Beyond Correlation: Extreme Co-movements Between Financial Assets *

Roy Mashal
Columbia University

Assaf Zeevi
Columbia University

This version: October 14, 2002

Abstract

This paper investigates the potential for extreme co-movements between financial assets by directly testing the underlying *dependence structure*. In particular, a t-dependence structure, derived from the Student t-distribution, is used as a proxy to test for this extremal behavior. Using likelihood ratio-based methods, we show that the presence of extreme co-movements is statistically significant in three asset markets (equities, currencies, and commodities), as well as across international (G5) equity markets. In addition, this likelihood ratio test indicates that the "correlation-based" Gaussian dependence structure is not supported on the basis of observed asset co-movements. The economic significance and consequences of these results are illustrated via several examples.

JEL Classification: C12, C15, C52, G11.

Keywords: asset returns, extreme co-movements, copulas, dependence modeling, hypothesis testing, international markets, pseudo-likelihood, portfolio models, risk management.

*The authors would like to thank Andrew Ang, Geert Bekaert, Mark Broadie, Loran Chollete, and Paul Glasserman for their helpful comments on an earlier version of this manuscript. Both authors are with the Columbia Graduate School of Business, Roy Mashal is also with Quantitative Credit Research at Lehman Brothers Inc. E-mail: rm586@columbia.edu, assaf@gsb.columbia.edu, current version available at http://www.columbia.edu/∼rm586

1 Introduction

Specification and identification of dependencies between financial assets is a key ingredient in almost all financial applications: portfolio management, risk assessment, pricing, and hedging, to name but a few. The seminal work of Markowitz (1959) and the early introduction of the Gaussian modeling paradigm, in particular dynamic Brownian-based models, have both contributed greatly to making the concept of *correlation* almost synonymous with that of dependence. Unfortunately, correlation is a rather limited notion of dependence and consequently provides partial and often misleading information on the actual underlying dependencies. This shortcoming is particularly evident when "extremal behavior" is to be modelled or inferred from financial data.

Recent years have witnessed an increasing interest in extremal phenomena involving financial assets including market linkage effects, contagion, "spillover," "breakdown of correlation" and *extreme co-movements*. These phenomena all bear significant implications on diversification strategies across international money markets, a topic that in itself has been the focus of many recent studies. Factors influencing joint movements in the US-Japan markets are identified by Karolyi and Stulz (1996) using regression methods, while Bae *et al.* (2002) consider contagion models. Forbes and Rigobon (2001) find evidence suggesting that contagion models are perhaps not the source of such observed co-movements, rather, the explanation lies in conditional correlation models. This line has also been advocated in the recent papers by Ang and Bekaert (2002), Ang and Chen (2002), and Ang *et al.* (2002) who provide ex post estimates of conditional correlation via regime shifts in correlation, or asymmetry in correlation based on return realizations. Extreme value theory provides another important set of tools and models by which joint extremal behavior can be examined. Longin and Solnik (1995), Longin (1996), and Longin and Solnik (2001) follow this approach in investigating the phenomena of large co-movements in international markets and changes in correlation driven by the state of the market. DanÍelsson and de Vries (1997) study tail index estimation via extremal statistics, while Starica (1999) develops extreme value theory for conditional correlation models. The limitations of simple correlation-based models are elucidated in the recent work of Embrechts *et al.* (2001b) and Hult and Lindskog (2001).

This study strives to contribute to this stream of research that focuses on the co-movement patterns and extremal phenomena exhibited by financial assets and markets. The methodology that this paper pursues, however, differs in a fundamental way from most of the methods described above. In particular, we propose to examine the potential for extreme co-movements via a direct test of the underlying *dependence structure*. The latter is a succinct and exact representation of the dependencies between underlying variables, irrespective of their marginal distributions.[1]

[1]To be precise, if X_i is, say, the ith asset return, $i = 1, \dots, d$, with cumulative distribution $F_i(\cdot)$, then $U_i := F_i(X_i)$ are distributed uniformly on the unit interval. The joint distribution of $(U_1, \dots, U_d)$ is called a *copula*; it retains the

Specifically, we focus on dependence structures derived from the multivariate Student t-distribution. The latter generalizes the ubiquitous Gaussian dependence structures and is more adequate for financial modeling, in the same spirit that the univariate Student t-distribution is a more adequate model than the Normal for individual asset returns.[2] In particular, the *t-dependence structure* has two salient features that the term "beyond correlation" alludes to. First, unlike the Gaussian structure where dependence is captured only via correlation, the t-dependence structure retains the use of correlation while introducing a single additional parameter, the degrees-of-freedom (DoF), that determines the potential for extreme co-movements. In particular, increasing the value of the DoF parameter decreases the tendency of underlying to exhibit extreme co-movements. Moreover, as the DoF tend to infinity one recovers the Gaussian dependence structure in a well defined limiting sense. Second, the t-dependence structure supports extreme co-movements regardless of the marginal behavior of the individual assets (e.g., this extremal behavior is manifestly present even if the individual assets follow "light-tailed" distributions). The t-dependence model can therefore serve two statistical testing objectives: (1) detect whether the presence of extreme co-movements is statistically significant, while concurrently indicating the extent of this extremal dependence via the aforementioned DoF parameter; and, (2) test the validity of the Gaussian dependence structure assumption that underlies most financial applications. Moreover, the t-dependence structure can still be easily implemented for purposes of calculating risk measures (e.g., VaR numbers) and pricing various instruments that are sensitive to extreme co-movements (such as multi-name credit derivatives and collateralized debt obligations).

To illustrate the flavor of our main results, as well as their financial significance, consider the G5 (France, Germany, Japan, UK and the US) equity market indices. We find that all pairs exhibit extremal behavior in a statistically significant manner, as manifested in a "small" value of the DoF parameter in the t-dependence structure. Two additional interesting observations can be made: (1) the Japanese index, while only moderately correlated with the rest (pair-wise correlations are roughly 35-40%), has a substantial tendency for extreme co-movements with the other indices; (2) a higher value of observed correlation may go hand-in-hand with a smaller potential for extreme co-movements (as in the UK-Japan pair). Thus, correlation is not representative, nor is it very informative of these extremal phenomena. The implications on risk exposure of international market portfolios is evident in value-at-risk (VaR) calculations and expected shortfall (conditional VaR). Both risk measures are seen to be substantially *over-optimistic* when the potential for extreme co-movements is neglected. In particular, we observe that VaR numbers, as well as expected shortfall,

original *dependence structure* while "normalizing" the univariate marginals, thus, effectively decoupling the two. Further details are given in Section 3.

[2]Mandelbrot (1963) and Fama (1965) were among the first studies to question the Gaussian assumption for individual stock return movements; subsequent studies find significant empirical support favoring the univariate t distribution, see, e.g., Praetz (1972), Blattberg and Gonedes (1974), and more recently Danİelsson and de Vries (1997).

can differ by a factor of two (100%) when extreme co-movements are not properly accounted for.

Our empirical study focuses on three representative financial time series of asset prices (commodities, equities, and foreign exchange rates), as well as a more detailed analysis of international equity markets across the G5 economies. Our main findings may be summarized informally as follows:

i.) **Empirical support for extreme co-movements.** All financial data sets indicate a non-trivial degree of dependence in the tails, and thus substantial potential for extreme co-movements.

ii.) **Evidence rejecting the Gaussian model.** A sensitivity analysis strongly suggests that the Gaussian, correlation-based, dependence structure should be rejected in all data sets with negligible error probability, when tested against the alternative t-dependence structure.

iii.) **Effects of dimensionality.** Statistical evidence supporting extreme co-movements is increasingly significant as the number of underlying assets increases.

The economic significance of these findings is subsequently illustrated via a close examination of the extreme co-movement potential across international (G5) equity markets, and the resulting implications on the value-at-risk of portfolios with holdings across these markets. We also indicate how the potential for extreme co-movements influences the price of multi-name credit derivatives; in this context extreme co-movements correspond to joint defaults of the underlyings.

In terms of methodology, we use maximum likelihood estimation and generalized likelihood ratio methods that exploit the natural nesting within the family of t-dependence models. (For example, the Gaussian dependence structure is seen to be a particular instance within the family of t-models.) This approach, is revealing in several ways. First, it allows us to obtain "sharp" results in support of the extreme co-movement potential, in particular, a sensitivity analysis of the likelihood ratio test statistic indicates a very small range of plausible values for the DoF parameter. (Recall, small values of the latter indicate a strong potential for extreme co-movements.) This, in turn, also provides significant evidence rejecting the Normal dependence structure. Second, we do not make any parametric assumptions on the individual asset distributions, rather, we use a semi-parametric estimation approach that treats the unknown marginal distributions as infinite dimensional "nuisance" parameters (indeed, these do not affect extreme *co-movements*). "Factoring out" the marginals introduces further dependence, which manifests itself in a slightly different limiting distribution of the likelihood ratio test statistic. In particular, this distribution is a scaled version of the familiar Chi-squared distribution. Since these results are, to the best of our knowledge, the first that are obtained in the context of an estimation theory for the t-dependence structure, we carry out a comprehensive Monte-Carlo simulation study that establishes the adequacy and validity of these methods. In particular, this study validates both the estimation procedure per se, as well as the asymptotic theory indicated above (i.e., consistency of the estimators and large sample properties of the test statistic).

Concentrating on the underlying dependence structure, in particular the t-based model, is inspired by the recent paper by Embrechts *et al.* (2001a) that focuses on copula models and discusses their implications on extreme tail dependence [see also, Bouyé *et al.* (2000)].[3] The copula-based approach completely side-steps the need to make any specific assumptions on the dynamic processes generating the individual movements. This stands in contrast with many recent papers, such as Bae *et al.* (2002) that use logistic regression methods, Karolyi and Stultz (1996) that rely on factor models and GARCH specifications, and the linear models considered in Forbes and Rigobon (2002). In addition, the copula-based approach and the use of a t-dependence structure introduces notions of dependence that extend and generalize correlation-based ones in an unconditional way. In contrast, both Forbes and Rigobon (2002) as well as the sequence of papers by Ang and Bekaert (2002), Ang and Chen (2002), and Ang *et al.* (2002) all propose conditional correlation-based models for coupled movements. The t-based model, on the other hand, explicitly hinges on measures that cannot be captured via correlation or conditional correlation-based models.

Perhaps the one study that is most akin to ours is the recent paper by Longin and Solnik (2001). The focus there is also on the dependence structure of joint asset returns, where a certain parameter serves as an indicator of extreme co-movement potential. In contrast to our work, their approach emphasizes extreme value theory both in terms of the choice of copula (Gumbel), and the choice of marginal specification (essentially a generalized Pareto distribution). Thus, when testing for extremal dependence, their results are influenced to some degree by the adequacy of their marginal distribution assumption. In fact, this can result in an inconsistent estimation of the "joint extremal parameter" and therefore their results are more restrictive (they become accurate only in a well-defined asymptotic sense). Moreover, their test for the Normal dependence structure essentially hinges on whether the extreme tail-index is zero or not. In contrast, we pursue this via the natural nesting of the Gaussian dependence structure within the family of t-based models. Consequently, we do not need to assume multivariate Normality, nor do we rely on any particular specification of the marginals as Longin and Solnik (2001) do. In addition, we test for empirical evidence of extreme co-movements between any number of given assets whereas Longin and Solnik (2001) only consider the bivariate case. (It turns out that this dimensionality actually plays an interesting role in making

[3] Copula methods have received increasing attention recently, mostly in the context of financial and risk modeling [for some applications see, e.g., Li (2000), Frey *et al.* (2001), and Frey and McNeil (2001)]. Testing for the underlying dependence structure, in particular estimating the underlying copula, is discussed in broad terms in Durrelman *et al.* (2001), and general estimation theory is developed in Genest *et al.* (1995). For further discussion on estimation in the context of copulas the reader is referred to Joe (1997), Genest and Rivest (1993), and Klasssen and Wellner (1997).

extreme co-movements more pronounced.) Finally, our study also indicates that the t-dependence structure is quite useful not only as inference tool, but also from an application standpoint. This is illustrated via out-of-sample prediction of extreme co-movement frequencies, VaR calculations for multi-national portfolios, and pricing credit derivatives. Thus, the t-dependence structure provides a parsimonious and tractable parametric model that allows to mitigate some of the mis-specification risk associated with the more restrictive Gaussian dependence structure.

The next three sections contain mostly preliminaries, including qualitative empirical evidence (Section 2), a discussion of basic dependence concepts (Section 3), and details of the testing procedure (Section 4). The reader who wishes to skip these preliminary discussions may head directly to Section 5 which presents empirical evidence of extreme co-movements, followed by a comprehensive discussion of the economic significance of these results in Section 6. The latter two sections constitute the heart of the paper, and are followed by some concluding remarks (Section 7). Technical details are deferred to three appendices: Appendix A discusses the theory of pseudo-likelihood asymptotics which are fundamental to the testing procedure; Appendix B contains a numerical analysis of the testing methodology, establishing its validity via a comprehensive Monte-Carlo study; and, Appendix C describes the implementation of the estimation procedure. Finally, the structure of the data is detailed in Appendix D.

2 A Qualitative Inspection of the Data

This section provides a brief qualitative analysis of three representative asset classes that serves to motivate the notions of dependence pursued in the sequel, (we postpone an examination of the G5 equity markets to Section 6.) The main data sets we consider span three different types of assets: equities, currencies, and commodities. For each of the above, a time-series of adjusted prices $S = (S_{i1}, S_{i2}, ..., S_{id})_{i=0}^{n}$ is given (the price adjustments accounting for splits etc.), where d is the number of assets in the "basket". The log-asset returns, X, are then defined in the usual way, viz,

$$X_{ij} = \log(\frac{S_{ij}}{S_{i-1j}}),\ i = 1, 2, ..., n,\ j = 1, 2, ..., d,$$

where $\log(\cdot)$ stands for the natural logarithm and $n+1$ is the number of trading days, thus, $X_i = (X_{i1}, X_{i2}, \ldots, X_{id})$ is a return observation for the i'th day. The number of assets in each basket is 30, 9, and 6 for the equity, currency, and commodity baskets respectively. The number of daily return observations for each asset is 2,526, 2,263, and 3,176 for the equity, currency, and commodity baskets respectively. Details on the specific tickers and date ranges are provided in Appendix D.

We begin our analysis with the univariate aspects of the data. Evidence that the distribution of univariate asset returns does not comply with the "Normality assumption" is quite conclusive, dating back to the pioneering work of Mandelbrot (1963) and Fama (1965), and explored further in the works of Praetz (1972) and Blattberg and Gonedes (1974).[4] It is common that univariate return distributions exhibit heavier tails and sharper peaks than the Normal distribution. These findings have triggered a search for alternative models for asset returns, candidates include the t-distribution, stable distribution, and characterizations derived from dynamic models such as jump diffusion and stochastic volatility models. The data sets that we analyze herein share the same characteristics; they exhibit, to varying extents, "fatter tails" than the Normal distribution and corresponding peakedness. Figure 1 displays these effects via Quantile-to-Quantile plots (Q-Q plots) of Alcoa equity and the Australian Dollar returns. In Panels A and B the quantiles of the standardized returns are plotted against quantiles of the Normal distribution. The deviation from the Normal quantiles (linear dashed line) represents the extra mass in the tails of the return distribution, corresponding also to positive kurtosis (extra "peakedness"). In Panels C and D the quantiles of the standardized returns are plotted against quantiles of t-distribution with the estimated DoF for each time series, these panels show that the univariate empirical data fit well the univariate t-distribution. [For the interpretation of such Q-Q plots and deviation from Normality see, e.g., Rencher (1995), p. 106.]

In contrast to the study of univariate distributions, this paper is concerned with the multivariate distribution of asset returns, and, more precisely, with the underlying dependence structure. From an intuitive standpoint, the univariate analysis concludes that fatter-than-Normal tails are present, thus, it would also seem plausible that the multivariate case would have more mass on the "joint tails." A slight caveat in this logic is that multivariate distributions can be easily constructed such that the marginal tails are "fat," but the joint dependence structure is Gaussian. In order to motivate the results of our study and to provide some intuitive understanding for the meaning of "fat joint tails" we show graphically in Figure 2 the difference between "light" and "heavy" mass on joint tails. As we will see in Section 5, the pair of currencies in Panel (b) is one for which the Gaussian dependence assumption is deemed inadequate, whereas in Panel (a) it is not. Panel (c) shows data simulated using a Normal-dependence structure while Panel (d) depicts points simulated with a t-dependence structure. All the four panels have the same number of points (2,263) and about the same linear correlation (16%, 22%, 18%, and 23% for Panels (a), (b), (c), and (d), respectively). Concentrating on the joint extreme realizations in the data, i.e. the upper right and lower left corners of the figures, the differences between Figures (a) and (c) to Figures (b) and (d) are evident; figures (b) and (d) have "many more" joint extreme realizations. The latter

[4] A comprehensive reference list can be found in Glasserman *et al.* (2002).

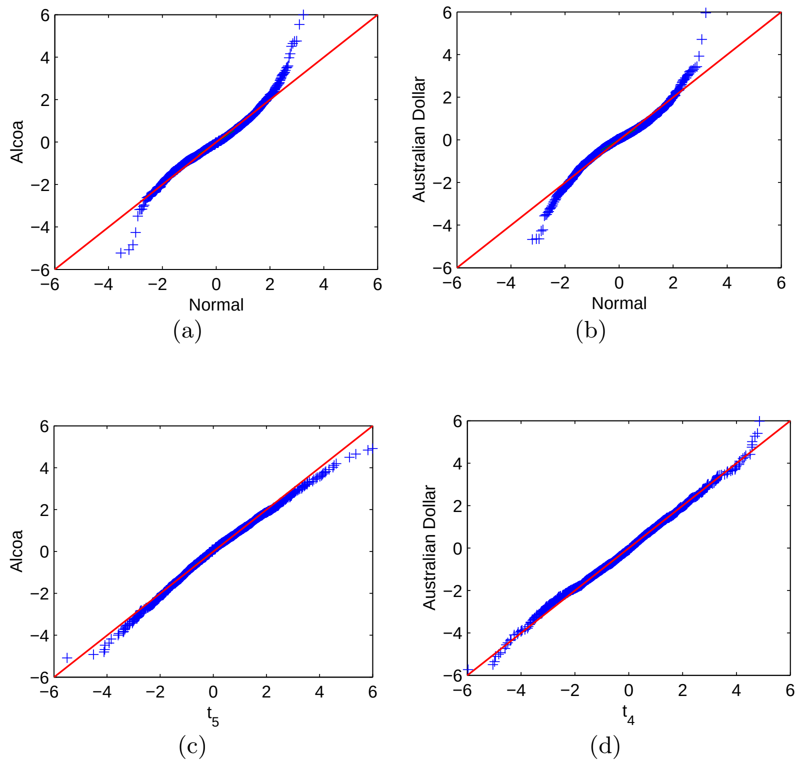


Figure 1: Q-Q plots of standardized univariate returns for Alcoa stock and the Australian Dollar.
The figure shows Quantile-to-Quantile plots of standardized returns of: Panel (a) Alcoa stock against the Normal distribution, Panel (b) Alcoa stock against a fitted t-distribution, Panel (c) the Australian Dollar against the Normal distribution, and Panel (d) the Australian Dollar against a fitted t-distribution.

is an indication of the extreme co-movements potential in the data, and the way it is captured via the t-model. Inspecting the figures, it is also clear that the t-model in Panel (d) bares more resemblance to the dependence exhibited in Panel (b). Indeed, our study shows that the evident differences between the panels are not explained by linear correlation, but by a different underlying dependence structure. (For further discussion see section 5.2.1 We emphasize that the data in the four panels were transformed to have exactly the same marginals, i.e., the differences among the panels are *only* due to the different underlying *dependence structures*. The next section explores these concepts, as well as that of extremal dependence, in more detail.

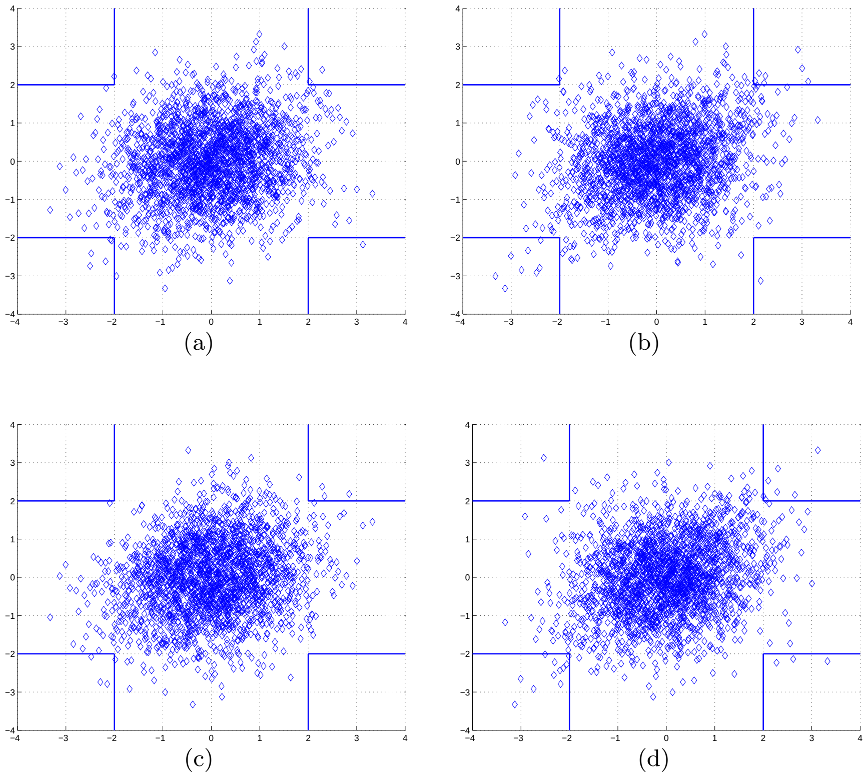


Figure 2: Scatterplots of bivariate normalized empirical and simulated returns.
The upper two panels show joint empirical returns: Panel (a) depicts the Italian Lira (ITL) against New-Zealand Dollar (NZD), Panel (b) depicts the Japanese Yen (JPY) against the New-Zealand Dollar. Panels (c) and (d) depict data simulated from a Normal-dependence structure and a t-one, respectively. The figure demonstrates how empirical observations of joint extreme events (depicted by points falling into the extreme four corners of the plot) indicate the potential for extreme co-movements, and statistical support for the appropriate underlying dependence structure; the Normal-model is seen to "match better" the ITL-NZD pair, whereas the JPY-NZD is better modelled using a t-dependence structure.

3 Beyond Correlation: Modeling Dependencies and Extreme Co-movements

This section provides a brief introduction to modeling data dependencies, in particular modeling the underlying *dependence structure* via *copula functions*. (The two terms will be used interchangeably throughout this paper depending on the context.) We provide this summary mostly to keep the paper as self-contained as possible and to facilitate the reading of subsequent sections that build on these ideas. The exposition follows in most parts the excellent recent survey paper by Embrechts *et al.* (2001a).[5]

[5] For a general introduction to copulas the reader is referred to the recent monographs by Joe (1997) and Nelsen (1999); for a more "financially" oriented exposition see, e.g., Frees and Valdez (1998), Bouyé *et al.* (2000), Frey and McNeil (2001), Embrechts *et al.* (2001a, 2001b), and references therein.

The departure point for this discussion is the inadequacy of linear correlation as a descriptor of data dependencies. Formally, the linear correlation coefficient between two random variables X, Y with finite second moments is given by $\rho := \mathbb{E}[XY]/(\mathbb{E}X^2\mathbb{E}Y^2)^{1/2}$. It is a well known fact that this measure fully characterizes statistical dependence only in the class of elliptical distributions [see, e.g., Embrechts *et al.* (2001b)], the most important example being the multivariate Normal distribution. Correlation has many flaws as a measure of dependence, some will be briefly discussed in what follows. One particular shortcoming concerns the adequacy of correlation as an indicator of potential extreme co-movements in the underlying variables. To this end, it is worth noting that correlation is essentially inadequate by construction, namely, it is a measure of central tendency involving only first and second moment information. As we shall see, *tail dependence* is a more representative measure that is used to summarize the potential for extreme co-movements. It turns out that this measure is closely related to (in fact, a property of) the underlying copula.

3.1 Modeling the dependence structure

The main concept that is used to capture the underlying *dependence structure* of a multivariate distribution is the *copula* function. The latter merely refers to the class of multivariate distribution functions supported on the unit cube with uniform marginals.

Definition 1 (Copula) A function $C : [0,1]^d \mapsto [0,1]$ is a d-dimensional copula if it satisfies the following properties :

1.) For all $u_i \in [0,1]$, $C(1,\ldots,1,u_i,1,\ldots,1) = u_i$.

2.) For all $u \in [0,1]^d$, $C(u_1,\ldots,u_d) = 0$ if at least one of the coordinates, u_i, equals zero.

3.) C is grounded and d-increasing, i.e., the C-measure of every box whose vertices lie in $[0,1]^d$ is non-negative.

The importance of the copula stems from the fact that it captures the dependence structure of a multivariate distribution. This can be seen more clearly from the following fundamental fact known as Sklar's theorem [see Sklar (1959)], adapted from Theorem 1.2 of Embrechts *et al.* (2001a).

Theorem 1 *Given a d-dimensional distribution function H with continuous marginal cumulative distributions $F_1,\ldots,F_d$, then there exists a unique d-dimensional copula $C : [0,1]^d \mapsto [0,1]$ such that*

$$H(x_1,\ldots,x_d) = C(F_1(x_1),\ldots,F_d(x_d)) \ . \tag{1}$$

From Sklar's Theorem we see that for continuous multivariate distribution functions, the univariate margins and the multivariate dependence structure can be separated, and the dependence structure can be represented by a copula. To spell out how this unique copula is related to the cumulative distribution function, we need the following definition. Let F be a univariate distribution function. The *generalized inverse* of F is defined as

$$F^{-1}(t) = \inf\{x \in \mathbb{R} : F(x) \geq t\}$$

for all $t \in [0, 1]$, using the convention $\inf\{\emptyset\} = \infty$.

Corollary 1 *Let H be a d-dimensional distribution function with continuous marginals $F_1, \ldots, F_d$ and copula C (where C satisfies (1)). Then for any $u \in [0, 1]^d$,*

$$C(u_1, \ldots, u_d) = H(F_1^{-1}(u_1), \ldots, F_d^{-1}(u_d)).$$

Without the continuity assumption, this may not hold [see, e.g., Nelsen (1999)]. Note, unlike correlation that captures the full dependence structure in multivariate Gaussian distributions (and more generally in the class of elliptical distributions), the copula summarizes this dependence structure for any multivariate distribution (with continuous marginals).

Perhaps one of the key properties of copulas, and one that elucidates the role played by copula functions in succinctly summarizing the dependence structure of a multivariate distribution, is the following invariance result adapted from Embrechts *et al.* (2001a).

Theorem 2 ***(copula invariance)*** *Consider d continuous random variables $(X_1, \ldots, X_d)$ with copula C. If $g_1, \ldots, g_d : \mathbb{R} \mapsto \mathbb{R}$ are strictly increasing on the range of $X_1, \ldots, X_d$, then $(g_1(X_1), \ldots, g_d(X_d))$ also have C as their copula.*

Note, this statement does not hold for correlation which is only invariant under linear transformations. The above also suggests how one might go about testing a copula structure per se, that is, if the marginals are known they can be "filtered out" by taking $g_i := F_i$. This observation lies at the heart of our statistical study. Armed with these results, we now proceed to introduce the two main copulas that play the central role in this study.

Definition 2 (Normal-copula) Let Φ denote the standard Normal cumulative distribution function and let Φ_Σ^d denote the multivariate Normal cumulative distribution function with zero mean, unit variance for each marginal, and linear correlation matrix $\Sigma \in \mathbb{R}^{d \times d}$, i.e., for $x \in \mathbb{R}^d$

$$\Phi_\Sigma^d(x) = \int_{-\infty}^{x} \frac{1}{(2\pi)^{d/2}|\Sigma|^{1/2}} \exp\{-\frac{1}{2} y^T \Sigma^{-1} y\} dy$$

where $|\Sigma|$ is the determinant of Σ, x^T denotes the transpose of the vector $x \in \mathbb{R}^d$ and the above integral denotes componentwise integration. Then, for $u = (u_1, \ldots, u_d) \in [0,1]^d$

$$\begin{aligned} C^G(u_1, \ldots, u_d; \Sigma) &= \Phi^d_\Sigma(\Phi^{-1}(u_1), \ldots, \Phi^{-1}(u_d)) \\ &= \int_{-\infty}^{\Phi^{-1}(u)} \frac{1}{(2\pi)^{d/2}|\Sigma|^{1/2}} \exp\{-\frac{1}{2} y^T \Sigma^{-1} y\} dy \end{aligned} \tag{2}$$

is the Gaussian or Normal-copula, parameterized by Σ. Here $\Phi^{-1}(u) := (\Phi^{-1}(u_1), \ldots, \Phi^{-1}(u_d))$, and $\Phi(\cdot)$ is the standard Normal cumulative distribution function.

This study focuses on a natural generalization of the Normal-copula, namely the Student t-copula.

Definition 3 (Student *t*-copula) Let t_ν denote the (standard) univariate Student-t cumulative distribution function with ν degrees-of-freedom, namely,

$$t_\nu(x_1) = \int_{-\infty}^{x_1} \frac{\Gamma((\nu+1)/2)}{\Gamma(\nu/2)(\nu\pi)^{1/2}} (1 + y_1^2/\nu)^{-(\nu+1)/2} dy_1.$$

Let $t^d_{\nu,\Sigma}$ denote the multivariate analogue in $\mathbb{R}^d$, with ν DoF and *shape parameter matrix* $\Sigma \in \mathbb{R}^{d\times d}$, namely, for $x \in \mathbb{R}^d$

$$t^d_{\nu,\Sigma}(x) = \int_{-\infty}^{x} \frac{\Gamma((\nu+d)/2)}{\Gamma(\nu/2)(\nu\pi)^{d/2}|\Sigma|^{1/2}} (1 + y^T\Sigma^{-1}y/\nu)^{-(\nu+d)/2} dy.$$

Then, for $u = (u_1, \ldots, u_d) \in [0,1]^d$

$$\begin{aligned} C(u_1, \ldots, u_d; \nu, \Sigma) &= t^d_{\nu,\Sigma}(t_\nu^{-1}(u_1), \ldots, t_\nu^{-1}(u_d)) \\ &= \int_{-\infty}^{t_\nu^{-1}(u)} \frac{\Gamma((\nu+d)/2)}{\Gamma(\nu/2)(\nu\pi)^{d/2}|\Sigma|^{1/2}} (1 + y^T\Sigma^{-1}y/\nu)^{-(\nu+d)/2} dy \ , \end{aligned} \tag{3}$$

is the t-copula, parameterized by (ν, Σ), and $t_\nu^{-1}(u) := (t_\nu^{-1}(u_1), \ldots, t_\nu^{-1}(u_d))$. If $\nu > 2$, the shape parameter matrix Σ can be interpreted as the linear correlation matrix.

In what follows, let $c(\cdot; \nu, \Sigma)$ denote the *density* of the t-copula given explicitly by

$$c(u; \nu, \Sigma) = \frac{\Gamma((\nu+d)/2)[\Gamma(\nu/2)]^{d-1}}{[\Gamma((\nu+1)/2)]^d|\Sigma|^{1/2}} \left[\prod_{i=1}^d (1 + y_i^2/\nu)^{(\nu+1)/2}\right] (1 + y^T\Sigma^{-1}y/\nu)^{-(\nu+d)/2} \ , \tag{4}$$

where $u = (u_1, \ldots, u_d)$, $y = (y_1, \ldots, y_d)$ and $y_i := t_\nu^{-1}(u_i)$. In the sequel, whenever a t-copula is used, we make the assumption that $\nu > 2$, so as to ensure that the variables in question have a finite second moment. The notation introduced above, taking $C^G(\cdot; \Sigma)$ for the Normal-copula and reserving the "superscript-less" $C(\cdot;\nu,\Sigma)$ for the t-copula, will be used throughout to distinguish between the two copulas.[6]

Recall that the multivariate t-distribution is a generalization of the multivariate Normal in the sense that the Normal distribution can be considered as a t-distribution with infinite degrees of freedom. The same is true for the respective copulas, viz, the Normal-copula can be viewed as a member of the parametric family of t-copula, where the degrees of freedom (DoF) are infinite. Fixing Σ we have that

$$\sup_{u\in[0,1]^d} |C(u;\nu,\Sigma) - C^G(u;\Sigma)| \to 0 \qquad \text{as } \nu \to \infty \tag{5}$$

where C^G is the Normal-copula given in (2). The above asymptotic can be refined to obtain rates of convergence, but, for our purposes, it suffices to note that the two copulas are very close for DoF that are greater that 100, and essentially indistinguishable for DoF that are greater that 1000 (for an illustration see Appendix B.3).

3.2 Extreme co-movements and tail dependence

The concept of *tail dependence* relates to the amount of dependence in the upper-right-quadrant tail or lower-left-quadrant tail of a bivariate distribution. It turns out that tail dependence between two continuous random variables X and Y is a copula property and hence the amount of tail dependence is invariant under strictly increasing transformations of X and Y.

Definition 4 (tail dependence) Let X and Y be random variables with continuous distributions F and G respectively. The (upper) tail dependence coefficient of X and Y is given by

$$\lambda_U := \lim_{u\uparrow 1} \mathbb{P}(Y > G^{-1}(u) | X > F^{-1}(u))$$

provided that the limit $\lambda_U \in [0,1]$ exists.

If $\lambda_U \in (0,1]$, X and Y are said to be *asymptotically dependent in the upper tail*; if $\lambda_U = 0$, X and Y are said to be *asymptotically independent in the upper tail.* Simple manipulations of the above definition lead to an alternative and equivalent definition (for continuous random variables) [cf. Joe (1997), p. 33],

$$\lambda_U = \lim_{u\uparrow 1} \frac{1 - 2u + C(u,u)}{1-u}$$

[6]Following standard practice, we use the above *standardized* formulation of copulas, [for further discussion on the canonical representation of copulas, see, e.g., Embrechts *et al.* (2001a)]. This amounts to taking the normalized covariance (respectively, shape parameter) matrix, and taking the means to be zero. Clearly any meaningful notion of dependence structure should not depend on shifting and scaling.

if the above limit exists. Thus, upper tail dependence is a copula property. The concept of lower tail dependence can be defined in a similar way; we omit the details and refer the interested reader to Joe (1997) or Embrechts *et al.* (2001a).

It is important to note that while independence of X and Y implies that $\lambda_U = 0$, the converse is not true in general. That is, $\lambda_U = 0$ does not necessarily imply that X and Y are statistically independent. Intuitively, when the tail independence $\lambda_U = 0$ we might expect that X and Y become independent, in the sense that the joint distribution factors out to the product of the marginals for sufficiently large values. However, the following simple, yet important, observation asserts that this intuition is false.

Proposition 1 ***(Tail dependence for Normal-copula)*** *Let $C^G(\cdot;\Sigma)$ be the Normal-copula over $[0,1]^2$, with correlation $\Sigma_{12} = \rho$. Then,*

$$\lambda_U = 2 \lim_{x\to\infty} (1 - \Phi(x\sqrt{1-\rho}/\sqrt{1+\rho})) \ .$$

For a proof see Embrechts *et al.* (2001a, §4.3). The above expression for upper tail dependence in the Normal-copula asserts that $\lambda_U = 0$ (zero tail dependence) for values of $\rho \in [0,1)$, and only in the trivial case where the two random variables are related linearly do we get non zero tail dependence ($\lambda_U = 1$).

This observation is important for the following reason. The concept of tail dependence reflects the tendency of X and Y to "move together," in particular it gives an asymptotic indication of how frequently we expect to observe joint extreme values. It is therefore important to consider copulas that posses the property of non-trivial tail dependence. To this end, for the t-copula we have,

Proposition 2 ***(Tail dependence for t-copula)*** *Let C be the $t_{\nu,\Sigma}$ copula over $[0,1]^2$, with correlation $\Sigma_{12} = \rho$. Then,*

$$\lambda_U = 2\left(1 - t_{\nu+1}\left(\frac{\sqrt{\nu+1}\sqrt{1-\rho}}{\sqrt{1+\rho}}\right)\right) \ . \tag{6}$$

For a proof see Embrechts *et al.* (2001a, §4.3). The above proposition indicates that for random variables that are linked via a t-copula, we can expect joint extreme movements to occur with non-negligible probability even when the random variables are "weakly dependent" (in the sense that their correlation is small). For example, with DoF = 4 and correlation $\rho = 0.5$ we have that $\lambda_U = 0.25$, and moreover, if we take $\rho = 0$, we still get $\lambda_U = 0.08$! This also illustrates a fundamental difference between the t-dependence structure and the Gaussian structure. In the latter $\rho = 0$ implies independence, in particular zero tail dependence, while for the t-dependence structure it is essentially the DoF parameter that controls the extent of tail dependence and tendency to exhibit extreme co-movements.

4 The Testing Methodology

The main idea that underlies the testing procedure may be described, informally, as follows. We set the null hypothesis to correspond to some fixed value of the DoF parameter (ν_0) in the t-dependence structure, while this is a free parameter in the alternative hypothesis. As we vary the null parameter ν_0 we can ascertain the range of values of the DoF which cannot be rejected, based on the corresponding p-values for the test. This analysis is essentially equivalent in spirit to producing confidence intervals for point estimates of ν at varying levels of confidence; we believe that the "hypothesis testing formulation" is much more transparent in this context. In particular, since the Gaussian dependence structure is nested within the t-family (DoF $= \infty$), we can use p-values obtained for arbitrary large values of ν_0 as a proxy that indicates whether a Gaussian dependence structure is likely to be supported on the basis of the observed empirical asset co-movements.

Consider a random sample $\mathcal{X}_n = \{X_i\}_{i=1}^n$, where the $X_i = (X_{i1}, \ldots, X_{id})$ are assumed to be mutually independent and distributed according to a common distribution function H with continuous univariate marginals, $F_1(\cdot), \ldots, F_d(\cdot)$. Now, if the marginal distributions were *known*, then, by Corollary 1, $U := F(X_i) \sim C(\cdot;\theta)$, where C is the copula function for H, and $F(X_i) := (F_1(X_{i1}), \ldots, F_d(X_{id}))$. Since the marginals are unknown and potentially do not even follow a parametric distribution, we treat them as infinite dimensional nuisance parameters (parameterizing the multivariate distribution H via Sklar's representation) leaving the parameters of the copula, in particular the DoF, as the main focus. To this end, the empirical distribution function is used as a surrogate for the unknown marginals, that is, $\hat{F}_j(\cdot) = \frac{1}{n}\sum_{i=1}^n \mathbb{I}\{X_{ij} \leq \cdot\}$ for $j = 1, \ldots, d$, where $\mathbb{I}\{\cdot\}$ is the indicator function. We then define $\hat{U}_i = (\hat{F}_1(X_{i1}), \ldots, \hat{F}_d(X_{id}))$ and let $\mathcal{U}_n = \{(\hat{U}_i\}_{i=1}^n$ denote the *pseudo-sample*. The transformation that takes the original sample $\mathcal{X}_n$ to the pseudo-sample $\mathcal{U}_n$ will be referred to as the *empirical marginal transformation*.[7]

Remark 1 (Properties of the empirical marginal transformation and the pseudo-sample) When the sample size approaches infinity, the Glivenko-Cantelli lemma [cf. §2.1.4 of Serfling (1980)] asserts that $\hat{F}$ converges to F uniformly on the real-line, almost surely. In fact more is true, as it is well known that $\hat{F}$ is, roughly speaking, asymptotically Normal and "centered" around the true distribution F [the precise meaning of this statement involves the functional central limit theorem; for details the reader is referred to §1.11 of Serfling (1980)]. Thus, $\hat{F}$ is a $\sqrt{n}$-consistent estimator for F, which is crucial in deriving asymptotics for the corresponding hypothesis tests. Note that the pseudo-sample is no longer comprised of mutually independent observations due to the data dependence inherent in the empirical marginal transformation. (Since the marginal distributions

[7] This approach follows the semi-parametric estimation framework pursued by Genest *et al.* (1995), and has also been suggested by Bouyé *et al.* (2000) and Durrelman *et al.* (2001) in the context of calibrating copulas to observed financial data. In the econometric literature an estimation procedure that uses this intermediate step is often referred to as "two step estimation," see , e.g., Amemiya (1984).

are typically unknown, this problem is intrinsic to any inference problem involving dependence structure parameters.)

Focusing on the t-dependence structure given formally by the t-copula (3), we fix a value of the DoF parameter ν_0 and consider the following hypotheses

$$H_0: \ \ \theta \in \Theta_0 \qquad \text{vs.} \qquad H_1: \ \ \theta \in \Theta \ \ ,$$

where $\Theta = \{(\nu, \Sigma) : \nu \in (2, \infty], \ \Sigma \in \mathbb{R}^{d\times d}$ is symmetric and positive definite$\}$, and $\Theta_0 = \{\theta \in \Theta : \nu = \nu_0\} \subseteq \Theta$. (Note that we allow for real values of the DoF parameter.) The two parameter sets index the associated dependence structure. For a given sample $\{X_i\}_{i=1}^n$, set the *pseudo* log-likelihood function to be

$$L_n(\theta) = \sum_{i=1}^{n} \log c(\hat{U}_i; \theta) \ \ ,$$

where $c(\cdot; \theta)$ is the t-copula density function (4). Note that this is obtained from a likelihood function that imposes a product form $\prod_{i=1}^n c(\hat{U}_i; \theta)$. This should be viewed as a structural assumption that is imposed on the objective function; the reader will recall that the variables $\{\hat{U}_i\}$ *are not* mutually independent. Put

$$\hat{\theta} := (\hat{\nu}, \hat{\Sigma}) \in \arg\max\{L_n(\theta) : \theta \in \Theta\} \ ,$$

the *pseudo maximum likelihood* (p-ML) estimator, and set the *pseudo-likelihood ratio test statistic* (p-LRT statistic) to be

$$\Lambda_n(\hat{\nu}|\nu_0) := -2 \log \frac{\sup_{\theta\in\Theta_0} \prod_{i=1}^n c(\hat{U}_i; \theta)}{\prod_{i=1}^n c(\hat{U}_i; \hat{\theta})} \ \ .$$

The precise details concerning the implementation of the ML estimation procedure, as well as computational aspects, are discussed in Appendix C.

To determine the adequacy of each value of ν_0, we need to characterize the distribution of the test statistic $\Lambda_n(\hat{\nu}|\nu_0)$. Since this distribution is not tractable, the standard approach is to derive the asymptotic distribution and use that as an approximation. The derivation of the asymptotic distribution of p-LRT test statistic is pursued on the basis of the p-ML theory developed in Genest *et al.* (1995). Specifically, the derivations in Appendix A suggest that under the null hypothesis,

$$\Lambda_n(\hat{\nu}|\nu_0) \Rightarrow (1+\gamma)\chi_1^2, \quad \text{as } n \to \infty \tag{7}$$

where "$\Rightarrow$" denotes convergence in distribution, and $\gamma > 0$ is a constant that depends on the null hypothesis [an explicit characterization of γ is given in (10) in Appendix A]. Here, and in the sequel, χ_1^2 denotes a r.v. distributed according to a Chi-squared law with one degree-of-freedom. Thus, the p-LRT test retains the familiar limiting distribution, which is now scaled by a factor of

$(1+\gamma)$. Using the approximation $\Lambda_n(\hat{\nu}|\nu_0) \approx (1+\gamma)\chi^2_1$, we may now calculate approximate p-values as a function of ν_0, given by

$$p\text{-value}(\nu_0) = \mathbb{P}\left(\chi^2_1 \geq \frac{\Lambda_n(\hat{\nu}|\nu_0)}{(1+\gamma)}\right) . \tag{8}$$

To summarize, for each value of ν_0 the corresponding p-value is determined and this, in turn, yields the range of values of the DoF parameter that are supported (respectively rejected) on the basis of the observed sample.

To rigorously justify the above asymptotic approximation, one needs to prove the weak limit given in (7). Appendix A provides a sketch of how this result is derived based on the asymptotic theory of p-ML estimation in the context of copulas. However, the technical regularity that support the p-ML asymptotic theory [see Appendix 1 in Genest *et al.* (1995)] are quite difficult to verify, in particular, in the context of the t-dependence structure and when considering simultaneous estimation of the DoF and the correlation matrix. Consequently, the derivation of the asymptotic (7) is done informally and serves to suggest the "right" large sample behavior of the p-LRT test statistic. The validity of the asymptotic theory described in Appendix A is established via a comprehensive numerical study that is discussed in Appendix B; the simulation study in Appendix B.1 serves to validate the estimation procedure, while the Monte Carlo study conducted in Appendix B.2 is used to validate the form of the limiting distribution. This study indicates that the t-dependence structure is amenable to the p-LRT asymptotic, in particular, the Chi-squared distribution is quite accurate in describing the behavior of the p-LRT test statistic, across various sample sizes and copula parameters (see, for example, the graphs in Figure 5 and the results depicted in Table 8).

Remark 2 (Testing the Gaussian dependence assumption) To use the above p-LRT test to determine whether the Gaussian dependence structure is supported on the basis of observed asset co-movements, we note that if $\nu_0 = \infty$, then the null hypothesis corresponds to a Gaussian dependence structure. Unfortunately, we cannot simply use the value $\nu_0 = \infty$ since then the restricted hypothesis $\Theta_0 \subseteq \Theta$ is derived from Θ by fixing a parameter value on the *boundary* of the feasible set.[8] To resolve this problem, we take a "large" value of the null *DoF* as a proxy for the infinite case. It turns out that much like the case of univariate distributions, the t-copula is essentially equivalent to a Normal-copula for moderate values of the DoF parameter. (For further details see Appendix B.3.) In particular, by setting the DoF> 1000 we have that the t-copula is indistinguishable from a Normal-copula. (To be conservative we set this value to be $\nu_0 = 10^5$.) Moreover, note that the tail dependence index (6) is essentially zero for such a large magnitude of

[8]In most settings the large sample results for LRT tests do not hold unless the restricted set Θ_0 is obtained from Θ by fixing a parameter value in the interior of the feasible set. For more details on the basic assumptions and derivations involved see, e.g., Serfling (1980) §4.4; for some recent results on testing when the restricted parameter is on the boundary of the parameter space see Andrews (2001).

the DoF parameter, irrespective of the magnitude of the correlation coefficient. This is consistent with the fact that the Gaussian dependence structure has zero tail dependence index. A similar argument regarding the extent of tail dependence underlies the test proposed recently by Longin and Solnik (2001).

5 Extreme Co-movements: Empirical Evidence

This section presents the empirical findings restricting attention to the three asset classes; the study of the G5 national equity markets is deferred to the following section. The p-values that we report represent the probability of making a type-I error, that is, rejecting the null hypothesis (the null having some fixed DoF) when it is true. For example, a p-value of 1% for a restricted version with DoF=50 reflects a 1% probability of rejecting a *true* null hypothesis of DoF=50 based on the given sample.[9] Testing the Gaussian dependence structure is always taken to mean a restricted null with $\nu_0 = 10^5$ as discussed in the previous section. (This restriction of the null is essentially indistinguishable from a Gaussian model and serves as a proxy for the latter.) In what follows we focus primarily on the estimation of the DoF, and for brevity omit estimation results pertaining to the correlation matrices that parameterize the dependence structure.

A short summary of the results is as follows.

- **Section 5.1**. The range of possible DoF consistent with the data is seen to be very narrow (ranges from 7 to 12). Consequently, tail dependence and extreme co-movements are seen to be statistically significant, and the Gaussian assumption is not supported.
- **Section 5.2**. Larger baskets of assets are seen to be characterized with more significant potential for exhibiting extreme co-movements. An increase in the number of assets results in increased statistical evidence rejecting the Gaussian dependence structure.

5.1 Determining the extent of tail dependence

This section presents results in the form of rejection and non-rejection ranges of DoF for every basket of assets. The main implications of these results is four-fold. First, the results show that the interval of non-rejected DoF is very narrow. Second, the estimation procedure is very "sharp," in the sense that little differences in the assumed DoF lead to significant changes in the likelihood and corresponding test statistic. Third, in the region of interest, i.e., $\nu_0 > \hat{\nu}$, the behavior of the

[9] Note that the calculations of the p-values throughout require the value of γ [recall, the p-LRT test statistic is distributed $(+\gamma)\chi^2_1$]; this value is estimated by Monte-Carlo methods as explained in Appendix B.2, and is typically found to be in the range [0, 0.1]. Throughout, we use $\gamma = 1.0$ as a conservative bound.

test statistic is *monotone* non-decreasing in ν_0, indicating that as the null DoF get larger there is increasing statistical evidence rejecting the null, which in turn approaches the Normal-copula. Note that at $\nu_0 = 10^5$ (used to approximate the Gaussian null), the value of the test statistic is more than two orders of magnitude above the nominal rejection level! Finally, the economic significance of these results is that extreme co-movements are statistically significant, and the extent of joint extremal events is substantial (due to relatively small values of the DoF). These consequences are discussed in Section 6.

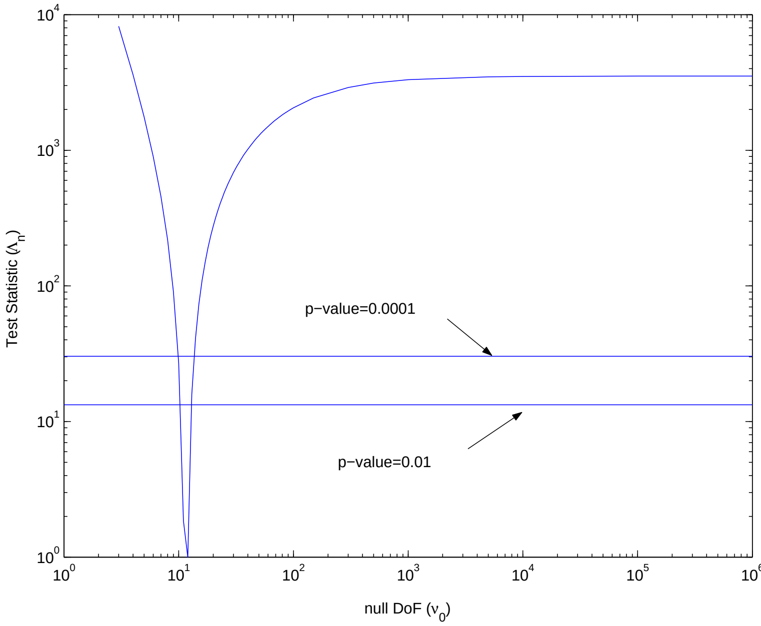
Figure 3 depicts graphically the sensitivity analysis for the three cases, it shows, for each basket, the behavior of the test statistic $\Lambda_n(\hat{\nu}\,|\nu_0)$ as a function of the null DoF (ν_0). The figure also displays the levels that correspond to p-values of 1% and 0.01%. Table 1 gives detailed results describing the "sharpness" of the test as indicated by p-values and likelihood values evaluated near the ML estimates for the three baskets. (For brevity, we refer to the estimator and objective as "ML" and "likelihood" except where confusion may arise.) The following is a short summary of the main findings.

- Equity basket [Figure 3(a) and Table 1(a)]. The ML estimator is 12; only DoF in the range 11-12 cannot be rejected at the 99% confidence level.
- Currency basket [Figure 3(b) and Table 1(b)]. The ML estimator is 7; only DoF in the range 7-8 cannot be rejected at the 99% confidence level.
- Metal basket [Figure 3(c) and Table 1(c)]. The ML estimator is 11; only DoF in the range 9-13 cannot be rejected at the 99% confidence level.

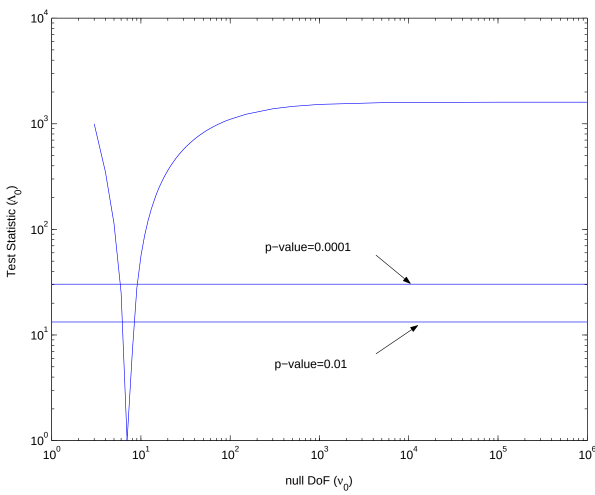
These results indicate that the Gaussian dependence structure ought to be rejected on the basis of observed co-movements. Indeed, when testing the proxy for the Normal-copula (using $\nu_0 = 10^5$) we obtain p-values that are equal to zero in all three financial data sets. (Due to computational limits, p-values that are smaller than 10^{-16} are rounded to zeros.)

5.2 Dimensionality effects

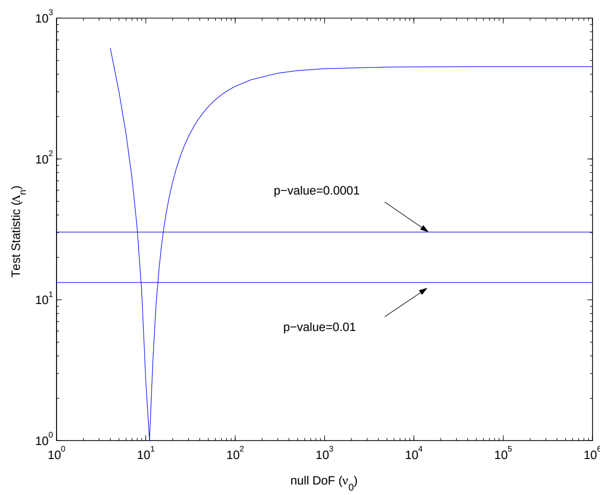
This section examines how the dimensionality of the basket (the number of underlying assets) affects the significance level of the estimated dependence structure, and the corresponding potential for extreme co-movements in the underlyings. Our main findings indicate that as the dimension increases, evidence supporting extreme co-movements is more significant. The basic intuition for this finding is clear: as the number of assets in the basket increases, there is higher likelihood to encounter large coupled movements among subsets of the assets in the basket. In particular,



(a)



(b)



(c)

Figure 3: Value of the test statistics for the three baskets as a function of the DoF underlying the null hypothesis (ν_0). Panels (a), (b), and (c) depicts the equity, currency, and metal baskets, respectively.
This figure depicts the behavior of the test statistic, $\Lambda_n\left(\hat{\nu}\,|\nu_0\right)$, for a range of DoF underlying the null hypothesis and approaching the Normal-copula. The two horizontal lines correspond to p-values of 0.01% and 1%. Values of the test-statistic below each p-value line represent the acceptable DoF with the respective confidence level. Only narrow range of DoF around the ML estimated DoF are statistically significant.

null DoF (ν_0)	**log likelihood function**	**test statistic: $\Lambda_n(12\|\nu_0)$**	**p-value**
10	11392.00	26.01	$3.11 \cdot 10^{-4}$
11	11404.57	0.850	0.514
12 (ML)	11405.00	0.00	1
13	11397.56	14.87	$6.38 \cdot 10^{-3}$
14	11384.95	40.10	$7.54 \cdot 10^{-6}$

(a)

null DoF (ν_0)	**log likelihood function**	**test statistic: $\Lambda_n(7\|\nu_0)$**	**p-value**
6	8923.02	24.19	$5.05 \cdot 10^{-4}$
7 (ML)	8935.12	0.00	1
8	8932.18	5.88	$8.64 \cdot 10^{-2}$
9	8907.76	54.72	$2.65 \cdot 10^{-4}$
10	8891.95	86.33	$1.68 \cdot 10^{-7}$

(b)

null DoF (ν_0)	**log likelihood function**	**test statistic: $\Lambda_n(11\|\nu_0)$**	**p-value**
8	2464.38	30.95	$8.36 \cdot 10^{-5}$
9	2474.67	10.35	$2.28 \cdot 10^{-2}$
10	2479.04	1.62	0.37
11 (ML)	2479.85	0.00	1
12	2478.48	2.75	0.24
13	2475.75	8.21	$4.3 \cdot 10^{-2}$
14	2472.19	15.33	$5.63 \cdot 10^{-3}$
15	2468.13	23.44	$6.18 \cdot 10^{-4}$
16	2463.79	32.12	$6.14 \cdot 10^{-5}$
17	2459.32	41.07	$5.86 \cdot 10^{-6}$
18	2454.81	50.09	$5.60 \cdot 10^{-7}$

(c)

Table 1: Sensitivity analysis for the three baskets.
The table gives the estimation results for the three complete baskets. Panel (a) presents the equity results, Panel (b) the currency results, and Panel (c) the metal results. Each panel depicts the log-likelihood function value, the corresponding test statistic, $\Lambda_n\left(\hat{\nu}\,|\nu_0\right)$, and p-value for different DoF, given that $\hat{\nu} = 12$, 7, and 11 (the ML estimates) for the three panels respectively. The table depicts only p-values that are greater than 10^{-7}. In each case only a narrow range of DoF is statistically significant.

these findings imply that the Gaussian dependence structure is less likely to provide an adequate description of the underlying dependencies.

5.2.1 Results for two and three-name baskets

We start by testing all possible sub-baskets of size two and then size three. For example, in the equity basket there are 435 pairs and 4060 three-asset combinations. Setting a rejection level of 1%, and testing the Normal-copula proxy results in the following findings.

- Equities: the Normal-copula proxy is rejected for 428 of the 435 pairs, and all 4060 triples. The median DoF for the non-rejected pairs is 7 in both cases.
- Currencies (FX): the Normal-copula proxy is rejected in 22 pairs of the possible 36 pairs, and all 84 triples. The median DoF for the non-rejected pairs is 8 in both cases.
- Metals: the Normal-copula proxy is rejected for 13 out of the 15 possible pairs, and for all 20 triples. The median DoF for the non-rejected pairs is 8 and 9, respectively.

Discussion. For the FX sub-baskets there are 14 pairs of currencies for which there is insufficient evidence to reject the null hypothesis. When setting a less conservative value for the asymptotic constant $\gamma = 0.1$ (consistent with the simulation-based estimates derived in Appendix B), there are still nine currency pairs for which statistical evidence of extreme co-movements is inconclusive. It is interesting to observe that all nine of these pairs include either the Australian Dollar or the New-Zealand Dollar. Specifically the Australian Dollar matched with the French Franc, German Mark, Italian Lira, and United Kingdom Pound. The New-Zealand Dollar is matched with the French Franc, German Mark, Italian Lira, Swiss Franc, and United Kingdom Pound. According to Section 2, one main distinguishing feature separating the t-dependence structure from the Gaussian is that the former supports joint extreme movements while the latter does not. Thus, the above results signal that the two currencies, namely, Australian Dollar and New-Zealand Dollar do not tend to have large co-movements with the listed currencies. Consequently, these currencies may serve for purposes of jump diversification.

This point is further illustrated in Figure 2 in Section 2. The figure presents the normalized empirical bivariate distribution of the Italian Lira vs. the New-Zealand Dollar [Panel (a)] and the Japanese Yen vs. the New-Zealand Dollar [Panel (b)]; the former is one of the pairs for which the proxy of the Normal-copula was not rejected and the latter is an example of a pair for which it was rejected. The scatter-plot in Panel (a) contains far fewer points in the far corners of the graph as opposed to the scatter-plot in Panel (b). Recall that the t-copula is distinguished from the

Normal-copula by placing more probability mass on joint extreme events, making it a more plausible candidate to model the data dependency observed in the right hand plot. To further demonstrate the point, Panel (c) displays normalized data that was simulated using a Normal-copula, whereas Panel (d) displays data simulated with a t-copula. It is easy to see that the scatterplot in Panel (b), which is the common empirical dependence, resembles much better the t-copula simulated data in Panel (d), rather than the Normal-copula depicted in Panel (c). Note that the correlation of the four pairs is similar (16%, 22%, 18%, and 23% for Panels (a), (b), (c), and (d), respectively).

5.2.2 Effects of dimensionality: summary

The previous sections provide preliminary evidence that a higher dimensionality of the basket effectively reduces the ability of the Normal-copula to "explain" observed dependencies in the data due to the increasing likelihood of extreme co-movements. To substantiate this observation we proceed to test all different baskets (of sizes 2, 3,..., 9) for the FX and Metal baskets (of sizes 2, 3,..., 6). The test statistics $\Lambda_n\left(\hat{\nu}\,|\nu_0\right)$ and the p-values for rejecting the proxy for the Gaussian dependence hypothesis were computed for each case. In the following we report the *minimum* test statistic and the *maximum* p-value over all various combinations of a particular size of basket. The decay of the p-values with the increase of the size of the basket is evident. Table 2 summarizes the results for the currencies basket and Figure 4 displays graphically the behavior of the minimal test statistic as the dimension of the basket increases. We note that only for the two-asset baskets the minimal test statistic cannot be rejected. Similar results were obtained for the metal baskets.

number of currencies in each basket	**number of baskets tested**	**maximum p-value**	**minimum value of the test statistic**
2	36	0.378	1.56
3	84	$1.19 \cdot 10^{-3}$	20.10
4	126	$4.04 \cdot 10^{-9}$	69.21
5	126	0	148.02
6	84	0	229.56
7	36	0	388.27
8	9	0	710.86
9	1	0	1597.63

Table 2: FX basket dimensionality study.
This table depicts the largest p-value and the smallest value of the test statistic, calculated over all possible sub-baskets (sizes=2, 3,..., 9) of the nine currencies. The p-values are calculated using $(1+\gamma)=2$, and the null DoF is $\nu_0 = 10^5$.

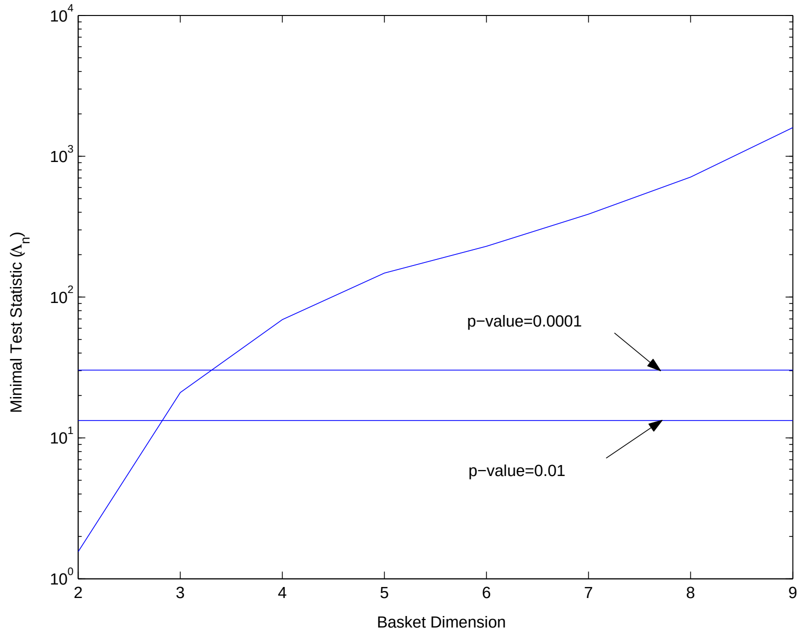


Figure 4: FX dimensionality study.
The figure shows the minimal values of the test statistic (calculated over all basket compositions), as a function of the number of currencies in the basket. The two horizontal lines correspond to p-values of 1% and 0.01%. Values of the test statistic below each line represent the non-rejection of the proxy for the Gaussian dependence hypothesis, with the respective p-value. (The null DoF is $\nu_0 = 10^5$.)

6 Economic Significance and Financial Implications

The estimation results in the previous section indicate that an accurate description of asset dependencies ought to be based on models that generalize the notions of correlation. In particular, the common Normality-based dependence assumptions must be abandoned in order to capture extreme co-movements that are inherent in financial data. In this section we provide some additional insights that attempt to elucidate these phenomena and illustrate their consequences and economic significance. The first example focuses on the behavior of international equity markets, specifically, cross-market coupled movements. The subsequent two examples illustrate some of the merits of the t-dependence model in the context of risk management and pricing. The examples serve two purposes: first, they illustrate the consequences of mis-specification intrinsic in the Normal (correlation based) model; second, they demonstrate the relative ease of implementing the t-based model in these contexts, and the potential gains in reducing the aforementioned mis-specification effects.

6.1 Coupling of International markets

Motivation. International diversification has always been considered to be a simple mechanism for reducing portfolio risk. However, recent changes in technology, information transmission, and legislation have resulted in increased dependence across markets [see, e.g., Longin and Solnik (1995) for evidence of the increase in international correlation]. Moreover, as economic borders become more and more transparent, portfolios are becoming more globalized and money shifts faster from one market to another. The testing methodology that this paper pursues differs in a fundamental way from the recent studies that focus on extreme value theory, conditional correlations, and regression-based methods.[10] In particular, focusing directly on the underlying dependence structure supports arbitrary marginal distributions, introduces a more refined notion of extremal dependence that blends in correlation with the extreme value index (DoF), and most importantly, provides an *unconditional* model of dependencies among markets, that is also easy to implement.

Numerical example. For estimation of international dependencies we use the Morgan Stanley Capital International (MSCI) country equity data. MSCI indices represent a broad aggregation of national equity markets and are the leading benchmarks for global portfolio managers. We analyze weekly data in local currency and US Dollar denominations for the G5 markets: France, Germany, Japan, United Kingdom, and United States. The data span the dates 1/10/90 until 5/29/02, a total of 646 return observations. Table 3(a) presents the pair-wise estimated DoF and correlation for the local currency data. All the pairs exhibit low estimated DoF indicating relatively high tail dependence. We stress that the estimation procedure is a semi-parametric one and therefore the "fat tails" of the univariate returns and their "peakedness" do not affect the estimated DoF of the dependence structure. Table 3(b) provides the estimated DoF and correlation for the MSCI *US Dollar* denominated data.

Comparing the results in Tables 3(a) and 3(b) it is evident that the dependencies between all countries except Japan do not change considerably when using local currency or US Dollar denominated data. However, it can be seen that the estimated DoF for pairs that include Japan exhibit significant changes. In order to illustrate the meaning of the different DoF we compute the frequency of large co-movements. For example, the frequency that both the Japanese index and the UK index simultaneously suffer an extreme down realization of 1.6 standard deviations is 1.86% in the local currency data, whereas it is 0.77% in the US Dollar denominated data. Note that the linear correlation estimates are similar for all pairs, in the local as well as the US Dollar denominated data. The implications of these results are illustrated in Table 4, which gives the empirical and simulated (model-based) probability of the UK-Japan indices having joint realizations below given thresholds that are measured in terms of standard deviations. We note that joint realizations below 2.4 standard deviations are not observed in the data. The simulated probabilities are based on the estimated models, viz, a t-dependence structure with 6 and 54 DoF, that correspond to the local currency data and US Dollar denominated data, respectively. The t-based dependence structure

[10] For some examples of the use of these methodologies, see, e.g., Longin (1996), Karolyi and Stulz (1996), DanÌelsson and de Vries (1997), Starica (1999), Straetmans (2000), Longin and Solnik (2001), Forbes and Rigobon (2001), Ang and Bekaert (2002), Ang and Chen (2002), Ang *et al.* (2002), and Bae *et al.* (2002), as well as references therein.

	Germany		Japan		UK		US	
France	6	[0.76]	7	[0.36]	5	[0.70]	8	[0.62]
Germany			7	[0.34]	5	[0.66]	7	[0.61]
Japan					6	[0.36]	10	[0.41]
UK							9	[0.64]

(a)

	Germany		Japan		UK		US	
France	5	[0.75]	15	[0.37]	5	[0.67]	11	[0.55]
Germany			10	[0.35]	4	[0.61]	6	[0.54]
Japan					54	[0.40]	28	[0.34]
UK							8	[0.56]

(b)

Table 3: Estimated DoF and [correlation] for MSCI Equity market pairs.
The table gives the estimated DoF and correlation values that were fitted based on the observed co-movements for each pair of indices, using a bivariate t-copula. Panel (a) presents estimation results for the local currency data whereas Panel (b) presents the estimation results for the US Dollar denominated data. The estimated DoF indicate that extreme co-movements are statistically significant in almost all cases (the exceptional cases are the UK-Japan and US-Japan pairs).

allows us to extrapolate extreme joint movements outside of those we observe in the sample.

	threshold in standard deviations	bivariate t_6 [Local Currency]		bivariate t_{54} [US Dollars]		bivariate Normal
in sample	**-1.6**	1.20%	[1.86%]	1.06%	[0.77%]	1.01%
	-1.8	0.79%	[0.93%]	0.58%	[0.15%]	0.54%
	-2.0	0.53%	[0.46%]	0.31%	[0.15%]	0.28%
	-2.2	0.36%	[0.31%]	0.16%	[0.15%]	0.13%
out of sample	**-2.4**	0.25%		0.08%		0.06%
	-2.6	0.17%		0.04%		0.03%
	-2.8	0.12%		0.02%		0.01%
	-3.0	0.09%		0.01%		0.01%

Table 4: Probability of large negative co-movements between the UK and Japanese indices.
The table depicts simulated and [empirical] probabilities for joint movements (measured in standard deviations) between the UK and Japanese indices. The probabilities correspond to joint realizations that fall, in each market, below the given threshold. Model-based calculations are done via Monte-Carlo simulations using the estimated parameters for the t-dependence structure and are given as a reference. The similarity in the joint tails of the empirical data and of the Normal and t-simulated data is evident, in particular, the t-model captures very well the joint extreme co-movements in the data. Both parametric dependence models allow out-of-sample prediction of large joint movements (below -2.4) which are not observed in the UK-Japan data set. Note that the t-model places up to 9 times more probability on joint extremes than the Normal model does, thus affecting tail risk measures such as VaR, or joint default rates.

Discussion. These initial results are interesting in several regards. First, the empirical observations lead to a relatively low estimated value of the DoF parameter in the t-dependence model. This implies high likelihood for extreme joint movements as manifested in the UK-JPY pair. (Note that a t-based model allows the extrapolation, out-of-sample, of joint extreme probabilities.)

The high likelihood of extreme co-movements is of paramount importance to investors who wish to diversify portfolios globally; the well documented "breakdown of correlation" during volatile markets (e.g., Longin and Solnik (1995), Solnik *et al.* (1996), and Rey (2000)] can be estimated and assessed using the t-model. For example, Table 4 shows that the t-model places as much as nine times more probability on extreme co-movements than the Normal model does. This coupling of international markets also suggests a partial explanation for the "investor home bias puzzle" [see, e.g., the literature review in Strong and Xu (2002)]. If global diversification is prone to breakdown at the time it is most needed, i.e., if the strong coupling between markets implies that they tend to experience concurrent extreme shocks, then cross market diversification may be of limited use as a means of reducing risk exposure. In contrast, the Gaussian-based models seem to under-emphasize this coupling and thus indicate incorrectly that there is more to be gained by cross-border holdings. A second observation concerns the relation between correlation numbers and the values of DoF. In particular, the former exhibits only minor variations, while the latter more significant. This is another indication that correlation as a stand-alone is a limited measure of dependence. Note that in the Japanese – UK example the correlation actually goes up from 0.36 in the local currency to 0.40 in the US Dollar denominated data, whereas the frequency of extreme joint movements actually *decreases*. Normal-based models would therefore place more probability on extreme co-movements in the US Dollar pair rather than in the local currency one. Finally, the dramatic changes in the Japanese pairs' DoF between local and US Dollar denominated data can be exploited in a model that would incorporate the two equity markets and the USD/JPY exchange rate. A possible explanation for this disparity could be that multinational market shocks are compensated by shocks in the Yen value, whereas the exchange rates among other pairs are not affected that much by market shocks.

In summary, modeling international co-movements using a dependence structure that supports extreme co-movements can affect almost all investment decisions in a multinational portfolio, and, as the next example illustrates, most measures of the associated risks.

6.2 Risk management

Motivation. Almost all risk management models; regulatory, market, and credit, build on the Gaussian dependence structure assumption to model asset returns or the associated risk factors [for example: the Basel accord II proposal (2001) and DanÌelsson *et al.* (2001); RiskMetrics, see Mina and Xiao (2001); CreditMetrics, see Gupton *et al.* (1997); and KMV, see Kealhofer (1998)]. The results in Section 5 cast doubt on the adequacy of the Gaussian dependence structure and its use as a model for the behavior of joint asset returns. To this end, a more realistic model, such as the t-dependence structure, would enable more probability mass to be "placed" on joint extreme co-movements. The latter may manifest themselves in the form of events where many instruments in the portfolio realize substantial losses, as had happened in several market crashes, e.g., the 1987 crash, or the collapse of Long Term Capital Management (LTCM). Jorion (2000) examines the collapse of LTCM, and an adaption of his findings shows that the event of collapse under Normality assumptions has a likelihood of occurring once in 1,000 years, whereas under a "heavy talied" (t-model with 4 DoF) assumption, it would be once in 8 years. We note that implementing a risk model based on the t-dependence structure is hardly different from the Gaussian counterpart, in particular, a t-factor model can be easily constructed and the estimation of its parameters would be similar to a Normal-based model [e.g., Frey *et al.* (2001)]. Moreover, the t-dependence model is relatively tractable, for example, Glasserman *et al.* (2002) demonstrate that efficient variance reduction techniques can be applied to risk factors having a t-copula.

Numerical example. In order to illustrate the above, we have constructed 6 examples of portfolios holding short positions in 50 call and 50 put options on 100 assets. (Each option strike is 100, maturity of two months and the short rate is flat at 5%.) Portfolios 1-2 correspond to at-the-money options (the stocks' initial price is 100); Portfolios 3-4 correspond to out-of-the-money options (the callable stocks' initial price is 95 and the putables 105), and; Portfolios 5-6 correspond to in-the-money options (the callable stocks' initial price is 105 and the putables 95). The univariate dynamics of each underlying asset follows the usual Black-Scholes model of geometric Brownian motion (GBM) with zero drift and 30% volatility. The pairwise correlation in portfolios 1,3,5 is 0% and 30% for 2,4,6. For each example we simulate 10^6 realizations of the underlying stock price changes for a time horizon of one week while the dependence structure among the names follows a t-model with DoF that is in the range 3–15; these values reflect the typical range found for the data sets in this study, in particular for the G5 pairs. [Note, this model with infinite DoF corresponds exactly to standard multivariate GBM.] Each realization of the stock prices gives the loss of the portfolio for the one-week horizon. The simulation of 10^6 realizations allows for a "fine" assessment of the *loss distribution*, $L(\cdot)$. We next define the Value-at-Risk, $\mathrm{VaR}_\alpha := L^{-1}(\alpha)$ and the conditional-VaR, $\mathrm{CVaR}_\alpha := \frac{1}{1-\alpha}\int_\alpha^1 L^{-1}(t)dt$. [For our purposes, the conditional-VaR and the Expected Shortfall (ES) are equivalent; for more on ES, see, e.g., Rockafellar and Uryasev (2001).] Table 5 provides the computed $\mathrm{VaR}_{0.99}$ and $\mathrm{CVaR}_{0.99}$ for the 6 examples. The differences between the standard assumption of a Gaussian dependence structure and the alternative t-dependence are evident. The VaR and the conditional-VaR numbers increase by a factor of 2 in some of the cases. Given the results of this study, the Gaussian-based risk assessments seem to be over optimistic, while the t-based calculations seem to provide a more realistic picture. We note that these results are amplified even more when one goes further into the tail of the distribution, i.e., increasing the quantile, or examining portfolios that are less correlated.

Discussion. This examples illustrates that the potential for extreme co-movements among

portfolio	t_3		t_6		t_9		t_{12}		t_{15}		Normal	
1	156	[213]	118	[156]	101	[129]	92	[115]	86	[107]	62	[72]
2	168	[227]	136	[181]	123	[163]	116	[153]	113	[148]	100	[127]
3	136	[186]	99	[132]	83	[107]	74	[94]	68	[86]	44	[50]
4	143	[196]	113	[153]	101	[134]	94	[125]	91	[121]	77	[100]
5	162	[219]	127	[165]	113	[143]	105	[130]	101	[123]	80	[92]
6	175	[236]	148	[196]	138	[179]	132	[170]	130	[165]	117	[145]

Table 5: Value-at-risk and conditional-VaR for the respective portfolios: $\mathrm{VaR}_{0.99}$ and [$\mathrm{CVaR}_{0.99}$] for the different cases.
The table gives the $\mathrm{VaR}_{0.99}$ and the $\mathrm{CVaR}_{0.99}$ numbers for the different portfolios and typical values of the DoF parameter. Portfolios 1-2 correspond to at-the-money options, 3-4 to out-of-the-money options, and 5-6 to in-the-money options. The t-dependence can result in approximately doubling the risk numbers in some scenarios, relative to numbers based on the Gaussian dependence structure. Note, most markets lead to a DoF estimate in the range 5-15 (excluding the exceptional pairs, like UK-Japan and US-Japan in the US Dollar denominated data).

financial instruments can have dramatic consequences on the level of risk exposure. Most of the estimated DoF for financial series (except for some currency pairs and some of the Japanese equity market pairs) are relatively small. For example, in the previous section, the estimated DoF of equity market pairs were mostly in the range 5–10, a brief examination of Table 5 reveals that the risk measures double in some of the portfolios when using a t-model with these DoF instead of a Normal model. The results also demonstrate that the t-model is quite easy to implement and requires relatively minor changes compared with a Gaussian-based implementation. By direct modeling of the dependence structure, one can maintain whatever marginal process that is desired for the univariate returns, without sacrificing the tractability and richness of the multivariate dependence structure. In particular, as the example illustrates, univariate returns can follow a GBM process so that consistency with single name risk models is maintained.

The results displayed above illustrate that using a Normal-based model implicitly leads to mis-specification of the potential for extreme co-movements, and therefore, by definition, carries more model risk than a t-based model.

6.3 Pricing

Motivation. Recent years have witnessed a continuing trend of trading multidimensional instruments, e.g., basket options, min/max options, spread options, and more importantly basket and index trading, as well as various structured products. Pricing models for these instruments are usually based on the assumed joint distribution of equity returns. Different distributional assumptions will carry different model risk and would lead to different prices of these instruments. This point is illustrated here via pricing of multi-name credit derivatives. This class of instruments has experienced extraordinary growth in recent years, for example, in 2001 the issuance of Collateralized

Debt Obligations (CDO) in their different forms was approximately 25% of corporates issuance. Multi-name credit derivatives' payoffs are contingent on the default realization in a portfolio of underlying credits. Dependent default models usually use asset returns as a proxy for default time dependency [see, e.g., KMV (Kealhofer 1998), CreditMetrics (Gupton, 1997), Li (2000), Hull and White (2001)]. Accurate specification of the multivariate return process is therefore an essential ingredient in these models.

Numerical example. In this example we consider prices (premiums) of first and second-to-default baskets on a portfolio of five names. The payoff of a first (respectively second)-to-default basket is triggered by the first (respectively second) default in the underlying portfolio, within a predefined time horizon. Therefore, the prices of these instruments are sensitive to extreme movements and co-movements; the natural proxy for defaults in the underlying portfolios. The yearly hazard rate for default is 1% for all the names, the recovery rate is assumed 50% and the discount curve is flat at 5%. Table 6 gives an example of the different prices of the baskets assuming a Gaussian and t-dependence structures with 9 DoF, using 0% and 20% values for the pair-wise correlations between the names. (The value chosen for the DoF is representative of the various data sets considered in this study.) For further details and related calculations see Mashal and Naldi (2002).

maturity	**dependence structure**	**0% correlation**		**20% correlation**	
		1st-TD	**2nd-TD**	**1st-TD**	**2nd-TD**
1 year	**Normal**	248	8	246	15
	$\boldsymbol{t_9}$	237	18	225	30
	price differential	-4%	260%	-9%	100%
3 years	**Normal**	251	13	230	29
	$\boldsymbol{t_9}$	239	26	212	39
	price differential	-5%	100%	-8%	34%

Table 6: Prices of first and second-to-default (TD) baskets, with valuations based on a t-copula with 9 DoF and a Normal-copula.
The table presents prices in basis points for first and second-to-default baskets on a five-name basket with flat discount curve at 5%, recovery rates of 50%, and constant yearly hazard rates of 1% for all the names. The correlation values correspond to pair-wise correlation for each pair of names in the basket.

Discussion. The estimation results throughout this paper suggest that co-movements of financial returns are captured much better with a t-based model than with a Normal one. In the above example we have used DoF and correlation that are commonly observed in equity returns. Time to default models build on Merton's idea that default results from the asset value declining below the debt of the firm. The multivariate generalization postulates that the dependence between defaults follow the dependence between asset values with equity values serving as proxies for assets. These models, therefore, model the dependence structure of default times in a direct manner. The difference between the t-based prices and the Normal-based ones for the second-to-default case are a direct consequence of the underlying dependence assumption since a second to default would be triggered when at least two names default within the horizon, or, in other words, at least two names experienced extreme realizations. The differences will be even more pronounced for third-to-default baskets or for higher order baskets, and similar differences will also occur in prices of CDO tranches. Consequently, valuations that build on Normal-based models imply mis-specification and, therefore, mis-pricing. This can be mitigated to some extent by resorting to a model that supports extremal behavior, such as the t-based model. Moreover, the latter is not a major departure from the tractability offered by Gaussian models, in particular, Glasserman *et al.* (2002) have recently demonstrated that efficient variance reduction techniques are still amenable in the t-model context. (The importance of this result is evident in that most risk management and pricing application rely on simulation methodology.)

7 Concluding Remarks

Statistical evidence that indicates the potential for extreme co-movements is obtained through a test for the underlying dependence structure. Our use of the t-dependence structure is well suited for this objective insofar as it provides a natural "first step" generalization of the correlation-based Gaussian dependence structure. In particular, it supports extremal behavior, and this phenomenon is seen to be statistically significant in all data sets analyzed in this study. One implication of these findings is that the Gaussian dependence hypothesis that underlies most modern financial applications should be deemed inadequate. In fact, the nesting of the Gaussian model within the family of t models enables us to test this hypothesis, which, in turn, is rejected with negligible p-values. This also implies, in particular, the rejection of the multivariate Normal distribution.

Dependence structure assumptions are the cornerstone in most multivariate models in finance. While our focus on the t-dependence structure has been driven by the aforementioned testing objectives, the results of this study also suggest that this model could be well suited for various financial modeling purposes. As the three examples in this paper indicate, the t-dependence assumption could lead to a more realistic assessment of the linkage between international markets, as well as more accurate risk management and pricing models, relative to the Gaussian assumption.

Given the results described in this study, we believe that the t-dependence structure is an important step towards a more realistic model of dependencies between underlying assets, while concurrently retaining a relatively high degree of tractability. The need therefore arises for a more thorough investigation of the various aspects of the t-copula (e.g., specification, estimation, implementation, and factorization) and, in particular, its applicability in various financial contexts.

A Asymptotics for the Pseudo Likelihood Ratio Test Statistic

The main objective is to derive (7), namely,

$$\Lambda_n(\hat{\nu}|\nu_0) \Rightarrow (1+\gamma)\chi^2_1, \quad \text{as } n \to \infty$$

where $\gamma > 0$, is a constant that depends on the null hypothesis, and will be identified in what follows. The above asymptotic is essentially a straightforward consequence of the p-ML asymptotic theory in the context of copulas [Genest *et al.* (1995) serves as our main reference], blended in with the standard machinery that is used to derive the LRT asymptotics [Serfling (1980) serves as our main reference]. The reader should note, as discussed previously in Section 4, that the derivation we present here assumes the p-ML theory to hold in the context of the t-copula model, although we have not been successful in verifying all the required technical regularity conditions given in Appendix 1 of Genest *et al.* (1995) in the context of the two-parameter t-copula model. The derivation should thus be viewed as an informal application of this theory.

We start by stating the asymptotics of the p-ML estimator $\hat{\theta}$, which are derived from §2 (Proposition 2.1) and §4 of Genest *et al.* (1995). Specifically, let the null hypothesis be given by $\theta^0 = (\nu_0, \Sigma_0)$ where ν_0 are the restricted DoF, and, with slight abuse of notation, we set $\theta^0 = (\theta_1, \ldots, \theta_m) \in \mathbb{R}^m$ to be the full parameter vector, with $m := 1 + d(d-1)/2$. Then, we have

$$\sqrt{n}(\hat{\theta} - \theta^0) \Rightarrow N\left(0, \Gamma\right) \ , \tag{9}$$

where

$$\Gamma = I^{-1}(\theta^0)\Delta I^{-1}(\theta^0) \ .$$

Here $I(\theta^0)$ is the Fisher information matrix evaluated at $\theta^0 \in \Theta_0$ whose ij entry is given by

$$\left(I(\theta^0)\right)_{ij} := \mathbb{E}^0\left[\frac{\partial \log c(U;\theta^0)}{\partial \theta_i}\frac{\partial \log c(U;\theta^0)}{\partial \theta_j}\right] \quad \text{for } i,j = 1,\ldots,m$$

where $U = (U_1, \ldots, U_d)$ is distributed according to the null hypothesis copula distribution, and $\mathbb{E}^0\{\cdot\}$ is the expectation operator with respect to this distribution. The matrix $\Delta \in \mathbb{R}^{m\times m}$ is a covariance matrix whose ij entry is given by

$$(\Delta)_{ij} = \mathrm{Cov}\left[\frac{\partial \log c(U;\theta^0)}{\partial \theta_i} + \sum_{k=1}^{d} W_{ki} \ , \ \frac{\partial \log c(U;\theta^0)}{\partial \theta_j} + \sum_{k=1}^{d} W_{kj}\right] \quad \text{for } i,j = 1,\ldots,m.$$

The random variables W_{ki} are given by

$$W_{ki} = \int_{u\in[0,1]^d} \mathbb{I}\{U_k \le u_k\} \frac{\partial^2}{\partial u_k \partial \theta_i} \log c(u;\theta^0) dC(u;\theta^0) \quad \text{for } k = 1,\dots,d; \text{ and } i = 1,\dots,m$$

where $C(u;\theta^0)$ is the (null) copula cumulative distribution function, $c(u;\theta^0)$ is the (null) copula density function, and $u = (u_1,\dots,u_d) \in [0,1]^d$.

As Genest *et al.* (1995) observe, the collection $\{W_{ki} : k = 1,\dots,d \text{ and } i = 1,\dots,m\}$ are uncorrelated with the Fisher score variables, consequently, we have that $\Delta = I(\theta) + \Omega$ where $\Omega \in \mathbb{R}^{m\times m}$ is the covariance matrix of $(\sum_{k=1}^d W_{k1},\dots,\sum_{k=1}^d W_{km})$. Consequently, we obtain a simplified expression for the limiting covariance of the estimator given by

$$\Gamma = I^{-1}(\theta^0) + I^{-1}(\theta^0)\Omega I^{-1}(\theta^0) \; .$$

Under technical regularity, Proposition 2.2 of Genest *et al.* (1995) yields that as $\Sigma \to I$ (in the sense of, say, pointwise convergence), where I is the identity matrix in $\mathbb{R}^{m\times m}$, then $\Omega \to 0$. Thus, diminishing correlation implies that the p-ML estimator becomes increasingly efficient in the sense that its asymptotic variance approaches the Cramér-Rao lower bound [cf. Corollary 5.23 in Schervish (1995)].

Put $e_1 = (1,0,\dots,0)$, the first coordinate basis vector in $\mathbb{R}^m$. Following Serfling (1980, §4.4.4), let $b_{\hat\theta} := e_1 \cdot \hat\theta = (\hat\nu - \nu_0)$, where $a \cdot b$ is the usual inner product operation. Now, using the p-ML asymptotic (9) we have that

$$\sqrt{n} b_{\hat\theta} \Rightarrow N\left(0, e_1 \cdot \Gamma e_1\right).$$

Put

$$Z_n = \sqrt{n} b_{\hat\theta} \left(e_1 \cdot I(\theta^0)^{-1} e_1\right)^{-1} b_{\hat\theta} \sqrt{n},$$

the so-called *Wald statistic*. The continuous mapping theorem, applied as in Serfling (1980, pp. 156-158), now gives

$$Z_n \Rightarrow (1+\gamma)\,\chi_1^2$$

where χ_1^2 is the Chi-squared distribution with one degree-of-freedom, and the constant γ is given by

$$\begin{aligned} (1+\gamma) &= (e_1 \cdot \Gamma e_1)\left(e_1 \cdot I(\theta^0)^{-1} e_1\right)^{-1} \\ &= 1 + e_1 \cdot I^{-1}(\theta^0)\Omega I^{-1}(\theta^0) e_1 \left(e_1 \cdot I(\theta^0)^{-1} e_1\right)^{-1} \end{aligned} \tag{10}$$

where the second equality follows from the above simplified expression for Γ. Moreover, $\Lambda_n(\hat\nu|\nu_0) = Z_n + o_p(1)$, follows from the standard arguments used, for example, in the proof of Theorem 4.4.

of Serfling (1980). Thus, by the converging together principle we have that

$$\Lambda_n(\hat{\nu}|\nu_0) \Rightarrow (1+\gamma)\chi_1^2, \quad \text{as } n \to \infty \ .$$

Thus, the p-LRT asymptotic still leads to the usual Chi-squared law, with the added scaling factor $(1+\gamma)$. An intuitive explanation for this is the following. Since $\Lambda_n(\hat{\nu}|\nu_0)$ is formulated in terms of the logarithm of the *product* distribution, we anticipate that the resulting p-ML estimators should be consistent and asymptotically Normally distributed. [This follows from the standard expansion arguments using Taylor's theorem, and by noting that $\hat{F}$ are $\sqrt{n}$-consistent estimators of F, which implies that the pseudo-variable $\hat{U}$ deviates from its ideal counterparts, $U = F(X)$, by order $n^{-1/2}$, in probability.] Since the key in establishing the limiting Chi-squared distribution for the likelihood ratio test statistic is the asymptotic Normality of the estimators, it is not surprising that this result carries through. In addition, the added dependence due to the use of the pseudo-sample $\mathcal{U}_n$ results in inefficiency of the pseudo-ML estimator, that is, the asymptotic variance does not achieve the Cramér-Rao lower bound [see, e.g., Schervish (1995), §5.1.2]. This, in turn, is manifested in the test statistic converging to a limiting random variable which is stochastically larger [by a factor of $(1+\gamma)$] than the usual Chi-squared limit.

Remark 3 (On the asymptotic constant γ)

1. Although an expression for γ is given in (10), it is not straightforward to evaluate this explicitly, nor have we been able to derive "tight" upper bounds on γ. We thus resort to a Monte-Carlo simulation-based estimation technique. The comprehensive study described in Appendix B.2 yields simulation-based estimates $\hat{\gamma} \in [0, 0.1)$ for all scenarios that were tested using bivariate and tri-variate copulas over a range of correlations, as well as larger "baskets" that emulate the three empirical data sets. The corresponding confidence intervals indicate that $\gamma = 0$ cannot be rejected based on the observed value, and moreover, we obtain bounds of the form $\gamma < 0.3$, that hold with very high probability (0.99) for the test cases in question. Given the above we take as a conservative estimate $\gamma \equiv 1$ for the purpose of p-value calculations. [The numerical study carried out by Genest *et al.* (1995) for the Clayton bi-variate family yields lack of asymptotic efficiency for the p-ML estimator, that would translate in the p-LRT setting to values of γ not exceeding 1.]

2. In the particular instance where the sample consists of uncorrelated r.v.'s (i.e., if under H_0, we have $\Sigma = I$, the identity matrix of dimension d), we can explicitly derive using the results in Appendix A that $\gamma \equiv 0$ and the test statistic is asymptotically distributed exactly as a Chi-squared with one DoF.

3. As Genest and Werker (2002) point out, the p-ML estimator is typically not efficient (i.e., asymptotic variance is larger than that associated with the ML estimator). Excluding the case of a bivariate Normal-copula [where efficiency of the correlation estimator was proved by Klaassen and Wellner (1997)], the necessary and sufficient condition derived in Genest and Werker (2002) strongly suggests that asymptotic efficiency in the Gaussian model is an exception to the rule. The inefficiency of the p-ML estimator manifests itself in the constant $\gamma > 0$.

B Numerical Analysis and Validation of the Testing Procedure

B.1 Validation of the Estimation Results

This section presents a comprehensive simulation study that serves to validate our estimation results. First we study the scenario when the null assumption is given by a Normal-copula, and subsequently we investigate the situation where the null assumption is a t-copula. The results strongly suggest the integrity of the estimation procedure.

B.1.1 Testing when the null hypothesis is a Gaussian dependence structure

Here we simulate data that has a Normal-copula dependence structure, fixing the size of basket (recall, these are 30, 9, 6 for the DJIA, FX, and Metals, respectively), the number of observations as in the historical data sets (2,526, 2,263, 3,176 for the DJIA, FX, and Metals, respectively), and using the historical correlation matrix in each case. We simulate 1,000 realizations for each asset type, for example, every realization of the DJIA data consists of 2,526 observations of the 30 stock returns. For every such realization we estimate the DoF using the p-ML method described in Section 4 and Appendix C. Since we simulate a Normal-copula (alternatively, a t-copula with infinite DoF), we would expect our average DoF estimate to be very large. For computational plausibility we set a cutoff level to be 100 DoF, i.e., any estimate that is above 100 is considered to be ∞, corresponding to a Normal-copula. The lowest estimated DoF for all the realizations over the three asset types is 71. Note, these DoF are considered "large" and the t-copula with 71 DoF is already quite "close" to a Normal-copula (see also Appendix B.3). In particular, this estimate in itself is very different than the estimated DoF based on the historical data of the three assets, for all practical purposes.

- **Equities simulation study**. In all 1,000 realizations the estimated DoF were greater than 100.
- **FX simulation study**. In 99.9% of the realizations, the estimated DoF were 100 or greater, the lowest realization was 90.

- **Commodity simulation study**. In 98.5% of the realizations, the estimated DoF were 100 or greater, the lowest realization was 71.

The above results provide some validation for the estimation results given earlier, while also indicating (again) that larger portfolios provide stronger support for rejecting the Normal-dependence structure proxy.

B.1.2 Testing when the null hypothesis is a t-dependence structure

Here we simulate data that has a t-copula dependence structure, fixing the size of the basket (30, 9, 6 for the DJIA, FX, and Metals, respectively), historical correlation structure, DoF (12, 7, 11 for the DJIA, FX, and Metals, respectively), and number of observations as in the three asset types (2,526, 2,263, 3,176 for the DJIA, FX, and Metals, respectively). We simulate 1,000 realizations for each asset type, for example, every realization of the DJIA basket consists of 2,526 observations, each one consisting of a vector of the 30 asset returns. For each realization, the DoF are estimated using the p-ML procedure outlined in Section 4 and Appendix C. For all three asset types, the estimated DoF in all realizations fall in a short interval (the longest consists of 6 different estimated DoF) around the true DoF used to generate the simulated data. These results provide further support for the p-ML estimation procedure (and p-LRT testing methodology) we employ throughout this paper.

- **Equities simulation study ($\boldsymbol{\nu_0 = 12}$)**. The estimated DoF were all in the range 10–13. With 95% in the range 11–12.
- **FX simulation study ($\boldsymbol{\nu_0 = 7}$)**. The estimated DoF were all in the range 6–8. With 73% equal 7.
- **Commodities simulation study ($\boldsymbol{\nu_0 = 11}$)**. The estimated DoF were all in the range 9–14. With 92% in the range 10–12.

B.2 Numerical Analysis of the Pseudo-Likelihood Method

The main objective here is to validate the asymptotic theory of the p-LRT procedure laid out in Section 4 and Appendix A. Recall that the suggested asymptotic for the p-LRT test statistic $\Lambda_n\left(\hat{\nu}\,|\nu_0\right)$, is a χ^2_1 distribution scaled by a factor of $(1+\gamma)$. The main goal here is two-fold: First, we would like to verify that the asymptotic theory outlined in Section 4 and developed in Appendix A is indeed applicable, in the sense that finite sample numerical results conform with, and support, this anticipated theoretical behavior. Second, since the expression for the constant γ is typically difficult to evaluate, we estimate the value of this constant via Monte-Carlo simulation methods.

In order to test the pseudo-likelihood and the p-ML estimators for the t-copula, we have simulated multivariate-t data with $\nu_0 = 100$, and transformed the data to the unit hypercube using two different methods: via the empirical marginal transformation (this is the transformation we use throughout the paper); and, using the t-marginal transformation [that is, using $U_i = t_\nu^{-1}(X_i)$]. For each transformation we compute the relevant test statistic:

$$\Lambda_n(\hat{\nu}\,|\nu_0) = -2\left(\log \frac{\prod_{i=1}^{n} c(\hat{U}_i;\nu_0,\Sigma)}{\prod_{i=1}^{n} c(\hat{U}_i;\hat{\nu},\Sigma)}\right) \quad .$$

Denote as $\Lambda_i^e(\hat{\nu}\,|\nu_0)$ the test statistic of the i'th realization of the empirical marginal transformation and by $\Lambda_i^t(\hat{\nu}\,|\nu_0)$ the test statistic of the i'th realization of the t-marginal transformation. We note that for the latter, standard likelihood estimation theory holds and therefore we would expect $\Lambda_n^t(\hat{\nu}|\nu_0) \approx \chi_1^2$, for large enough samples. We also note that the same set of simulated observations is used in both transformations (to maintain common randomness). We generate $N = 500$ simulated realizations and determine the constant γ using a first moment ratio estimator[11]:

$$1 + \hat{\gamma} := \frac{\frac{1}{N}\sum_{i=1}^{N} \Lambda_i^e(\hat{\nu}\,|\nu_0)}{\frac{1}{N}\sum_{i=1}^{N} \Lambda_i^t(\hat{\nu}\,|\nu_0)} \quad .$$

A standard central limit theorem for this estimator is the basis for establishing confidence intervals for the unknown parameter γ.

We also compare graphically the distribution of the test statistics to a χ_1^2 distributed random variable. Overall we observe that there is little difference between the results obtained using the two transformations. The implication of this result is that using the empirical marginal transformation for the marginals in the case of the t-copula, although it introduces dependence among the observations, has little impact on the resulting limiting distribution for the test statistic. We have conducted tests for bi-variate and tri-variate cases that are detailed below, as well as for a 6-dimensional case using the metals historical data. The largest value obtained for $(1+\hat{\gamma})$ in all cases is 1.09. In practice, taking $(1+\gamma) = 1.1$ as an upper bound makes little difference in the calculated p-values compared with $(1+\gamma) = 1.0$ (standard asymptotic theory). Table 7 makes this

[11] We have used the ratio of the means rather than the mean itself because of the limited accuracy intrinsic in the calculations involving the limiting behavior for large DoF. The ratio estimator exhibits better stability, since the denominator serves as a normalizing factor.

explicit in the context of the DJIA basket, while also displaying p-values for the more conservative value of the asymptotic constant $[(1+\gamma) = 2.0]$ used to calculate the p-values throughout the paper.

DoF	**log likelihood function**	**test statistic**	**p-value using $(1+\gamma) = 1.0$**	**p-value using $(1+\gamma) = 1.1$**	**p-value using $(1+\gamma) = 2.0$**
10	11392.001	26.005	$3.4061 \cdot 10^{-7}$	$1.16 \cdot 10^{-6}$	$3.11 \cdot 10^{-4}$
11	11404.578	0.850	0.356	0.379	0.515
12 (ML)	11405.003	0	1	1	1
13	11397.566	14.876	$1.15 \cdot 10^{-4}$	$2.36 \cdot 10^{-4}$	$6.39 \cdot 10^{-3}$

Table 7: Effects of the asymptotic scaling constant $(1 + \gamma)$ on the computed p-values.
The table depicts the p-values for the equity basket example calculated for: (i). $\gamma = 0$, corresponding to the "usual" chi-squared assumption; (ii). $\gamma = 0.1$, the upper bound derived using bootstrap-based estimation; and, (iii). $\gamma = 1$, the conservative upper bound used for all p-value calculations in the paper. The effect of the asymptotic constant is seen to be relatively small.

B.2.1 Bi-variate results

For different values of correlation we have simulated 500 realizations, each consists of 20,000 observations. Table 8(a) illustrates that $(1 + \hat{\gamma}) < 1.1$ for all tested correlation values (these simulations require substantial computing time). In addition, we also calculate confidence intervals using a standard central limit theorem for the simulation-based ratio estimator $(1 + \hat{\gamma})$. Note, in all cases below this confidence interval (at 95% confidence) covers 1, i.e., we cannot see statistically significant evidence that $\gamma > 0$. Recall also that $\gamma \equiv 0$ for $\rho = 0$ is asserted by the p-LRT theory.

The contrast between the distribution of the simulation-based empirical transformation test statistic, along with the theoretical χ^2_1 are depicted in Figure 5(a). This is further illustrated in the Q-Q plot of the distribution of the empirical transformation test statistic against the theoretical χ^2_1 in Figure 5(b). Clearly the χ^2_1 limiting distribution is a very good approximation to the observed sampling distribution of the test statistic.

B.2.2 Tri-variate results

The results for the tri-variate case are similar to the bivariate case. Table 8(b) gives the estimated values of $(1 + \hat{\gamma})$, as well as the associated confidence intervals. Note, in all cases below this confidence interval (at 95% confidence) covers 1, i.e., we cannot see statistically significant evidence that $\gamma > 0$. Note that the negative correlations only run to -0.4 since for -0.5 and lower values the correlation matrix is no longer positive definite. (Recall, the correlation value is taken to hold pair-wise for all pairs.) Figure 5(c) depicts the distributions of the test statistic $\Lambda^e_n(\hat{\nu} \,|\nu_0)$, based on the empirical marginal transformation and the theoretical χ^2_1 cumulative distribution function

(CDF), again the close resemblance of the two is evident. This is further depicted in the QQ-plot of the two distributions in Figure 5(d).

correlation	$\mathbf{1+\hat{\gamma}}$	**95% confidence interval**
-0.8	1.07	[0.9565, 1.1835]
-0.6	1.02	[0.8968, 1.1432]
-0.4	1.01	[0.8889, 1.1311]
-0.2	1.00	[0.8913, 1.1087]
0.0	1.00	[0.8767, 1.1233]
0.2	1.01	[0.8921, 1.1279]
0.4	1.03	[0.9043, 1.1557]
0.6	1.03	[0.8911, 1.1689]
0.8	1.06	[0.9082, 1.2118]

(a)

correlation	$\mathbf{1+\hat{\gamma}}$	**95% confidence interval**
-0.4	1.08	[0.9492, 1.2108]
-0.3	1.01	[0.8782, 1.1418]
-0.1	1.00	[0.8724, 1.1276]
0.1	1.00	[0.8819, 1.1181]
0.3	1.00	[0.8767, 1.1233]
0.5	1.00	[0.8567, 1.1433]
0.7	1.09	[0.9523, 1.2277]

(b)

Table 8: Simulation-based estimates of $(1+\hat{\gamma})$ for different correlation values of the bivariate [Panel 8(a)] and tri-variate [Panel 8(b)] simulated data.
Each entry in the table was computed using 500 simulation runs, out of which the distribution of the test statistic was computed. The estimate $1+\hat{\gamma}$ is derived from the ratio estimator and the 95% confidence interval is based on the CLT for the aforementioned estimator. Note, the two sided confidence interval extends to values smaller than 1, however, $\gamma \geq 0$ by definition. The correlation values in the table represent the pair-wise correlation between all pairs.

B.3 The Accuracy of the t-Copula Approximation to the Normal

The basis for using the t-copula with large DoF as a proxy for the Normal-copula rests on the following observation. Fixing Σ we have that

$$\sup_{u\in[0,1]^d} |C(u;\nu,\Sigma) - C^G(u;\Sigma)| \to 0 \quad \text{as } \nu \to \infty \ , \tag{11}$$

where C and C^G are the t-copula and Normal-copula given in (2) and (3) respectively. The above asymptotic can be refined to obtain rates of convergence; the following examples demonstrate

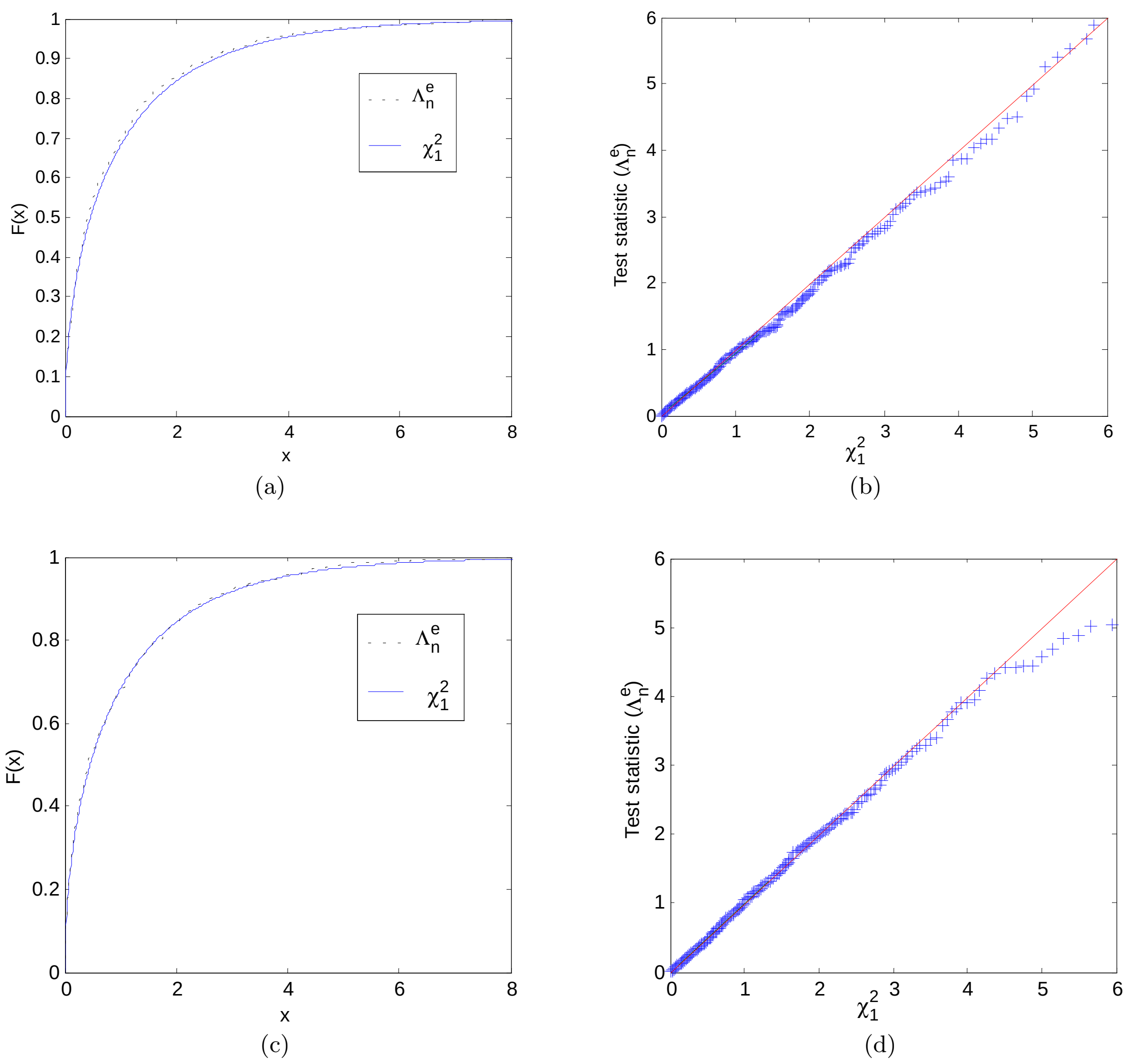


Figure 5: The estimated limiting distribution for the test statistics of the empirical transformation, $\Lambda_n^e\left(\hat{\nu}\,|\nu_0\right)$, vs. χ_1^2.

The figure contrasts the distribution of the test statistic against a χ_1^2 for two cases, bivariate 20% correlation [Panels (a) and (b)] and tri-variate with 10% correlation [Panels (c) and (d)]. Panels (a) and (c) depict the CDF of the test statistics based on the empirical marginal transformation (Λ_n^e) in reference to a χ_1^2. The differences between the two distributions in both panels are quite minor. Panels (b) and (d) present Quantile-to-Quantile plots for the distribution of the test statistics against the theoretical quantiles of χ_1^2. This is another confirmation for the resemblance of the distributions.

that the basic "rules-of-thumb" are similar to those that apply for the Normal and Student-t distributions. We consider two examples that are worked out by sampling the unit hypercube at a uniform, d-dimensional grid, $\Gamma_n = \left\{ \left(\frac{1}{n+1}, ..., \frac{n}{n+1} \right)^d \right\}$ and computing the copula function values at these grid points. Fix $\Sigma \in \mathbb{R}^{d \times d}$, the correlation matrix and compute for each point in $u \in \Gamma_n$ the t-copula pdf for different DoF (ν) and the Normal-copula pdf, where the t-copula pdf is defined in (4) and the Normal-copula pdf is:

$$c(u_1, ..., u_d; \Sigma) = \frac{1}{|\Sigma|^{1/2}} \exp \left(-\frac{1}{2} y^T \left(\Sigma^{-1} - I \right) y \right) \quad ,$$

where $y_i = \Phi^{-1}(u_i)$, $y = (y_1, \ldots, y_d)$, I is the identity matrix of dimension d, and $\Phi^{-1}(\cdot)$ is the inverse Normal transformation. For each value of DoF (ν) we compute the maximal difference:

$$d_\infty(\nu) := \max_{u \in \Gamma_n} \left| c(u; \nu, \Sigma) - c^G(u, \Sigma) \right| \quad ,$$

for fixed Σ. The first example examines the above convergence for a bivariate copula with 0.4 correlation, and, the second example considers a tri-variate copula with uniform correlation of 0.3, i.e., all pair-wise correlations are equal to 0.3. Example 1 was calculated using a grid with n=100, while example 2 uses a grid with n=21. Note that both grids contain roughly 10^4 sampling points. The results are depicted in Figure 6. In both cases for DoF$> 10^3$ the two copulas are essentially indistinguishable.

C Computational Aspects and Implementation

The purpose of this appendix is to explain in detail the estimation procedure that we employ in the paper. Recall from Section 4 that we are maximizing $L_n(\theta) = \prod_{i=1}^{n} c\left(\hat{U}_i; \theta\right)$, where, $\hat{\mathcal{U}}_n = \left\{\hat{U}_i\right\}_{i=1}^{n}$ is the pseudo-sample, $\theta = (\nu, \Sigma)$, and $c(\cdot)$ is the t-copula given in (4). The likelihood function $L_n(\theta)$ should be maximized simultaneously with respect to both ν and Σ. The simultaneous maximization is quite involved [see, e.g., Johnson and Kotz (1972 §37.4) for a discussion of the multivariate-t, and Bouye *et al.* (2000) for a discussion in the context of the t-copula]. In particular, it can be shown that the ML estimator of Σ is given by an implicit solution to a fixed point equation [see Bouye *et al.* (2000), §4.3.1.2]. An iterative procedure may be used as in Bouye *et al.* (2000), however, this procedure is computationally intensive and suffers from numerical stability problems arising from the inversion of close-to-singular matrices. Consequently we propose to use a rank correlation estimator, specifically, the Kendall τ estimator. Based on this estimator we then maximize the pseudo-likelihood function with respect to the DoF (ν).

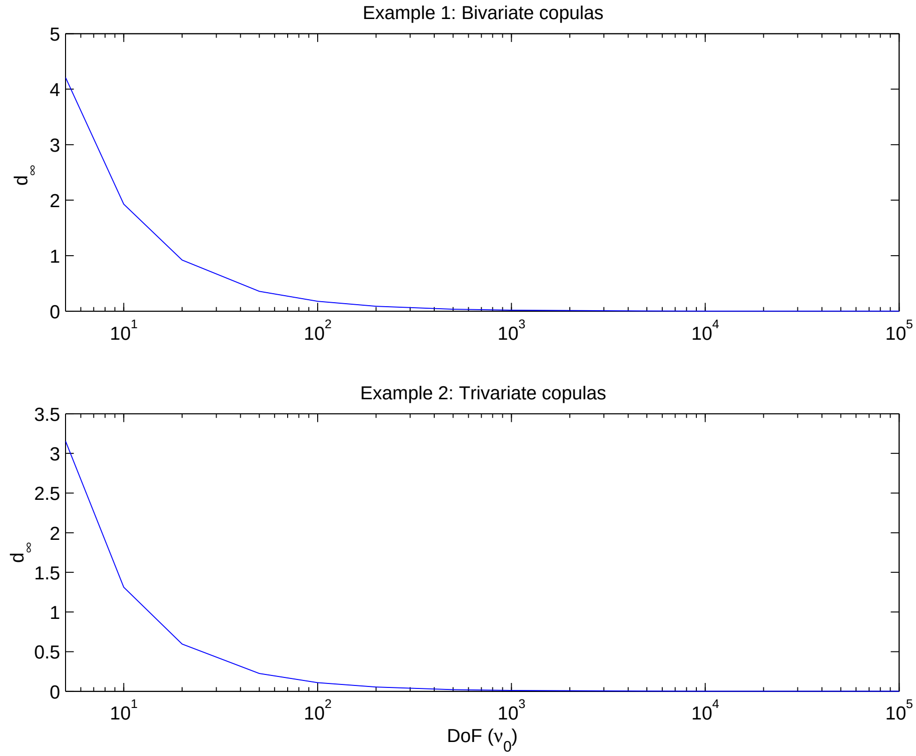


Figure 6: Uniform convergence rate for the t-copula to the Normal-copula.
The figure depicts the convergence rate for the t-copula pdf to the limiting Normal-copula as the DoF grow, measured in the uniform metric d_∞. Example 1 concerns bivariate copulas with 40% correlation and Example 2 concerns tri-variate copulas with pair-wise correlation 30%. When the DoF$>10^3$ the pdfs of the two copulas are almost identical.

The following definition of Kendall's τ for a random vector (X, Y) is taken from Embrechts *et al.* (2001a), §3.4:

$$\tau(X, Y) = P\left\{\left(X - \tilde{X}\right) \cdot \left(Y - \tilde{Y}\right) > 0\right\} - P\left\{\left(X - \tilde{X}\right) \cdot \left(Y - \tilde{Y}\right) < 0\right\} \quad , \tag{12}$$

where $\left(\tilde{X}, \tilde{Y}\right)$ is an independent copy of (X, Y), and X, Y are random vectors of length n, say. The sample estimator $\hat{\tau}$ is given by the following [taken from Lindskog (2000)]. In a sample of length n there are $n(n-1)/2$ pairs of points $\{(x_i, y_i), (x_j, y_j)\}$. Let c denote the number of pairs such that $(x_i - x_j) \cdot (y_i - y_j) > 0$ and d denote the number of pairs such that $(x_i - x_j) \cdot (y_i - y_j) < 0$. Then the estimator for τ is given by:

$$\hat{\tau} = \frac{c - d}{c + d} \ . \tag{13}$$

Note that a slight adjustment is needed in case of ties see, e.g., Lindskog (2000).

For the family of elliptical distributions, and more generally, for the family of distributions with elliptical copulas, the relationship

$$\Sigma = \sin\left(\frac{\pi}{2}\tau\right) \quad , \tag{14}$$

holds. Exploiting this relationship, the obvious "plug-in" approach uses a non-parametric estimator for Kendall's τ and applies relation (14). In particular, Kendall's τ can be estimated using (13). To summarize, the estimator is given by $\hat{\Sigma} = \sin\left(\frac{\pi}{2}\hat{\tau}\right)$. This estimator turns out to be efficient and robust [see, e.g., Embrechts *et al.* (2001a), Lindskog (2000), Hult and Lindskog (2001), and Lindskog *et al.* (2001)]. In our experience the difference between the iterative procedure described in Bouye *et al.* (2000) and the plug-in estimator using $\hat{\tau}$ is negligible. In addition, the latter estimator has three computational advantages: first, it is of order $\mathrm{O}\left(n^2\right)$ for each correlation coefficient, which is computationally expensive but strictly better than applying the iterative procedure; second, it involves rudimentary mathematical operations and is therefore computationally stable; finally, since the rank correlation is invariant under strictly increasing transformations of the marginals we anticipate that the empirical marginal transformation ought to work well for large enough financial time-series. Note, the estimate $\hat{\tau}$ does not use information on the actual DoF (ν) or any estimate of the latter based on the given sample. We also note in passing that the Pearson estimator is a potential alternative to Kendall's τ in our set-up, since it is an efficient estimator within the family of elliptical distributions [see, e.g., Fang and Anderson (1990 §17)]. To summarize, the estimation algorithm proceeds as follows.

Step 1. Transform the raw data, $\mathcal{X}_n$ to the pseudo-sample, $\mathcal{U}_n$, using the empirical marginal transformation.

Step 2. Estimate the correlation matrix $\hat{\Sigma}$ via the plug-in estimator using Kendall's τ, i.e., $\hat{\Sigma}_{ij} = \sin\left(\frac{\pi}{2}\hat{\tau}_{ij}\right)$.

Step 3. Perform a numerical search for $\hat{\nu}$, i.e., $\hat{\nu} = \underset{\nu \in (2,\infty]}{\arg\max} \left[\sum_{i=1}^{n} \log\left(c\left(\hat{U}_i; \nu, \hat{\Sigma}\right)\right)\right]$.

D Data Description

Equities. The equity data consists of the thirty stocks that were included in the Dow Jones Industrial Average (DJIA), on February 1, 2002: Alcoa (AA), American Express (AXP), AT&T (T), Boeing (BA), Caterpillar (CAT), Citigroup (C), Coka Cola (KO), Du Pont (DD), Eastman Kodak (EK), Exxon Moblie (XOM), General Electric (GE), General Motors (GM), Home Depot (HD), Honeywell (HON), Hewlett-Packard (HWP), Intel (INTC), International Business Machines (IBM), International Paper (IP), JPMorgan-Chase (JPM), Johnson and Johnson (JNJ), McDonalds (MCD), Merck (MRK), Microsoft (MSFT), Minnesota Mining (MMM), Philip Morris (MO), Procter & Gamble (PG), Southwestern Bell (SBC), United Technologies (UTX), WalMart (WMT), and Walt Disney (DIS). The data source is the Center for Research in Security Prices (CRSP). The dataset consists of daily returns for the 30 aforementioned stocks, across 2,527 trading days that span 01/02/1991 to 12/28/2000.

Foreign Exchange (FX). The FX data contain the following nine currencies: Australian Dollar (AUD), Canadian Dollar (CAD), French Franc (FRF), German Mark (DEM), Italian Lira (ITL), Japanese Yen (JPY), New Zealand Dollar (NZD), Swiss Franc (CHF), and the United Kingdom Pound (UKP). The data source is the historical database of the Federal Reserve Board. The dataset consists of 2,263 observations of daily returns for each currency in the period between 01/02/1990 to 12/31/1998. We note that the reported asset returns are computed in the different currencies values against the US Dollar, which serves as the *baseline currency*.

Metals. The Metals data consists of the following traded commodities: Aluminum, Copper, Nickel, Lead, Tin, and Zinc. The major market trading in metals is the London Metals Exchange and this is also the source of the data. The data spans the period between 07/03/1989 to 01/31/2002, a total of 3176 daily return observations for each metal type.

G5 equity data. The international equity markets data, both in local and US Dollar denominations, for the G5 countries (France, Germany, Japan, United Kingdom, and United States) was downloaded from Morgan Stanley Capital International (`www.msci.com`). The data consists of 646 weekly observations (Wednesday) and span the dates 1/10/90 until 5/29/02.

References

[1] Amemiya, T., 1984, *Advanced Econometrics*, Harvard university press.

[2] Andrews, D.W.K., 2001, Testing when a parameter is on the boundary of the maintained hypothesis, *Econometrica*, **69**:683–734.

[3] Ang, A. and G. Bekaert, 2002, International asset allocation with regime shifts, Forthcoming *Review of Financial Studies*.

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

[5] Ang, A., J. Chen, and Y. Xing, 2002, Downside correlation and expected stock returns, Working paper, Columbia Business School.

[6] Bae, K. H., G. A. Karolyi, and R. M. Stulz, 2002, A new approach to measuring financial contagion, Working paper, Ohio State University.

[7] The Basel accord II proposal, 2001, Available at `www.bis.org`.

[8] Blattberg, R. C. and N. J. Gonedes, 1974, A comparison of the stable and student distributions as statistical models for stock prices, *Journal of Business* **47**:244–280.

[9] Bouye, E., V. Durrleman, A. Nikeghbali, G. Riboulet, and T. Roncalli, 2000, Copulas for finance a reading guide and some applications, Working paper, Credit Lyonnais.

[10] DanÌelsson, J., P. Embrechts, C. Goodhart, C. Keating, F. Muennich, O. Renault, and H.-S. Shin, 2001, An academic response to Basel II, Working paper, LSE Financial Markets group, ESRC research center.

[11] DanÌelsson, J. and C.G. de Vries, 1997, Tail index and quantile estimation with very high frequency data, working paper, Tinbergen Institute, Rotterdam, The Netherlands.

[12] Durrleman, V., A. Nikeghbali, and T. Roncalli, 2001, Which copula is the right one? Working paper, Credit Lyonnais.

[13] Embrechts, P., F. Lidskog, and A. J. McNeal, 2001a, Modeling dependence with copulas and applications to risk management, Working paper, Risklab ETH Zurich.

[14] Embrechts, P., A. J. McNeal, and D. Straumann, 2001b, Correlation and dependence in risk management: properties and pitfalls, Working paper, Risklab ETH Zurich.

[15] Fama, E.F, 1965, The behavior of stock market prices. *Journal of Business*, **38**:34–105.

[16] Fang, K.T. and T. W. Anderson, 1990, *Statistical inference in elliptically contoured and related distributions*, Allerton press, New York.

[17] Forbes, K. J. and R. Rigobon, 2002, No contagion, only interdependence: measuring stock market co-movements, Forthcoming *The Journal of Finance*.

[18] Frees, E. W. and E. Valdez, 1998, Understanding relationships using copulas, *North American Actuarial Journal volume* **2(1)**:1–25.

[19] Frey, R. and A. J. McNeil, 2001, Modelling dependent defaults, Preprint, University and ETH Zurich.

[20] Frey, R., A. J. McNeil, and M. Nyfeler, 2001, Copulas and credit models, *RISK*, October.

[21] Genest, C., K. Ghoudi and L.-P. Rivest, 1995, A semiparametric estimation procedure of dependence parameters in multivariate families of distributions, *Biometrika*, **82:**543–552.

[22] Genest, C. and L.-P. Rivest, 1993, Statistical inference procedures for bivariate Archimedean copulas, *J. Amer. Statist. Assoc.*, **88**:1034–1043.

[23] Genest, C. and B.J.M. Werker, 2002, Conditions for the asymptotic semiparametric efficiency of an omnibus estimator of dependence parameters in copula models, *Proceedings of the Conference on Distributions with Given Marginals and Statistical Modelling*, Barcelona, Spain, 103–112.

[24] Glasserman, P., P. Heidelberger, and P. Shahabuddin, 2002, Portfolio value-at-risk with heavy-tailed risk factors, *Mathematical Finance*, **12(3)**:239–269.

[25] Gupton, G. M., C. C. Finger, and M. Bhatia, 1997, *CreditMetrics - technical document*, Available from `www.creditmetics.com`.

[26] Hull, J. and A. White, 2001, Valuing credit default swaps II: modeling default correlation, *Journal of Derivatives*, **8**:12–22.

[27] Hult H. and F. Lindskog, 2001, Multivariate exremes, aggregation and dependence in elliptical distributions, Working paper, Risklab ETH Zurich.

[28] Joe, H., 1997, Multivariate models and dependence concepts, Volume 37 of *Monographs on statistics and applied probability*, Chapman and Hall, London.

[29] Johnson, N. L. and S. Kotz, 1972, *Continuous multivariate distributions.* Wiley, New York.

[30] Jorion, P., 2000, Risk management lessons from Long-Term Capital Management, *European Financial Management*, **6**:277–300.

[31] Karolyi, G. A. and R. M. Stulz, 1996, Why do markets move together? An Investigation of U.S.-Japan Stock Return Comovements, *The Journal of Finance*, **51(3)**:951–986.

[32] Kealhofer, S., 1998, Portfolio management of credit risk, Working paper, KMV, Available from `www.kmv.com`.

[33] Klaassen, C.A.J. and J.A. Wellner, 1997, Efficient estimation in the bivariate normal copula model: normal margins are least favorable, *Bernoulli*, **3:**55–77.

[34] Li, D. X., 2000, On default correlation: a copula function approach, *Journal of Fixed Income*, **9**:43–54.

[35] Lindskog, F., 2000, Linear correlation estimation, Working paper, RiskLab, ETH Zurich.

[36] Lindskog F., A. McNeil, and U. Shmock, 2001, Kendall's tau for elliptical distributions, Working paper, Risklab ETH Zurich.

[37] Longin, F., 1996, The asymptotic distribution of extreme stock market returns, *The Journal of Business*, **63**:383–408.

[38] Longin, F. and B. Solnik, 1995, Is the correlation in international equity returns constant: 1960-1990 ?, *Journal of International Money and Finance*, **14(1)**:3–26.

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

[40] Mandelbrot, B., 1963, The variation of certain speculative prices, *Journal of Business*, **36**:394–419.

[41] Markowitz H.M., 1959, *Portfolio selection: efficient diversification of investments*, Yale University Press.

[42] Mashal R., and M. Naldi, 2002, Generalizing asset dependency - application to credit derivatives pricing, Working paper, Columbia Business School.

[43] Mina, J. and H. Xiao, 2001, Return to RiskMetrics: the evolution of a standard, RiskMetrics group Inc.

[44] Nelsen, R. B., 1998, *An Introduction to Copulas*, Vol. 139 of Lecture Notes in Statistics, Springer, Berlin.

[45] Praetz, P.D., 1972, The distribution of share price changes, *Journal of Business*, **45**:49–55.

[46] Rencher, A. C., 1995, *Methods of multivariate analysis*, John Wiley & Sons, Inc., New York.

[47] Rey, D.M., 2000, Time-varying stock market correlations and correlation breakdown, Working paper, University of St. Gallen, Switzerland.

[48] Rockafellar, T. and S. Uryasev, 2001, Conditional value-at-risk for general loss distributions, University of Florida, Research report.

[49] Schervish, M.J., 1995, *Theory of statistics*, Springer-Verlag.

[50] Serfling, R.A., 1980, *Approximation theorems of mathematical statistics*, Wiley, New York.

[51] Sklar, A., 1959, Fonctions de répartition à n dimensions et leurs marges, *Publications de l'Institut de Statistique de l'Université de Paris*, **8**:229–231.

[52] Solnik, B., C. Boucrelle, and Y. Le Fur, 1996, International market correlation and volatility, *Financial Analyst Journal* **Sep/Oct**:17-34.

[53] Starica, C., 1999, Multivariate extremes for models with constant conditional correlations, *Journal of Empirical Finance* **6**:515-553.

[54] Straetmans S., 2000, Extremal spillovers in equity markets, Working paper, University of Maastricht, the Netherlands.

[55] Strong, N. and X. Xu, 2002, Understanding the equity home bias: evidence from survey data, Working paper, University of Manchester, England.