Tail Dependence in Financial Markets: Evidence from Stock, Foreign Exchange, and Commodities Markets

Ann Marie Hibbert, Mbodja Mougoué & Joel Sango*

PRELIMINARY AND INCOMPLETE. PLEASE DO NOT CITE OR CIRCULATE WITHOUT THE PRIOR PERMISSION OF THE AUTHORS

* Ann Marie Hibbert is an Associate Professor of Finance in the College of Business and Economics, West Virginia University, P.O. Box 6205, Morgantown, WV 26505. Tel: (304) 293-2447. amhibbert@mail.wvu.edu

Mbodja Mougoué is an Associate Professor of Finance in the Mike Ilitch School of Business, Wayne State University, 5201 Cass Avenue, Detroit, MI 48202. (313) 577-4538. ad4906@wayne.edu

Joel Sango is a Methodologist and Analyst at the Méthodologiste et Analyste (Statistique Canada), 257 Rue Robert Martial, Gatineau, Québec, J9J 2V1. Tel: (438) 938-2007. joellesango@gmail.com
1. Introduction

The importance of uncommon events in the process of portfolio selection and asset pricing has been widely investigated in recent empirical studies. Examples of such rare events include large swings in asset returns, natural catastrophes such as earthquakes and hurricanes, and asset market crashes. Rare events are known to occur in the tails of returns distributions and, thus, directly influence the size of all moments. Poon, Rockinger, and Tawn (2004) show that extreme value theories may be the only effective way of analyzing these rare events. In this study we first examine the extremal dependence structure of the returns for selected international stock market indices. We then extend our analysis to test for the existence of extreme-value dependence between stock market indices and exchange rate markets for several countries. Finally, we investigate the existence of extreme-value dependence between stock market indices and commodities prices.

The finance literature on stock market extreme events may be broadly grouped into two strands. First there are those papers that try to measure the likelihood of these events and identify predictors. A second set of papers are devoted to modeling the impact of these events on asset returns, and the corresponding implications for issues such as systemic risk and the stability of international financial markets, portfolio diversification, credit risk modeling, option valuation, etc. Our paper lies in this second strand of the literature and builds largely on the work of Poon, Rockinger, and Tawn (2004), who are
the first to develop a generalized framework to identify and measure the joint-tail distribution of risky-asset returns. Most of the early literature assumes asymptotic dependence, i.e. very large (and small) values of each variable tend to occur together (King and Wadhwani, 1990). However, Poon, Rockinger, and Tawn (2004) show that international stock market returns tend to be asymptotically independent. Using data on five major stock markets, i.e. the US, UK, Germany, France and Japan, they demonstrate how ignoring this asymptotic independence can lead to overestimation of portfolio risk. In line with the existing literature (e.g. Longin and Solnik, 2001), Poon et al., 2004 confirm an asymmetry in the asymptotic dependence of stock markets, where left-tail dependence is stronger than right tail dependence, especially in European markets. In line with the earlier findings of Forbes and Rigobon (2002), Poon et al., 2004 show that much of this dependence is due to correlation in the conditional volatilities, which increases during bear markets.

Jondeau and Rockinger (2006) use a Copula-GARCH model to measure the dependency between the US, UK, German and French stock markets and show that the conditional dependency depends on past realizations for European market pairs only. Specifically, in contrast to the findings of Longin and Solnik (2001), they show that the correlation between pairs of European markets increase significantly following both market crashes and boom. Similar to Poon, Rockinger, and Tawn (2004), they attribute much of the increased correlation to an increase in the conditional volatilities. In a further test of the asymmetry of the asymptotic dependence between international stock markets, Okimoto (2008) combine copula theory with a Markov switching model and confirm the existence of two separate regimes and a variety of asymmetric dependence between pairs
of the G7 countries.\textsuperscript{1} For example, in the bear regime, they find asymmetry in the lower tail dependence structure between the US and UK, but they find no asymmetry in the US-Japan dependence. Unlike these studies that focus mainly on developed markets, Christoffersen et al. (2012) also find significant evidence of asymmetry in the tail dependence among emerging market economies.

Most of the papers that investigate the joint tail distribution of stock returns focus on equity markets, but some researches have investigated the linkage between other asset markets. For example, Hartmann, Straetmans, and Vries (2004) investigate the linkage between stock markets and government bond markets, during crisis periods. They find that among the G-5 countries, increased interdependence between the stock and bond markets, or stock-bond contagion, occurs just about as often as a stock market crash causing a boom in bond markets, the phenomenon referred to as a \textit{flight to quality}.

Ning (2010) was one of the first to use copulas to study tail dependence in the stock and exchange markets. Using data on the G5 countries, i.e. the US, UK, Germany, France and Japan, Ning (2010) finds significant evidence of dependence in both the upper and lower tails. However, she finds no evidence of asymmetry. In contrast, Wang, Wu, and Lai (2013) use a generalized copula model that nests reports asymmetry in the tail dependence under a “negative correlation regime”. While these studies focus on developed economies, Cumparayot, Keijzer, and Kouwenberg (2006) use a simultaneous equations probit model instead to study the link between the currency and stock markets of 26 countries during extreme events and find a one-way spillover effect from the stock markets to the currencies of most emerging markets during crisis periods. Reboredo,

\textsuperscript{1} Okimoto (2008) consider all G7 countries except Italy.
Rivera-Castro, and Ugolini (2016) also find evidence consistent with the flight to quality by investigating the stock and exchange markets co-movements and spillover effects for eight emerging market economies. By employing copulas to model the dependency structure and using Value-at-Risk (VaR) and Conditional Value-at-Risk (CoVaR) to measure the spillover effects, they find a positive relation between the equity returns and currencies of emerging markets, relative to both the US and the euro.

Even though there is ongoing debate on the level and structure of the tail dependency between asset returns, and how best to model this dependency, what seems to be inarguable is that ignoring this dependency has important consequence for portfolio managers and policy makers, among others. In a recent paper, Elkamhi and Stefanova (2015) develop a model that considers the various extremal dependence structures in asset returns and show the significant economic loss that can occur when these extreme co-movements are ignored. They demonstrate that tail dependence affects the investor’s portfolio allocation decision, irrespective of whether the motive is hedging or diversification.

2. The dependence measure for multivariate extreme

To study the multivariate dependence structure, the influence of the marginal behavior is first removed by transforming the raw data to a common marginal distribution. According to Sklar theorem (Sklar, 1959), each joint distribution can be decomposed into its marginal distributions and its dependence structure (also called Copula). After removing the marginal aspects, the remaining differences between the distributions are only due to dependence aspects (Embrechts, McNeil and Strautman, 1999). Given a
bivariate return $(X,Y)$, we transform them into a standard Frechet marginals $(S,T)$, as follows:

$$S = -1/logF_X(X) \text{ and } T = -1/logF_Y(Y)$$  \hspace{1cm} (1)

Where $F_X(X)$ and $F_Y(Y)$ are marginal distribution functions of $X$ and $Y$ respectively. As a consequence of this transformation, \( \Pr(S > s) = \Pr(T > s) = s^{-1} + O(s^{-2}) \) as \( s \to \infty \) and the variables $(S,T)$ have the same dependence structure as $(X,Y)$ (Poon et al., 2003, 2004).

2.1. Asymptotic dependence

Because of the characteristics of $(S,T)$, for an extreme value, the events $S > s$, and $T > s$ are also extreme for both variables. Coles, Heffernan, and Tawn (1999), Poon, Rockinger, and Tawn (2003, 2004) defined the non-parametric measure of dependence $\gamma$ as:

$$\gamma = \lim_{s \to \infty} \frac{\Pr(T > s, S > s)}{\Pr(S > s)}$$  \hspace{1cm} (2)

If $\gamma > 0$, then $S$ and $T$ are asymptotically dependent. There is perfect dependence if $\gamma = 1$. If $\gamma = 0$, then $S$ and $T$ are asymptotically independent.

2.2. Asymptotic independence

Given the fact that two random variables which are asymptotically independent ($\gamma = 0$) may show different degree of dependence for finite level of $s$, Coles, Heffernan and Tawn (1999) proposes the following measure of asymptotic dependence.
\[
\bar{y} = \lim_{s \to \infty} \frac{2\log\{\Pr(S > s)\}}{\Pr(T > s, S > s)} - 1
\]  

(3)

According to Poon et al. (2003, 2004), it measures the rate at which \( \Pr(T > S|S > s) \) approaches zero and values of \( \bar{y} > 0, \bar{y} = 0, \) and \( \bar{y} < 0 \) are approximate measure of positive dependence, exact dependence and negative dependence in the tail respectively.

It is also known that under normality, \( \bar{y} \) is identical to the Pearson correlation coefficient.

The degree and form of extremal dependence can then be characterized with the two dependence measures \((\gamma, \bar{y})\). More precisely, the bivariate distribution is tested for \( \bar{y} = 1 \) before quantifying the bivariate asymptotic dependence using \( \gamma \). For asymptotically dependent variables, \((\bar{y} \neq 1)\) and the degree of dependence is given by \( \bar{y} \).

If they are asymptotically dependent \((\bar{y} = 1)\), the degree of dependence is measured by the value of \( \gamma \).

3. Non-parametric estimation of \( \gamma \) and \( \bar{y} \)

3.1. The Hill estimator

To model the behavior of extreme values above a threshold (peaks over threshold or POT), the excess distribution above a threshold \( u \) is given by the conditional probability distribution:

\[
F_u(x) = Pr(X - u \leq x|X > u) = \frac{F(x + u) - F(u)}{1 - F(u)}, \quad x > 0
\]

(4)

Under some regularity conditions and for a large enough value of \( u \), the distribution (4) can be well approximated by a Generalized Pareto Distribution (GDP), with shape
parameter $k$, location parameter $\theta$ and scale parameter $\sigma = \sigma(u) > 0$ (see, for example, Coles, 2000; or Embrechts, Klüpperberg and Mikosch, 1997).

$G_{k,\theta,\sigma(u)}(x)$

$$= \begin{cases} 
1 - \left\{ 1 + \frac{k(x - \theta)}{\sigma(u)} \right\}^{-\frac{1}{k}} & \text{for } (x > \theta \text{ and } k > 0) \text{ or } (\theta < x < \theta - \frac{\sigma(u)}{k} \text{ and } k < 0) \quad (5) \\
1 - e^{\left\{ \frac{x - \theta}{\sigma(u)} \right\}} & \text{for } k = 0 \text{ and } x > \theta 
\end{cases}$$

If $k = 0$, and $\theta = 0$ the GPD is equivalent to the exponential distribution. If $k > 0$ and $\theta = \sigma/k$, then the generalized Pareto distribution is equivalent to the Pareto distribution with a scale parameter equal to $\sigma/k$ and a shape parameter equal to $1/k$. The distribution $F_u(\cdot)$ is said to be from the Frechet family when $k > 0$. For the particular case of GDP from the Frechet family, Poon et al (2003, 2004) showed that for a random variable $X$, the tail above a high threshold $u$ can be approximated as:

$$1 - F_X(x) = Pr(X > x) \sim x^{-1/k}L(x), \quad x > u \quad (6)$$

Where $k > 0$ and $L(x)$ is a slowly varying function of $x$, that is, $L(tx)/L(x) \to \infty$ as $t \to \infty$. Taking $L(x)$ as constant for $x > u$ and assuming the observations to be iid, the maximum likelihood estimator of $k$ also known as Hill’s estimator (Hill, 1975), and $L(x)$ are:
\hat{k} = \frac{1}{n_u} \sum_{j=1}^{n_u} \log \left( \frac{x(j)}{u} \right),
\quad (7)

\overline{L(x)} = \frac{n_u}{n} u^{1/\hat{k}},
\quad (8)

Where \( n \) is the total number of observations, \( x(1), \ldots, x(n_u) \) are the \( n_u \) observations above the threshold, \( u \).

The tail index is determined by the following steps:

i. Plot the mean excess plot and detect the value from which the curve is linear.

ii. Plot the Hill plot and detect the range of values from which there is a certain stability of the curve.

iii. With the threshold values that were detected from the previous plots, estimate the parameters of the GPD.

iv. Assess the goodness of fit of the fitted GDP with three methods.

v. Plot of the empirical distribution of the tail and the estimated GPD distribution.

vi. Perform the QQ plot test of the fitted GPD and the empirical distribution of the tail.

vii. Perform the Chi square, the Kolmogorov Smirmov and the Cramer Von Mises adequation test.

Among the threshold candidates of the first two steps, select the one that provides the smallest standard error of estimators and the best goodness of fit. The more the fitted tail distribution is adequate, it will be close to the empirical distribution on the graph, the QQ plot will be linear and the p-value the Chi square, Cramer Von Mises and Kolmogorov Smirmov test will be higher.
3.2 Estimation of $\gamma$ and $\tilde{\gamma}$

Defining $Z = \min(S, T)$, the tail dependence test relies on the fact that:

$$\Pr(Z > z) \sim L(z)z^{-1/k} \quad \text{as} \quad z > u$$  \hspace{1cm} (9)

For some high threshold $u$ where $L(z)$ is a slow varying function. This result was shown by Ledford and Tawn (1996, 1998). The constant, $k$, $0 < k \leq 1$ which is called the coefficient of tail dependence can then be computed by the use of the Hill estimator constrained to the interval $(0,1]$. It follows that:

$$\hat{\gamma} = 2\hat{k} - 1 = \frac{2}{n_u} \sum_{j=1}^{n_u} \log \left( \frac{X(j)}{u} \right) - 1, \quad \text{var}(\hat{\gamma}) = \frac{(\hat{\gamma} + 1)^2}{n_u}$$  \hspace{1cm} (10)

Where $\hat{\gamma}$ is asymptotically normal.

If $\hat{\gamma} + 1.96\sqrt{\text{var}(\hat{\gamma})} < 1$, then the null hypothesis of asymptotic dependence ($\tilde{\gamma} = 1$) is rejected and the two random variables are asymptotically independent with degree of dependency measured by $\tilde{\gamma}$ (Fernandez, 2006). If the null hypothesis cannot be rejected, then $\gamma$ is estimated under the assumption $\tilde{\gamma} = k = 1$, where:

$$\hat{\gamma} = \frac{un_u}{n}, \quad \text{var}(\hat{\gamma}) = \frac{un_u(n - n_u)}{n^3}.$$  \hspace{1cm} (11)

4. Empirical analysis

In this section, we examine the asymptotic dependence and investigate the degree of tail dependence of the ordinary-daily returns of the following markets: CAC40, SPTX, DAX, NKY, SPX, SHCOMP and UKX. A GARCH filter is also used to check whether the asymptotic dependence between these markets is related to heteroscedasticity present in
the logarithm return series. The daily levels of all indices are obtained from Bloomberg. The sample period is from December 19, 1990 to March 01, 2017, a total of 5616 daily observations for each of the market. For each series, daily index returns are calculated by taking the first difference of the logarithm of the index levels.

4.1 Descriptive statistics and univariate tail behavior

Table 1 presents summary statistics for the stock index returns. For each market, we present the mean, variance, skewness and excess kurtosis. Following Poon et al. (2003, 2004), these statistics are evaluated using generalized method of moment to ensure robustness to heteroscedasticity (see appendix A1). A filtered version of the stock returns was also considered to eliminate the variation in volatility from the extreme returns and to remove clustering of extreme values caused by volatility persistence. The filter used is an asymmetric GARCH (1, 1) model described in appendix A2. Table 1 includes Hill estimators for both left and right tails of the unfiltered and filtered univariate data series.
References


Appendix A1 : The General Method of Moment for mean, variance, skewness and kurtosis

The mean, the variance, the skewness and the kurtosis are determined with the general method of moment (GMM). Given a random variable, let's \( \mu_X \) the mean, \( \sigma_X \) the standard deviation, \( \gamma_X \) the skewness and \( \kappa_X \) the kurtosis of \( X \). They are given by the equation :

\[
\begin{align*}
E(X - \mu_X) &= 0; \\
E((X - \mu_X)^2) &= \sigma_X^2; \\
E\left(\frac{X - \mu_X}{\sigma_X}\right)^3 &= \gamma_X; \\
E\left(\frac{X - \mu_X}{\sigma_X}\right)^4 &= \kappa_X. \\
\end{align*}
\]

This provides the moment equation :

\[
E\left(\frac{X - \mu_X}{\sigma_X}\right)^k - E\{g_k(X, \mu_X, \sigma_X, \gamma_X, \kappa_X)\} = 0.
\]

With a random sample of a size \( n \), \( X_1, \ldots, X_n \), we have the following equation :

\[
G_n(X, \theta) = \frac{1}{n} \sum_{i=1}^{n} g_k(X_i, \theta) = 0.
\]

where \( \theta = (\mu_X, \sigma_X, \gamma_X, \kappa_X)^T \).

starting with an initial estimator \( \hat{\theta}_0 = (\mu_0, \sigma_0, \gamma_0, \kappa_0)^T \):  
— have the initial estimator by the optimization program \( \hat{\theta}_1 = \arg\min \{G_n^2(X, \theta) \mid G_n(X, \theta)\} \);  
— build the covariance matrix estimate \( \hat{W} = G_n(X, \hat{\theta}_1)^T G_n(X, \hat{\theta}_1) \);  
— get the final estimator : \( \hat{\theta}_2 = \arg\min \{G_n^2(X, \theta) \hat{W}^{-1} G_n(X, \theta)\} \).

The final estimator \( \hat{\theta}_2 \) is known to autocorrelation and heteroskedasticity consistent.

Appendix A2 : Asymmetric Garch model (Agarch) and estimation

The Asymmetric GARCH model (AGARCH(1, 1)) is used to filter the raw time series. Given a sample \( X_1, \ldots, X_n \), the AGARCH(1, 1) is defined as followed.

\[
\begin{align*}
X_t &= \omega - \sqrt{h_t} \epsilon_t; \\
h_t &= \alpha_0 + \alpha \epsilon_{t-1}^2 D_{t-1} + \gamma \epsilon_{t-1}^2 D_{t-1} + \beta h_{t-1},
\end{align*}
\]
where $\theta = (\omega, \alpha_0, \alpha^+, \alpha^-, \beta, \gamma)^T$ and $D_\phi$ is the indicator function.

The $AGARCH(1,1)$ model is estimated by the mean of the maximum likelihood. The likelihood of the sample $X_1, \ldots, X_n$, is defined by:

$$l(\theta) = \sum_{t=1}^n \left[ C_{te} - \frac{1}{2} \log(h_t) - \frac{\varepsilon_t^2}{2h_t} \right]$$

- Starting with an initial estimator $\theta_0 = (\omega_0, \alpha_0, \alpha^+, \alpha^-)^T$:
  - get $h_1 = \alpha_0 + (\alpha^+ - \alpha^-)/(2(1 - \beta_0))$;
  - set $x_1 = (X_1 - \omega_0)/\sqrt{h_1}$;
  - calculate $l_1 = -\frac{1}{2} \log(h_1) - \frac{\varepsilon_1^2}{2h_1}$;
  - $h_t = \alpha_0 + \alpha^+ \varepsilon_{t-1}^2 1_{D_{log}>0} + \alpha^- \varepsilon_{t-1}^2 1_{D_{log}<0} + \beta h_{t-1}$;
  - set the estimator $\theta_1 = \underset{\theta}{\text{ArgMin}} \left( \sum_{t=1}^n l_t \right)$;
- repeat the process until convergence.
Table 1: Summary Statistics of Stock index returns

<table>
<thead>
<tr>
<th></th>
<th>Mean</th>
<th>Standard deviation</th>
<th>Skewness</th>
<th>Kurtosis</th>
<th>Raw data</th>
<th>Filtered data</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>left tail</td>
<td>right tail</td>
</tr>
<tr>
<td><strong>CAC40</strong></td>
<td>0.0164</td>
<td>1.3970</td>
<td>-0.1689</td>
<td>8.1920</td>
<td>0.0802(0.0184)</td>
<td>0.1328(0.0324)</td>
</tr>
<tr>
<td><strong>SPTX</strong></td>
<td>0.0475</td>
<td>1.5426</td>
<td>-1.4733</td>
<td>47.8932</td>
<td>0.1572(0.0330)</td>
<td>0.2293(0.0313)</td>
</tr>
<tr>
<td><strong>DAX</strong></td>
<td>0.0242</td>
<td>1.2343</td>
<td>-0.1565</td>
<td>10.3308</td>
<td>0.0271(0.0288)</td>
<td>0.1015(0.0242)</td>
</tr>
<tr>
<td><strong>NKY</strong></td>
<td>0.0179</td>
<td>1.3088</td>
<td>-0.4083</td>
<td>12.3667</td>
<td>0.0224(0.0193)</td>
<td>0.1425(0.0434)</td>
</tr>
<tr>
<td><strong>SPX</strong></td>
<td>0.0247</td>
<td>1.1829</td>
<td>-0.4910</td>
<td>23.5928</td>
<td>0.0936(0.0136)</td>
<td>0.1519(0.0195)</td>
</tr>
<tr>
<td><strong>SHCOMP</strong></td>
<td>0.0544</td>
<td>2.3599</td>
<td>5.2186</td>
<td>151.2010</td>
<td>0.2183(0.0286)</td>
<td>0.2588(0.0312)</td>
</tr>
<tr>
<td><strong>UKX</strong></td>
<td>0.0227</td>
<td>1.1042</td>
<td>-0.4563</td>
<td>12.5004</td>
<td>0.1325(0.0386)</td>
<td>0.1484(0.0396)</td>
</tr>
</tbody>
</table>