb{P}$ and that this asymmetry is even more pronounced under $\mathbb{Q}$:

$$\rho^{u,\mathbb{Q}} < \rho^{u,\mathbb{P}} < \rho^{d,\mathbb{P}} < \rho^{d,\mathbb{Q}}.$$

Consequently, Table 4 reports that the average down CRP θ^d is negative (-0.186), while the average up CRP θ^u is positive (0.138). Both risk premia are highly significant and economically

[18] It is common to estimate the realized and implied variances at a given time-t from the information available at that time. That is, the implied variance is estimated from options available at time-t, while the realized variance is estimated from past historical returns up to time-t. Specifically, Bollerslev, Tauchen, and Zhou (2009) define the variance risk premium as "the difference between this ex-ante risk-neutral expectation of the future return variation over the $[t, t+1]$ time interval and the ex-post realized return variation over the $[t-1, t]$."

[19] The magnitude of the CRP for *sectors* is generally smaller than that for *individual stocks*. In particular, Driessen, Maenhout, and Vilkov (2013), Buss, Schönleber, and Vilkov (2017, 2019a) report the CRP for the stocks in the S&P 500 index of -0.106, -0.100, and -0.099, respectively. However, it is important to stress that correlations for sectors are not directly comparable to those for stocks. Sectors are diversified portfolios with little idiosyncratic risk remaining. The average pairwise correlations between sectors are considerably higher than between stocks (0.7 versus 0.3, under $\mathbb{P}$) and the CRP is not as strong.

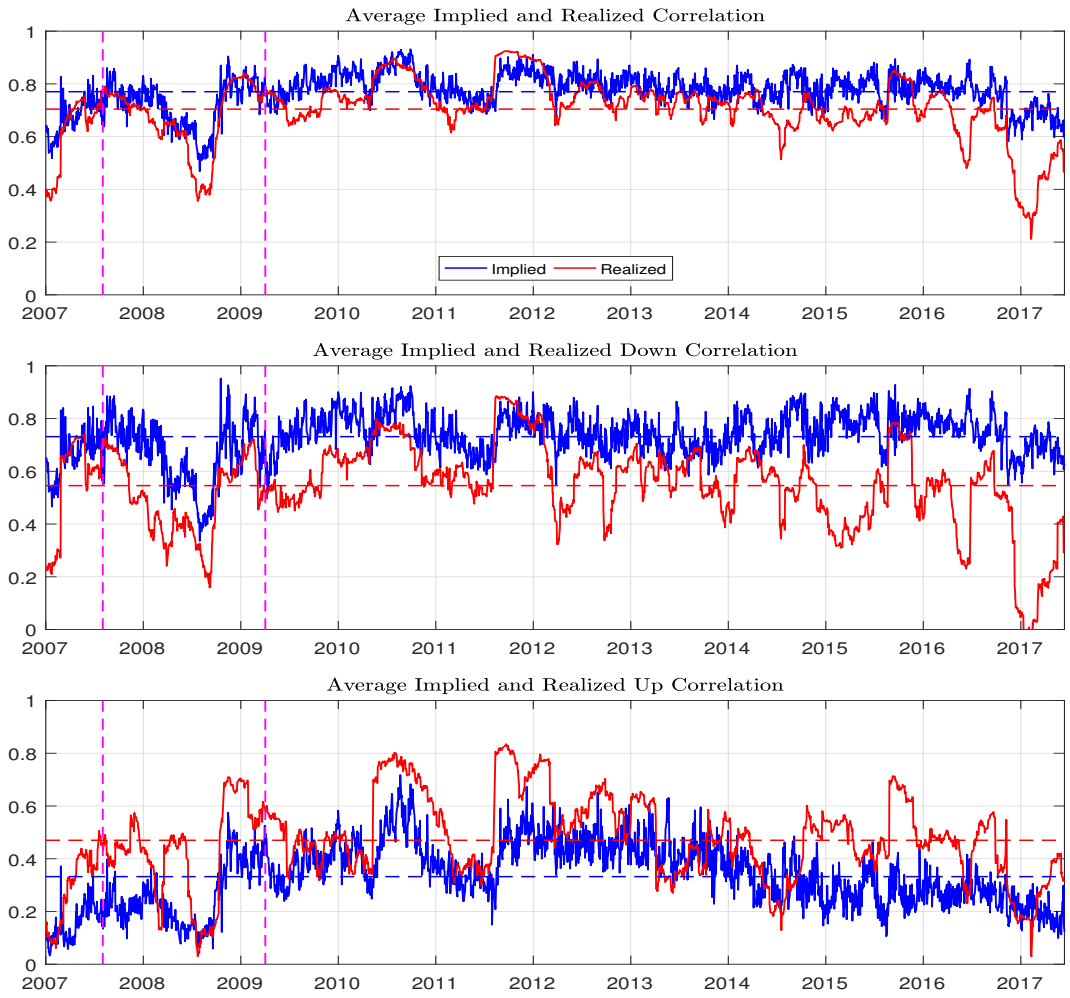

Figure 10: **Implied and Realized Correlations**. Implied correlations (blue) are computed from option-implied dependences; realized correlations (red) are computed from sector index returns. The corresponding means of the two series are shown with the horizontal dashed lines. The pink vertical lines indicate the period corresponding to the financial crisis.

very large. They are consistent with the economic intuition that investors are mainly concerned with the loss of diversification when the market falls. Consequently, they are willing to pay a considerable premium to hedge against increases in the down correlation. However, investors actually *prefer* high correlation when the market rallies. That is, investors view the down correlation as "bad" and the up correlation as "good." The net effect of the negative premium for the down correlation and the positive premium for the up correlation is a negative premium for the global correlation.

The last row of Table 4 focuses on the correlation spread (the difference between the down and up correlations):

$$\Delta\rho^{\mathbb{M}} = \rho^{d,\mathbb{M}} - \rho^{u,\mathbb{M}},$$

where $\mathbb{M} \in \{\mathbb{Q}, \mathbb{P}\}$ indicates the probability measure under which the expectations are evaluated. Unsurprisingly, the corresponding risk premium, $\Delta\rho^{\mathbb{P}} - \Delta\rho^{\mathbb{Q}}$, is negative and highly significant (-0.324, t-stat = -14.6). Its magnitude is about 5 times larger than the CRP for the global

	N_{obs}	Under $\mathbb{P}$	Under $\mathbb{Q}$	Premium	t-stat
Global	2584	0.704	0.770	-0.066	-6.0
Down	2584	0.546	0.731	-0.186	-9.6
Up	2584	0.470	0.332	0.138	8.8
Down-Up	2584	0.076	0.399	-0.324	-14.6

Table 4: **Correlation Risk Premium.** The table reports statistics for the risk premia θ, θ^d, and θ^u computed for the average global, down, and up correlations. The last row is the correlation spread, $\Delta\rho = \rho^d - \rho^u$. The last column shows the Newey-West t-statistics computed with 63 lags.

correlation: the magnitude of risk premium for the down (up) correlation is approximately 3 (2) times larger than for the global correlation.

These observations motivate us to define the down minus up correlation swap, or DUC swap for short. At time-T, this swap has a payoff (the variable leg) equal to the difference between the *realized* down and up correlations, or $\Delta\rho^{\mathbb{P}}$. That is, the daily returns realized over the life of the swap are split into two equal groups, when the market return is low and high. The average weighted correlations are computed for each group and their difference gives the payoff of the swap. At time-0, the DUC swap has the initial cost (the fixed leg) equal to the difference between the *risk-neutral* down and up correlations, or $\Delta\rho^{\mathbb{Q}}$. Historically, selling the DUC swap would have been very profitable, as it takes advantage of both selling the "expensive" down correlation and buying the "cheap" up correlation. However, two caveats are in order. The DUC swap is not yet tradable. We assume that historically it would have been traded at the price implied by our BRA approach. Furthermore, the transaction costs, likely to be substantial, are not accounted for.[20]

5.3 Marginal Distributions or Dependence?

As shown in the previous subsection, time-varying correlation is priced and its risk premium changes the sign between the down to up correlations. Since correlations are affected jointly by the margins and the dependence, the risk premium could potentially stem from either one. Disentangling the respective roles of the margins and the dependence is an important but hard question. Ideally, for each trading day we need to estimate the two components under both the risk-neutral and physical measures. Under $\mathbb{Q}$, this is feasible, as MFDR yields the full joint distribution. However, under $\mathbb{P}$, we only observe one draw from the unobserved (and time-varying) distribution.

To better understand the roles of the margins and the dependence, we perform two exercises. In the first exercise, we focus on $\mathbb{Q}$ measure only, for which we have complete information. We formally investigate the causes of the enormous asymmetry between the down and up correlations under $\mathbb{Q}$ in Table 4, i.e., a large positive correlation spread $\Delta\rho^{\mathbb{Q}} = \rho^{d,\mathbb{Q}} - \rho^{u,\mathbb{Q}}$. As argued earlier, this asymmetry is inconsistent with a multivariate normal (MVN) distribution, for which the down and up correlations must be the same. The observed asymmetry may be caused by nonnormality of the margins or nonnormality of the dependence. For example, the observed asymmetry could at least partially be due to the heavier left tails of the margins compared to the right tails. To disentangle the effects of the margins and of the dependence in explaining

[20] We leave for future research development of a tradable strategy, which could proxy the DUC swap.

the correlation spread $\Delta\rho^{\mathbb{Q}}$, we contrast four cases, labeled NN, EN, NE, and EE, where the first letter denotes the type of the margins (Normal or Empirical) and the second letter denotes the type of the copula (again, Normal or Empirical). The EE case corresponds to the output distribution obtained using our model-free BRA algorithm (Empirical margins and Empirical copula). The NN case corresponds to Normal margins and Normal copula (i.e., an MVN distribution). Specifically, the standard deviations for the normal margins match those of the empirical margins, and the normal copula is calibrated to a constant correlation matrix in such a way that the model preserves the average pairwise correlation among the nine sectors. The NE case uses Normal margins joined with the Empirical dependence from the BRA algorithm (i.e., misspecified margins and correct dependence). Finally, the EN case uses the Empirical margins joined with a Normal copula (i.e., correct margins and misspecified dependence).

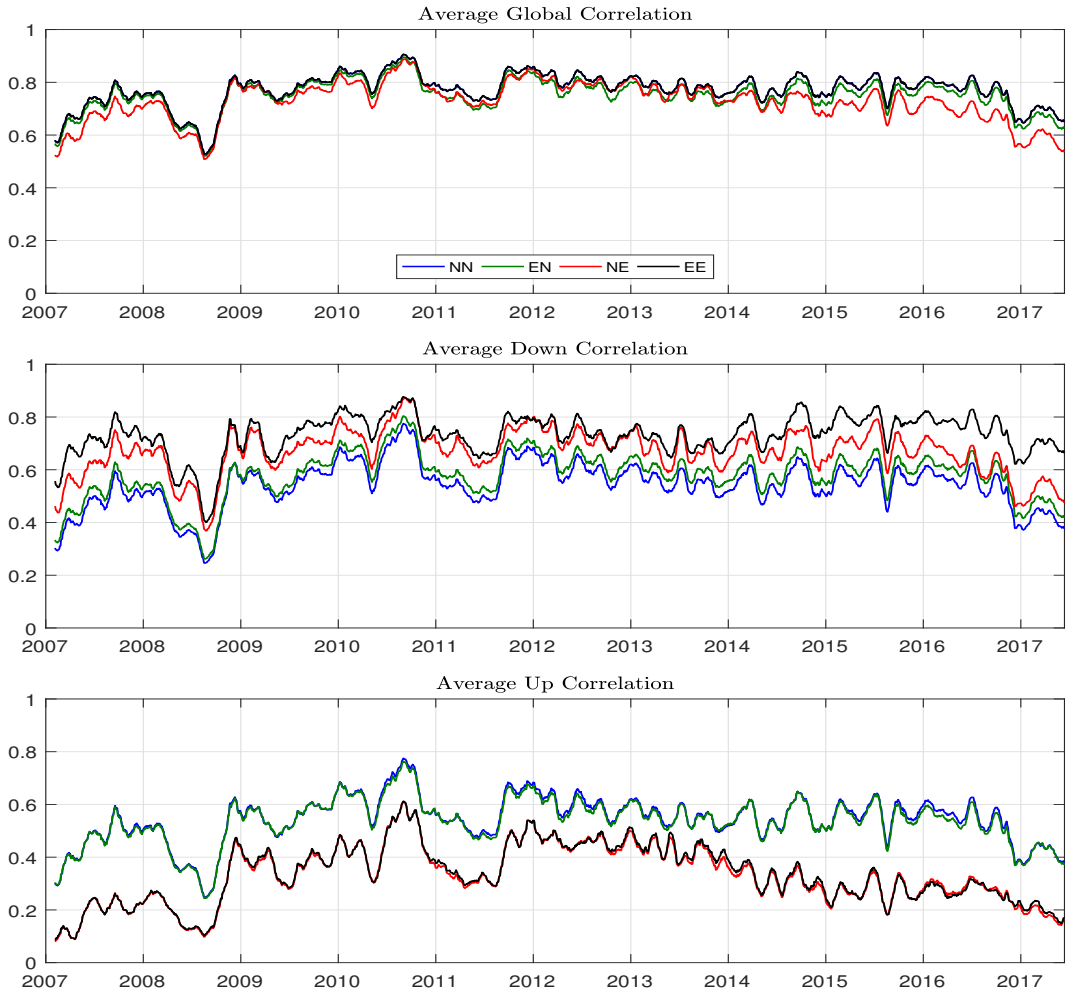

Figure 11: **Implied Correlations**. Average implied global, down, and up correlations are computed for the four cases (NN, EN, NE, EE), where the first letter denotes the type of margins (Normal or Empirical) and the second letter denotes the type of copula (Normal or Empirical). For example, EN denotes Empirical margins joined with a Normal copula. Statistics are plotted as 1-month moving averages.

For all four cases, we compute the average pairwise global, down, and up correlations and display them as time series in the three panels of Figure 11: NN (blue), EN (green), NE (red), and EE (black). In Table 5, we report the time-series averages of each quantity. In the first panel

	NN	EN	NE	EE
Global	0.770	0.749	0.726	0.770
Down	0.542	0.577	0.667	0.731
Up	0.542	0.535	0.328	0.332
Down-Up	0.000	0.042	0.339	0.399

Table 5: **Implied Correlations.** Average implied global, down, and up correlations are computed for the four cases (NN, EN, NE, EE), where the first letter denotes the type of margins (Normal or Empirical) and the second letter denotes the type of copula (Normal or Empirical). For example, EN denotes Empirical margins joined with a Normal copula. All statistics are under the $\mathbb{Q}$ measure. The last row is the correlation spread, $\Delta\rho^{\mathbb{Q}} = \rho^{d,\mathbb{Q}} - \rho^{u,\mathbb{Q}}$.

of Figure 11 (the average global correlation), the four lines are all close to each other. In fact, the average global correlations for EE and NN are identical by construction. In the second and third panels (the average down and up correlations), the four lines are farther apart. Furthermore, the EN and NN lines are close to each other, as are the NE and EE lines. This is especially true for the up correlation. We thus conclude that the correlation spread $\Delta\rho^{\mathbb{Q}}$ is mainly driven by the type of copula and not by the potential nonnormality of the margins. This conclusion is reinforced by the last row of Table 5. The correlation spread is 0 for the multivariate normal case (NN), 0.042 for the EN model, and 0.399 for the EE model, which fits perfectly options on individual sectors and the index. Even though the EN model has symmetric copula, its down and up correlations are not the same due to the skewed margins. When comparing the NN and EN models, we can intuitively see that adopting the nonnormal margins explains only about 11% of the true correlation spread (0.042/0.399). The rest can be attributed to the nonnormality of the dependence.

In the second exercise, we focus on both $\mathbb{Q}$ and $\mathbb{P}$ measures and attempt to disentangle the respective roles of the margins and the dependence on the CRP. We repeat the analysis of Table 4 but now using Spearman (instead of Pearson) correlations. Spearman's is a rank correlation and it has the advantage of being unaffected by the margins. To compute Spearman correlation under $\mathbb{Q}$, however, the complete joint distribution is required, even for the case of the global correlation. Thus, our BRA methodology is crucial.

The results are reported in Table 6. We observe that, compared to Table 4, all correlations are now slightly lower under both measures. However, the correlation spread $\Delta\rho$ remains positive under both $\mathbb{P}$ and $\mathbb{Q}$. Importantly, the risk premia for the three types of correlations (global, down, and up) remain quantitatively very similar, but now they are not confounded by the dynamics of the marginal distributions. Therefore, by contrasting the results for Pearson and Spearman correlations, we conclude that the CRP is mainly driven by the dependence and not by the margins.

	N_{obs}	Under $\mathbb{P}$	Under $\mathbb{Q}$	Premium	t-stat
Global	2584	0.664	0.704	-0.040	-4.8
Down	2584	0.475	0.649	-0.174	-12.2
Up	2584	0.418	0.284	0.134	9.7
Down-Up	2584	0.057	0.365	-0.308	-16.5

Table 6: **Correlation Risk Premium, Spearman.** The table reports statistics for the risk premia θ, θ^d, and θ^u computed for the average global, down, and up correlations. It is the same as Table 4, except now the computations are based on Spearman (instead of Pearson) correlations. The last row is the correlation spread, $\Delta\rho = \rho^d - \rho^u$. The last column shows the Newey-West t-statistics computed with 63 lags.

5.4 Tail Indices

Observe that any risk indicators estimated using historical returns can now be estimated using option-implied information using MFDR. In this paper, we have constructed two new dependence indicators in (18). Their definitions are based on the down and up correlations, or correlations conditional on the index return being below or above its median. One natural generalization is to consider more detailed indicators that condition on other quantiles of index returns (e.g., quartiles). This would permit a better understanding of the dependence deeper in the tails and extend for instance the work of Longin and Solnik (2001) who perform a study in this spirit under the real-world measure. Let us illustrate how we are now in position to construct similar option-implied forward-looking measures. We take our original idea to the limit and attempt to estimate the left and right tail indices. Given a joint distribution of two random variables (X, Z), the tail dependence could be measured as follows. If X_q and Z_q denote q-quantiles for X and Z, then a left-tail index is defined as

$$LT(q) := \frac{2\log(q)}{\log(\text{Prob}(X \leqslant X_q\,,\, Z \leqslant Z_q))} - 1.$$

This index is equal to 0 when the two variables are independent and 1 when they are comonotonic (perfectly correlated). The right tail index $RT(q)$ is defined similarly. In finite samples, estimates of $LT(q)$ and $RT(q)$ are not guaranteed to be positive. Therefore, we choose q to be not too small, such as $q = 0.1$). In Figure 12, we show the indices $LT(q)$ and $RT(q)$ over time and observe that their empirical behavior is qualitatively similar to that of the down and up correlations, with the left tail index generally being larger than the right tail index. Quantitatively, both tail indices are higher than the corresponding correlations. There are caveats as mentioned for instance by Longin and Solnik (2001), who warn against interpreting blindly down and up correlation indicators (see the first paragraph on page 650 of their paper). Similarly to the analysis of Section 5.3, a careful check must be performed to better understand whether the asymmetry is driven by the properties of the marginal distributions or of the dependence.

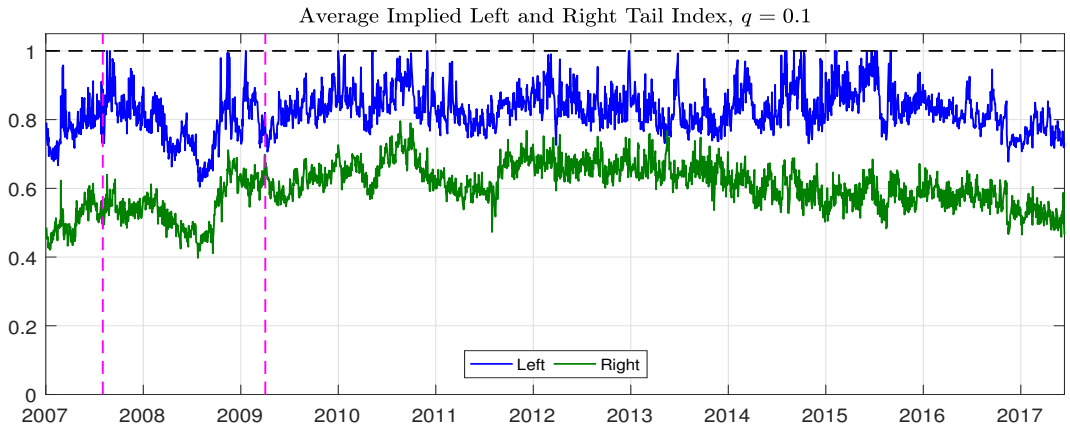

Figure 12: **Implied Left and Right Tail Index.** Tail indices $LT(q)$ (blue) and $RT(q)$ (green) are computed for $q = 0.1$. The pink vertical lines indicate the period corresponding to the financial crisis.

5.5 Implications for the Literature and Hybrid Model

The results in the previous subsection highlight the importance of proper modeling of the dependence, especially under $\mathbb{Q}$. The large positive correlation spread $\Delta\rho^{\mathbb{Q}}$ is mainly driven by the nonnormality of the dependence and the CRP is mainly a compensation for the dependence risk. Therefore, a financial economist who wants to build a multi-asset model that matches the salient features of the option data must pay a special attention to the dependence. She could pursue two distinct strategies: (a) to model the joint distribution directly, (b) to model the margins and the dependence separately and then combine the two pieces to build the joint distribution.

The first strategy is more common because of its tractability. As an example of this strategy, consider the single-factor model of Chang, Christoffersen, Jacobs, and Vainberg (2012), or CCJV henceforth. In this model, the log-return of asset j is represented under $\mathbb{Q}$ as

$$x_j = \alpha_j + \beta_j x_m + \varepsilon_i, \tag{22}$$

where x_j and x_m are the log-returns of asset j and the market, ε_i is the idiosyncratic shock, which has zero mean and variance u_j^2 and which is independent of the market return. In this model, the distribution for the market return can be flexible, but the distribution for the asset is obtained via a rigid linear transformation. CCJV suggest to estimate the key model parameters, β_j and u_j, from the moments of the RNDs by matching the following theoretical restrictions:[21]

$$\beta_j = corr_{j,m}\left(\frac{var_j}{var_m}\right)^{\frac{1}{2}} = \left(\frac{skew_j}{skew_m}\right)^{\frac{1}{3}}\left(\frac{var_j}{var_m}\right)^{\frac{1}{2}}, \tag{23}$$

$$u_j^2 = var_i - \beta_j^2 \cdot var_m. \tag{24}$$

For internal consistency, several conditions must be satisfied: (1) the skewness of the asset's return must be of the same sign as the skewness of the market, (2) in the absolute value, the skewness of the asset should be smaller than that of the market (or correlation $corr_{j,m}$ will be greater than 1), and (3) the beta coefficient estimated in (23) should not be too large (or the

[21] Given β_j and u_j, another restriction $E[e^{x_j}] = 1$ allows one to recover α_j.

variance u_j^2 could become negative). We find that in our empirical application of the nine sectors these conditions can sometimes be violated. The first condition is violated least frequently. The skewness of the market and asset are typically negative. In this respect, it is critical that the model is formulated in terms of log-returns, as opposed to simple returns, because it could happen that the RND for the log-return x_j is negatively skewed while the RND for the corresponding simple return X_j is positively skewed. The violations of condition (2) are more prevalent and violations of condition (3) are even more common. Because sectors are diversified portfolios, the idiosyncratic component is generally much smaller than for individual stocks and violations of condition of (3) are more likely. For these reasons, we estimate CCJV model differently, that is, we search for the parameters β_j and u_j, which minimize the distance between the two implied volatility curves for asset j, actual and fitted, with the distance defined as in (26). In other words, instead of perfectly matching the two moments in (23)-(24), we (imperfectly) fit the whole implied volatility curve. Although conditions (1)-(3) are now satisfied, we find that the model is still misspecified. The model is too rigid to simultaneously fit different parts of the RND (OTM put, ATM options, and OTM calls). The overall fit can be improved if parameters β_j and u_j are allowed to differ depending on whether the market return is low or high. Specifically, consider the following generalization of (22):

$$x_j = \alpha_j + \left(\beta_j^d x_m + \varepsilon_j^d\right) \mathbf{1}^d + \left(\beta_j^u x_m + \varepsilon_j^u\right) \mathbf{1}^u, \tag{25}$$

where $\mathbf{1}^d$ is the indicator variable equal one if the market return is below its median and $\mathbf{1}^u = 1 - \mathbf{1}^d$ is equal to one if the market return is above its median. The above mixture model allows for two different betas and levels of idiosyncratic shock (with standard deviations u_j^d and u_j^u, respectively). As a result, the generalized model is now able to generate a strongly asymmetric dependence, which is much stronger in the left tail than in the right tail and, which implies a large positive correlation spread $\Delta\rho^{\mathbb{Q}}$.

We now introduce an alternative approach to model the multivariate joint distribution, which is based on above strategy (b) and which offers a number of advantages. First, it fits perfectly all sector marginal distributions. That is, the model by construction matches exactly the prices of all sector options. Second, it automatically enforces the additional consistency condition that the average beta of the nine sectors (computed for simple returns) must be equal to 1.[22] Third, the model provides a much tighter overall fit to option data. We refer to this model as *hybrid* because it combines (i) fully nonparametric margins extracted from the individual options and (ii) a parsimonious parametric copula. The proposed copula is based on the homogeneous multivariate skewed normal distribution driven by two parameters only. Full details for the hybrid model are provided in Appendix E.

Despite its simplicity, the model captures reasonably well the most salient features of the option-implied dependence.[23] As already mentioned, the hybrid model fits perfectly the sector options, but it could potentially misprice the index options. The mispricings, however, are much smaller than when a multivariate normal copula is used. Of course, in terms of fitting option prices the hybrid model cannot compete with the BRA approach, because the latter produces (essentially) a perfect fit. Instead, the hybrid model offers different advantages: it is transparent, intuitive, and easy to implement. Our primary motivation for developing this model is two-fold.

[22]The model in (22) and its generalization in (25) ignore information about portfolio weights ω_j. The model parameters for each asset are estimated independently of the other assets.

[23]The hybrid model has only 2 free parameters and, thus, is relatively easy to estimate. For comparison, the models in (22) and (25) have $2 \times 9 = 18$ and $4 \times 9 = 36$ parameters to estimate.

First, because the hybrid model does not rely on the somewhat opaque BRA methodology, it provides an alternative confirmation of our key empirical results, including the analysis of the CRP. Second, we believe that the hybrid model could prove useful in other applications where the BRA methodology cannot be implemented due to data limitations.[24]

6 Conclusions and Futures Directions

We propose a novel methodology to estimate the risk-neutral dependence among several assets that is consistent with market prices of options on these assets and on their index. Termed MFDR, it offers two critical advantages compared to the existing methods. First, the methodology is completely model-free and requires no parametric assumptions. Second and most importantly, it yields a full dependence structure, not just the average correlation coefficient. To achieve so, the methodology matches a continuum of moments on the risk-neutral distributions, as opposed to satisfying just one restriction in the existing methods.

In the empirical application, we implement our methodology to the nine economic sectors comprising the S&P 500 index. We document that the option-implied dependence for the nine sectors is highly asymmetric and time-varying. We study the correlation risk premium and find that it is negative for the down correlation and positive for the up correlation. These findings are consistent with the economic intuition that investors are mainly concerned with the loss of diversification when the market falls and that they actually prefer high correlation when the market rallies. That is, investors view the down correlation as "bad" and the up correlation as "good". While it might be possible to rationalize the negative risk premium for the down correlation with disappointment aversion preferences, the positive risk premium for the up correlation presents a bigger challenge for the theoretical literature.

We anticipate that our model-free methodology has numerous promising applications and that our empirical findings has implications for asset allocation and portfolio construction. For instance, Bollerslev, Patton, and Quaedvlieg (2020) propose a novel extension of the CAPM model by decomposing the traditional market beta into four semibetas depending on the signed covariation between the market and individual asset returns, extending the work of Ang, Chen, and Xing (2006). Bollerslev, Patton, and Quaedvlieg (2020) show that their decomposition into the four semibetas offers superior cross sectional predictions compared to those obtained by the traditional beta estimates. The four semibetas are estimated from historical returns, either daily or intraday. It would be of interest to use our methodology to estimate the forward-looking option-implied semibetas and to contrast them to the semibetas estimated under $\mathbb{P}$.

Secondly, optimal portfolio construction is another interesting area of potential application. Our approach allows us to estimate not just the average pairwise correlation, but the full risk-neutral covariance matrix. The existing literature has already established that the moments of the univariate implied distribution can improve portfolio choice. Superior asset allocation can be achieved by using the risk-neutral moments estimated from options rather than their realized counterparts estimated from stock returns (e.g., Kostakis, Panigirtzoglou, and Skiadopoulos (2011), DeMiguel, Plyakha, Uppal, and Vilkov (2013)). Therefore, a promising avenue for future

[24] Our approach can be compared to Jackwerth and Vilkov (2019) who estimate the joint distribution of the S&P 500 return and its volatility. They also estimate a hybrid model, where model-free margins are combined with a parametric copula. Specifically, Jackwerth and Vilkov (2019) find that Frank copula works best for their application. Our problem is quite different and of higher dimension. The skewed normal copula provides a parsimonious way to model the strong asymmetric dependence observed in our application.

research would be to investigate whether these results can be extended to multidimensional portfolios, where the asset allocation is not limited to the risk-free assets and one risky asset but rather could include many risky assets. Furthermore, a large body of the literature uses the option-implied variance, correlation, or skewness to predict future market returns (e.g., Amaya, Christoffersen, Jacobs, and Vasquez (2015), Buss and Vilkov (2012), Buss, Schönleber, and Vilkov (2019b), Jondeau, Zhang, and Zhu (2019), Martin and Wagner (2019), Stilger, Kostakis, and Poon (2017)). Applying MFDR can provide new insights as to which specific option-implied factors drive predictability. For example, an intriguing questions is whether the down and up correlations can predict future returns better than the global correlation.

Thirdly, our methodology is designed to find a joint distribution, which perfectly reproduces both the individual and index options. Importantly, this methodology can also inform us when no feasible joint distribution exists. This happens when the prices of individual options are inconsistent with the index options, implying that there exists an arbitrage opportunity. Typically, such a situation arises when a specific portfolio of individual options is too cheap compared to the index option. Our methodology can be used to detect potential arbitrage opportunities and to verify whether they can survive realistic trading costs.This idea generalizes dispersion arbitrage, which is based solely on global correlations.

Fourthly, our methodology could shed more light on the phenomenon studied in Kelly, Lustig, and Van Nieuwerburgh (2016). Specifically, they use the so-called put spread for the financial industry as a measure of tail dependence. The spread is defined as the difference in costs between OTM puts for individual banks and OTM puts for the financial sector index. Kelly, Lustig, and Van Nieuwerburgh (2016) find that during the 2007–2009 financial crisis, the put spread for the financial industry was extraordinary large, which implies a very low implied correlation among the banks and is consistent with the perceived sector-wide government bailout guarantee. When implemented at the industry level, our approach would potentially be able to recover the full joint distribution for the banks (not just one specific partial measure of tail behavior) and to disentangle the effects due to changes in the margins (e.g., the volatility of the banks) and the dependence (e.g., interaction between banks).

References

Aït-Sahalia, Y., and A. W. Lo (2000): "Nonparametric risk management and implied risk aversion," *Journal of Econometrics*, 94(1), 9–51.

Alcock, J., and A. Hatherley (2017): "Characterizing the asymmetric dependence premium," *Review of Finance*, 21(4), 1701–1737.

Alcock, J., and P. Sinagl (2020): "The Price of Asymmetric Dependence: International Evidence," *Available at SSRN 3568496*.

Amaya, D., P. Christoffersen, K. Jacobs, and A. Vasquez (2015): "Does realized skewness predict the cross-section of equity returns?," *Journal of Financial Economics*, 118(1), 135–167.

Ang, A., and J. Chen (2002): "Asymmetric correlations of equity portfolios," *Journal of Financial Economics*, 63(3), 443–494.

Ang, A., J. Chen, and Y. Xing (2006): "Downside risk," *Review of Financial Studies*, 19(4), 1191–1239.

Azzalini, A., and A. D. Valle (1996): "The multivariate skew-normal distribution," *Biometrika*, 83(4), 715–726.

Bakshi, G., N. Kapadia, and D. Madan (2003): "Stock return characteristics, skew laws, and the differential pricing of individual equity options," *Review of Financial Studies*, 16(1), 101–143.

Banz, R., and M. Miller (1978): "Prices for state-contingent claims: Some estimates and applications," *Journal of Business*, 51, 653–672.

Barone-Adesi, G., and R. E. Whaley (1987): "Efficient analytic approximation of American option values," *Journal of Finance*, 42(2), 301–320.

Bernard, C., O. Bondarenko, and S. Vanduffel (2018): "Rearrangement algorithm and maximum entropy," *Annals of Operations Research*, 261(1-2), 107–134.

——— (2021): "A model free approach to multivariate option pricing," *Review of Derivatives Research*.

Bernard, C., and D. McLeish (2016): "Algorithms for Finding Copulas Minimizing Convex Functions of Sums," *Asia-Pacific Journal of Operational Research*, 33, 1650040–26.

Bollerslev, T., A. J. Patton, and R. Quaedvlieg (2020): "Realized semibetas: Signs of things to come," *Available at SSRN 3528276*.

Bollerslev, T., G. Tauchen, and H. Zhou (2009): "Expected stock returns and variance risk premia," *Review of Financial Studies*, 22(11), 4463–4492.

Bollerslev, T., and V. Todorov (2011): "Tails, fears, and risk premia," *Journal of Finance*, 66(6), 2165–2211.

Bondarenko, O. (2000): "Recovering risk-neutral densities: A new nonparametric approach. Working Paper, University of Illinois at Chicago," *SSRN Working paper presented at EFA 2000, London.*

Bondarenko, O. (2003): "Estimation of risk-neutral densities using positive convolution approximation," *Journal of Econometrics*, 116(1), 85–112.

Bondarenko, O. (2014): "Variance trading and market price of variance risk," *Journal of Econometrics*, 180, 81–97.

Breeden, D. T., and R. H. Litzenberger (1978): "Prices of state-contingent claims implicit in option prices," *Journal of Business*, 51(4), 621–651.

Britten-Jones, M., and A. Neuberger (2000): "Option prices, implied price processes, and stochastic volatility," *Journal of Finance*, 55(2), 839–866.

Buraschi, A., R. Kosowski, and F. Trojani (2013): "When there is no place to hide: Correlation risk and the cross-section of hedge fund returns," *Review of Financial Studies*, 27(2), 581–616.

Buraschi, A., F. Trojani, and A. Vedolin (2014): "When uncertainty blows in the orchard: Comovement and equilibrium volatility risk premia," *Journal of Finance*, 69(1), 101–137.

Buss, A., L. Schönleber, and G. Vilkov (2017): "Option-implied correlations, factor models, and market risk," *INSEAD Working Paper No. 2017/20/FIN. Available at SSRN, Feb 16, 2017: https://ssrn.com/abstract=2906484*.

——— (2019a): "Expected correlation and future market returns," *Available at SSRN: https://ssrn.com/abstract=3114063 and CDI working paper*.

——— (2019b): "Expected correlation and future market returns: The sum of parts is more than the whole," *Presentation at the 2019 Conference on Derivatives and Volatility, Financial Management Association*.

Buss, A., and G. Vilkov (2012): "Measuring equity risk with option-implied correlations," *Review of Financial Studies*, 25(10), 3113–3140.

Campbell, R. A., C. S. Forbes, K. G. Koedijk, and P. Kofman (2008): "Increasing correlations or just fat tails?," *Journal of Empirical Finance*, 15(2), 287–309.

Carr, P., and D. Madan (2001): "Optimal positioning in derivative securities," *Quantitative Finance*, 1(1), 19–37.

Carr, P., and L. Wu (2009): "Stock options and credit default swaps: A joint framework for valuation and estimation," *Journal of Financial Econometrics*, 8(4), 409–449.

Chabi-Yo, F., S. Ruenzi, and F. Weigert (2018): "Crash sensitivity and the cross section of expected stock returns," *Journal of Financial and Quantitative Analysis*, 53(3), 1059–1100.

Chang, B.-Y., P. Christoffersen, K. Jacobs, and G. Vainberg (2012): "Option-implied measures of equity risk," *Review of Finance*, 16(2), 385–428.

Chicago Board Options Exchange (2009): "CBOE S&P 500 implied correlation index," *White paper*.

DeMiguel, V., Y. Plyakha, R. Uppal, and G. Vilkov (2013): "Improving portfolio selection using option-implied volatility and skewness," *Journal of Financial and Quantitative Analysis*, 48(6), 1813–1845.

Driessen, J., P. J. Maenhout, and G. Vilkov (2009): "The price of correlation risk: Evidence from equity options," *Journal of Finance*, 64(3), 1377–1406.

Driessen, J., P. J. Maenhout, and G. Vilkov (2013): "Option-implied correlations and the price of correlation risk," *Working paper, INSEAD*.

Embrechts, P., G. Puccetti, and L. Rüschendorf (2013): "Model uncertainty and VaR aggregation," *Journal of Banking & Finance*, 37(8), 2750–2764.

Engle, R., and S. Figlewski (2014): "Modeling the dynamics of correlations among implied volatilities," *Review of Finance*, 19(3), 991–1018.

Farago, A., and R. Tédongap (2018): "Downside risks and the cross-section of asset returns," *Journal of Financial Economics*, 129(1), 69–86.

Faria, G., R. Kosowski, and T. Wang (2018): "The correlation risk premium: International evidence," *Working Paper of Imperial College Business School*.

Feunou, B., M. R. Jahan-Parvar, and C. Okou (2018): "Downside variance risk premium," *Journal of Financial Econometrics*, 16(3), 341–383.

Figlewski, S. (2018): "Risk-neutral densities: A review," *Annual Review of Financial Economics*, 10, 329–359.

Gul, F. (1991): "A theory of disappointment aversion," *Econometrica: Journal of the Econometric Society*, pp. 667–686.

Harvey, C. R., Y. Liu, and H. Zhu (2016): "... and the cross-section of expected returns," *Review of Financial Studies*, 29(1), 5–68.

Hong, Y., J. Tu, and G. Zhou (2006): "Asymmetries in stock returns: Statistical tests and economic evaluation," *Review of Financial Studies*, 20(5), 1547–1581.

Jackwerth, J., and G. Vilkov (2019): "Asymmetric volatility risk: Evidence from option markets," *Review of Finance*, 23(4), 777–799.

Jackwerth, J. C., and M. Rubinstein (1996): "Recovering probability distributions from option prices," *Journal of Finance*, 51(5), 1611–1631.

Jaynes, E. T. (2003): *Probability theory: The logic of science*. Cambridge University Press.

Jiang, L., K. Wu, and G. Zhou (2018): "Asymmetry in stock comovements: An entropy approach," *Journal of Financial and Quantitative Analysis*, 53(4), 1479–1507.

Jondeau, E., Q. Zhang, and X. Zhu (2019): "Average skewness matters," *Journal of Financial Economics*.

Kelly, B., and H. Jiang (2014): "Tail risk and asset prices," *Review of Financial Studies*, 27(10), 2841–2871.

Kelly, B., H. Lustig, and S. Van Nieuwerburgh (2016): "Too-systemic-to-fail: What option markets imply about sector-wide government guarantees," *American Economic Review*, 106(6), 1278–1319.

Kilic, M., and I. Shaliastovich (2019): "Good and bad variance premia and expected returns," *Management Science*, 65(6), 2522–2544.

Kostakis, A., N. Panigirtzoglou, and G. Skiadopoulos (2011): "Market timing with option-implied distributions: A forward-looking approach," *Management Science*, 57(7), 1231–1249.

Linders, D., J. Dhaene, and W. Schoutens (2015): "Option prices and model-free measurement of implied herd behavior in stock markets," *International Journal of Financial Engineering*, 02(02), 1550012.

Longin, F., and B. Solnik (2001): "Extreme correlation of international equity markets," *Journal of Finance*, 56(2), 649–676.

Mardia, K. (1975): "Assessment of multinormality and the robustness of Hotelling's T 2 test," *Applied Statistics*, pp. 163–171.

Mardia, K. V. (1970): "Measures of multivariate skewness and kurtosis with applications," *Biometrika*, 57(3), 519–530.

——— (1974): "Applications of some measures of multivariate skewness and kurtosis in testing normality and robustness studies," *Sankhyā: Indian Journal of Statistics, Series B*, pp. 115–128.

Mardia, K. V., J. T. Kent, and J. M. Bibby (1980): *Multivariate analysis*. Academic press.

Martin, I. (2017): "What is the expected return on the market?," *Quarterly Journal of Economics*, 132(1), 367–433.

Martin, I., and C. Wagner (2019): "What is the expected return on a stock?," *Journal of Finance*, 74(4), 1887–1929.

Mueller, P., A. Stathopoulos, and A. Vedolin (2017): "International correlation risk," *Journal of Financial Economics*, 126(2), 270–299.

Orłowski, P., P. Schneider, and F. Trojani (2020): "On the nature of jump risk premia," *Swiss Finance Institute Research Paper*, 19-31.

Patton, A. J. (2004): "On the out-of-sample importance of skewness and asymmetric dependence for asset allocation," *Journal of Financial Econometrics*, 2(1), 130–168.

Pollet, J. M., and M. Wilson (2010): "Average correlation and stock market returns," *Journal of Financial Economics*, 96(3), 364–380.

Puccetti, G., and L. Rüschendorf (2012): "Computation of sharp bounds on the distribution of a function of dependent risks," *Journal of Computational and Applied Mathematics*, 236(7), 1833–1840.

Ross, S. (1976): "Options and efficiency," *Quarterly Journal of Economic*, 90, 75–89.

Routledge, B. R., and S. E. Zin (2010): "Generalized disappointment aversion and asset prices," *Journal of Finance*, 65(4), 1303–1332.

Rubinstein, M. (1994): "Implied binomial trees," *Journal of Finance*, 49(3), 771–818.

Schneider, P., and F. Trojani (2015): "Fear trading," *Working paper of the Swiss Finance Institute, Research Paper 15-03*.

Schreindorfer, D. (2020): "Macroeconomic tail risks and asset prices," *Review of Financial Studies*, 33(8), 3541–3582.

Stilger, P. S., A. Kostakis, and S.-H. Poon (2017): "What does risk-neutral skewness tell us about future stock returns?," *Management Science*, 63(6), 1814–1834.

Stutzer, M. (1996): "A simple nonparametric approach to derivative security valuation," *Journal of Finance*, 51(5), 1633–1652.

Todorov, V. (2010): "Variance risk-premium dynamics: The role of jumps," *Review of Financial Studies*, 23(1), 345–383.

Appendix

A Implications of the Restriction in (8)

It might be instructive to consider in more detail a few specific choices for function $g(z,s)$. When $g(z,s) = z$ or $g(z,s) = z^2$, we recover conditions on the first and second central moments discussed above. Similarly, the choice of $g(z,s) = z^3$ delivers a condition on the average measure of coskewness. However, let us consider some simple "cross-moments" when $g(z,s)$ depends on both z and s. When $g(z,s) = zs$, we obtain the condition similar to (5):

$$\sum_{j=1}^{d} \omega_j \sigma_j \rho_{jS} = \sigma_S.$$

With the additional assumption of *equal correlations with the index*, $\rho_{j,S} = \rho_S$, it leads to a new identifying equation for the implied correlation:

$$\rho_S = \frac{\sigma_S}{\sum_{j=1}^{d} \omega_j \sigma_j}$$

The above is *not* the same as the equation used in the existing literature. This is because the two auxiliary assumptions of constant correlations $\rho_{jk} = const$ (between two assets) and $\rho_{jS} = const$ (between an asset and the index) are not equivalent when σ_j and ω_j are not constant.

Finally, consider expectations conditional on S. Let $g(z,s) = z\,\mathbb{I}(s \leqslant K)$, for some level K. We obtain the following restriction:

$$\sum_{j=1}^{d} \omega_j E[X_j \,|\, S \leqslant K] = E[S \,|\, S \leqslant K].$$

That is, conditional on S being below some critical level K, the average value of the portfolio of the d assets must be equal to that of the index itself.

Generally, there will be many copulas C that are consistent with (4). This situation is not uncommon given that in incomplete markets, the risk-neutral measure is not unique. However, although they could differ in "micro" details, all solutions will agree on broad, "aggregate" features. In particular, all solutions will imply the same average pairwise correlation ρ obtained by the existing approaches. More generally, they will all agree on moments $E[g(Z,S)]$ for any function $g(z,s)$. In this sense, any solution will provide very valuable information.

B Inferring Dependence

B.1 Toy Example of BRA

We provide full details of the BRA algorithm for the oversimplified example from Section 3.1. There are $d = 3$ assets and $n = 5$ equiprobable. The matrix $\mathbf{M}$ in (10) represents the values of Y_1, Y_2, Y_3, and $-S$. To simplify presentation, in what follows, we denote $Y_4 = -S$. For the original matrix, we have $V := \mathrm{var}(Y_1 + Y_2 + Y_3 + Y_4) = 126$.

The objective is to rearrange the values within each column such that the row sums of the four columns of $\mathbf{M}$ are all equal to zero, which is equivalent to $\mathrm{var}(Y_1 + Y_2 + Y_3 + Y_4) = 0$. Recall that rearrangements within columns affect the dependence among the Y_j but not their marginal distributions. In contrast, swapping values among columns would affect the marginal distributions and is not allowed.

Step 1: Rearranging Y_1,

$$Y_2 + Y_3 + Y_4 = \begin{bmatrix} -18 \\ -8 \\ -3 \\ 2 \\ 10 \end{bmatrix} \qquad \mathbf{M}^{(1)} = \begin{bmatrix} 6 & 1 & 0 & -19 \\ 5 & 2 & 3 & -13 \\ 3 & 3 & 4 & -10 \\ 2 & 5 & 5 & -8 \\ 1 & 7 & 9 & -6 \end{bmatrix} \qquad V = 58.$$

Step 2: Rearranging Y_2,

$$Y_1 + Y_3 + Y_4 = \begin{bmatrix} -13 \\ -5 \\ -3 \\ -1 \\ 4 \end{bmatrix} \qquad \mathbf{M}^{(2)} = \begin{bmatrix} 6 & 7 & 0 & -19 \\ 5 & 5 & 3 & -13 \\ 3 & 3 & 4 & -10 \\ 2 & 2 & 5 & -8 \\ 1 & 1 & 9 & -6 \end{bmatrix} \qquad V = 12.4.$$

Step 3: Rearranging Y_3,

$$Y_1 + Y_2 + Y_4 = \begin{bmatrix} -6 \\ -3 \\ -6 \\ -4 \\ -4 \end{bmatrix} \qquad \mathbf{M}^{(3)} = \begin{bmatrix} 6 & 7 & 9 & -19 \\ 5 & 5 & 0 & -13 \\ 3 & 3 & 5 & -10 \\ 2 & 2 & 3 & -8 \\ 1 & 1 & 4 & -6 \end{bmatrix} \qquad V = 4.$$

In this case, the order of the fourth and fifth rows for Y_3 is arbitrary, and the rearrangement is not unique. Both lead to the same new variance of 4.

Step 4: Rearranging Y_4,

$$Y_1 + Y_2 + Y_3 = \begin{bmatrix} 22 \\ 10 \\ 11 \\ 7 \\ 6 \end{bmatrix} \qquad \mathbf{M}^{(4)} = \begin{bmatrix} 6 & 7 & 9 & -19 \\ 5 & 5 & 0 & -10 \\ 3 & 3 & 5 & -13 \\ 2 & 2 & 3 & -8 \\ 1 & 1 & 4 & -6 \end{bmatrix} \qquad V = 2.8.$$

Step 5: Rearranging the block $[Y_1 \quad Y_2]$ does not reduce the variance, as it is already (weakly) antimonotonic. We can keep it unchanged or swap the first and second rows. We do the latter to illustrate the algorithm:

$$Y_1 + Y_2 = \begin{bmatrix} 13 \\ 10 \\ 6 \\ 4 \\ 2 \end{bmatrix} \qquad Y_3 + Y_4 = \begin{bmatrix} -10 \\ -10 \\ -8 \\ -5 \\ -2 \end{bmatrix} \qquad \mathbf{M}^{(5)} = \begin{bmatrix} 5 & 5 & 9 & -19 \\ 6 & 7 & 0 & -10 \\ 3 & 3 & 5 & -13 \\ 2 & 2 & 4 & -8 \\ 1 & 1 & 3 & -6 \end{bmatrix} \qquad V = 2.8.$$

Step 6: Rearranging the block $[Y_1 \quad Y_3]$ does not help either, as it is already antimonotonic:

$$Y_1 + Y_3 = \begin{bmatrix} 14 \\ 6 \\ 8 \\ 6 \\ 4 \end{bmatrix} \qquad Y_2 + Y_4 = \begin{bmatrix} -14 \\ -3 \\ -10 \\ -6 \\ -5 \end{bmatrix} \qquad \mathbf{M}^{(6)} = \begin{bmatrix} 5 & 5 & 9 & -19 \\ 6 & 7 & 0 & -10 \\ 3 & 3 & 5 & -13 \\ 2 & 2 & 4 & -8 \\ 1 & 1 & 3 & -6 \end{bmatrix} \qquad V = 2.8.$$

Step 7: Rearranging the block $[Y_1 \quad Y_4]$, we need to swap the second, fourth, and fifth rows:

$$Y_1 + Y_4 = \begin{bmatrix} -14 \\ -4 \\ -10 \\ -6 \\ -5 \end{bmatrix} \qquad Y_2 + Y_3 = \begin{bmatrix} 14 \\ 7 \\ 8 \\ 6 \\ 4 \end{bmatrix} \qquad \mathbf{M}^{(7)} = \begin{bmatrix} 5 & 5 & 9 & -19 \\ 1 & 7 & 0 & -6 \\ 3 & 3 & 5 & -13 \\ 6 & 2 & 4 & -10 \\ 2 & 1 & 3 & -8 \end{bmatrix} \qquad V = 2.$$

Step 8-11: Now, we go back to Step 1 and look again at the respective columns Y_1 (which is already antimonotonic), Y_2 (in matrix $\mathbf{M}^{(11)}$, which decreases the variance to 1.6), Y_3 (which is already antimonotonic in $\mathbf{M}^{(11)}$), and Y_4 (which is already antimonotonic in $\mathbf{M}^{(11)}$). After rearranging these four columns sequentially, we obtain

$$\mathbf{M}^{(11)} = \begin{bmatrix} 5 & 5 & 9 & -19 \\ 1 & 7 & 0 & -6 \\ 3 & 3 & 5 & -13 \\ 6 & 1 & 4 & -10 \\ 2 & 2 & 3 & -8 \end{bmatrix} \qquad V = 1.6.$$

Step 12: Next, we apply again the rearrangement on block $[Y_1 \;\; Y_2]$. We switch rows 2 and 3 and find that the variance is equal to 0. The algorithm has converged and the final matrix is

$$\mathbf{M}^{(12)} = \begin{bmatrix} 5 & 5 & 9 & -19 \\ 3 & 3 & 0 & -6 \\ 1 & 7 & 5 & -13 \\ 6 & 1 & 4 & -10 \\ 2 & 2 & 3 & -8 \end{bmatrix} \qquad V = 0.$$

B.2 Formal Description of BRA

The method we propose for inferring dependence is inspired by the so-called rearrangement algorithm (RA) of Puccetti and Rüschendorf (2012) and of Embrechts, Puccetti, and Rüschendorf (2013), which was originally introduced to deal with the assessment of model risk, adjusted by Bernard and McLeish (2016) and Bernard, Bondarenko, and Vanduffel (2018) to make it suitable for inferring dependence, and termed the block rearrangement algorithm (BRA).

The objective is to find a rearrangement of the first d columns such that the row sums of the $d+1$ columns of $\mathbf{M}$ are all equal to zero. In other words, the opposite of the last column is the sum of the previous columns, i.e., $S = \sum_{j=1}^{d} Y_j$. Denote this last column by $Y_{d+1} = -S$. This procedure is equivalent to finding a rearrangement of the matrix $\mathbf{M}$ such that $Y_1 + \cdots + Y_{d+1}$ is identically equal to zero and thus such that $\mathrm{var}(Y_1 + \cdots + Y_{d+1}) = 0$. We allow for rearrangements within columns, as doing so affects the dependence among Y_j, $j = 1, \ldots, d$ but not their respective marginal distributions. By contrast, swapping values among columns will affect the marginal distributions and is not allowed. Clearly, for $Y_1 + \cdots + Y_{d+1}$ to have the smallest possible variance, it must hold that for all $\ell = 1, \ldots, d+1$, Y_ℓ is as negatively correlated as possible with $\sum_{j=1, j\neq\ell}^{d+1} Y_j$ (Puccetti and Rüschendorf 2012, Theorem 2.1), i.e., is antimonotonic. This observation lies at the core of this rearrangement method.

In fact, it must actually hold that for any decomposition of $\{1, \ldots, d+1\} = I_1 \cup I_2$ into two disjoint sets I_1 and I_2, the sums $S_1 := \sum_{k\in I_1} Y_j$ and $S_2 := \sum_{k\in I_2} Y_j$ are antimonotonic and not only for singleton sets of the form $I_1 = \{j\}$. This observation makes it possible to generalize the standard RA by rearranging "blocks of columns" instead of one column at a time: the columns in the first set I_1 are stacked into a matrix (block) $\mathbf{Y}_1$, and its rows are rearranged (i.e., entire rows are swapped) such that the row sums of $\mathbf{Y}_1$ (reflecting S_1) are in increasing order. As for the matrix $\mathbf{Y}_2$ that is formed by stacking the remaining columns, the rows are rearranged such that the row sums (reflecting S_2) are in decreasing order.

Algorithm for inferring dependence

1. Select a random sample of n_{sim} possible partitions of the columns $\{1, \ldots, d+1\}$ into two nonempty subsets $\{I_1, I_2\}$. In our case, $d = 9$ and it is feasible to consider all nontrivial partitions, thus we take $n_{sim} = 2^d - 1$.
2. For each of the n_{sim} partitions, create the matrices (blocks) $\mathbf{Y}_1$ and $\mathbf{Y}_2$ with corresponding row sums S_1 and S_2 and rearrange rows of $\mathbf{Y}_2$ so that S_2 is antimonotonic to S_1.
3. If there is no improvement in var $\left(\sum_{j=1}^{d+1} Y_j\right)$, output the current matrix $\mathbf{M}$; otherwise, return to step 1.

At each step of this algorithm, we ensure that the variance decreases or remains the same: the columns, say Y_j before rearranging and $\widetilde{Y}_j$ after rearranging, satisfy the inequality[25]

$$\text{var}\left(\sum_{j=1}^{d+1} Y_j\right) \geqslant \text{var}\left(\sum_{j=1}^{d+1} \widetilde{Y}_j\right).$$

B.3 Illustration of BRA for $d = 2$

There are two assets, whose returns X_1 and X_2 are normally distributed with zero mean and standard deviations $\sigma_1 = 0.2$ and $\sigma_2 = 0.4$. The two marginal distributions are fixed, but we vary the distribution for the weighted sum $S = \frac{1}{2}X_1 + \frac{1}{2}X_2$. We consider three cases:

(1) S is $N(0, \sigma_S^2)$, where σ_S is chosen such that the implied correlation is 0;

(2) S is $N(0, \sigma_S^2)$, where σ_S is chosen such that the implied correlation is 0.97;

(3) S follows a skewed distribution with a heavy left tail, modeled by a mixture of two normals.

The three cases (no dependence, strong dependence, and asymmetric dependence) are shown in Figures 13–15. In each figure, the top panels show PDFs and CDFs of X_1, X_2 and S, which are all centered at zero. We discretize the three CDFs into $n = 1,000$ equiprobable states (only ten of them are shown on the plots) and run BRA to extract the dependence and joint distribution between X_1 and X_2. Those are represented by 1,000 dots and are shown in the two bottom panels.

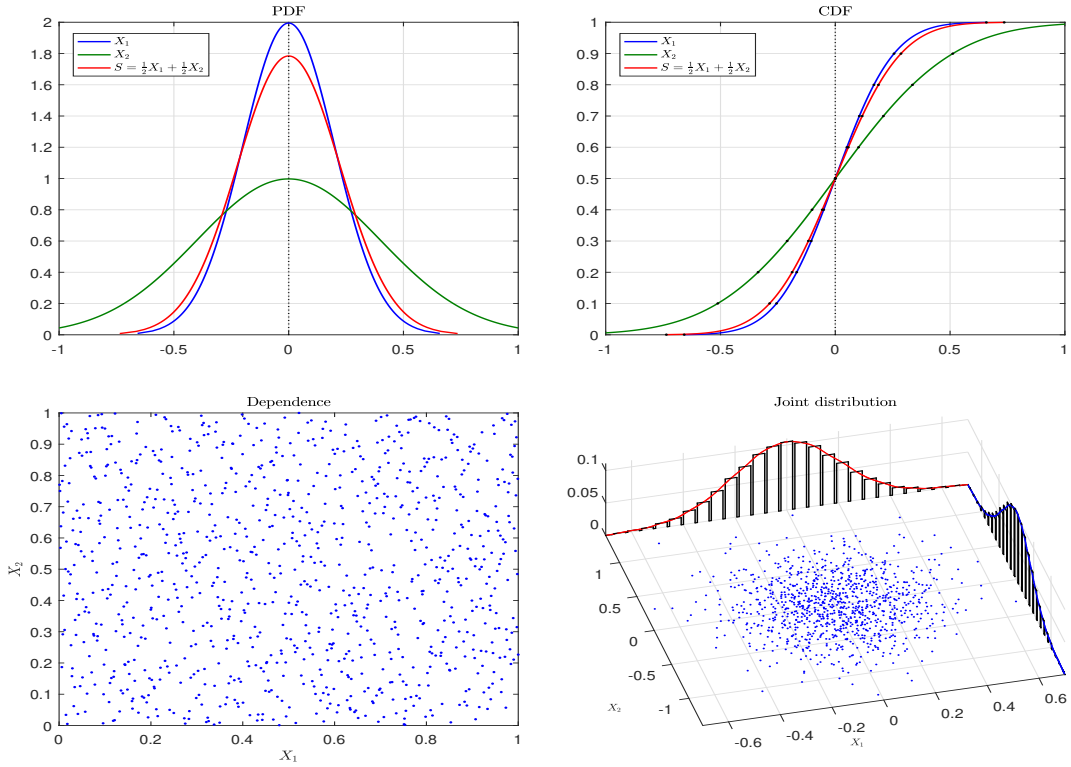

Figure 13: **No Dependence.**

[25] Indeed, $\text{var}\left(\sum_{k=1}^{d+1} Y_k\right) = \text{var}\left(Y_j + \sum_{k\neq j} Y_k\right)$, and a necessary condition for $\text{var}\left(\sum_{k=1}^{d+1} Y_k\right)$ to become minimum is that each Y_j is antimonotonic with $\sum_{k\neq j} Y_k$.

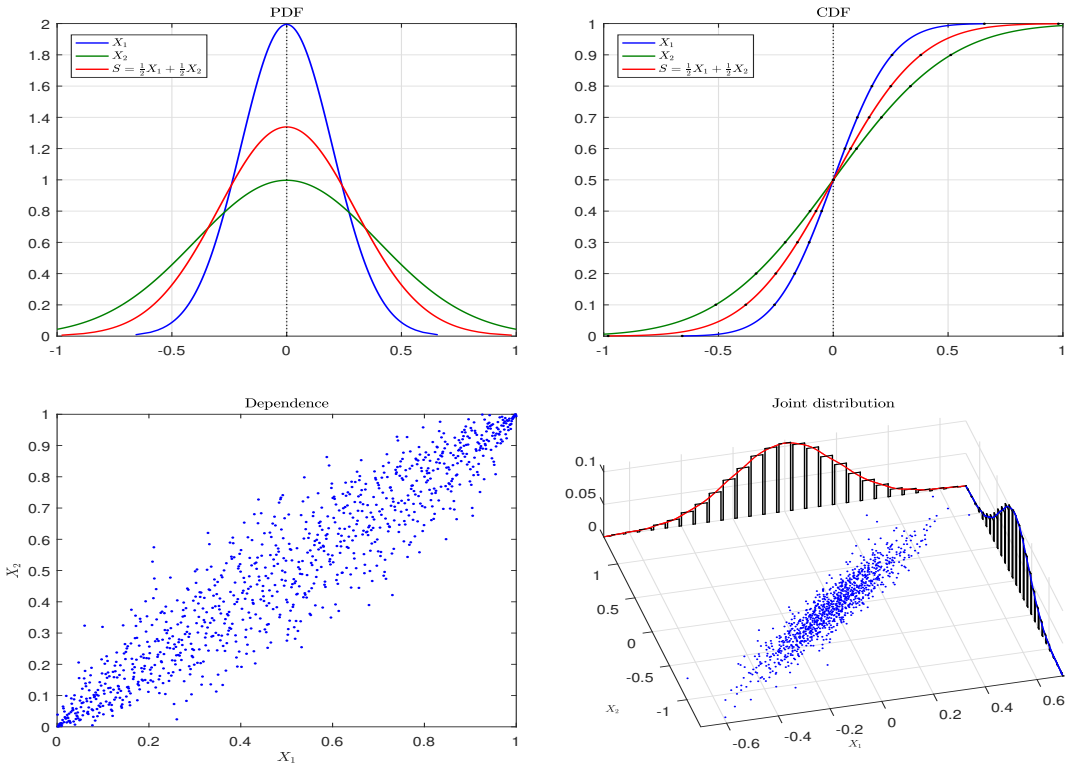

Figure 14: **Strong Dependence.**

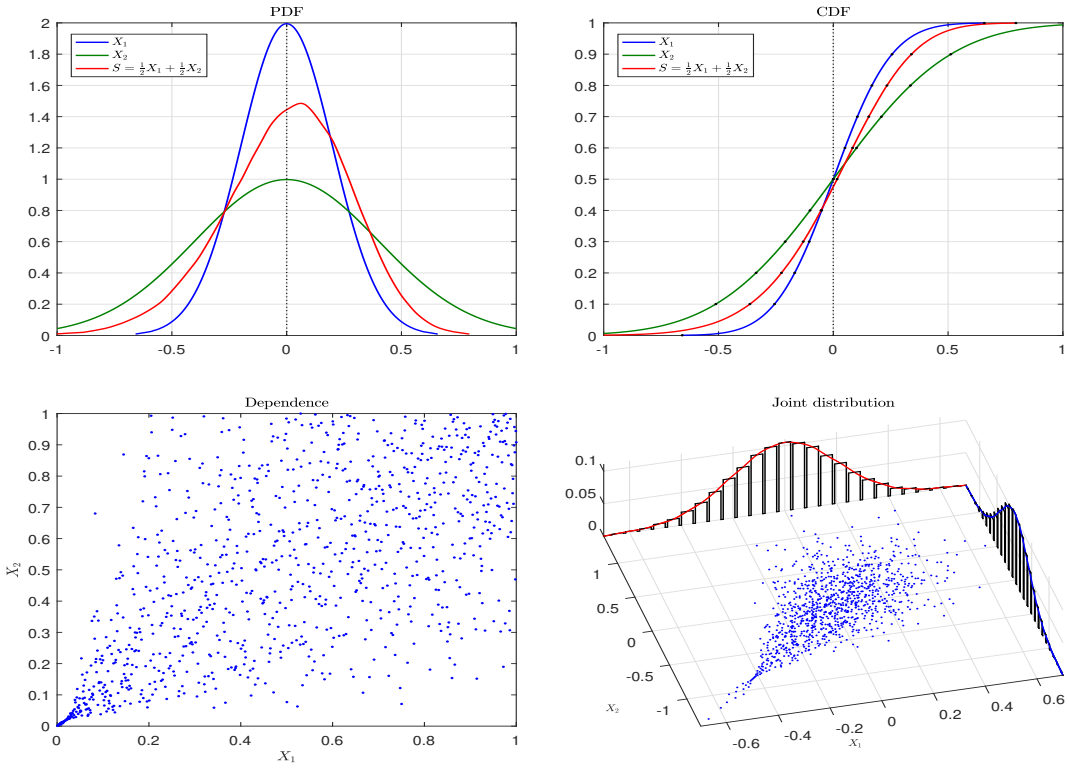

Figure 15: **Asymmetric Dependence.**

C CBOE Options

We use CBOE options on the SPDR ETFs for the nine Select Sectors and the S&P 500 itself. The ETFs are managed by State Street Global Advisors. The stocks in the S&P 500 index are divided into eleven industry sectors, but Information Technology and Telecommunications are combined in a single ETF (ticker XLK), while Financial and Real Estate are combined in another ETF (ticker XLF). The ETF options are physically settled and have an American-style exercise. Their underlying assets pay quarterly dividends. The option contract size is 100 shares of the corresponding ETF. The minimum price movement is 0.05. The strikes are multiples of $1. Sector options all expire on the same day. On any given trading day, we estimate option-implied RNDs with a constant time to maturity of 91 days. The details of the procedure are summarized below.

Dataset Construction

1. We use closing quotes to compute midpoint prices. In the dataset, we match all puts and calls by trading date t, maturity T, and strike K. For each pair (t, T), we drop very low (high) strikes with zero bids. We approximate the risk-free rate r over $[t, T]$ by the rate of 3-month Treasury bills.

2. Because sector options are American-style, their prices $P_t^A(K)$ and $C_t^A(K)$ could be slightly higher than the prices of the corresponding European options $P_t(K)$ and $C_t(K)$. The difference, however, is small for the short maturities on which we focus. This is particularly true for OTM and ATM options. To infer the prices of European options $P_t(K)$ and $C_t(K)$ on a given underlying X_t and maturity $\tau = T - t$, we proceed as follows. First, we discard all ITM options. That is, we use put prices for $K/Z_t \leqslant 1.00$ and call prices for $K/Z_t \geqslant 1.00$, where $Z_t := X_t e^{(r-\delta)\tau}$ is the forward price. Prices of OTM and ATM options are both more reliable and less affected by the early exercise feature. Second, we correct American option prices $P_t^A(K)$ and $C_t^A(K)$ for the value of the early exercise feature by using Barone-Adesi and Whaley (1987) approximation. Third, we compute the prices of ITM options through the put-call parity relationship:

$$C_t(K) - P_t(K) = (Z_t - K)e^{-r\tau}.$$

3. We check option prices for violations of the no-arbitrage restrictions. To preclude arbitrage opportunities, European call and put prices must be monotonic and convex functions of the strike. In particular, the call pricing function $C_t(K)$ must satisfy

$$(a) \quad C_t(K) \geqslant (F_t - K)^+ e^{-r\tau}, \qquad (b) \quad -e^{-r\tau} \leqslant C_t'(K) \leqslant 0, \qquad (c) \quad C_t''(K) \geqslant 0.$$

In real data, however, restrictions (a)-(c) can sometimes be violated and we enforce them by running the *Constrained Convex Regression* (CCR) introduced in Bondarenko (2000). Intuitively, CCR searches for the smallest (in the sense of least squares) perturbation of option prices that restores the no-arbitrage restrictions. The procedure is also useful for identifying possible recording errors or typos.

4. We construct prices of synthetic options with constant time to maturity $\tau_c = T_c - t = 91$ days. Specifically, we start with two available time to maturities τ_1 and τ_2 which bracket the target time to maturity τ_c. Cleaned European options for τ_1 and τ_2 are converted into implied volatilities and linearly interpolated with respect to the normalized moneyness m, defined in (14). The interpolated implied volatilities for τ_c are then converted back into option prices, which are used for the RND estimation.[26]

5. For each trading day t, we estimate the RND corresponding to time to maturity τ_c using the method of *Positive Convolution Approximation* (PCA) developed in Bondarenko (2003). The method allows one

[26] For most days, constructing options with constant time to maturity does not require extrapolation, as there exist two maturities such that $\tau_1 \leqslant \tau_c \leqslant \tau_2$. However, in the earlier part of our sample, there are a few days for which the shortest available maturity τ_1 exceeds τ_c and extrapolation is unavoidable. We discard days which require extensive extrapolation, that is, when the weight for the shortest maturity is negative and less than -0.5.

to infer the RND $f_t(x)$ and RNCD $F_t(x)$ through the relationships in (1) and (2). The method directly addresses the important limitations of option data that (a) options are only traded for a discrete set of strikes, as opposed to a continuum of strikes, (b) very low and very high strikes are unavailable, and (c) option prices are recorded with substantial measurement errors, which arise from nonsynchronous trading, price discreteness, and the bid-ask bounce. The PCA method is fully nonparametric, always produces arbitrage-free estimators, and controls against overfitting while allowing for small samples. We implement a version of PCA for which extreme left and right tails of the RND are extended in accordance with power laws.

6. Given the estimated RNCDs for the nine sectors and the index, we discretize each of them into $n = 1,000$ equally probable points, collect them into $n \times (d+1)$ matrix $\mathbf{M}$, and apply the BRA method described in Section 3. The output of the BRA is given by another $n \times (d+1)$ matrix that describes the joint distribution between the nine sectors compatible with all marginal distributions. Armed with the full joint distribution, we are able to compute various statistics of interest, including global, down, and up pairwise correlation for sector returns.

D Stability of the Algorithm

To assess the stability of our MFDR, we conduct a Monte-Carlo study. We distinguish between two separate effects: (i) instability of the optimization algorithm (BRA) itself, and (ii) instability of the inputs (i.e., measurement errors in the marginal distributions F_j and F_S estimated in the first step of MFDR). The first effect (unrealistically) assumes that the inputs are perfectly measured and focuses on how much noise is introduced by the optimization algorithm. Since the BRA is an heuristic approach, it finds an *approximate* solution. In addition, since the algorithm relies on initial randomization, each run actually yields a *different* approximate solution. However, as we show next, the first effect is extremely small and can easily be ignored in practice. Moreover, the effect can be made arbitrarily small by increasing the number of states n. The second effect is not specific to our method only. The existing approaches need estimates of MFIV or $\sigma^{\mathbb{Q}}$, and those too require the knowledge of the whole distribution. Options prices are recorded with considerable measurement errors (due to illiquidity, non-synchronous trading, and other market imperfections) and these errors affect the estimation of the corresponding risk-neutral distributions and derived measures. The second effect is typically much larger than the first one.

To assess the first effect, we proceed as follows. For each day in our sample, we run the algorithm $N_{sim} = 10$ times for fixed inputs. As before, we use the distributions F_j and F_S estimated from midpoint prices of available options. This gives us N_{sim} joint distributions, which are not identical due to noise introduced by the BRA. From each joint distribution we compute the corresponding correlations ρ^g, ρ^d, and ρ^u and then bootstrap 50,000 times to obtain the confidence intervals for the time-series averages. The results for this experiment are reported in the first three columns in Table 7. The confidence intervals are extremely narrow. For example, the widest 1%–99% interval is for the up correlation and it is only 0.0002. As it turns out, the confidence intervals for a *single* day are already very narrow, but the intervals calculated for the *sample averages* become negligible.[27]

The last three columns in Table 7 report on the second effect. We conduct a similar experiment, except now joint distributions are estimated from *perturbed* option prices. We want to add noise to option prices and assess how this affects the estimates of the three types of correlations. It is not completely obvious how to model realistic measurement errors. Our approach is to assume that the true, but unobserved, option price is uniformly distributed between the bid and the ask. That is, instead of the midpoint prices, we use simulated prices. This leads to N_{sim} sets of perturbed marginal distributions F_j and F_S and, thus, perturbed joint distributions. The last three columns in Table 7 demonstrate that the second effect is at least 10 times larger than the first effect, but it is still quite small in economic terms.

[27] As mentioned earlier, the accuracy of the BRA can be further improved by increasing the number of states, to, say, $n = 10,000$. This only increases the computation time but has no real effect on the empirical results.

Again, the confidence intervals for the up correlation are widest because the information on the right tail of RND is typically less precise than for the left tail (there are more OTM puts available than OTM calls). The 1%–99% interval for the up correlation is now 0.0026.

	Fixed Inputs			Perturbed Inputs		
	$P_{0.90}$-$P_{0.10}$	$P_{0.95}$-$P_{0.05}$	$P_{0.99}$-$P_{0.01}$	$P_{0.90}$-$P_{0.10}$	$P_{0.95}$-$P_{0.05}$	$P_{0.99}$-$P_{0.01}$
Global	0.000005	0.000084	0.000110	0.0011	0.0013	0.0015
Down	0.000007	0.000107	0.000141	0.0014	0.0017	0.0018
Up	0.000010	0.000151	0.000203	0.0019	0.0023	0.0026

Table 7: **Stability of MFDR.** The table reports the size of the confidence intervals for the average global, down, and up correlations. The confidence intervals are obtained via bootstrap for the two cases: (1) when the BRA inputs are fixed (the first three columns), and (2) when the BRA inputs are perturbed (the last three columns). The confidence intervals are computed as the difference between percentiles at 10% and 90%, 5% and 95%, and 1% and 99%, respectively.

E Hybrid Model

The definition in (3) says that the weighted sum Z can be constructed from the known margins F_j and the copula C. By changing the copula C, we obtain different distributions of Z and, ideally, want to match the known distribution of S as in (4). We need some measure of goodness-of-fit. One intuitive approach is to compare the two implied curves IV^Z and IV^S, which correspond to Z and S. Specifically, we define the distance as the root mean squared relative error (RMSRE):

$$D := D(Z,S) = \sqrt{\frac{1}{L}\sum_{k=1}^{L}\left(\frac{IV_k^Z - IV_k^S}{IV_k^S}\right)^2}, \tag{26}$$

where we use the $L = 20$ moneyness levels $k = 1, \ldots, L$, equally spaced between 1% and 99%.

We consider several choices for the copula C and thus for the corresponding weighted sum Z. One choice corresponds to the output of our model-free approach, which we, as before, denote as EE. Alternatively, we can construct the copula from some parametric model. A common choice in the literature is the normal copula, but this model is not likely to perform well, given our finding on the very strong asymmetry of the implied dependence. Therefore, we consider a more flexible generalization, a multivariate skewed normal (SN) copula with two parameters ρ and δ, denoted by $SN(\rho, \delta)$ (see Section E.1 for the formal definition). We focus on the homogeneous case, where the pairwise correlations are identical and the skewness parameter is also identical across all assets.

In Table 8 and Figure 16, we contrast different models for September 8, 2008, one of the two dates studied in Section 4. Initially, we choose the parameters of the SN copula to match the standard deviation of the index $\sigma_S^{\mathbb{Q}}$. Specifically, for three (arbitrary) choices of the skewness parameter δ, we vary the remaining free parameter ρ to match $\sigma_S^{\mathbb{Q}}$. These three cases, denoted as SN_1, SN_2, and SN_3, are presented in the first three rows of Table 8. The second and third columns in Table 8 report the parameters of the SN copulas. Next, we estimate the "optimal" SN copula, for which both parameters are chosen to minimize the distance D in (26). That model, denoted SN^*, is presented in the fourth row. Finally, the EE model from our MFDR approach is presented in the last row.

For each of the five models, Table 8 also reports the average correlations (global, down, and up). Since models SN_1, SN_2, and SN_3 all match $\sigma_S^{\mathbb{Q}}$, the global correlation is forced to be the same by construction. That is, any approach based solely on matching $\sigma_S^{\mathbb{Q}}$, as is typically done in the literature, would yield the

exact same global correlation if the index distribution were given by any of the three models SN_1–SN_3. In other words, the approach would not have been able to distinguish between the three cases even though their down and up correlations are drastically different.

Model	δ	ρ	$\sigma^{\mathbb{Q}}$	ρ^g	ρ^d	ρ^u	$\Delta\rho$	$D\times 100$
SN_1	-0.87	0.129	0.454	0.579	0.429	0.124	0.305	2.78
SN_2	0.00	0.589	0.454	0.579	0.311	0.284	0.027	8.13
SN_3	0.80	0.353	0.454	0.579	0.233	0.389	-0.156	14.90
SN^*	-0.83	0.230	0.454	0.580	0.425	0.157	0.268	0.87
EE			0.454	0.578	0.438	0.172	0.266	0.19

Table 8: **Five Fitted Models.** The table reports statistics for the five models estimated on September 8, 2008. SN_1, SN_2, and SN_3 are the skewed normal copulas fit to match the standard deviation of the index $\sigma_S^{\mathbb{Q}}$; SN^* is the SN copula fit to match the whole IV curve by minimizing the distance D in (26); EE is the empirical copula obtained from our algorithm. Here, δ and ρ are the parameters of the SN models; ρ^g, ρ^d, and ρ^u are the average global, down, and up correlations; $\Delta\rho$ is the correlation spread.

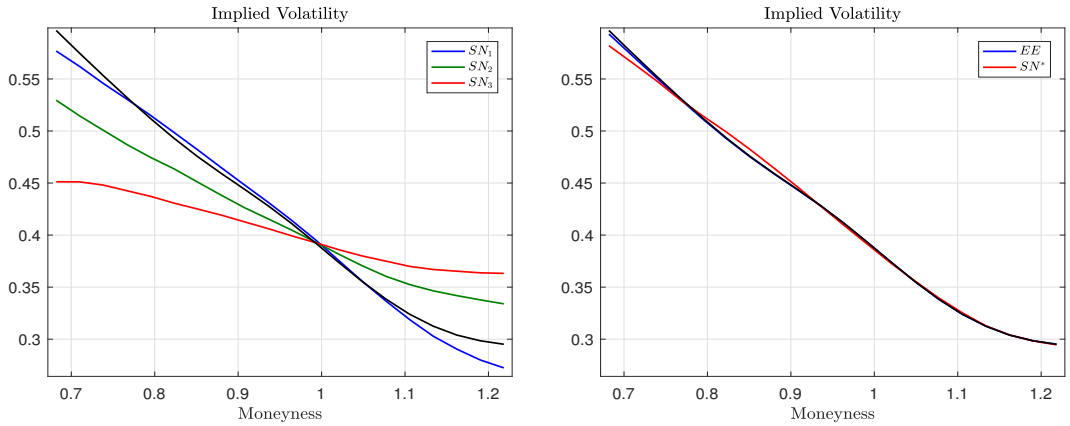

Figure 16: **Implied Volatilities for the Five Models on September 8, 2008.** Left panel: Skewed normal copulas SN_1–SN_3. Right panel: Skewed normal copula SN^* and empirical copula (from MFDR) EE. The black curves in both panels show the true implied volatilities (for the index) IV^S.

Figure 16 plots the implied volatility curves IV^Z for the five models SN_1, SN_2, SN_3, SN^*, and EE. For reference, also shown is the implied volatility curve for the index IV^S (the black curve). The left panel of Figure 16 shows the three models SN_1–SN_3, which are fit solely to match $\sigma^{\mathbb{Q}}$ and thus (5). They are very far from the observed IV^S (the black curve). The right panel of Figure 16 shows models SN^* and EE, with the latter one matching IV^S almost perfectly. Note that model SN^* does not match $\sigma^{\mathbb{Q}}$ perfectly, as it is fit by minimizing a different objective function (the distance D). Even for the optimal SN model, the distance D is about 4.6 times larger than for the EE model, which does not even aim at minimizing the distance D as its objective. Visually, the fit with model SN^* seems quite good on September 8, 2008, but this is not always the case. Table 9 and Figure 17 repeat the same analysis, but now for a different day, June 12, 2017. For that day, even the optimal SN model is too inflexible to approximate the implied volatility curve well. The distance D is now about 38 times as large as that for the EE model.

The above illustrations focus on two specific dates. To generalize the analysis, we fit a multivariate normal copula (with one parameter ρ) and a multivariate skewed normal copula (with two parameters ρ and δ) for all days in our sample. Specifically, we search for the normal copula N^* and the skewed normal

copula SN^* that minimize the distance D in (26) on each day. In Table 10, we report the time-series averages of the fitted parameters and distance D. It is clear that adding the skewness parameter greatly reduces that distance (by a factor of more than 2), but the distance for SN^* is still much larger than for the model-free EE (by a factor of 9). Note also that the correlation coefficient in the normal copula ρ is not directly comparable to the coefficient ρ in the skewed normal copula. For SN^*, the average global correlation matches exactly that for EE (0.770), even though ρ is on average equal to only 0.414 (combined with a very negative skewness parameter $\delta = -0.875$). Recall that the three models do not have to match the average global correlation, as the objective is not to match $\sigma^{\mathbb{Q}}$ but instead to minimize the distance D.[28]

Model	δ	ρ	$\sigma^{\mathbb{Q}}$	ρ^g	ρ^d	ρ^u	$\Delta\rho$	$D\times 100$
SN_1	-0.87	0.180	0.255	0.616	0.496	0.159	0.336	10.23
SN_2	-0.40	0.621	0.255	0.616	0.409	0.350	0.060	14.21
SN_3	0.80	0.486	0.255	0.616	0.344	0.487	-0.143	19.18
SN^*	-0.84	0.220	0.253	0.606	0.490	0.163	0.327	9.50
EE			0.255	0.610	0.629	0.124	0.506	0.25

Table 9: **Five Fitted Models.** The table reports statistics for the five models estimated on June 12, 2017. SN_1, SN_2, and SN_3 are the skewed normal copulas fit to match the standard deviation of the index $\sigma_S^{\mathbb{Q}}$; SN^* is the SN copula fit to match the whole IV curve by minimizing the distance D in (26); EE is the empirical copula obtained from our algorithm. Here, δ and ρ are the parameters of the SN models; ρ^g, ρ^d, and ρ^u are the average global, down, and up correlations; $\Delta\rho$ is the correlation spread.

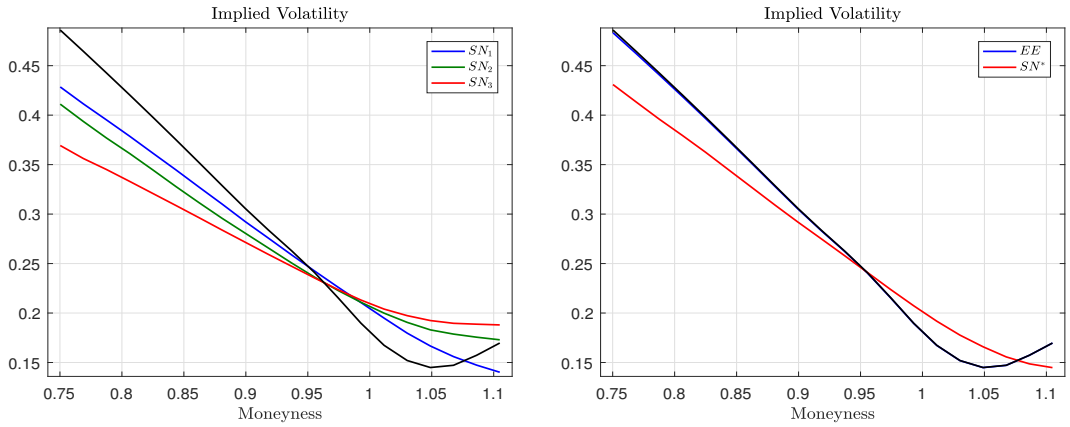

Figure 17: **Implied Volatilities for the Five Models on June 12, 2017.** Left panel: Skewed normal copulas SN_1–SN_3. Right panel: Skewed normal copula SN^* and empirical copula (from MFDR) EE. The black curves in both panels show the true implied volatilities (for the index) IV^S.

Table 10 shows that the normal copula is unable to reproduce the asymmetry between the down and up correlations that is exhibited by the model-free approach EE. Specifically, using the EE approach, the correlation spread $\Delta\rho = \rho^d - \rho^u = 0.399$. Even though a normal copula is symmetric, the down and up correlations are not identical because the margins are skewed (the heavy left tail). For model N^*, the correlation spread $\Delta\rho = 0.048$, or approximately 12% of that for EE. Intuitively, about 12% of the correlation spread can be attributed to nonnormality of the margins. This complements the results

[28]Furthermore, recall that the parameter ρ of the normal copula can only be interpreted as a correlation coefficient if the margins are normal. Thus, the coefficient ρ does not match the average global correlation.

of Table 5 in Section 5.3. On the other hand, the skewed normal model with only one extra parameter can reproduce the down and up correlations considerably better, now accounting for approximately 76% of the correlation spread (0.305/0.399). This also means that many of our conclusions regarding the down and up CRP can be (approximately) confirmed by adopting the hybrid model SN^*, for which nonparametric margins are joined by the SN copula.

Model	ρ^g	ρ^d	ρ^u	$\Delta\rho$	ρ	δ	$D\times 100$
N^*	0.744	0.572	0.524	0.048	0.760		7.27
SN^*	0.770	0.684	0.379	0.305	0.414	-0.875	3.09
EE	0.770	0.731	0.332	0.399			0.34

Table 10: **Three Models over the Full Sample.** The table reports time-series averages for three models estimated daily from January 1, 2007, to June 12, 2017. N^* is the normal copula with parameter ρ, and SN^* is the skewed normal copula with two parameters, ρ and δ. They are fit to match the whole IV curve by minimizing the distance D in (26). EE is the model-free copula from our algorithm. The first three columns are the average global, down, and up correlations; the forth column is the correlation spread $\Delta\rho = \rho^d - \rho^u$.

E.1 Multivariate Skewed Normal Model

We follow the procedure of Azzalini and Valle (1996) to simulate a d-dimensional skewed normal copula. We are interested in the special case of constant pairwise correlation and constant skewness.

Generally, the joint pdf of $(X_1, X_2, \ldots, X_d)$ is a skewed normal distribution with mean parameter $\mathbf{0}$, correlation matrix R and skewness parameter $\boldsymbol{\lambda} = (\lambda_1, \ldots, \lambda_d)$ if

$$f(\mathbf{x}) = 2\phi_d(\mathbf{x}, \Sigma)\Phi(\boldsymbol{\alpha}^t\mathbf{x}),$$

where $\phi_d(\mathbf{x}, \Sigma)$ is the pdf of an MVN distribution with mean $\mathbf{0}$ and covariance matrix Σ evaluated at $\mathbf{x}$ and

$$\boldsymbol{\alpha}^t = \frac{\boldsymbol{\lambda}^t R^{-1}\Delta^{-1}}{\sqrt{1+\boldsymbol{\lambda}^t R^{-1}\boldsymbol{\lambda}}}, \quad \Sigma = \Delta(R + \boldsymbol{\lambda}\boldsymbol{\lambda}^t)\Delta,, \quad \Delta = diag\left(\sqrt{1-\delta_1^2}, \sqrt{1-\delta_2^2}, \ldots, \sqrt{1-\delta_d^2}\right),$$

$$\lambda_j = \frac{\delta_j}{\sqrt{1-\delta_j^2}}, \quad \text{for some} \quad \delta_j \in (-1, 1). \tag{27}$$

The simulation procedure can be summarized as follows:

a) Simulate Z as a standard normal $N(0,1)$ and simulate $(Y_1, \ldots, Y_d)$ as an MVN vector with mean $\mathbf{0}$ and correlation matrix R.

b) Define $X_j = \delta_j|Z| + \sqrt{1-\delta_j^2}Y_j$. Then each X_j is the standard skewed normal with parameter λ_j, and the d-dimensional vector $(X_1, \ldots, X_d)$ follows $SN_d(R, \boldsymbol{\lambda})$.

c) Obtain the d-dimensional skewed normal copula by replacing simulated values of $(X_1, \ldots, X_d)$ with their ranks.

In our case, $d = 9$, and we fit a skewed normal distribution with only two free parameters. One parameter, ρ, is the constant pairwise correlation in the correlation matrix R. The other parameter is a constant skewness coefficient (i.e., $\delta_1 = \ldots = \delta_d = \delta$ where δ is linked to λ by (27)).[29]

[29]Schreindorfer (2020) has recently used another multivariate skewed distribution to model asymmetric tail dependence. Similarly to the skewed normal distribution, it is also constructed as a mixture, but in step b) of the above simulation an exponential distribution is used instead of the absolute value of a standard normal distribution ($|Z|$) and ρ is set to zero.