Working Paper Series

No 12 / June 2016

Extreme risk interdependence

by
Arnold Polanski
Evarist Stoja

ESRB
European Systemic Risk Board
European System of Financial Supervision

Abstract

We define tail interdependence as a situation where extreme outcomes for some variables are informative about such outcomes for other variables. We extend the concept of multi-information to quantify tail interdependence, decompose it into systemic and residual interdependence and measure the contribution of a constituent to the interdependence of a system. Further, we devise statistical procedures to test: a) tail independence, b) whether an empirical interdependence structure is generated by a theoretical model and c) symmetry of the interdependence structure in the tails. We outline some additional extensions and illustrate this framework by applying it to several datasets.
Key words: co-exceedance, Kullback-Leibler divergence, multi-information, relative entropy, risk contribution, risk interdependence

JEL: C12, C14, C52

1. Introduction

The recent intense interest in (tail) interdependence is driven by its importance in economics, finance, insurance and in many other areas of applied probability and statistics. Research has documented that dependence has a complex nature, is strongly non-normal, with a time-varying strength and shape (e.g., Patton, 2009). Simultaneously capturing these characteristics has proved to date difficult.

In economics and finance, dependence is paramount for many important applications such as portfolio decisions (e.g., Ang and Bekaert, 2002), risk management (e.g., Embrechts et al., 2002), multidimensional options (e.g., Cherubini and Luciano, 2002), credit derivatives, collateralised debt obligations and insurance (e.g., Hull and White 2006; Kalemanova et al., 2007; Su and Spindler, 2013), contagion, spillovers and economic crises (Bae et al., 2003; Zheng, et al., 2012) and market integration (e.g., Bartram et al., 2006).

The literature contains several notions of dependence (e.g., Li, 2009; Colangelo et al., 2005; Joe, 1997). The most widely applied dependence measure, the Pearson's correlation coefficient, is an inadequate measure in many situations as it captures only the linear dependence between pairs of random variables (see e.g., Embrechts et al., 2002, Longin and Solnik, 2001). Alternatively, dependence has been captured by copulas (e.g., Patton, 2009; Giacomini et al., 2009). However, while copulas have useful properties such as analytic measures of dependence and the invariance of dependence under increasing and continuous transformations, they are still based on parametric assumptions that may not hold in practice, e.g. imposing specific marginal probability density functions (PDFs) and a copula on the data. (Multivariate) extreme value theory (EVT) has also been applied to extreme interdependence (see, for example, Jansen and de Vries, 1991; Hartmann et al., 2000). However, EVT only provides asymptotic results and it relies heavily on parametric models (see Longin and Solnik, 2001). Ledford and Tawn (1996) propose models characterizing the asymptotic dependence of distributions while Coles et al (1999) propose diagnostics for such dependence. Heffernan (2001) provides a directory of coefficients of tail dependence. However, these studies are typically focused on bivariate distributions and it is not clear whether they can be extended easily to higher dimensions.

A difficulty in measuring dependence in financial data is asymmetry. In the univariate case, the leverage or feedback effects, where the magnitude of a negative return following bad news is larger than the magnitude of a positive return following good news of the same nature, has motivated the asymmetric GARCH literature (see Engle, 2002; Bollerslev, 2009, and the references therein). Similarly, in the multivariate case, the magnitude of co-movement in negative returns following bad news is larger than the magnitude of co-movement in positive returns following good news of the same nature. This phenomenon has motivated the literature of asymmetric return dependence (see, for example, Longin and Solnik, 2001; Ang and Bekaert, 2002; Bae et al., 2003). As Hong et al. (2007) point out, accounting for asymmetries is important as otherwise they can cause severe problems with hedging and portfolio diversification. In particular, the standard advice to hold a well diversified portfolio might be questionable if all stocks tend to fall as the market experiences an extreme drop. However, accounting for asymmetric dependence requires care (see, for example, Boyer et al., 1999; Forbes and Rigobon, 2002). Formal tests to assess the existence of asymmetric correlations have been developed by Ang and Chen (2002) and Hong et al. (2007).

In this paper we focus on co-exceedances - counts of joint occurrences of extreme outcomes. We compute for all subsets of variables their observed co-exceedances and compare them to co-exceedances expected under a hypothesized model. Formally, we define the *tail interdependence* as a situation where the tail events of some random variables are informative about such events for other variables. Conversely, under *independence*, tail events in any subset of variables do not convey any information about tail events of other variables. Further, while we often use the terms interdependence and dependence interchangebly, we distinguish between the two concepts as follows. Dependence refers to the relationship between two random variables whereas interdependence refers to the relationship among $n \geq 2$ variables. Hence, the latter nests the former concept of dependence in that it is more general and encompassing. Similar to Longin and Solnik (2001), Ang and Chen (2002), Bae et al. (2003) and Hong et al. (2007) among others, we treat positive co-exceedances (upper tails) separately from the negative co-exceedances (lower tails). This separation allows for testing whether the dependence in the lower and the dependence in the upper tails are symmetric.

Our approach to measuring interdependence, similarly to Joe (1989), relies on the concept of (relative) entropy or multi-information. Entropy is used in many areas of natural sciences and has recently been productively employed in economics and finance (see, Van Nieuwerburgh and Veldkamp, 2010; Backus et al., 2014).

We make the following contributions. We propose a non-parametric measure of tail interdependence, the coefficient of tail interdependence (CTI). This measure follows naturally from the concept of multi-information, is generic and can be applied to an array of problems. Then, we decompose total interdependence into systemic *and* residual interdependence and measure the contributions of constituents (e.g., assets) to the interdependence of a system (e.g., portfolio). Further, we provide a natural framework for statistical tests of independence in the tails; a goodness-of-fit test assessing the compatibility of the observed tail interdependence structure with the one generated by a hypothesized model; and dependence symmetry between the lower and the upper tails (or any two tails). These tests can be employed unconditionally and, importantly, conditionally to distinguish between different models of conditional dependence such as multivariate GARCH or time-varying copulas. Moreover, this framework can easily be applied to generate synthetic data with the same tail interdependence structure as that observed in an actual dataset. This is particularly useful in applications, where the tail interdependence is the overriding concern such as risk management. In the Appendix, drawing on the insights developed in information theory and the related areas of natural sciences, we discuss additional interesting extensions that arise naturally in the relative entropy/multi-information framework.

To illustrate the potential and flexibility of this methodology for providing insights into tail interdependence, we apply it (conditionally and unconditionally) to daily returns of equity indices of G7 countries, high-frequency returns for six European markets and daily returns of Dow Jones Industrial Average (DJ30) index constituents. Our empirical findings confirm some well-known and uncover a few new stylized facts on extreme returns. For example, standard asset pricing factors account for most of the interdependence of the DJ30 stock returns in the center but not in the tails of the distribution - a result of their own high interdependence in the tails.

The paper proceeds as follows. In Section 2 we introduce the joint tails and the tail interdependence structure as the fundamental tools of our framework. Using this concept, in Section 3 we define the coefficient of tail interdependence and introduce statistical tests of independence, goodness-of-fit and interdependence symmetry. We illustrate the flexibility and potential of the framework in Section 4. Section 5 summarizes the paper and offers some concluding remarks. In the Appendix, we prove some of the results presented in the paper and discuss some extensions of the tail interdependence framework.

2. Joint tails and the tail interdependence structure

Let $\mathcal{N} = \{1, ..., n\}$ be a finite set and $F = F_{\mathcal{N}}$ a continuous joint CDF (PDF $f = f_{\mathcal{N}}$) of a vector $X = (X_1, ..., X_n)$ of n random variables with the support on a convex and full-dimensional set $\Omega \subseteq R^n$. For the strictly increasing marginal CDF F_i, $i \in \mathcal{N}$, the value at risk (VaR) at the nominal level $\alpha \in (0, 1)$ is the α-quantile $F_i^{-1}(\alpha)$. For $i \in \mathcal{N}$, we define the (lower) univariate tail $S_i(\alpha) = \{x \in \Omega : x_i < F_i^{-1}(\alpha)\}$ as a set of outcomes in Ω with the i-th component below the quantile $F_i^{-1}(\alpha)$. For the tail probabilities it holds that $f(S_i(\alpha)) = \alpha$, where the notation $f(S)$ stands for the probability of the set S under the PDF f. We define the (lower) joint tail (JT) at the nominal level α as follows: for a subset $C \subseteq \mathcal{N}$, a JT $T_C(\alpha)$ contains outcomes $x \in \Omega$ such that x_i exceeds $F_i^{-1}(\alpha)$ for

$i \in C$ and x_i does not exceed $F_i^{-1}(\alpha)$ for $i \in \mathcal{N} \backslash C$,

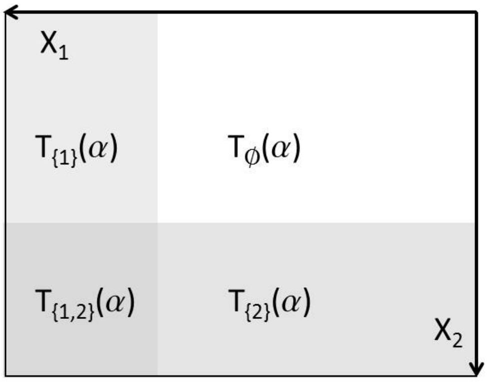
$$T_C(\alpha) = \{x \in \Omega : x_i < F_i^{-1}(\alpha) \quad \forall i \in C \quad \& \quad x_i \geq F_i^{-1}(\alpha) \quad \forall i \in \mathcal{N} \backslash C\}. \quad (1)$$

Note that the univariate tail $S_i(\alpha)$ is the union of all JTs where X_i exceeds its VaR,

$$S_i(\alpha) = \bigcup_{C \subseteq \mathcal{N} : i \in C} T_C(\alpha).$$

Importantly for our purposes, the joint tails $T_C(\alpha)$ and $T_B(\alpha)$ are disjoint if $C \neq B$. Therefore, the superset $\mathcal{T}(\alpha) = \{T_C(\alpha) : C \subseteq \mathcal{N}\}$ partitions the outcome space Ω into 2^n (the number of all subsets of $\mathcal{N}$) regions. In other words, the disjoint sets in $\mathcal{T}(\alpha)$ cover the entire outcome space Ω. Figure 1 illustrates the partition of Ω into $\mathcal{T}(\alpha)$.

Figure 1: The Partition of the Outcome Space into Joint Tails



Notes: The figure illustrates the partition of the outcome space into joint tails $T_C(\alpha)$ for $n = 2$.

The subsets in $\mathcal{T}(\alpha)$ depicted in Figure 1 could be given interesting interpretations. For example, for a low α, the JT $T_\emptyset$ captures the dependence in returns in the spirit of CAPM or APT - the dependence of the expected returns of an asset on the expected return of the market or another asset. The JT $T_{\{1,2\}}$ could be interpreted as dependence in risk - the dependence of an extreme event for asset 1 on the extreme events on asset 2 and vice versa. Similarly, the JTs $T_{\{1\}}$ and $T_{\{2\}}$ could be interpreted as return-risk dependence - the dependence of the return of asset 2 (1) on the extreme risk of asset 1 (2) respectively.

Different users may only be interested in particular subsets of $\mathcal{T}(\alpha)$ and overlook others. For example, a properly hedged investor may only be interested in $T_\emptyset$ while a regulator may only be interested in $T_{\{1,2\}}$. Similarly, only $T_{\{1\}}$ and $T_{\{2\}}$ may be relevant for the pricing of exotic securities or insurance products.

For a partition $\mathcal{T}(\alpha)$ of the outcome space Ω and a PDF $f : \Omega \to R_+$, we define the *tail interdependence structure* (TIS) $u(f, \alpha) = \{u_C(f, \alpha)\}_{C \subseteq \mathcal{N}}$ as an 2^n-dimensional vector, where

$$u_C(f, \alpha) = f(T_C(\alpha)) = \int_{\tau \in T_C(\alpha)} f(\tau) d\tau, \tag{2}$$

is the probability mass of the JT $T_C(\alpha)$ under f. When there is no risk of confusion, we omit the reference to f and α in $u(f, \alpha)$ and write u instead. Clearly, u is a (discrete) PDF as $\mathcal{T}(\alpha)$ is a partition of the sample space. Generally, the information content of the discrete PDF p defined on the domain $\mathcal{D}$, is measured by its entropy (Shannon, 1948),

$$H(p) := -\sum_{i \in \mathcal{D}} p_i \ln p_i, \tag{3}$$

where $\ln(.)$ is the natural logarithm and, by convention, $0 \ln 0 = 0$. For example, when the marginal probability distribution of VaR exceedances is given by $p^\alpha := (\alpha, 1 - \alpha)$ (i.e. VaR is exceeded with probability α and not exceeded with probability $1 - \alpha$) then $H(p^\alpha) = -\ln(\alpha^\alpha (1 - \alpha)^{1-\alpha})$. The entropy $H(p^\alpha)$ depends only on α and plays an important role in the ensuing analysis.

The TIS u contains all the relevant information regarding the joint exceedances in the lower JTs, e.g., joint losses of some assets. In other cases, the focus of the investigation may be on joint gains or, more generally, on the tail interdependence of some linear combinations (portfolios) of the random vector X,

$$Y_i = A_{i1} X_1 + ... + A_{in} X_n, \quad i = 1, ..., m.$$

For an $m \times n$ real matrix[1] $A = (A_{ij})$, we can compute the density function $g(y)$ of the random vector $Y = (Y_1, ..., Y_m)$ and, hence, the TIS $u(g, \alpha)$ by the change of variables theorem,

$$Y = AX \Rightarrow g(y) = \frac{1}{|\det A|} f(A^{-1}y).$$

In particular, we can use the latter formula to compute the TIS $u(g, \alpha)$ when A is a rotation matrix,

$$Y = AX \Rightarrow g(y) = f(A^T y), \quad as \quad A^T = A^{-1} \quad \& \quad |\det A| = 1.$$

For example, by setting $A = -I$, where I is the identity matrix, we obtain the TIS for the upper tails. Rotations will allow us to compute the TIS not only for the lower and the upper tails but also for the mixed tails, i.e., among the lower univariate tails for some variables and the upper univariate tails for others.

3. Measurement and statistical testing of tail interdependence

3.1. Coefficient of tail interdependence

The interdependence of the VaR exceedances of n discrete random variables with the joint PDF u and with marginals p^a is fully defined by the *multi-information* (MI) (Cover and Thomas, 2006),

$$I(u) = D(u||\pi^{\alpha}) = \sum_{C \subseteq \mathcal{N}} u_C \ln \frac{u_C}{\pi_C^{\alpha}} \tag{4}$$

where the probability of the JT $T_C(\alpha)$ under independence is

$$\pi_C^{\alpha} = \Pr(T_C(\alpha)) = \prod_{i \in C} \Pr(S_i) \prod_{i \in \mathcal{N} \setminus C} (1 - \Pr(S_i)) = \alpha^{\#C} (1-\alpha)^{n-\#C},$$

[1] The matrix A can be interpreted, for example, as the exposure of the investor or the financial system to each of the $X_1 ... X_n$ assets or financial institutions.

and $\#C$ is the cardinality of set C. MI is non-negative and equals zero in case of independence only, i.e., if and only if $u = \pi^\alpha$. In statistics, $D(u||\pi^\alpha)$ is known as the Kullback-Leibler (KL) divergence between the PDFs u and π^α. MI quantifies the *total* amount of interdependence among random variables that arises from pairwise, triplet or more complex interactions. It is widely used in, for example, physics (Schneidman et al., 2003) and biosciences (Wennekers and Ay, 2003; Schneidman et al., 2006). In particular, it allows for the study of the global statistical structure of a system as a whole, the total dependence between subsystems, and the temporal statistical structure of each subsystem (Chicharro and Ledberg, 2012). Importantly, MI can also be represented as the difference between the sums of individual (marginal) entropies and the joint entropy (Schneidman et al., 2003),

$$I(u) = D(u||\pi^\alpha) = \textstyle\sum_{i=1}^n H(p^a) - H(u). \tag{5}$$

Intuitively, $H(u)$ is a measure of uncertainty in the joint distribution u of the exceedances. Thus, the lower the uncertainty $H(u)$ the higher the MI $I(u)$. This interpretation reveals an important inverse relationship between interdependence and uncertainty (entropy).

We use the MI (4) to measure tail interdependence. Specifically, we define the *coefficient of tail interdependence* (CTI) as,

$$\kappa(\alpha, u) = \frac{D(u||\pi^\alpha)}{(n-1)H(p^a)} = \frac{nH(p^a) - H(u)}{nH(p^a) - H(p^a)}. \tag{6}$$

The CTI has many desirable properties. In the Appendix, we show that the CTI lies in the unit interval. In particular, $\kappa(\alpha, u) = 0$ when all exceedances are mutually independent and $\kappa(\alpha, u) = 1$ in case of perfect dependence, i.e., when all n variables always exceed together their respective thresholds.[2] Secondly, the CTI is scale invariant under strictly increasing transformations of the underlying variables in X. Specifically, if each $\xi_i(X_i)$

[2]Perfect dependence occurs when $H(u) = H(p^a)$, i.e., when the TIS u carries the same information as one marginal p^a.

is an increasing and continuous function, then the CTI computed from the transformed variables $\xi(X) = (\xi_i(X_i))_{i=1,...,n}$ is the same as that computed from X. This property follows by the construction of the TIS from the quantiles of the variables in X as the same events fall into a JT $T_C(\alpha)$ under X and under $\xi(X)$. Further, by the construction of the TIS (2), the CTI is robust to outliers and is invariant under the permutation of the random variables in X. The CTI can also be decomposed into a systemic and a residual component (see subsection 3.2) and it can be used as a test statistic to test tail independence (see subsections 3.3 and 3.4). It is important to note that the CTI does not measure the overall interdependence among random variables. Instead, it quantifies the interdependence of extreme events, where the parameter α defines the severity of the extreme events and a rotation matrix specifies their directions. Although the probability $u_\emptyset$ of the no-exceedance event $T_\emptyset$ is used in the computation the CTI, this probability is fully determined by the probabilities of the other joint tails (because all tail probabilities sum up to one). In this sense, $u_\emptyset$ does not contain any independent information and the computation of the CTI relies exclusively on the information in the probabilities of "genuine" joint tails with at least one exceedance.

Interestingly, the CTI allows for interpreting joint exceedances of the n variables in X as joint exceedances of a smaller number of mutually independent "factors". Specifically, writing (6) as

$$H(u) = (n - n\kappa + \kappa)H(p^a)$$

makes it obvious that the TIS u conveys the same information as $n - n\kappa + \kappa$ independent marginals $p^\alpha = (\alpha, 1 - \alpha)$. In particular, for $\kappa = 1$ ($\kappa = 0$) the information in u is equivalent to that in 1 (n) marginal(s). We can think then of the exceedances in the data generating process X as being driven by $n - n\kappa + \kappa$ independent binary "factors", each having the same distribution p^α as the exceedances of X_i. Moreover, as the CTI effectively relates the information in the TIS u to the information in the n marginals p^a, it allows for comparing the strength of interdependence for different levels of α. An examination of the number $n - n\kappa + \kappa$ of factors over time may be informative regarding the strength of the interdependence of assets or financial institutions and hence, may shed light into the dynamics of the diversification benefits or the financial fragility.

3.2. Interdependence Decomposition

MI (4) is equal to the *total* KL divergence $D(u||\pi^{\alpha})$ between the TIS $u = \{u_C\}_{C \subseteq \mathcal{N}}$ and the PDF $\pi^{\alpha} = \{\pi^{\alpha}_C\}_{C \subseteq \mathcal{N}}$ that holds under tail independence. In some applications however, it is optimal to focus on the aggregate or systemic component of $D(u||\pi^{\alpha})$. Specifically, we define the *aggregate TIS* as the $(n+1)-$dimensional vector $\widetilde{u} = \{\widetilde{u}_k\}_{k=0}^{n}$ where,

$$\widetilde{u}_k = \textstyle\sum_{C \subseteq \mathcal{N}: \# C = k} u_C,$$

and the corresponding $(n+1)-$dimensional vector of JT probabilities under independence as $\widetilde{\pi}^{\alpha} = \{\widetilde{\pi}^{\alpha}_k\}_{k=0}^{n}$, where,

$$\widetilde{\pi}^{\alpha}_k = \textstyle\sum_{C \subseteq \mathcal{N}: \# C = k} \pi^{\alpha}_C.$$

Hence, $\widetilde{u}$ and $\widetilde{\pi}^{\alpha}$ are discrete probability distributions of observing $k = 0, ..., n$ exceedances under the PDF f and under the tail independence, respectively. From the TIS u, we compute the conditional probability $u^k = (u_C / \widetilde{u}_k)_{C \subseteq \mathcal{N}: \# C = k}$ given that k exceedances have ocurred.[3] Similarly, we compute the conditional probability $\pi^{\alpha,k} = (\pi^{\alpha}_C / \widetilde{\pi}^{\alpha}_k)_{C \subseteq \mathcal{N}: \# C = k}$ from the PDF π^{α} for each $k = 0, ..., n$. In the Appendix, we show that the total KL divergence $D(u||\pi^{\alpha})$ can be decomposed as follows,

$$D(u||\pi^{\alpha}) = D(\widetilde{u}||\widetilde{\pi}^{\alpha}) + \textstyle\sum_{k=0}^{n} \widetilde{u}_k D(u^k||\pi^{\alpha,k}). \tag{7}$$

The measure $D(\widetilde{u}||\widetilde{\pi}^{\alpha})$ quantifies the *systemic* or *aggregate* tail interdependence, i.e., the divergence between the distributions of the observed and the expected (under tail inde-

[3] For example, in the bivariate case $u^1_{\{2\}} = u_{\{2\}} / \widetilde{u}_1 = u_{\{2\}} / (u_{\{1\}} + u_{\{2\}})$ is the conditional probability of X_2 exceeding when $k = 1$, i.e., when exactly one exceedance has occurred.

pendence) number of exceedances. On the other hand, each KL divergence $D(u^k||\pi^{\alpha,k})$ quantifies the conditional interdependence among subsets of variables, given that k exceedances have occurred. Thus, while $D(\widetilde{u}||\widetilde{\pi}^{\alpha})$ measures the dependence that is jointly generated by all constituents, the weighted total on the r.h.s. of (7) sums up the intra-systemic dependence among subsets of constituents. Due to the limited importance of the latter to the interdependence of the system, we refer to it as residual interdependence.

In analogy to the CTI (6), we define the *systemic* and *residual CTIs* as, respectively,

$$\widetilde{\kappa}(\alpha, u) = \frac{D(\widetilde{u}||\widetilde{\pi}^{\alpha})}{(n-1)H(p^a)}, \quad \kappa^k(\alpha, u) = \frac{D(u^k||\pi^{\alpha,k})}{(n-1)H(p^a)}, \tag{8}$$

and show in the Appendix that

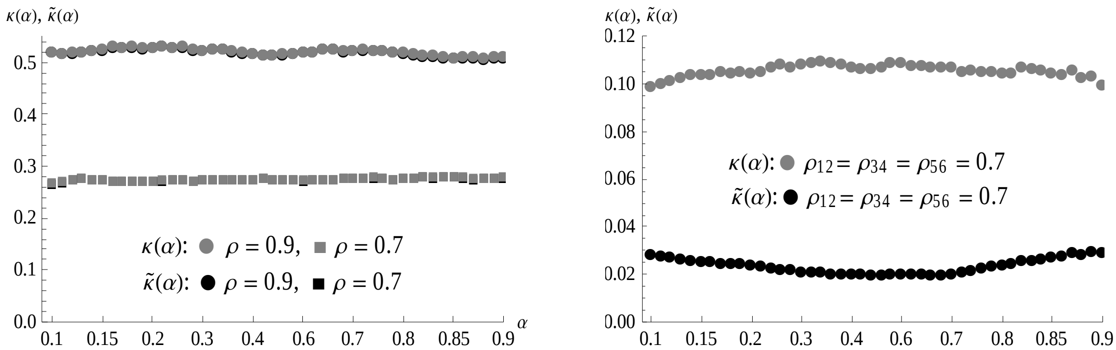
$$\begin{aligned} \kappa(\alpha, u) &= \widetilde{\kappa}(\alpha, u) + \textstyle\sum_{k=0}^{n} \widetilde{u}_k \kappa^k(\alpha, u), \\ 0 &\leq \widetilde{\kappa}(\alpha, u) \leq \kappa(\alpha, u) \leq 1, \end{aligned}$$

with $\widetilde{\kappa}(\alpha, u) = \kappa(\alpha, u) = 0$ in the case of tail independence and $\widetilde{\kappa}(\alpha, u) = \kappa(\alpha, u) = 1$ for perfect dependence (i.e., when all exceedances always occur together).

In high dimensions, the total divergence $D(u||\pi^{\alpha})$, and thus the aggregate CTI $\kappa(\alpha, u)$, may not be estimated accurately when there are no sufficient observations in all joint tails. However, this is not a problem for the systemic interdependence measure $D(\widetilde{u}||\widetilde{\pi}^{\alpha})$ and the systemic CTI $\widetilde{\kappa}(\alpha, u)$, Therefore, a practical advantage of the decomposition (7) is that it efficiently addresses the curse of dimensionality. Moreover, our extensive empirical analysis suggests that conclusions drawn from $\kappa(\alpha, u)$ and $\widetilde{\kappa}(\alpha, u)$ are almost identical in most applications.

The left panel of Figure 2 shows the CTI (6) of a standardized $n = 6$ dimensional multinormal X with $corr(X_i, X_k) = \rho$ for all $i, k = 1, ..., 6$, $i \neq k$. In particular, we observe that for $\rho = 0.9$ the joint exceedances in X are driven by $n - n\kappa + \kappa = 6 - 6 \cdot 0.52 + 0.52 \approx 3$ independent binary "factors". In other words, they carry the same information as approximately 3 marginal distributions of X_i-exceedances. Note the striking feature that the tail interdependence from multinormal samples (with a fixed correlation for all pairs of variables) is constant across the entire range of α. Hence, the interdependence in this case neither increases nor decreases as the tails become more extreme. Moreover, the total and the systemic CTIs are identical for all α implying that the residual CTI is close to zero in this case. The right panel of Figure 2 shows the results when the correlation is the same for three pairs but zero for the remaining pairs ($corr(X_1, X_2) = corr(X_3, X_4) = corr(X_5, X_6) = 0.7$ and zero for all other pairs). In this case, while the patterns of the total and the systemic CTIs are similar for all α, $\kappa(\alpha, u)$ is about three times larger than $\widetilde{\kappa}(\alpha, u)$ confirming that interdependence originates primarily in interactions within subsets of variables.

Figure 2: Coefficient of Tail Interdependence



Notes: The left panel of this figure shows the total ($\kappa(\alpha, u)$) and systemic ($\widetilde{\kappa}(\alpha, u)$) CTIs computed for a sample of $10,000$ obs. from a standardized $n = 6$ dimensional multinormal X with $corr(X_i, X_k) = \rho$ for all $i, k = 1, ..., 6$, $i \neq k$. The right panel shows the total $\kappa(\alpha, u)$ and systemic $\widetilde{\kappa}(\alpha, u)$ CTIs computed for a sample of $10,000$ obs. from a standardized $n = 6$ dimensional multinormal X with $corr(X_1, X_2) = corr(X_3, X_4) = corr(X_5, X_6) = \rho$ and all other correlations equal to zero.

3.3. Goodness-of-Fit and Independence Tests

Recall that $\mathcal{T}(\alpha)$ is a partition of the sample space of the $n-$dimensional random variable $X = (X_1, ..., X_n)$ into 2^n joint tails. We compute the empirical TIS $\widehat{u}(\widehat{f}, \alpha) = \{\widehat{u}_C(\widehat{f}, \alpha)\}_{C \subseteq \mathcal{N}}$ by formula (2), where the difference is that we employ an empirical PDF $\widehat{f}$ rather than the theoretical PDF f. The vector $\widehat{u}(\widehat{f}, \alpha)$ contains, then, the relative frequencies of observations that fall into the JTs $T_C(\alpha) \in \mathcal{T}(\alpha)$. When there is no risk of confusion, we omit the reference to $\widehat{f}$ and α in $\widehat{u}(\widehat{f}, \alpha)$. We use $\widehat{u}_C$ to test whether the observed interdependence structure comes from a hypothesized PDF f, which produces u_C. For this purpose, we compute the KL divergence $D(\widehat{u}||u)$,

$$D(\widehat{u}||u) = \textstyle\sum_{C \subseteq \mathcal{N}} \widehat{u}_C \ln \frac{\widehat{u}_C}{u_C}. \tag{9}$$

If exceedances are mutually independent under f, this procedure boils down to a test of tail independence. In the case of independence, the hypothesized TIS is π^α and (9) is proportional to the CTI (6),

$$D(\widehat{u}||\pi^\alpha; \alpha) = (n-1)H(p^\alpha)\kappa(\alpha, \widehat{u}). \tag{10}$$

Our goodness-of-fit test with the mutual independence test as a special case, is conditional on sufficient statistics estimated from the data (e.g., on the estimates of quantiles in the sample). For the conditional test, the asymptotic distribution of the test statistic $2 \cdot T \cdot D(\widehat{u}||u)$, where T is the sample size, follows the χ^2-distribution with d degrees of freedom (e.g., McCullagh, 1986). For the degrees of freedom, we observe that we have 2^n outcomes (JTs) and $n+1$ restrictions on probabilities or frequencies of these outcomes: these probabilities must sum up to one and, moreover,

$$\textstyle\sum_{C \subseteq \mathcal{N}: i \in C} u_C = \sum_{C \subseteq \mathcal{N}: i \in C} \widehat{u}_C = \alpha, \quad \forall i = 1, ..., n.$$

Therefore, we apply $d = 2^n - n - 1$ degrees of freedom in our goodness-of-fit tests.

Similarly, we can use the systemic CTI to compute the statistic,

$$\widetilde{D}(\widehat{u}||\pi^{\alpha};\alpha) = (n-1)H(p^{a})\widetilde{\kappa}(\alpha,\widehat{u}), \tag{11}$$

for testing the systemic independence. In this case, the statistic $2 \cdot T \cdot \widetilde{D}(\widehat{u}||u;\alpha)$ is distributed approximately as χ^2-variable with d degrees of freedom. As there are $n+1$ outcomes (total number of exceedances) and two restrictions on probabilities or frequencies of these outcomes,

$$\sum_{k=0}^{n} \widetilde{u}_k = 1, \quad and \quad \sum_{k=0}^{n} k\widetilde{u}_k = n\alpha,$$

we apply $d = n - 1$ degrees of freedom in tests based on the systemic CTI.

3.4. Interdependence Symmetry Test

Another interesting question is whether two tail interdependence structures (e.g., lower and upper tails) are symmetric. Specifically, let $\widehat{u}^{+}$ and $\widehat{u}^{-}$ be two empirical (aggregate) TISs with the same cardinality $K \leq 2^n$. Our objective is to test whether $\widehat{u}^{+}$ and $\widehat{u}^{-}$ were generated by a process with an identical tail interdependence structure. In order to test the null $u^{+} = u^{-}$, we apply the Kullback–Leibler test statistic,

$$KL^{\pm} = \sum_{k=1}^{K} T^{+}\widehat{u}_k^{+} \ln \frac{\widehat{u}_k^{+}}{\widehat{u}_k} + \sum_{k=1}^{K} T^{-}\widehat{u}_k^{-} \ln \frac{\widehat{u}_k^{-}}{\widehat{u}_k},$$
$$where, \quad \widehat{u}_k = \frac{(T^{+}\widehat{u}_k^{+} + T^{-}\widehat{u}_k^{-})}{T^{+} + T^{-}},$$

and T^{+} (T^{-}) is the size of the sample from which $\widehat{u}^{+}$ ($\widehat{u}^{-}$) have been computed. The asymptotic distribution of $2 \cdot KL^{\pm}$ follows the χ^2-distribution with $K - 1$ degrees of freedom (e.g., Quine and Robinson, 1985). We refer to this procedure as the interdependence symmetry test.

3.5. *Modeling an empirical TIS*

Modeling multidimensional dependence of random variables is inherently difficult. A standard approach is the multivariate GARCH class of models (see Engle, 2002; Bollerslev, 2009 and the references therein) or copulas (e.g., Chen, 2007; Giacomini et al., 2009; Patton, 2009). Here, we address the simpler task of replicating the observed TIS. Clearly, this approach is only appropriate when the overriding concern is the tail interdependence and the user overlooks other characteristics such as co-moments. Specifically, we construct a PDF that replicates the TIS $\widehat{u}$ estimated from a sample of multidimensional data. First, we estimate from a given sample a multidimensional PDF $\widehat{f}$ with a simple yet flexible parametric form (such as multinormal or multivariate-t). Then, the mixture,

$$m(x) = \sum_{C\subseteq\mathcal{N}} \widehat{u}_C \cdot \widehat{f}(x|T_C(\alpha)), \quad x \in \Omega, \tag{12}$$

assigns the desired probability mass $\widehat{u}_C$ to each JT $T_C(\alpha)$. Intuitively, the mixture (12) selects first the JT $T_C(\alpha)$ with probability $\widehat{u}_C$ and then, draws an observation $x \in T_C(\alpha)$ from the conditional PDF$\widehat{f}(x|T_C(\alpha))$. Although (12) will have, in general, different co-moments and marginals than those estimated from the sample, the fact that it draws (after selecting the tail) each observation from $\widehat{f}(x|.)$ suggests that the synthetic data will be close to the sample.[4]

4. Empirical Illustrations

There are many interesting issues on which the tail interdependence framework can shed light. As an example and illustration of the ideas introduced above, we now present an array of short empirical studies. In all statistical tests that follow, we say that the null is strongly rejected (or rejected with a high significance) if the p-value of the relevant test

[4]We compared the performance of this technique relative to multivariate GARCH and copulas and find that it performs significantly better than them in modeling tail dependence. To preserve space, we do not present the results. They are available upon request.

does not exceed 0.01. A simple rejection occurs with a p-value below 0.1. If we (do not) reject the null for all tail probabilities α, this implies that we tested the null for $\alpha \in \{0.1, 0.15, ..., 0.85, 0.9\}$.

4.1. Daily Returns in G7 Equity Markets

This subsection illustrates the tail interdependence framework in the context of the daily returns of the equity indices for G7 countries (Italy, Canada, France, Germany, Japan, UK and US). We compute the daily returns between 2 January 1973 and 26 July 2013 ($N = 10,584$ synchronized observations obtained from Datastream). While a lower frequency would account better for different opening times across G7 countries and for microstructure effects, it would result in a dramatic loss of observations.[5] Summary statistics are reported in Table 1. In particular, we observe that the returns are highly leptokurtic and negatively skewed.

Table 1: Summary Statistics for G7 Equity Index Daily Returns

	Italy	Canada	France	Germany	Japan	UK	USA
Mean	0.000	0.000	0.000	0.000	0.000	0.000	0.000
SD	1.357	0.984	1.187	1.069	1.129	1.086	1.09
Skewness	-0.232	-0.824	-0.251	0.053	-0.404	-0.273	-1.045
Kurtosis	7.9	16.56	8.459	20.22	14.94	11.7	28.84

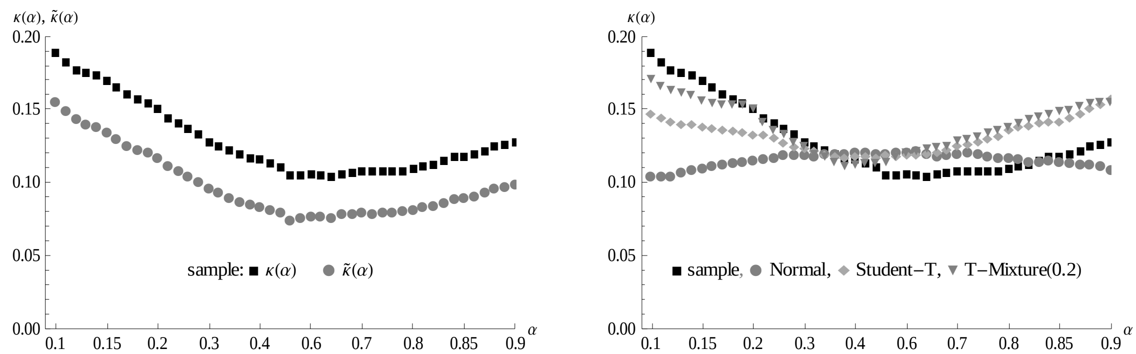
Notes: The table reports the mean, standard deviation, skewness, kurtosis for the synchronized daily log returns for G7 equity indices (Italy, Canada, France, Germany, Japan, UK and US) for the sample period from 2 January 1973 to 26 July 2013. The sample was obtained from Datastream and contains 10,584 synchronized daily observations.

[5]However, we conducted the analysis accounting for the different opening times of the G7 equity indices. To preserve space, we do not present the results. They are available upon request.

4.1.1. *Goodness-of-fit test*

Multivariate normal, or more generally, multivariate eliptical distributions are essential assumptions in many financial applications such as portfolio allocations and risk management. However, the empirical evidence in support of such assumptions is mixed and the tail interdependence framework can easily be applied to examine whether such assumptions are appropriate for the application at hand. The left panel of Figure 3 shows the total ($\kappa(\alpha)$) and systemic ($\widetilde{\kappa}(\alpha)$) CTIs computed in the lower and the upper JTs for the empirical distribution. The results are shown for α ranging between 0.1 and 0.9. The values $\alpha \in [0.1, 0.5]$ correspond to the lower joint tails in $\mathcal{T}(\alpha)$ and the values $\alpha \in [0.5, 0.9]$ to the upper joint tails in $\mathcal{T}(1-\alpha)$. For example, the CTI for the upper JTs in $\mathcal{T}(0.4)$ is computed for $\alpha = 0.6$. There is a strong asymmetry between the lower and the upper tails in the sample. In particular, the interdependence in the lower tails is higher relative to the upper tails for both CTIs. This is confirmed by our interdependence symmetry test that strongly rejects the null of the same interdependence structure, at both the total and systemic level, for $\alpha \leq 0.35$ but not for higher α. Therefore, negative extreme returns are indeed more interdependent than their positive counterparts. Moreover, the total CTI is clearly larger than the systemic CTI, which indicates a pronounced tail interdependence among groups of countries.

Figure 3: Tail Interdependence for G7 Equity Index returns



Notes: The left panel of the figure shows the total ($\kappa(\alpha, u)$) and systemic ($\widetilde{\kappa}(\alpha, u)$) CTIs computed in the lower ($T^{-}(\alpha)$) and the upper ($T^{+}(\alpha)$) joint tails for the empirical distribution. The right panel shows the total CTI ($\kappa(\alpha, u)$) computed in the lower ($T^-(\alpha)$) and the upper ($T^+(\alpha)$) joint tails for the empirical distribution, the simulated multinormal, the simulated multivariate-t and the t-mixture (12) with $\alpha = 0.2$ and parameters estimated from the sample. The results are shown for α ranging between 0.1 and 0.9. The values $\alpha \in [0.1, 0.5]$ correspond to the lower joint tails in $T^-(\alpha)$ and the values $\alpha \in [0.5, 0.9]$ to the upper joint tails in $T^+(1 - \alpha)$.

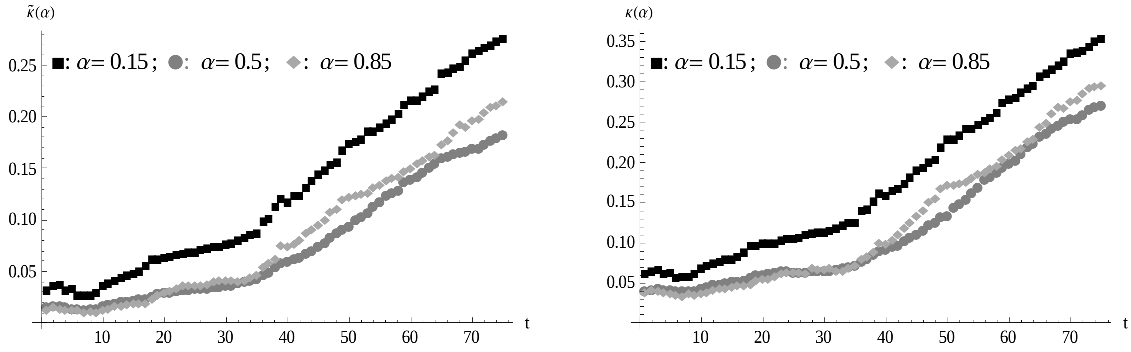
The right panel of Figure 3 shows the total CTI $\kappa(\alpha)$ for the the empirical distribution, the simulated multinormal and the simulated multivariate-t with parameters estimated from the sample. The panel depicts also the total CTI generated by the mixture (12) where the estimated multivariate-t plays the role of the parametric PDF $\widehat{f}(x)$ and the empirical TIS $\widehat{u}$ is computed from the data for $\alpha = 0.2$ (lower tails). The figure shows that the empirical interdependence exceeds the interdependence generated by the multinormal and by the multivariate-t in the lower tails (for $\alpha < 0.35$) while in the upper tails the empirical CTI is below the multivariate-t and, for $1 - \alpha < 0.82$, below the multinormal. Tests of mutual independence and of compatibility of the observed interdependence structure with the multinormal and multivariate-t are strongly rejected for all α. Identical inferences are made from the systemic CTI. We observe the significantly improved fit of the mixture for $\alpha \in (0.15, 0.5)$. The goodness-of-fit test does not reject the null that the sample has been generated from this mixture for α's in this interval. Therefore, the mixture successfully replicates the TIS of the sample locally.

4.1.2. Integration of G7 equity markets

Christoffersen et al. (2012) find that the interdependence among the equity market returns in G7 countries has increased substantially over the past. In this subsection, we address questions pertaining to market integration by examining the evolution of their tail interdependence over time. We compute the CTI (6) in the windows $[t - 2500, t]$ for $t = 2501, 2601, ..., T$ and $\alpha \in \{0.15, 0.5, 0.85\}$. The right panel of Figure 4 shows that the tail interdependence among the G7 countries has increased significantly over time.

Interestingly, the figure indicates that while the dependence of the extreme positive returns ($\alpha = 0.85$) has considerably increased, it remains consistently below the dependence of the extreme negative returns ($\alpha = 0.15$). Moreover, it appears that the gap between the two CTIs has increased somewhat suggesting the asymmetry has got stronger. This is further confirmed by the systemic interdependence $\widetilde{\kappa}(\alpha)$ for $\alpha = 0.15$ which has got even stronger over time relative to the dependence for $\alpha = 0.85$ as shown in the left panel.

Figure 4: Evolution of the CTI in G7 Returns over Time



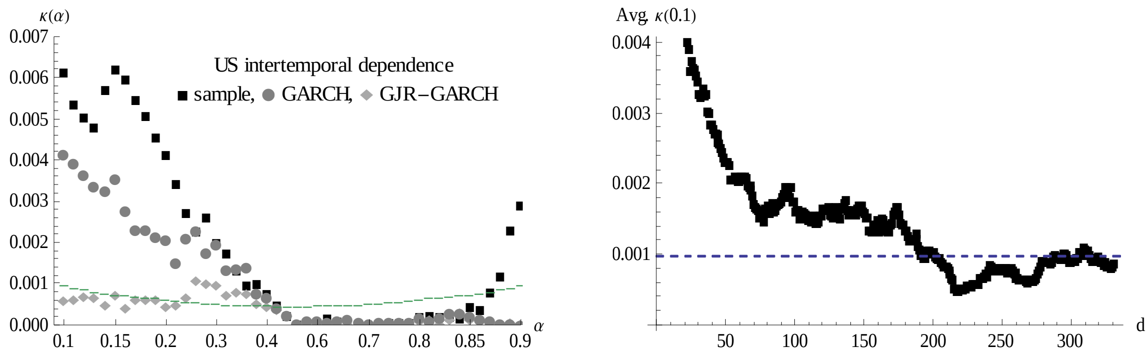
Notes: The left and right panels of this figure show the evolution of the systemic ($\widetilde{\kappa}(\alpha, u)$) and total ($\kappa(\alpha, u)$) CTIs respectively in G7 equity index returns from 2 January 1973 to 26 July 2013 in the windows $[t - 2500, t]$ for $t = 2501, 2601, ..., N$ and $\alpha \in \{0.15, 0.5, .85\}$.

4.1.3. Persistence of intertemporal dependence

There is a large literature that goes back to Mandelbrot (1963) documenting persistence in volatility (see Bollerslev, 2009). It is therefore natural to enquire whether intertemporal dependence displays any features of persistence. In the simple bivariate setting, we transform $T = 10,584$ unidimensional returns $\{r_t\}_{t=1}^{T}$ of the US equity index S&P500 into $T - d$ two-dimensional observations $\{r_{t-d}, r_t\}_{t=d+1}^{T}$ and compute the CTI for the latter series. The results are presented in Figure 5, where the lines mark the 1% critical values for the test statistic (10) in the test of intertemporal independence. All values of $\kappa(\alpha)$ above the line lead to the strong rejection of the null of independence. For $d = 1$ in the left panel, we note a stark asymmetry between the left and the right tail, which indicates that the violation of the intertemporal independence is more likely for (extreme) negative returns. Hence, the intertemporal independence in the US market is rejected for $\alpha < 0.4$ and $\alpha \geq 0.87$. This finding is reminiscent of the well-documented volatility clustering as it results from the tendency of extreme (negative) returns to be followed by such returns in the next period. It may be that the failure to reject intertemporal dependence is due to GARCH effects but once these effects are taken into account, the returns are intertemporally independent. To address this concern, we estimate the CTI for the GARCH(1,1)- and GJR-GARCH(1,1)-standardized returns. Although GARCH effects account for a large amount of intertemporal dependence, the latter is not completely eliminated for the GARCH(1,1) standardization in the negative tails. The intertemporal dependence for GARCH(1,1)- and GJR-GARCH(1,1)-standardized returns is even more pronounced and strongly significant for the other G7 indices.

The right panel in Figure 5 reports the CTI at level $\alpha = 0.1$ as a function of the lag d. Specifically, we compute the CTI for each time series $\{r_{t-\delta}, r_t\}_{t=\delta+1}^{T}$, where $\delta \in \{d, ..., 20 + d\}$, and report the average of these CTIs for each $d = 1, ..., 400$. As the figure indicates, if we applied our test to these averages, it would robustly reject the null of intertemporal independence for roughly $d \leq 180$ days. Thus, the return generating process appears to have a long memory for returns in the lowest decile.

Figure 5: Intertemporal dependence in S&P 500



Notes: This figure shows the persistence of intertemporal dependence in S&P 500 index returns. We transform the $N = 10,584$ unidimensional daily returns $\{r_t\}_{t=1}^{N}$ of the S&P 500 index into $N - d$ two-dimensional observations $\{r_{t-d}, r_t\}_{t=d+1}^{N}$ and compute the total CTI ($\kappa(\alpha)$) for the latter series. The lines mark the 1% critical values for the test statistic (10) in the test of intertemporal independence. Left panel ($d = 1$): The CTI as a function of the tail α. The intertemporal dependence is computed for returns as well as the returns standardized by GARCH(1,1) and GJR-GARCH(1,1) models. Right panel ($\alpha = 0.1$): The CTI as a function of the lag d. We compute a CTI for each time series $\{r_{t-\delta}, r_t\}_{t=\delta+1}^{N}$, where $\delta \in \{d, ..., 20 + d\}$, and report the average of these CTIs for each $d = 1, ..., 400$. Our test robustly rejects the null of intertemporal independence for roughly $d \leq 180$ days.

4.2. *High Frequency Returns in European Equity Markets*

In this section, we illustrate the tail interdependence framework with a dataset of high frequency returns on six European equity markets covering UK, Switzerland, Italy, Germany, France and Spain. The sample contains returns at 5 minute frequency and spans the period from 2 January 2004, 8:00 AM through 15 May 2006, 12:10 ($65,532$ synchronized observations obtained from the Bank of America). Summary statistics are reported in Table 2. For all six indices, 5-minute log returns are zero, negatively skewed and leptokurtic.

Table 2: Summary Statistics for 6 European Equity Index High Frequency Returns

	UK	Switzerland	Italy	Germany	France	Spain
Mean	0.000	0.000	0.000	0.000	0.000	0.000
SD	0.055	0.065	0.063	0.085	0.073	0.065
Skewness	-0.113	-0.504	-0.873	-0.737	-0.585	-2.199
Kurtosis	50.681	77.527	73.003	74.706	74.937	113.194

Notes: The table reports the mean, standard deviation, skewness, kurtosis for the synchronized 5-minute log returns for 6 European equity indices (UK, Switzerland, Italy, Germany, France and Spain) for the sample period from 2 January 2004 (08:00) to 15 May 2006 (12:10). The sample was obtained from Bank of America and contains 65,532 synchronized 5-minute observations.

4.2.1. *Interdependence dynamics across return measurement frequency*

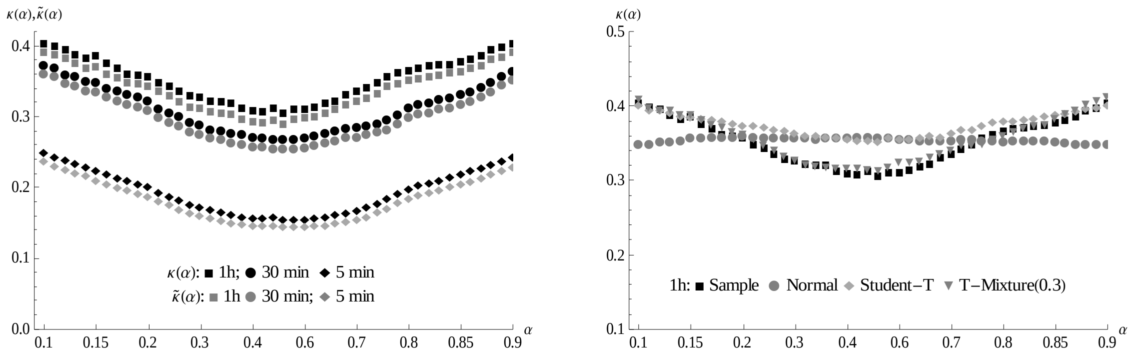
First, we illustrate how the tail interdependence framework could be employed to examine dependence dynamics across frequencies. Figure 6 shows the CTIs computed from the returns at different frequencies and from the simulated multinormal, multivariate-t and the t-mixture (12) with $\alpha = 0.3$. The parameters of all three distributions are estimated from the sample. The results are shown for α ranging between 0.1 and 0.9 where, as before, the values $\alpha \in [0.1, 0.5]$ correspond to the lower joint tails and the values $\alpha \in [0.5, 0.9]$ to the upper joint tails.

In the left panel, we observe that the interdependence decreases in frequency. This effect is particularly pronounced when the frequency increases from 30 to 5 minutes. Our symmetry test rejects the null of the same interdependence structure for 30- and 60-minute returns at 10% confidence level, while the same null for 5- and 30-minute returns is rejected with 1% confidence. We interpret this finding as a manifestation of the Epps effect (Epps, 1979) that reflects the information aggregation process. At high frequencies, idiosynchratic or market-specific news drive returns and there is a time lag before the information spreads to related markets. As frequency decreases (i.e. the time available to gather and process information increases), then returns are affected not only by their market-specific news but also by news in other markets thereby increasing their interdependence.

In contrast to the daily returns of the G7 countries, we cannot reject the null of symmetry of the lower and upper tails for the frequencies 5, 30 and 60 minutes and for all α. Thus, whereas low-frequency dependence is rotated J- or L-shaped, high frequency dependence seems to be U-shaped. Further, the total and systemic CTIs have identical patterns for all frequencies and all α but the latter is very marginally lower. Therefore, it appears that the residual interdependence is insignificant and it does not vary with α for high frequency returns.

Turning to the right panel, there is apparently a good fit of the multivariate-t in the extreme tails of the data. Indeed, for $\alpha \leq 0.15$ and $\alpha \geq 0.85$ we cannot reject the null that the corresponding tails have been generated from this distribution. The good approximation in both tails comes as a result of the symmetry of dependence in the tails for high frequency returns. Thus, a user interested only in the tails such as a regulator or creditor could overlook the failure of the multivariate-t distribution to approximate the central part of the distribution and exploit the good fit in the tails. However, a user interested in modeling the entire distribution may use the mixture (12) where the estimated multivariate-t plays the role of the parametric PDF $\widehat{f}(x)$ and the empirical TIS $\widehat{u}$ is computed from the data for $\alpha = 0.3$. The mixture approximates the data well for all α. As high frequency return interdependence is symmetric, good fit around $\alpha = 0.3$ implies similarly good fit around $\alpha = 0.7$, thus leading to a good approximation overall.

Figure 6: CTI for Different Frequencies and Parametric Distributions



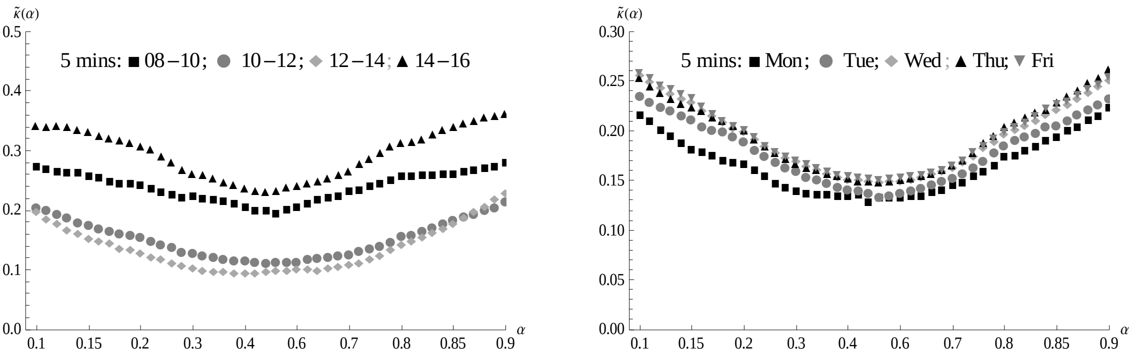
Notes: The left panel of this figure shows the total ($\kappa(\alpha)$) and systemic ($\widetilde{\kappa}(\alpha)$) CTIs for returns at different frequencies. The right panel shows the total CTI ($\kappa(\alpha)$) computed from the sample, the simulated multinormal, multivariate-t, and the mixture (12) with $\alpha = 0.3$ and parameters estimated from the sample. The results are shown for α ranging between 0.1 and 0.9 where the values $\alpha \in [0.1, 0.5]$ correspond to the lower joint tails and the values $\alpha \in [0.5, 0.9]$ to the upper joint tails.

4.2.2. Seasonality in interdependence

Andersen and Bollerslev (1998) find a strong seasonality effect in volatility and therefore it is natural to ask whether dependence is stronger during different times of the day or week. In Figure 7, we investigate the impact of the daytime and of the weekday employing the systemic CTI. The left panel suggests that the interdependence is lowest between 10:00 and 14:00. Before 10:00 and after 14:00 it increases significantly for all α. A possible explanation of this phenomenon could be related to the impact of Asian and US markets on European markets. The latter start each trading day similarly influenced by the shared information revealed in the Asian markets and, hence, display a relatively high level of interdependence. Gradually, idiosyncratic shocks arrive during the day pulling European markets apart resulting in a lower interdependence. In the afternoon, the six European markets react similarly to the shared information revealed by the opening of the focal US markets, which again leads to a higher interdependence.

The right panel, on the other hand, suggests that the interdependence increases during the week. A possible explanation could be related to the dissipation of information. Since the interdependence of the six markets is the inverse of the information revealed in these markets (cf. 5), we observe that the latter decreases as the week progresses. At the beginning of each week, a relatively large amount of idiosyncratic news arrives which is progressively (and partially) incorporated into the market prices resulting in more similarities in market movements i.e., in less joint uncertainty or, equivalently, in higher interdependence. Moreover, systemic CTIs in both time-of-the-day and day-of-the-week cases have identical U-shaped patterns to those of the total CTIs and are only marginally lower. This implies that the size of the residual interdependence is quite small and flat for all α. Therefore, seasonality affects only the systemic interdependence.

Figure 7: CTI across Different Trading Hours and Days



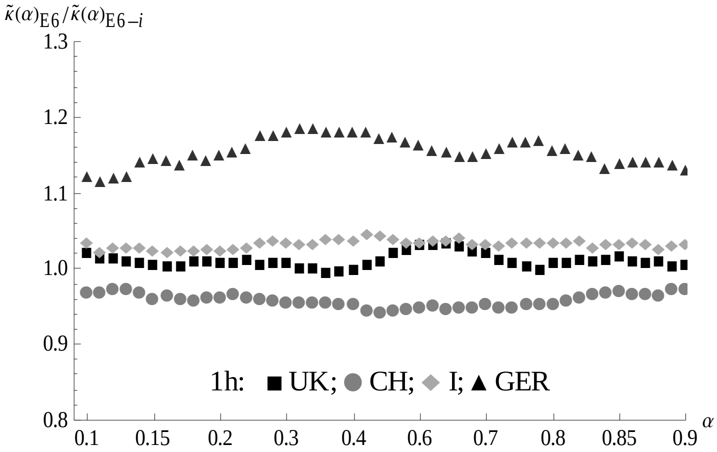
Notes: This figure shows the systemic CTI ($\widetilde{\kappa}(\alpha)$) of 5-minute returns across different trading hours (left panel) and trading days (right panel). The results are shown for α ranging between 0.1 and 0.9 where the values $\alpha \in [0.1, 0.5]$ correspond to the lower joint tails and the values $\alpha \in [0.5, 0.9]$ to the upper joint tails.

4.2.3. Contribution to interdependence in European equity markets

It is important for the study of spillovers and contagion to isolate the impact or contribution of an individual institution or country to the overall interdependence of the system (see, Bank of England, 2013; Diebold and Yilmaz, 2014). The interdependence contribution may be computed by different measures such as the Shapley value, an idea which we discuss further in the Appendix. However, here we simply compute the interdependence contribution of a variable as the ratio of the CTIs that include and exclude that particular variable. This measure is intuitively appealing and computationally efficient.

Figure 8 depicts the systemic interdependence contribution for UK, Switzerland, Italy and Germany for 5- and 60-minute returns computed as $\widetilde{\kappa}_{E6}/\widetilde{\kappa}_{E6\setminus i}$ where $i \in \{UK, CH, I, GER\}$. We observe that Switzerland (Germany) has the lowest (highest) contribution to interdependence. This would suggest that the Swiss equity index may be an effective diversification asset in European equity portfolios.We can also apply our interdependence symmetry test to assess the significance of the exclusion of particular countries. For example, the symmetry tests strongly reject for the 5-minute returns the null that the CTI after excluding Germany is the same as the CTI after excluding Switzerland. For 60 minute returns the null is also rejected (except for $\alpha = 0.1$) with a lower significance. Similar results are obtained when testing for the exclusion of Germany and UK, respectively. Finally, the contributions to the total CTIs are almost identical in both shape and size suggesting that the contributions to the residual interdependence are insignificant and flat for all α at the high frequency (not shown but available upon request).

Figure 8: Interdependence Contribution



Notes: The figure shows the percentage contributions to the systemic interdependence computed for UK, Switzerland, Italy and Germany computed as $\tilde{\kappa}_{E6}/\tilde{\kappa}_{E6\setminus i}$ where $i \in \{UK, CH, I, GER\}$ at the one-hour frequency.

4.3. Stock and Factor Interdependence

In this section, we illustrate the tail interdependence framework with a dataset of daily frequency returns on 30 constituent stocks of Dow Jones Industrial Average (DJ30) equity index and relate their returns to the Fama-French-Carhart (FFC) factors. The data spans the period 1 January 1990 - 21 November 2012 (5770 synchronized observations obtained from Datastream, while the FFC factors for the same period were obtained from Keneth French's website. Summary statistics are reported in Table 3. For all four factors (and the DJ index constituents, which are not shown) daily log returns are zero, negatively skewed and leptokurtic.

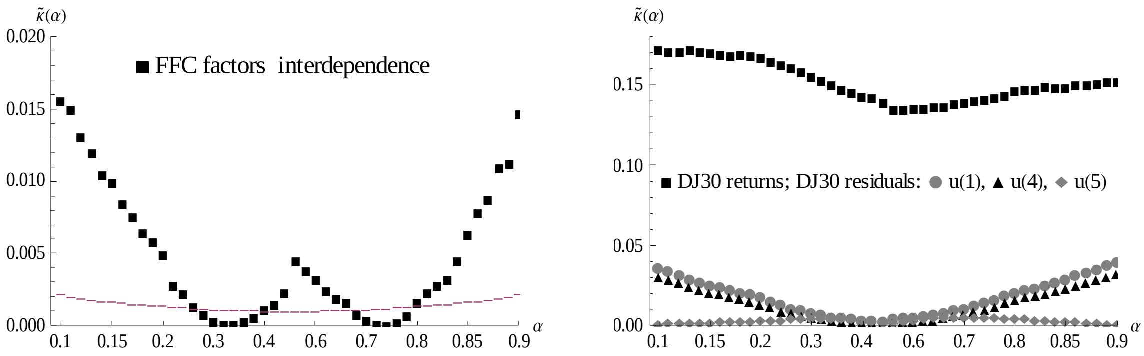
Table 3: Summary Statistics for the Fama-French-Carhart Factor Returns

	RP_m	SMB	HML	MOM
Mean	0.000	0.000	0.000	0.000
SD	0.012	0.006	0.006	0.009
Skewness	-0.105	-0.268	0.108	-0.956
Kurtosis	10.99	7.163	9.337	14.69

Notes: The table reports the mean, standard deviation, skewness, kurtosis for the Fama-French-Carhart (Market Risk Premim, Small minus Big, High minus Low, Momentum) factor returns. The data spans the period from 1 January 1990 through 21 November 2012 (5770 observations obtained from Keneth French's website.

Due to the curse of dimensionality, total CTI is unreliable because of the high number of JTs containing no observations. Thus, in the ensuing discussion we focus on the systemic CTI which is robust to the curse of dimensionality. The right panel of Figure 9 shows that the DJ30 returns are highly interdependent and asymmetric.While the FFC factors account for a high degree of this interdependence in the central part of the distribution, the factors are unable to account for the strong dependence of the DJ30 returns in the tails of the distribution. Moreover, comparing the interdependence of the residuals $u(1)$ of a regression of the DJ30 index constituent returns on the first FFC factor returns (market risk premium) with the interdependence of the residuals $u(4)$ of the same dependent variables on all four FFC factor returns, it appears that most of the interdependence is accounted for by the market risk premium. This comparison makes it clear that the remaining three FFC factors (SMB, HML and MOM) account for very little of the interdependence of the residuals. The inability of the FFC factors to account for the inderdependence of the DJ30 returns in the tails is a direct manifestation of the interdependence of the factors themselves. The systemic CTI depicted in the left panel of Figure 9 reveals that the FFC factors are highly interdependent for $\alpha < 0.2$ and $\alpha > 0.8$ but not for $\alpha \in [0.2, 0.8]$.

Figure 9: Interdependence of Fama-French-Carhart factors and DJ30 index constituent stocks



Notes: The left panel of this figure shows the systemic interdependence for the Fama-French-Carhart (FFC) factors. The line marks the 1% critical values for the test statistic (10) in the test of intertemporal independence. The right panel shows the systemic interdependence for the DJ30 index constituent returns as well as for the residuals $u(1)$ of a regression of the DJ30 returns on the first FFC factor (the market risk premium), the residuals $u(4)$ of a regression of the DJ30 returns on all FFC factors and the residuals $u(5)$ of a regression of the DJ30 returns on all FFC factors plus an additional multiplicative factor, the market dispersion F_d.

As a potential additional factor that accounts for the strong interdependence of the residuals in the tails, we explore *market dispersion* F_d. We estimate F_d by computing the standard deviation of the DJ30 constituents for every day in the sample. Then, we compute the residuals $u(5)$ by normalizing $u(4)$ with these estimates,

$$u_i(5) = u_i(4)/F_d, \quad i = 1, .., 30.$$

As the systemic CTI of $u(5)$ shows in the right panel of Figure 9, F_d accounts for a large part of the interdependence in the JTs for $\alpha \leq 0.3$ and $\alpha \geq 0.7$. Although the residuals $u_i(5)$ are not independent, their interdependence is overwhelmingly reduced.

5. Conclusion

In this paper, we present a new and flexible framework focused on the interdependence of extreme events. This framework aims to address several issues that have recently attracted significant attention such as the testing of the independence of extreme events, the symmetry of (extreme) positive and negative outcomes and the modeling of the dynamics of the tail interdependence. In particular, we develop a new dependence measure, which captures the magnitude of the departure from independence and propose a technique to generate synthetic data that exactly match the tail interdependence structure of a particular dataset. The framework also allows for computing the contributions of individual variables to tail interdependence and can be adapted to examine other extreme event-related questions.

A complementary consideration to our non-parameteric approach is the modelling of the observed dependence structure in the data. The literature addresses this issue mainly via VAR-(multivariate) GARCH models with the innovations following a particular distribution such as multivariate normal or t and, more recently, via copulas (see Bollerslev, 2009; Chen, 2007; Patton, 2009 and the references therein). However, multivariate models suffer from model misspecification, thus necessitating goodness-of-fit testing. A number of tests exist for this purpose such as Cramer-von Mises, Anderson-Darling and Kolmogorov-Smirnov tests which are based on comparing the cumulative distribution function (CDF) of the hypothesized model to the empirical one while independence is typically tested with Pearson's chi-square test. We discuss the related issue of parameter estimation uncertainty and its relevance for our study in the Appendix.

In the empirical part, we illustrate the tail interdependence framework with an array of applications and confirm some known stylized facts and uncover a few new and intriguing features of multidimensional extreme events. Our financial data shows, in particular, that the tail interdependence increases for more extreme events and is stronger in the lower than in the upper tails (except at high frequencies). We think that these are important findings with vital practical implications (e.g., for systemic risk monitoring and hedging). The CTI captures these phenomena in a clear and precise way. It would be interesting to investigate the potential of the CTI, e.g. in portfolio construction, hedging and derivativebased trading strategies. We intend to pursue these avenues in future research.

6. References

Andersen, T.G. and Bollerslev, T. 1998 Deutsche Mark–Dollar Volatility: Intraday Activity Patterns, Macroeconomic Announcements, and Longer Run Dependencies, Journal of Finance, 53, 219-265.

Ang, A. and Bekaert, G. 2002, International Asset Allocation With Regime Shifts, Review of Financial Studies, 15, 1137-1187.

Ang, A. and Chen, J. 2002, Asymmetric correlations of equity portfolios, Journal of Financial Economics, 63, 443-94.

Backus, D. Chernov, M. and Zin, S. 2014. Sources of entropy in representative agent models, Journal of Finance, 69, 51-99.

Bae, K-H., Karolyi, G.A. and Stulz, R.M., 2003, Review of Financial Studies, 16, 717-763.

Bank of England, 2013, A framework for stress testing the UK banking system, Discussion Paper 1013.

Bartram, S.M., Taylor, S.J. and Wang, Y.-H. 2006, The Euro and European Financial Market Dependence, Journal of Banking and Finance, 31, 1461-1481.

Bollerslev, T. 2009, Glossary to ARCH (GARCH), Research paper 49, CRATES.

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

Chen, Y.-T., 2007, Moment-Based Copula Tests for Financial Returns, Journal of Business & Economic Statistics, 25(4), 377-397.

Cherubini, U. and Luciano, E. 2002, Bivariate option pricing with copulas, Applied Mathematical Finance 9, 69–85.

Chicharro, D., Ledberg, A., 2012, Framework to study dynamic dependencies in networks of interacting processes, Physical Review E, 86(4).

Christoffersen, P.F, Errunza, V., Jacobs, K., Langlois, H., 2012, Is the Potential for International Diversification Disappearing? A Dynamic Copula Approach, Review of Financial Studies 25, 3711-3751.

Colangelo, A., Scarsini, M. and Shaked, M., 2005, Some notions of multivariate positive dependenceInsurance: Mathematics and insurance, 37, 13 - 26.

Coles, S.G.,Heffernan, J.E. and Tawn, J.A. 1999, Dependence Measures for Extreme Value Analyses, Extremes, 2, 339-365.

Cover, T.M., Thomas, J.A., 2006, Information Theory, 2nd ed., Wiley & Sons.

Diebold, F.X. and Yilmaz, K. 2014, On the Network Topology of Variance Decompositions: Measuring the Connectedness of Financial Firms, Journal of Econometrics, 182, 119–134.

Embrechts, P., McNeil, A., and Straumann, D., 2002, Correlation and dependence properties in risk management: properties and pitfalls, in M. Dempster, ed., Risk Management: Value at Risk and Beyond, Cambridge University Press.

Engle R.F., 2002, Dynamic conditional correlation - a simple class of multivariate GARCH models, Journal of Business and Economic Statistics, 20, 339-350.

Epps, T.W., 1979, Comovements in Stock Prices in the Very Short Run, Journal of the American Statistical Association, 74, 291–298

Forbes, K. and Rigobon, R. 2002, No contagion, only interdependence: measuring stock market co-movements, Journal of Finance, 57, 2223-61.

Giacomini, E., Haerdle, W., Spokoiny, V., 2009, Inhomogeneous Dependence Modeling with Time-Varying Copulae, Journal of Business & Economic Statistics, 27:2, 224-234.

Hartmann, P., Straetmans, S.T. and de Vries, C.G. 2000, Asset Market Linkages in Crisis Periods., paper presented at the Centre for Financial Studies Conference .Liquidity Risk: Rethinking Risk Management., 30 June-1 July 2000, Frankfurt.

Heffernan, J.E. 2001, A Directory of Coefficients of Tail Dependence, Extremes, 3, 279-290.

Hong, Y., Tu, J., and Zhou, G., 2007, Asymmetries in Stock Returns: Statistical Tests and Economic Evaluation, Review of Financial Studies, 20, 1547-1581.

Hull, J.C. and White, A.D. 2006, Valuing Credit Derivatives Using an Implied Copula approach, Journal of Derivatives, Winter, 8. 28.

Jansen, D.W., de Vries, C.G. 1991, On the Frequency of Large Stock Returns: Putting Booms and Busts into Perspective, Review of Economics and Statistics, 73, 18-24.

Joe, H., 1989, Relative Entropy Measures of Multivariate Dependence, Journal of the American Statistical Association, 84(405), 157-164.

Joe, H., 1997, Multivariate Models and Dependence Concepts, Chapman & Hall, London.

Kalemanova, A., Schmid, B. and Werner, R. 2007, The normal inverse gaussian distribution for Synthetic CDO Pricing, Journal of Derivatives 14(3), 80–93.

Ledford, A.W. and Tawn, J.A. 1996, Statistics for Near Independence in Multivariate Extreme Values, Biometrica, 83, 169-187.

Li, H., 2009, Orthant tail dependence of multivariate extreme value distributions, Journal of Multivariate Analysis 100: 243—256.

Longin, F. and Solnik, B. 2001 Extreme Correlation of International Equity Markets, Journal of Finance, 562, 649-676.

Mandelbrot, B., 1963, The variation of certain speculative prices, Journal of Business, 36, 394-419.

McCullagh, P., 1986, The Conditional Distribution of Goodness-of-Fit Statistics for Discrete Data, Journal of American Statistical Association, 81, 104-107.

Patton, A.J., 2009, Are Market Neutral Hedge Funds Really Market Neutral?, Review of Financial Studies, 227, 2495-2530.

Quine, M. P. and Robinson, J. 1985, Efficiencies of chi-square and likelihood ratio goodness-of-fit tests. Ann. Statist. 13, 727–742

Schneidman, E., Still, S., Berry, M. and Bialek, W. 2003, Network Information and Connected Correlations, Physical Review Letters 91(23).

Schneidman, E., Berry, M., Segev, R. and Bialek, W. 2006, Weak pairwise correlations imply strongly correlated network states in a neural population, Nature 440, 1007-1012.

Schreiber, T., 2000, Measuring Information Transfer, Phys. Rev. Lett., 85:461.

Shannon, C. E. 1948, A Mathematical Theory of Communication, Bell System Technical Journal, 27, 379-423.

Shapley, L.S., 1953. A Value for n-person Games. In Contributions to the Theory of Games, vol. II, H. Kuhn, A. Tucker (eds), Princeton University Press.

Su, L. and Spindler, M. 2013 Nonparametric Testing for Asymmetric Information, Journal of Business & Economic Statistics, 31, 208-225.

Van Nieuwerburgh, S., and Veldkamp, L. 2010, Information acquisition and portfolio underdiversification, Review of Economic Studies 77, 779–805.

Wennekers, T. and Ay, N., 2003, Spatial and temporal stochastic interaction in neuronal assemblies. Theory in Biosciences, 122, 5–18.

Young, H. P., 1985, Monotonic Solutions of Cooperative Games, International Journal of Game Theory 14, 65–72.

Zheng, S., Shi, N., and Zhang, Z. 2012, Generalized Measures of Correlation for Asymmetry, Nonlinearity, and Beyond, Journal of the American Statistical Association, 107, 1239-1252.

7. Appendix: Proofs

In order to prove the decomposition (7), we calculate,

$$\begin{aligned} D(u||\pi^\alpha) - D(\widetilde{u}||\widetilde{\pi}^\alpha) &= \textstyle\sum_{C\subseteq\mathcal{N}} u_C \ln \frac{u_C}{\pi^\alpha_C} - \sum_{k=0}^n \widetilde{u}_k \ln \frac{\widetilde{u}_k}{\widetilde{\pi}^\alpha_k} \\ &= \textstyle\sum_{k=0}^n \sum_{C\subseteq\mathcal{N}:\#C=k} u_C \ln \frac{u_C}{\pi^\alpha_C} - \sum_{k=0}^n \widetilde{u}_k \ln \frac{\widetilde{u}_k}{\widetilde{\pi}^\alpha_k} \\ &= \textstyle\sum_{k=0}^n \widetilde{u}_k \left(\sum_{C\subseteq\mathcal{N}:\#C=k} \left(\frac{u_C}{\widetilde{u}_k} \ln \frac{u_C}{\pi^\alpha_C} \right) - \ln \frac{\widetilde{u}_k}{\widetilde{\pi}^\alpha_k} \right) \\ &= \textstyle\sum_{k=0}^n \widetilde{u}_k \left(\sum_{C\subseteq\mathcal{N}:\#C=k} \left(\frac{u_C}{\widetilde{u}_k} \ln \frac{u_C}{\pi^\alpha_C} - \frac{u_C}{\widetilde{u}_k} \ln \frac{\widetilde{u}_k}{\widetilde{\pi}^\alpha_k} \right) \right), \end{aligned}$$

where the last equality follows from the fact that $\sum_{C\subseteq\mathcal{N}:\#C=k} u_C/\widetilde{u}_k = 1$. We can write now the last expression as,

$$\textstyle\sum_{k=0}^n \widetilde{u}_k \sum_{C\subseteq\mathcal{N}:\#C=k} \left(\frac{u_C}{\widetilde{u}_k} \ln \frac{u_C/\widetilde{u}_k}{\pi^\alpha_C/\widetilde{\pi}^\alpha_k} \right) = \sum_{k=0}^n \widetilde{u}_k D(u^k||\widetilde{\pi}^{\alpha,k}),$$

which completes the proof of (7). Dividing both sides of (7) by $(n-1)H(p^a) > 0$ for $0 < \alpha < 1$ yields the decomposition of the CTI,

$$\kappa(\alpha, u) = \widetilde{\kappa}(\alpha, u) + \textstyle\sum_{k=0}^n \widetilde{u}_k \kappa^k(\alpha, u), \tag{13}$$

We note that $\kappa(\alpha, u) \geq \widetilde{\kappa}(\alpha, u) \geq 0$ follows from the non-negativity of $\widetilde{\kappa}(\alpha, u)$ and $\kappa^k(\alpha, u)$ as the KL divergence and entropy are always non-negative (Cover and Thomas, 2006). Finally, Cover and Thomas (2006) show that $H(p^a) \leq H(u) \leq nH(p^a)$, which implies that

$$\frac{nH(p^a) - H(u)}{(n-1)H(p^a)} = \frac{D(u||\pi^\alpha)}{(n-1)H(p^a)} = \kappa(\alpha, u) \leq 1.$$

8. Appendix: Extensions of the TIS framework

In this section, we present some extensions and generalizations that arise naturally from the tail interdependence framework.

8.1. Directional CTI

The CTI measures the strength of interdependence among the tails of random variables but it does not specify its direction. The latter can be quantified by the expected number of exceedances under the distribution u in excess of the expected number of exceedances under mutual independence, given that at least an exceedance has occurred,

$$\phi(\alpha, u) = \sum_{C \subseteq \mathcal{N}} (\#C) \cdot (\frac{u_C}{1 - u_\emptyset} - \frac{\alpha^{\#C}(1-\alpha)^{n-\#C}}{1-(1-\alpha)^n}).$$

Generally speaking, when $\phi(\alpha, u) > 0$ (positive interdependence) exceedances tend to occur together and are more likely than under mutual independence while $\phi(\alpha, u) < 0$ (negative interdependence) means that joint exceedances are less likely than under mutual independence. It is important to note that $\phi(\alpha, u)$ itself is not a good measure of tail interdependence as, for example, it can take the value of zero when $\kappa(\alpha, u) > 0$, i.e., when variables are actually tail interdependent. Therefore, we define *directional coefficients of tail interdependence* as,

$$\vec{\kappa}(\alpha, u) = sign(\phi(\alpha, u)) \cdot \kappa(\alpha, u), \quad (14)$$

where $sign(x) = 1$ when $x \geq 0$ and $sign(x) = -1$ when $x < 0$. In the context of financial data, in particular the data in our empirical part, the tail interdependence turns out to be strongly positive.

8.2. Interdependence Contribution Measure

For the TIS u calculated from theoretical or empirical exceedances of n random variables by (2), we can obtain the overall contribution $\varphi_i(u)$ of the variable $i \in \mathcal{N} = \{1, ..., n\}$ to the JT interdependence $I(u)$ as a (weighted) average of marginal contributions of this variable to the interdependence in subsets of other variables. Specifically, we compute $\varphi_i(u)$ by the game-theoretical concept of Shapley value (Shapley, 1953),

$$\varphi_i(u) = \sum_{C \subseteq \mathcal{N} \setminus \{i\}} \frac{(\#C)!(n - \#C - 1)!}{n!} \{I(u^{C \cup i}) - I(u^C)\}, \quad (15)$$

where u^C is the marginal of the TIS u for random variables with indices in the set $C \subseteq \mathcal{N}$. The Shapley value has many desirable properties. For example, Young (1985) shows that Shapley value is the unique efficient and symmetric measure that is a function of marginal contributions only. Here, efficiency requires that all $\varphi_i(u)$ sum up to the total interdependence $I(u)$ while symmetry demands that $\varphi_i(u) = \varphi_k(u)$ whenever two variables, indexed by i and k, make the same contribution to $I(u^C)$ for any subset $C \subseteq \mathcal{N} \setminus \{i, k\}$. Moreover, each contribution $\varphi_i(u)$ is non-negative as $I(u^{C \cup i}) \geq I(u^C)$ for each C and i by the properties of the MI (Chicharro and Ledberg, 2012). Estimating the contribution of an asset to the interdependence of a portfolio or a system can reveal the main contributor to interdependence and risk. This is particularly useful in studies of crises and contagion as well as market integration.

8.3. Measuring the direction of information flow

Multi-information (4) in its standard format cannot inform on the direction of information flow. However, a simple modification to the CTI framework can be employed to reveal the dynamics in information flow between markets or institutions. For a stationary Markov process of order t, the probability of observing the process I in state $i_{\eta+1}$ at time $\eta+1$ is independent of states $i_{\eta-t}, i_{\eta-t-1}, \ldots$ Thus,

$$p(i_{\eta+1}|i_\eta, \ldots, i_{\eta-t+1}, i_{\eta-t}, i_{\eta-t-1}, \ldots) = p(i_{\eta+1}|i_\eta, \ldots, i_{\eta-t+1}) = p(i_{\eta+1}|i_\eta^{(t)}).$$

Schreiber (2000) proposes to measure the direction of information between processes I and J by the deviation from the Markov property $p(i_{\eta+1}|i_\eta^{(t)}) = p(i_{\eta+1}|i_\eta^{(t)}, j_\eta^{(k)})$ where k is the order of the stationary Markov process J. When there is no information flow from J to I, the previous k observations of J have no impact on the transition probabilities of I, which can be measured with a modified KL divergence as

$$T_{J \longrightarrow I}(t,k) = \sum p(i_{\eta+1}, i_\eta^{(t)}, j_\eta^{(k)}) \cdot \ln \frac{p(i_{\eta+1}|i_\eta^{(t)}, j_\eta^{(k)})}{p(i_{\eta+1}|i_\eta^{(t)})}, \tag{16}$$

where natural choices for k are $k = t$ or $k = 1$. Therefore, $T_{J \longrightarrow I}$ measures the information flow from process J to I. $T_{I \longrightarrow J}$, the information flow from I to J, can be measured in an analogue way. Note that measure (16) is asymmetric. Hence, by comparing $T_{J \longrightarrow I}$ to $T_{I \longrightarrow J}$ we can infer the dominant direction of the information flow - useful in studies of price discovery and market linkages or in examining how contagion spreads through markets.

8.4. A finer partition of the outcome space

In the discussion above, the TIS is defined for a partition of the outcome space Ω into 2^n regions (i.e., for a bi-partition of the outcome space of each variable X_i). This partition may be particularly relevant for a regulator or a creditor who is interested in the downside vulnerability of the system or debtor company but has little interest in its upside potential. A typical investor, on the other hand, is not just interested in the downside exposure of his portfolio but its upside potential too. In this case, we could partition the outcome space Ω into 3^n regions, such that for each variable X_i the two tail regions capture extreme losses and gains while the central part captures the average, day-to-day performance when little of importance happens. More generally, the partition could be made arbitrarily fine. In particular, for an infinitely fine partition, the MI (4) would take the form of the total correlation for continuous variables,

$$\int_{x \in \Omega} f(x) \ln \frac{f(x)}{f(x_1) \ldots f(x_n)} dx, \quad x = (x_1, \ldots, x_n).$$

For a finite partition, the construction of the CTI and the inference based on it would then simply generalize the bi-partition case.

8.5. Parameter Estimation Uncertainty

The tail interdependence framework is particularly suited to measure and test interdependence by applying it directly to the data. In this case, the issue of parameter estimation uncertainty would not arise. However, the flexibility of the tail interdependence framework means that it can be applied to an estimated model. For example, the focus of the investigation may be such that a researcher must impose a parametric density function e.g. multivariate t-distribution for the purpose of forecasting or hypothesis testing. In this case, the mean, variance and degrees of freedom parameters must be estimated. However, the presence of estimated parameters may complicate test inference. For example, the Kolmogorov test can be difficult to apply in the presence of estimated parameters, particularly for multivariate data with many parameters (see, for example, Bai and Chen, 2008).

Following other scholars (Diebold and Mariano, 1995; Christoffersen, 1998; Diebold et al. 1998, 1999; Clements and Smith, 2000, 2002), when required to estimate parametric densities, we consider them as primitives and ignore the method employed to obtain them. In many situations this may be an acceptable practice. Firstly, many densities are not based on estimated models. For example, the large-scale market risk models at many financial institutions combine estimated parameters, calibrated parameters and ad-hoc modifications that reflect the judgment of management. Another example is the density forecasts of inflation of the Survey of Professional Forecasters (see Diebold et al., 1998). Moreover, previous research suggests that parameter estimation uncertainty is of second-order importance when compared to other sources of inaccuracies such as model misspecification (Chatfield, 1993). Further, Diebold et al. (1998) find that the effects of parameter estimation uncertainty are immaterial in simulation studies geared toward the relatively large sample sizes employed in financial studies such as the present one.

When parameter estimation cannot be ignored, the problem can be approached as follows. Firstly, for time-invariant multidimensional densities, suitable estimators can often be found that lead to pivotal test statistics e.g., the "super-efficient" estimators (see Watson, 1958; Birch, 1964). Secondly, an important class of models comprises a time-varying hypothesised distribution with a well-defined structure on the co-evolution of the variables e.g. VAR and GARCH models. In this case, one way of accounting for parameter estimation uncertainty is to apply the K-transformation (Khmaladze, 1981), which allows for the construction of a distribution-free test statistic. In principle, the K-transformation can be applied in the tail interdependence framework along the lines of the V-test in Bai (2003) and Bai and Chen (2008). Its computation, however, may be cumbersome for non-standard multidimensional densities. Finally, in the case of arbitrary time-varying multidimensional densities parameter estimation is infeasible as only one observation is drawn from the multidimensional density at each date. As such, the only practical solution is to assume that the hypothesised model is correct under the null.

Acknowledgements

We have benefited from discussions with Richard Harris, George Bulkley, Frank Windmeijer, Michael Moore, Nick Taylor, Karin Thorburn, Svein-Arne Persson, Tore Leite, Marcin Kacperczyk, Lauren Cohen, Robin Greenwood, Michael Faulkender, Peter Pope, Fernando Vega Redondo, Raphael Markellos and the seminar and conference participants at University of Bristol, University of East of Anglia, Chinese University of Hong Kong, Cambridge University Isaac Newton Institute, Systemic Risk Centre at LSE, Bank of England, European Central Bank,World Finance Conference (Buenos Aires, 2015) and thank Europlace Institute of Finance (EIF) for research funding. Parts of this paper were written while Evarist Stoja was a Houblon-Norman Fellow at the Bank of England whose hospitality is gratefully acknowledged. The views expressed here are solely our own and do not necessarily reflect those of the Bank of England.

Arnold Polanski
School of Economics, University of East Anglia, Norwich, United Kingdom; email: **a.polanski@uea.ac.uk.**

Evarist Stoja
University of Bristol, School of Economics, Finance and Management, Bristol, Bank of England, London, United Kingdom; email: **e.stoja@bristol.ac.uk**

Postal address	60640 Frankfurt am Main, Germany
Telephone	+49 69 1344 0
Website	**www.esrb.europa.eu**

ISSN	2467-0677 (online)
ISBN	978-92-95081-39-0 (online)
DOI	10.2849/75773 (online)
EU catalogue No	DT-AD-16-012-EN-N (online)