Dependence structure and extreme comovements in international equity and bond markets

René Garcia a, Georges Tsafack b,∗

a EDHEC Business School, 393/400 Promenade des Anglais, BP 3116, 06202 Nice, Cedex 3, France
b Sawyer Business School, Suffolk University, Boston, MA, United States

ARTICLE INFO

Article history:
Received 9 September 2009
Accepted 13 January 2011
Available online 26 January 2011

JEL classification:
C32
C51
G15

Keywords:
Asymmetric correlation
Asymmetric dependence
Copula
Tail dependence
GARCH
Regime switching

ABSTRACT

Common negative extreme variations in returns are prevalent in international equity markets. This has been widely documented with statistical tools such as exceedance correlation, extreme value theory, and Gaussian bivariate GARCH or regime-switching models. We point to limits of these tools to characterize extreme dependence and propose an alternative regime-switching copula model that includes one normal regime in which dependence is symmetric and a second regime characterized by asymmetric dependence. We apply this model to international equity and bond markets, to allow for inter-market movements. Empirically, we find that dependence between international assets of the same type is strong in both regimes, especially in the asymmetric one, but weak between equities and bonds, even in the same country.

© 2011 Elsevier B.V. All rights reserved.

1. Introduction

There is ample evidence that negative returns are more dependent than positive returns in international equity markets. This phenomenon known as asymmetric dependence has been reported by many previous studies including Erb et al. (1994), Longin and Solnik (2001), Ang and Bekaert (2002), Ang and Chen (2002), Das and Uppal (2004), Patton (2004), and references therein. This asymmetric dependence has important implications for portfolio allocation, but to appreciate its full actual effects on portfolio returns or volatility suffer from some conditioning bias. Correlation asymmetry may therefore appear spuriously if these biases are not accounted for. To avoid these problems, Longin and Solnik (2001) use extreme value theory (EVT) by focusing on the asymptotic value of exceedance correlation. The benefit of EVT resides in the fact that the asymptotic result holds regardless of the distribution of returns. By the same token, as emphasized by Longin and Solnik (2001), EVT

The exceedance correlation between two series of returns is defined as the correlation for a subsample in which the returns of both series are simultaneously lower (or greater) than the corresponding thresholds θ1 and θ2. Formally, exceedance correlation of variables X and Y at thresholds θ1 and θ2 is expressed by

\[ \text{corr}(X, Y|X < θ_1, Y < θ_2) - \text{corr}(X, Y|X > θ_1, Y > θ_2), \]

for θ1 ≥ 0 and θ2 ≥ 0. Longin and Solnik (2001) use θ1 = 0 and θ2 = 0, while Ang and Chen (2002) use θ1 = (1 + θ)X and θ2 = (1 + θ)Y, where X and Y are the means of Y and X respectively.

Extreme value theory (EVT) is used to characterize the distribution of a variable conditionally to the fact that its values are beyond a certain threshold, and the asymptotic distribution is obtained when this threshold tends to infinity. Hartmann et al. (2004) also use extreme-value analysis to capture the dependence structure between stock and bond returns for pairs of the G5 countries.

0378-4266/$ - see front matter © 2011 Elsevier B.V. All rights reserved.
doi:10.1016/j.jbankfin.2011.01.003
cannot help to determine if a given return-generating process is able to reproduce the extreme asymmetric exceedance correlation observed in the data.

To overcome this shortcoming, we propose a model based on copulas that allows for tail dependence in lower returns and keeps tail independence for upper returns as suggested by the findings of Longin and Solnik (2001). Copulas are functions that build multivariate distribution functions from their unidimensional marginal distributions. The tail dependence coefficient can be seen as the probability of the worst event occurring in one market given that the worst event occurs in another market. Contrary to exceedance correlation, the estimation of the tail dependence coefficient is not subject to the problem of choosing an appropriate threshold and the use of extreme value distributions such as the Pareto distribution.

Another difference is that tail dependence is completely defined by the dependence structure and is not affected by variations in marginal distributions.

The disentangling between marginal distributions and dependence helps overcoming the curse of dimensionality associated with the estimation of models with several variables. For example, in multivariate GARCH models, the estimation becomes intractable when the number of series being modeled is high. The CCC of Bollerslev (1990), the DCC of Engle (2002), and the RSDC of Pelleter (2006) deal with this problem by separating the variance-covariance matrix in two parts, one part for the univariate variances of the different marginal distributions, another part for the correlation coefficients. This separation allows them to estimate the model in two steps, first the marginal parameters on each individual series then the correlation parameters. Copulas offer a tool to generalize this separation while extending the linear concept of correlation to nonlinear dependence.

Thanks to the tail dependence formulation of asymptotic dependence, we show analytically that the multivariate GARCH or regime-switching (RS) models with Gaussian innovations that have been used to address asymmetric dependence issues (see Ang and Bekaert, 2002; Ang and Chen, 2002) cannot in fact reproduce extreme asymmetric dependence. The key point is that these classes of models can be seen as mixtures of symmetric distributions and cannot produce asymptotically asymmetric dependence. The asymmetry produced by these models at finite distance disappears asymptotically. When we go far in the tails, we obtain a similar dependence for the upper and lower tails. Moreover, the asymmetry in RS models comes from the asymmetry created in the marginal distributions with regime switching in the mean. Hence it is not separable from the marginal asymmetry or skewness. This is a fundamental issue that also affects the statistical extreme-value analysis that have been conducted to study extreme dependence.

We use our regime-switching copula model to investigate the dependence structure between international equity and bond markets. The model allows for a switching between a normal state where markets will be linearly and symmetrically correlated and an asymmetric dependence state to capture common crashes. In a normal regime it is difficult to make a difference between the level of dependence for joint positive moves and joint negative moves. When the economy is in the asymmetric regime, even with a stable correlation, a downside move in one market will increase the probability of a similar event in another market. The rise in the level of dependence during market downturns is characterized by asymmetry in the dependence structure. This regime can be interpreted as contagion since bad news spread quickly between markets. This crash dependence can coexist with low correlation and implies a reduction of an apparent diversification benefit.

We separately analyze dependence between the two leading markets in North-America (US and Canada) and two major markets of the Euro zone (France and Germany). Our empirical analysis shows that dependence between international assets of the same type is strong in both the symmetric and the asymmetric regimes, while dependence between equities and bonds is low even in the same country. Another finding is that the presence of a regime with extreme asymmetric dependence makes the correlation in the normal regime differ from the unconditional correlation. We also provide some evidence that exchange rate volatility seems to contribute to asymmetric dependence. With the introduction of a fixed exchange rate the dependence between France and Germany becomes less asymmetric and more normal than before. High exchange rate volatility is associated with a high level of asymmetry. These results are consistent with those of Cappiello et al. (2006) who find an increase in correlation after the introduction of the Euro currency.

The rest of this paper is organized as follows. Section 2 reformulates the empirical facts about exceedance correlation in terms of tail dependence and shows how classical GARCH or regime-switching models fail to capture these facts. In Section 3 we develop a model with two regimes that clearly disentangles dependence from marginal distributional features and allows asymmetry in extreme dependence. As a result, we obtain a model with four variables that features asymmetry and a flexible dependence structure. Empirical evidence on the dependence structure is examined in Section 4, while conclusions are drawn in Section 5.

2. Extreme asymmetric dependence and modeling issues

In this section we present empirical facts about exceedance correlation in international equity market returns put forward by Longin and Solnik (2001) and the related literature. We next argue that these facts can be equivalently reformulated in terms of tail dependence. The latter formulation will allow us to explain why classical return-generating processes such as GARCH and regime-switching models based on a multivariate normal distribution fail to reproduce these empirical facts.

2.1. Empirical facts

Longin and Solnik (2001) investigate the structure of correlation between various equity markets in extreme situations by testing the equality of exceedance correlations, one obtained under a joint normality assumption and the other one computed using EVT. For the latter distribution, they model the marginal distributions of equity index returns with a generalized Pareto distribution (GPD) and capture dependence through a logistic function. Ang and Chen (2002) develop a test statistic based on the difference between exceedance correlations computed from the data and those obtained from GARCH or RS models.

4 The theory of this useful tool dates back to Sklar (1959) and a clear presentation can be found in Nelsen (1999). Well designed to analyze nonlinear dependence, copulas were initially used by statisticians for nonparametric estimation and measure of dependence of random variables (see Genest and Rivest, 1993 and references therein).

5 Ang and Chen (2002) conclude that even if regime-switching models perform best in explaining the amount of correlation asymmetry reflected in the data, these models still leave a significant amount of correlation asymmetry in the data unexplained.

6 Some related research focus on the dependence structure of bond and equity in European markets (see Kim et al., 2006; Cappiello et al., 2006; Abad et al., 2010). Panchenko and Wu (2008) investigate stock and bond return time-varying comovement in emerging markets, while Markwitz et al. (2009) and Kumar and Okimoto (2011) analyze international integration in a global perspective.

7 They define a test statistic $H = \sum_{i=1}^{N} \left( \hat{\rho}(\theta_i) - \hat{\rho}(\theta_0) \right)^2 / \hat{\sigma}^2$ which is the distance between exceedance correlations obtained from the normal distribution $(\hat{\rho}(\theta_1), \ldots, \hat{\rho}(\theta_N))$ and exceedance correlations estimated from the data $(\hat{\rho}(\theta_1), \ldots, \hat{\rho}(\theta_N))$ for a set of $N$ selected thresholds $(\theta_1, \ldots, \theta_N)$. In the same way they define $H^+$ and $H^-$ by considering negative points for $H^+$ and nonnegative points for $H^-$ such that $H^+ = (H^+)^2 + (H^-)^2$. They can therefore conclude to asymmetry if $H^-$ differs from $H^+$. 

8 The disentangling between dependence and tail dependence comes from the asymmetry created in the marginal distributions and cannot help to determine if a given return-generating process is able to reproduce the extreme asymmetric exceedance correlation observed in the data.
These two studies conclude that there exists asymmetry in exceedance correlation, that is large negative returns are more correlated than large positive returns. However, their results rely on choosing a set of thresholds for computing exceedance correlation and can only account for asymmetry at finite distance. Crashes are more in the nature of extreme events and involve measuring dependence for thresholds very far in the tail. Longin and Solnik (2001) confirm with an asymptotic test that exceedance correlation is positive and statistically different from zero for very large negative returns and not different from zero for very large positive returns. However they do not provide a model that is able to reproduce this fact. Ang and Chen (2002) as well as Ang and Bekaert (2002) find that regime-switching models can reproduce the asymmetry in exceedance correlation, but this result does not hold for extreme events as we will show later and the measured asymmetry amalgamates skewness in the marginal distributions and asymmetric dependence.

We illustrate these facts and the capacity of models to reproduce them in Fig. 1 with US and Canadian returns. We specify thresholds in term of quantiles: \( \theta_1 = F_X^{-1}(0.2) \) and \( \theta_2 = F_Y^{-1}(0.8) \) where \( F_X \) and \( F_Y \) are the cumulative distribution functions of \( Y \) and \( X \) respectively. Following Longin and Solnik (2001) and Ang and Chen (2002) exceedance correlations are symmetric if \( \text{Excorr}(Y; \theta_1, \theta_2) = \text{Excorr}(Y, X; 1 - \theta_1, 1 - \theta_2) \) for \( \theta \in (0, 1) \). Correlations of return exceedances exhibit the typical shape put forward in Longin and Solnik (2001) for the US equity market with various European equity markets. For the models, we chose to retain the multivariate normal, as a benchmark case to show that correlations go to zero as we move further in the tails, as well as a normal regime-switching model, as in Ang and Chen (2002). The latter model produces some asymmetry in correlations for positive and negative returns but not nearly as much as in the data. We also exhibit the exceedance correlations estimated with the procedure used by Longin and Solnik (2001). It is evidently much closer to the data. Finally, we also report the correlations obtained from a rotated Gumbel copula for the dependence function (see Appendix A for a definition), with Gaussian marginal distributions. The graph is very close to the Longin and Solnik (2001) one.

Since asymptotic exceedance correlation is zero for both sides of a bivariate normal distribution, Longin and Solnik (2001) interpreted these findings as rejection of normality for large negative returns and non-rejection for large positive returns. In the conclusion of their article, Longin and Solnik stress that their approach has the disadvantage of not explicitly specifying the class of return-generating processes that fail to reproduce these two facts. The difficulty in telling which model can reproduce these facts is the lack of analytical expressions for the asymptotic exceedance correlation and its intractability even for classical models such as Gaussian GARCH or regime-switching models. In order to investigate this issue, we introduce the concept of tail dependence. This will help us show analytically that some classes of models previously used in the literature cannot reproduce these asymmetries in extreme dependence and then propose a model that succeeds in doing so.

### 2.2. Tail dependence

To measure the dependence between an extreme event on one market and a similar event on another market, we define two dependence functions one for the lower tail and one for the upper tail, with their corresponding asymmetric tail dependence coefficients. For two random variables \( X \) and \( Y \) with cumulative distribution functions \( F_X \) and \( F_Y \) respectively, we call the lower tail dependence function (TDF) the conditional probability \( t^L(x) = \frac{\Pr(X \leq F_X^{-1}(x) | Y \leq F_Y^{-1}(y))}{\Pr(Y \leq F_Y^{-1}(y))} \) for \( x, y \in (0, 1/2) \) and similarly, the upper tail dependence function is \( t^U(x) = \frac{\Pr(X > F_X^{-1}(1-x) | Y > F_Y^{-1}(1-y))}{\Pr(Y > F_Y^{-1}(1-y))} \). The tail dependence coefficient (TDC) is simply the limit (when it exists) of this function when \( x \) tends to zero. More precisely lower TDC is \( t^L(x) = \lim_{y \to 1} t^L(x, y) \) and upper TDC is \( t^U(x) = \lim_{y \to 0} t^U(x, y) \). As in the case of joint normality, we have lower tail independence when \( t^L = 0 \) and upper tail independence for \( t^U = 0 \).

Compared to exceedance correlation used by Longin and Solnik (2001), Ang and Chen (2002), Ang and Bekaert (2002), and Patton (2004), a key advantage of TDF and corresponding TDCs is their invariance to modifications of marginal distributions that do not affect the dependence structure. Fig. 2 gives an illustration of this invariance. We simulate a bivariate Gaussian distribution \( N(0, I_p) \), where \( I_p \) is the bi-dimensional matrix with standard deviations equal to one on the diagonal and a correlation coefficient \( \rho \) equal to 0.5. Both exceedance correlation and tail dependence measures show a symmetric behavior of dependence in extreme returns. However, when we replace one of the marginal distributions \( N(0, 1) \) by a mixture of normals, a \( N(0, 1) \) and a \( N(4, 4) \) with equal weights, and let the other marginal distribution and the dependence structure unchanged, the TDF remains the same while the exceedance correlation is affected. In fact, the correlation coefficient and the exceedance correlation are a function of the dependence structure and of the marginal distributions while the tail dependence is a sole function of the dependence structure, regardless of the marginal distributions. Another problem with asymptotic exceedance correlation estimation using extreme value is the sample bias since fewer data points are available when we...
move further into the tails of the distribution. With tail dependence, the estimation can be done using all data points in the sample and the estimators of the tail coefficients are unbiased. By observing that for the logistic function used by Longin and Solnik (2001), the zero value for the asymptotic correlation coefficient is exactly equivalent to tail independence, we can reformulate their asymptotic result as follows: lower extreme returns are tail-dependent, while upper extreme returns are tail-independent.

This reformulation presents at least two main advantages. Compared to exceedance correlation, the tail dependence coefficient is generally easier to compute and analytical expressions can be obtained for almost all distributions. This is not the case for exceedance correlation even for usual distributions. Moreover, we can easily derive the tail dependence of a mixture from the tail dependence of the different components of the mixture. The last property will be used below to investigate which model can or cannot reproduce the results of Longin and Solnik (2001).

### 2.3. Why classical multivariate GARCH and RS model cannot reproduce asymptotic asymmetries

Ang and Chen (2002) and Ang and Bekaert (2002) try to reproduce asymmetric correlations facts with classical models such as GARCH and RS based on a multivariate normal distribution. After examining a number of models, they found that GARCH with constant correlation and fairly asymmetric GARCH cannot reproduce the asymmetric correlations documented by Longin and Solnik. However, they found that a RS model with Gaussian innovations is better at reproducing asymmetries in exceedance correlation. They clearly reproduce asymmetric correlations at finite distance. However, their finite-distance asymmetric correlation comes from gime switching in means, as suggested by the simulation in the previous section. Therefore it becomes difficult to distinguish asymmetries in dependence from asymmetry in marginal distributions. This is a problem of practical relevance since most return series exhibit asymmetry in volatility.

By reinterpreting Longin and Solnik (2001) results in term of TDC instead of asymptotic exceedance correlation, we show analytically that all these models cannot reproduce asymptotic asymmetry even if some can reproduce finite-distance asymmetry. These results are extended to the rejection of more general classes of return-generating processes. The key point of this result is the fact that many classes of models including Gaussian (or Student) GARCH and RS can be seen as mixtures of symmetric distributions. We establish the following result.

**Proposition 2.1**

(i) Any GARCH model with constant mean and symmetric conditional distribution has a symmetric unconditional distribution and hence a symmetric TDC.

(ii) If the conditional distribution of a RS model has a zero TDC, then the unconditional distribution also has a zero TDC.

(iii) From a multivariate distribution with symmetric TDC, it is impossible to construct an asymmetric TDC with a mixture procedure (as GARCH, RS or any other) by keeping all marginal distributions unchanged across mixture components.

**Proof.** See Appendix A.

This proposition allows us to argue that the classical GARCH or RS models cannot reproduce asymmetries in asymptotic tail dependence. Therefore, the classical GARCH models (BEKK, CCC or DCC) with constant mean can be seen as a mixture of symmetric distributions with the same first moments and therefore exhibit a symmetric tail dependence function as well as a symmetric TDC. When the mean becomes time-varying as in the GARCH-M model the
unconditional distribution can allow asymmetry in correlation (Ang and Chen, 2002), but this asymmetry comes from the mixture of the marginal distributions. The resulting skewness cannot be completely disentangled from the asymmetric correlation, since correlations are affected by marginal changes. Similarly, the classical RS model with Gaussian innovations is a discrete mixture of normal distributions which has a TDC equal to zero on both sides. Therefore, by (ii) we argue that both its TDCs are zero. However, at finite distance, when the mean changes with regimes, the exceedance correlation is not symmetric. This asymmetry is found by Ang and Chen (2002) and Ang and Bekaert (2002) in their RS model, but it disappears asymptotically and it comes from the asymmetry created in the marginal distributions by regime switching in means. Hence, the asymmetries in correlation are not separable from the marginal asymmetry, exactly like in the GARCH–M case. The part (iii) of Proposition 2.1 extends this intuition in terms of more general multivariate mixture models based on symmetric innovations. Actually when the marginal distributions are the same across all symmetric TDC components of a mixture, it is impossible to create asymmetry in TDCs.

Two relevant issues arise from the above discussion. First, how can we separate the marginal asymmetries from the asymmetry in dependence? Second, how can we account not only for asymmetry at finite distance but also for asymptotic dependence? In the next section, we propose a flexible model based on copulas that addresses these two issues.

3. A copula model for asymmetric dependence

Our model aims at capturing the type of asymmetric dependence found in international equity markets. Our discussion in the last section showed that it is important to disentangle the marginal distributions from the dependence structure. Therefore, we need to allow for asymmetry in tail dependence, regardless of the possible marginal asymmetry or skewness. Copulas, also known as dependence functions, are an adequate tool to achieve this aim.

3.1. Disentangling the marginal distributions from dependence with copulas

Estimation of multivariate models is difficult because of the large number of parameters involved. Multivariate GARCH models are a good example since the estimation becomes intractable when the number of series being modeled is high. The CCC of Bollerslev (1990), the DCC of Engle (2002), and the RSDC of Pelletier (2006) deal with this problem by separating the variance–covariance matrix into two parts, one for the univariate variances of the different marginal distributions, the other for the correlation coefficients. This separation allows them to estimate the model in two steps. In the first step, they estimate the marginal parameters and use them in the estimation of the correlation parameters in a second step. Copulas offer a tool to generalize this separation while extending the linear concept of correlation to nonlinear dependence.

Copulas are functions that build multivariate distribution functions from their unidimensional marginals. Let \( X \equiv (X_1, \ldots, X_n) \) be a vector of \( n \) univariate variables. Denoting \( F \) the joint \( n \)-dimensional distribution function and \( F_1, \ldots, F_n \) the respective marginals of \( X_1, \ldots, X_n \). Then the Sklar theorem states that there exists a function \( C \) called copula which joins \( F \) to \( F_1, \ldots, F_n \) as follows:

\[
F(x_1, \ldots, x_n) = C(F_1(x_1), \ldots, F_n(x_n)) \tag{3.1}
\]

This relation can be expressed in term of densities by differentiating with respect to all arguments. We can therefore write (3.1) equivalently as

\[
f(x_1, \ldots, x_n) = c(F_1(x_1), \ldots, F_n(x_n)) \prod_{i=1}^n f_i(x_i) \tag{3.2}
\]

where \( f \) represents the joint density function of the \( n \)-dimensional variable \( X \) and \( f_i \) the density function of the variable \( X_i \) for \( i = 1, \ldots, n \). The copula density function is naturally defined by \( c(u_1, \ldots, u_n) = \prod_{i=1}^n f_i(u_i) \). Writing the joint distribution density in the above form, we understand why it can be said that a copula contains all information about the dependence structure.\(^{15}\)

We now suppose that our joint distribution function is parametric and we separate the marginal parameters from the copula parameters. So the relation (3.2) can be expressed as:

\[
f(x_1, \ldots, x_n; \delta, \vartheta) = c(u_1, \ldots, u_n; \delta) \prod_{i=1}^n f_i(x_i; \delta_i); \tag{3.3}
\]

where \( \delta = (\delta_1, \ldots, \delta_n) \) are the parameters of the different margins and \( \vartheta \) denotes the vector of all parameters that describe dependence through the copula. Therefore, copulas offer a way to separate margins from the dependence structure and to build more flexible multivariate distributions.

More recent works allow some dynamics in dependence. In a bivariate context, Rodriguez (2007) introduces regime switching in both the parameters of marginal distributions and the copula function;\(^{16}\) Jondeau and Rockinger (2006) capture the time-varying volatilities of the individual equity index return series by a GARCH model and introduce Markov-switching Student-\( t \) copulas for pairs of countries. Recently, Chollette et al. (2009) propose a model of multivariate regime-switching copulas to capture asymmetric dependence in international financial returns. Ang and Bekaert (2002, 2004) allow all parameters of the multivariate normal distribution to change with the regime. The extension of these models to a large number of series faces the above-mentioned curse of dimensionality. Since the switching variable is present in both the margins and the dependence function, separation of the likelihood function into two parts is not possible and the two-step estimation cannot be performed. Pelletier (2006) uses the same separation as in the CCC or DCC and introduces the regime switching variable only in the correlation coefficients. By doing so, he can proceed with the two-step procedure to estimate the model while limiting the number of parameters to be estimated.\(^{17}\) We carry out a similar idea but for nonlinear dependence.

Therefore, we separate the modeling of marginal distributions from the modeling of dependence by using univariate GARCH models for the marginal distributions and introducing changes in regime in the copula dependence structure. The pattern of the model with four variables (two countries, two markets in our following application) is illustrated in Fig. 3. The four marginal distributions are linked through a dependence function with two regimes, one symmetric, the other asymmetric.

\(^{15}\) The tail dependence coefficients are easily defined through a copula as \( t^1 = \lim_{u \to 1} c(u, u) \) and \( t^0 = \lim_{u \to 0} c(u, u) \).

\(^{16}\) The models proposed by Rodriguez (2007) in his analysis of contagion can reproduce asymmetric dependence but it cannot distinguish between skewness and asymmetry in the dependence structure. In fact, a change in regime produces both skewness and asymmetric dependence, two different features that must be characterized separately. The analysis is limited to pairs of stock markets in Asia and Latin America.

\(^{17}\) Since Pelletier (2006) uses the normal distribution with constant mean, the resulting unconditional distribution is symmetric and cannot reproduce asymmetric dependence.
3.2. Specification of the marginal distributions

For marginal distributions, we use a M-GARCH (1, 1) model similar to Heston and Nandi (2000):

$$x_{it} = \mu_i + \lambda_i \sigma_{it}^2 + \sigma_{it} z_{it}; \quad z_{it} \sim N(0, 1); \quad i = 1, \ldots, 4$$ (3.4)

$$\sigma_{it}^2 = \omega_i + \beta_i \sigma_{it-1}^2 + \gamma_i (x_{it-1} - \mu_i)^2$$ (3.5)

The variables $x_{it}$ and $x_{at}$ represent the log returns of equities and bonds respectively for the first country while $x_{at}$ and $x_{at}$ are the corresponding series for the second country; $\sigma_{it}^2$ denotes the conditional variance of $x_{it}$; $\lambda_i$ can be interpreted as the price of risk and $\gamma_i$ captures potential asymmetries in the volatility effect. In the Heston and Nandi (2000) interpretation, $\mu_i$ represents the interest rate. The parameters of the marginal distributions are grouped into one vector $\delta \equiv (\delta_1, \ldots, \delta_4)$, with $\delta_i = (\mu_i, \lambda_i, \omega_i, \beta_i, \gamma_i)$.

3.3. Specification of the dependence structure

Our dependence model is characterized by two regimes, one Gaussian regime in which dependence is symmetric ($C_G$) and a second regime that can capture the asymmetry in extreme dependence ($C_A$). The conditional copula is given by:

$$C(u_1, \ldots, u_4; \rho^S, \rho^A | s_t) = s_t C_G(u_1, \ldots, u_4; \rho^S) + (1 - s_t) C_A(u_1, \ldots, u_4; \rho^A)$$ (3.6)

where $u_{it} = F_p(x_{it} - \delta_i)$ with $F_p$ denoting the conditional cumulative distribution function of $X_{it}$ given the past observations. The variable $s_t$ follows a Markov chain with a constant transitional probability matrix.

$$M = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \quad P = \Pr(s_t = 1 | s_{t-1} = 1) \quad Q = \Pr(s_t = 0 | s_{t-1} = 0)$$ (3.7)

The normal regime ($s_t = 1$) corresponds to the symmetric regime where the conditional joint normality can be supported and the asymmetric regime ($s_t = 0$) corresponds to the asymmetric regime in which markets are strongly more dependent for large negative returns than for large positive returns.

The Gaussian copula $C_G$ is defined straightforwardly by (3.1) where the joint distribution $F = \Phi_{p^S}$ is the four-dimensional normal cumulative distribution function with all diagonal elements of the covariance matrix equal to one, i.e. $C_G(u_1, \ldots, u_4; \rho^S) = \Phi_{p^S}$.

The asymmetric components of the copula are illustrated in Fig. 4. The first one is characterized by independence between the two countries, but possibly extreme dependence between equities and bonds for each country. The second one is characterized by independence between equity and bond markets but allows for extreme dependence between equity returns and bond returns separately. The third one allows for possible extreme dependence between bonds in one country and equities in another country but supposes independence for the rest.

Formally, the asymmetric copula is the mixture of these three components and is expressed as follows:

$$C_A(u_1, \ldots, u_4; \rho^A) \equiv \pi_1 C_{CGS}(u_1, u_2; t_1^A) \times C_{CGS}(u_3, u_4; t_2^A) + \pi_2 C_{CGS}(u_1, u_2; t_1^A) \times C_{CGS}(u_3, u_4; t_4^A) + (1 - \pi_1 - \pi_2) C_{CGS}(u_1, u_4; t_3^A) \times C_{CGS}(u_2, u_3; t_5^A)$$ (3.8)

with $\rho^A = (\pi_1, \pi_2, t_1^A, t_2^A, t_3^A, t_4^A, t_5^A)$, and the bivariate component is the Gumbel survival copula given by

$$C_{CGS}(u, v; t^A) = u + v - 1 + \exp \left[ \frac{(1 - \log(1 - u))^{\theta(t^A)}}{\exp(t^A)} + \frac{(1 - \log(1 - v))^{\theta(t^A)}}{\exp(t^A)} \right]$$ (3.9)

where $\theta(t^A) = \frac{\log(2)}{\log(\exp(t^A))}. t^A \in [0, 1]$ is the lower TDC and the upper TDC is zero. As illustrated by Fig. 1, the choice of this particular copula is justified by its ability to replicate the pattern of dependence observed in the data. The shape of the exceedance correlation for the rotated Gumbel copula is more similar to the asymmetric shape observed in the data, while as expected Gaussian or RS with Gaussian distributions fail to replicate this shape.

One can notice that our asymmetric copula specification implies some restrictions in the dependence structure. For three different couples from different components of this copula, the sum of their TDC is lower than one. Without any restrictions this sum may reach 3. For example, the TDC between bonds and equities in the first country is $\pi_1 t_1^A$, between equities of two countries $\pi_2 t_3^A$, and between equities in the first country and bonds in the second country $(1 - \pi_1 - \pi_2) t_5^A$. Therefore, the sum is $\pi_1 t_1^A + \pi_2 t_3^A + (1 - \pi_1 - \pi_2) t_5^A \leq 1$, since $t_1^A \leq 1$, $t_3^A \leq 1$, and $t_5^A \leq 1$. We explore the empirical implications of these restrictions in Section 4.4. A major problem in building multivariate distributions is how to

---

19 The condition $\beta_i + \gamma_i < 1$ is sufficient to have the stationarity of the process $x_{it}$ with finite unconditional mean and variance (see Heston and Nandi, 2000).

19 Here we keep $\mu_i$ as a free parameter to give more flexibility to our model.
construct multivariate copulas with specific bivariate marginal distributions. A theorem by Genest et al. (1995) states that it is not always possible to construct multivariate copulas with given bivariate margins. Therefore, even if in the bivariate case we can have a nice asymmetric copula with lower tail dependence and upper tail independence as Longin and Solnik (2001) suggest, some problems remain when we contemplate more than two series. Most existing asymmetric tail-dependent copulas are in the family of Archimedean copulas and the usual straightforward generalization in multivariate copulas constrains all bivariate marginal copulas to be the same. This is clearly not admissible in the context of our analysis. In the above model, we allow each of the six couples of interest to have different levels of lower TDC. As $C_4$ is constructed, it is easy to check that it is a copula since each component of the mixture is a copula and the mixture of copulas is a copula.

It is important to notice that, in this model, the labeling of each regime is defined ex-ante. The normal regime ($s_t = 1$) corresponds to the symmetric regime where the conditional joint normality can be supported and the asymmetric regime ($s_t = 0$) corresponds to the asymmetric regime in which markets are strongly more dependent for large negative returns than for large positive returns.

3.4. An adapted parsimonious model

Given our application, we impose an additional constraint: $\pi_1 + \pi_2 = 1$. This means that we neglect the asymmetric cross-dependence between equities in one country and bonds in another country. However, it should be stressed that we maintain cross-country dependence through the normal regime. In Section 4.4, we investigate the effect of this constraint on the overall level of dependence measured by the Spearman rho, a nonparametric measure of dependence. The mixed copula becomes:

$$C_{\rho}(u_1, \ldots, u_4; \rho) \equiv \pi C_{C_{0S}}(u_1, u_2; \tau_1^2) \times C_{C_{0S}}(u_3, u_4; \tau_2^2) + (1 - \pi) C_{C_{0S}}(u_1, u_3; \tau_1^2) \times C_{C_{0S}}(u_2, u_4; \tau_2^2)$$  \hspace{1cm} (3.10)

Therefore, the asymmetric copula is now characterized by just five parameters $\rho = (\pi, \tau_1^2, \tau_2^2, \tau_3^2, \tau_4^2)$.

3.5. Estimation

As already mentioned, our structure allows for a two-step estimation procedure. The likelihood function must be evaluated unconditionally to the unobservable regime variable $s_t$ and decomposed in two parts. Let us denote the sample of observed data by $\mathbf{X} = \{X_t, \ldots, X_T\}$, where $X_t = \{x_{1t}, \ldots, x_{4t}\}$. The log likelihood function is given by:

$$L(\delta, \theta; \mathbf{X}_t) = \sum_{t=1}^{T} \log f(X_t; \delta, \theta | X_{-1})$$  \hspace{1cm} (3.11)

where $X_{-1} = \{X_{1t}, \ldots, X_{-1}\}$ and $\theta$ is a vector including the parameters of the copula and the transition matrix. Hamilton (1989) describes a procedure to perform this type of evaluation. With $\mathbf{z}_t = (s_t, 1 - s_t)$ and denoting

$$\eta_t = f(X_t; \delta, \theta | X_{-1}, s_t = 1) / f(X_t; \delta, \theta | X_{-1}, s_t = 0)$$  \hspace{1cm} (3.12)

the density function conditionally to the regime variable $s_t$ and the past returns can be written as:

$$f(X_t; \delta, \theta | X_{-1}, s_t) = \mathbf{z}_t^\top \eta_t$$  \hspace{1cm} (3.13)

Since $s_t$ (or $\mathbf{z}_t$) is unobservable, we integrate on $s_t$ and obtain the unconditional density function:

$$f(X_t; \delta, \theta | X_{-1}) = \Pr[s_t = 1 | X_{-1}; \delta, \theta] \times f(X_t; \delta, \theta | X_{-1}, s_t = 1) + \Pr[s_t = 0 | X_{-1}; \delta, \theta] \times f(X_t; \delta, \theta | X_{-1}, s_t = 0)$$  \hspace{1cm} (3.14)

The conditional probabilities of being in different regimes at time $t$ conditional on observations up to time $t - 1$, denoted by $\tilde{z}_{t, t-1}$, are computed through the Hamilton filter. Starting with the initial value $\tilde{z}_{1,0}$, the optimal inference and forecast for each date in the sample is given by the iterative equations:

$$\tilde{z}_{t, t-1} = [\tilde{\mathbf{z}}_{t, t-1} \eta_t]^{-1} \tilde{\mathbf{z}}_{t, t-1} \circ \eta_t$$  \hspace{1cm} (3.15)

$$\tilde{z}_{t, t-1} = M^{-1} \tilde{z}_{t, t-1}$$  \hspace{1cm} (3.16)

where $\circ$ denotes element-by-element multiplication. Finally, the unconditional density can be evaluated with the observed data as:

$$L(\delta, \theta; \mathbf{X}_t) = \sum_{t=1}^{T} \log \tilde{z}_{t, t-1} \circ \eta_t$$  \hspace{1cm} (3.17)

21 Nelsen (1999, p. 86) mentions that it may be the most important open question concerning copulas. Aas et al. (2007) propose an approach to build multiple dependence based on copula decomposition. Their approach proceeds by a hierarchical incorporation of more variables in the conditioning sets. This procedure provides a nice way to build flexible multivariate copula. In practice it is important to make a good choice of copulas that should be used in the first level of the hierarchy since a limitation of the procedure is that the copulas after the first level of the hierarchy are based on conditional copulas.

22 A copula can be seen as the cdf of a multidimensional variable with uniform $[0,1]$ margins. If we consider two bivariate independent variables with uniform margins the copula linking the four variables is simply the product of the two bivariate copulas. Hence, such a product is always a copula.

23 A related study by Hartmann et al. (2004) using extreme value theory tends to support this restriction. Analyzing stock and bond returns for G-5 countries, they find that extreme dependence between stocks and bonds is much lower than extreme dependence between stock markets or bond markets. This is especially the case for cross-country dependence between stocks in one country and bonds in another country.

24 A general presentation can be found in Hamilton (1994, chapter 22).
To perform the two-step procedure, we decompose the log likelihood function into two parts: the first part includes the likelihood functions of all margins, while the second part represents the likelihood function of the copula.

**Proposition 3.2 (Decomposition of the log likelihood function).** The log likelihood function can be decomposed into two parts including the margins and the copula

\[
L(\delta, \theta; X_t) = \sum_{i=1}^{4} L_i(\delta_i; X_t) + L_C(\delta, \theta; X_t)
\]

where

\[
X_t = \{x_1, \ldots, x_t\}
\]

\[
L_i(\delta_i; X_t) = \sum_{t=1}^{T} \log f_i(x_t; \delta_i|x_{t-1})
\]

\[
L_C(\delta, \theta; X) = \sum_{t=1}^{T} \log \left( \frac{\hat{z}_{t+1}}{\hat{z}_{t+1}^c} \right)
\]

with

\[
\eta_{i \alpha} = \left[ \begin{array}{c} c(u_{1,1}(\delta_1), \ldots, u_{n,1}(\delta_1); \theta | S_t = 1) \\ c(u_{1,1}(\delta_1), \ldots, u_{n,1}(\delta_1); \theta | S_t = 0) \end{array} \right]; \quad u_{i2}(\delta_i) = F_i(x_t; \delta_i|x_{t-1})
\]

and \( \hat{z}_{t+1} \) filtered from \( \eta_{i \alpha} \) as

\[
\hat{z}_{t+1} = \left( \frac{\hat{z}_{t+1}^c}{\eta_{i \alpha}} \right)^{-1} \left( \frac{\hat{z}_{t+1}}{\eta_{i \alpha}} \right)
\]

\[
\hat{z}_{t+1}^c = M^t \cdot \hat{z}_{t+1}
\]

**Proof.** See Appendix A.

Several options are available for the estimation of the initial value \( \hat{z}_{10} \). One approach is to set it equal to the value of the unconditional probabilities, which is the stationary transitional probability of the Markov chain. Another simple option is to set \( \hat{z}_{10} = N^{-1}1_N \). Alternatively it could be considered as another parameter, which will be estimated subject to the constraint that \( 1_N^t \hat{z}_{10} = 1 \). We will use the first option here.

Through the above decomposition, we notice that each marginal log likelihood function is separable from the others. Therefore, even if the estimation of all margins is performed in a first step, we can estimate each set of marginal parameters separately into this step. The first step is then equivalent to a single estimations of univariate distributions. The two-step estimation is formally written as follows:

\[
\hat{\lambda} = \arg \max_{\lambda \in \Theta} \sum_{i=1}^{4} L_i(\delta_i; X_t),
\]

\[
\hat{\Theta} = \arg \max_{\delta, \theta} L_C(\delta, \theta; X).
\]

The estimator for the parameters of the marginal distributions is then \( \hat{\delta} = (\hat{\delta}_1, \ldots, \hat{\delta}_4) \), with \( \hat{\delta}_i = (\hat{\mu}_i, \hat{\lambda}_i, \hat{\omega}_i, \hat{\beta}_i, \hat{\gamma}_i) \); and \( \hat{\theta} = (\hat{p}_N, \hat{p}_A, \hat{p}_B, \hat{Q}) \) includes all estimators of the parameters involved in the dependence structure. \( \lambda \) and \( \Theta \) represent the sets of all possible values of \( \delta \) and \( \theta \) respectively.

**3.6 Tests of dependence structure**

We perform three different tests for dependence structure. The first one is the Longin and Solnik (2001) test for exceedence correlation. For this test, we focus on the lower tail dependence to compare exceedence correlation of return data, the simulated return from two RS with Gaussian distributions, and the simulated return from a Gumbel copula model.

We also test the asymmetry in dependence. While the proposed copula model has the potential to capture asymmetry in dependence, we need to test formally for the presence of such asymmetric dependence. The natural way to examine whether dependence is asymmetric is to test the null hypothesis of one normal copula regime \( (H_0: (P = 1 \text{ and } Q = 0)) \), where \( P \) and \( Q \) are the parameters of the transition probability matrix, against the alternative hypothesis of two-copula regimes including the normal one and the asymmetric one. This test faces the general problems found in testing in RS models. In particular, under the null hypothesis, nuisance parameters are unidentified and the scores are identically zero.

Maximized Monte Carlo (MMC) tests of Dufour (2006), which are a generalization of classical Monte Carlo (MC) tests of Dwass (1957) and Barnard (1963), are adapted for tests facing these problems. The MC tests of Dwass (1957) and Barnard (1963) are performed by doing many replications (with the same sample size as the data sample) under the null hypothesis, and compute the test statistic for each replication. The distribution of the test statistic is therefore approximated by the distribution of the obtained values. One can therefore compute the value of the test statistic with the data and deduce from the MC distribution the \( p \)-value of the test. The classical MC test does not deal with the presence of nuisance parameters under the null hypothesis. The MMC of Dufour (2006) addresses the problem of nuisance parameters under the null. When the test statistic involves the nuisance parameters as in the case of the likelihood ratio test under the alternative, the values of these parameters are needed to compute the test statistic on simulated data.

The MMC technique is the maximization of the \( p \)-values given all the possible values of the nuisance parameters. This test is computationally very demanding. However, Dufour (2006) proposes a simplified version that focuses on the estimated values of the nuisance parameters and shows that it works under the assumptions of uniform continuity, and convergence over the nuisance parameter space. Our model satisfies these assumptions of uniform continuity and convergence. Therefore, we can apply this simpler version also known as parametric bootstrap test.

Finally, we compare the goodness-of-fit of this model with two different structures. Chen and Fan (2005) propose a test to compare non-nested copula models. We apply this test to compare our model with a regime switching model where both regimes display Gaussian copulas. We also compare our model with a model with just one regime (of Gumbel copula) which displays an asymmetric dependence behavior to make sure that it is necessary to use in addition, a symmetric regime (of Gaussian copula).

**4. Dependence structure in international bond and equity markets: an empirical investigation**

**4.1 Data**

We will consider the same model for two pairs of two countries. First, we model the equity and bond markets in the United States and Canada. The US equity returns are based on the SP 500 index, while the Canadian equity returns are computed with the Datastream index. The bond series are indices of 5-year government bonds computed by Datastream. These bond indices are available.
daily and are chain-linked allowing the addition and removal of bonds without affecting the value of the index. We also consider France and Germany as a pair of countries. An additional interest here will be to see how the introduction of the European common currency changed the dependence structure between the asset markets in these two countries. The bond indices are the Datastream 5-year government bond indices, while the equity indices are the MSCI series. All returns are total returns and are expressed in US dollars on a weekly basis from January 01, 1985 to December 21, 2004, which corresponds to a sample of 1044 observations. Descriptive statistics are reported in Table 1.

Sharpe ratios appear to be of the same magnitude for both equities and bonds, around 0.6 on average for the first and slightly above 1 for the second. The United States exhibits the highest ratios among the four countries. All return series present negative skewness except for the French bond index. Both mean returns and return volatility are higher in France and Germany than in the US and Canada. The volatility of returns in France and Germany is more than 23%, while it is only 18% in the US and Canada.

Unconditional correlations are reported in Table 2. The US and Canadian markets exhibit relatively high correlations, 0.72 for equities and 0.5 for bonds. The same is true for the France–Germany pair, although the bond markets are tightly linked, with a correlation of 0.94. The North-American equity markets are less correlated with European equity markets (around 0.2) than their bond counterparts (around 0.32). The cross-correlations between equity and bond markets vary from country to country. On average the two markets seem to move independently in the United States, while they are more closely related in Canada (0.44) and in Europe (around 0.3 for both France and Germany). Cross-correlations between equities and bonds in two different countries are not very high for US and Canada, and of the same order of magnitude than within-country cross-correlations (0.3) for France and Germany.

### 4.2. Marginal distributions

The estimates of the marginal parameters are reported in Table 3. The large values for the $\beta$ parameters (around 90%) capture the high persistence in volatility. The high degree of significance for the parameter $\lambda$ indicates that asset returns are skewed.

One important assumption for these GARCH models is that the error terms are i.i.d. Therefore, to verify if the assumption is fulfilled, we perform some tests of independence and normality on the residuals. The test results in Table 4 suggest that the independence assumption of residuals cannot be rejected for all series with a good degree of confidence.

### 4.3. Dependence structure in bond and equity markets

Three main conclusions emerge from the empirical results. First, there appears to be a large extreme cross-country dependence for both the equity and bond markets, while there is little dependence between equities and bonds in the same country. Second, the dependence structure exhibits a strong nonlinearity. Third, there seems to be a link between exchange rate volatility and asymmetry of dependence.

#### 4.3.1. US–Canada dependence structure

In Table 5, we report the results of estimating the dependence model described in Section 3.4. The cross-country extreme dependence is large in both equity and bond markets, but the dependence across the two markets is relatively low in both countries. In the asymmetric regime, the TDCs are larger than 54% in both bond–bond and equity–equity markets, while both equity–bond TDCs in US and Canada are lower than 2%. This observation has an important implication for international diversification. The fact that extreme dependence in international equity and bond markets is larger than national bond–equity dependence can have a negative effect on the gain of international diversification and encourage the switching from equity to the domestic bond or risk-free asset in case of bear markets.

The average absolute value of correlation in the normal regime is larger than 39% for cross-country dependence and lower than 41% for equity–bond dependence. In the last case the correlation between bonds and equities in Canada is unusually high. The results underline the differences between unconditional correlation and the correlation in the normal regime. In fact, the presence of extreme

### Table 1

Summary statistics of weekly bond and equity index returns for the four countries. All returns are in US dollars, from January 01, 1985 to December 21, 2004, which corresponds to a sample of 1044 observations. Sharpe ratio represents the ratio of the mean over the standard deviation of return.

<table>
<thead>
<tr>
<th></th>
<th>Mean</th>
<th>Std</th>
<th>Kurtosis</th>
<th>Skewness</th>
<th>Min</th>
<th>Max</th>
<th>Sharpe ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td>US equity</td>
<td>13.67</td>
<td>17.51</td>
<td>17.00</td>
<td>–1.55</td>
<td>–680.36</td>
<td>311.10</td>
<td>0.78</td>
</tr>
<tr>
<td>US bond</td>
<td>7.57</td>
<td>4.69</td>
<td>0.67</td>
<td>–0.06</td>
<td>–66.91</td>
<td>58.81</td>
<td>1.61</td>
</tr>
<tr>
<td>CA equity</td>
<td>11.24</td>
<td>16.72</td>
<td>13.62</td>
<td>–1.67</td>
<td>–610.87</td>
<td>225.15</td>
<td>0.67</td>
</tr>
<tr>
<td>CA bond</td>
<td>8.81</td>
<td>8.15</td>
<td>1.13</td>
<td>–0.24</td>
<td>–130.55</td>
<td>118.07</td>
<td>1.08</td>
</tr>
<tr>
<td>FR equity</td>
<td>14.72</td>
<td>23.43</td>
<td>7.18</td>
<td>–0.09</td>
<td>–582.12</td>
<td>512.16</td>
<td>0.63</td>
</tr>
<tr>
<td>FR bond</td>
<td>11.52</td>
<td>11.16</td>
<td>0.92</td>
<td>0.04</td>
<td>–142.02</td>
<td>166.68</td>
<td>1.03</td>
</tr>
<tr>
<td>DE equity</td>
<td>12.57</td>
<td>24.97</td>
<td>8.01</td>
<td>–0.46</td>
<td>–574.96</td>
<td>463.08</td>
<td>0.50</td>
</tr>
<tr>
<td>DE bond</td>
<td>10.44</td>
<td>11.56</td>
<td>0.82</td>
<td>–0.01</td>
<td>–142.54</td>
<td>171.39</td>
<td>0.90</td>
</tr>
</tbody>
</table>

* Annualized percent.

### Table 2

Unconditional correlations between bonds and equity for US, Canada (CA), France (FR), and Germany (DE).

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>US bond</td>
<td>0.0576</td>
<td>0.0116</td>
<td>0.4769</td>
<td>0.4392</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CA equity</td>
<td>0.7182</td>
<td>0.0116</td>
<td>0.4769</td>
<td>0.4392</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CA bond</td>
<td>0.1783</td>
<td>0.0116</td>
<td>0.4769</td>
<td>0.4392</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>FR equity</td>
<td>0.1957</td>
<td>–0.0182</td>
<td>0.1974</td>
<td>0.1065</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>FR bond</td>
<td>–0.0499</td>
<td>0.3386</td>
<td>–0.0080</td>
<td>0.2433</td>
<td>0.3066</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>DE equity</td>
<td>0.2089</td>
<td>–0.0536</td>
<td>0.1995</td>
<td>0.1009</td>
<td>0.8099</td>
<td>0.2625</td>
<td></td>
<td></td>
</tr>
<tr>
<td>DE bond</td>
<td>–0.0832</td>
<td>0.3081</td>
<td>–0.0234</td>
<td>0.2143</td>
<td>0.3084</td>
<td>0.9403</td>
<td>0.2847</td>
<td></td>
</tr>
</tbody>
</table>
dependence in the negative returns explains this difference since the multivariate Gaussian distribution has independence in the tails of returns regardless of the level of correlation.

The separation of the distribution into two parts, including the normal regime and the asymmetric regime, allows to capture the strong nonlinear pattern in the dependence structure. Moreover, it is interesting to see that for a high unconditional correlated couple such as the US and Canada equity markets, this separation gives not only an extreme dependence for the asymmetric regime, but also a high correlation in the normal regime (87%) that appears larger than the unconditional correlation (72%). This result may seem counter-intuitive if we take the unconditional correlation as a “mean” of the correlations in the two regimes. Of course, one must realize that the asymmetric regime can be characterized by a low correlation but by a large TDC. This demonstrates the importance of distinguishing between correlation and extreme dependence. The mixture model is better able to capture this distinction in fitting the data. A normal distribution may be a good approximation for measuring finite distance dependence, but an appropriate copula structure is necessary for characterizing extreme dependence.

### 4.3.2. France–Germany dependence structure

The estimation results are shown in Table 6. Due to a high cross-country unconditional correlation in both markets, the results for France and Germany are more eloquent. The dependence between equities and bonds is low, while the dependence between assets of the same type is large in both regimes. For France and Germany, equity–equity correlation and bond–bond correlation are larger than 90% while bond–equity correlations are lower than 21% in the same country as well as between the two countries. In the asymmetric regime, the TDC are larger than 67% between assets of the same country as well as between the two countries. In the normal regime, the TDC are larger than 90% while bond–equity correlations are lower than 21% in the same country as well as between the two countries. In the asymmetric regime, the TDC are larger than 67% between assets.
of the same type and lower than 2% between bond and equities in both France and Germany.  

To analyze the effect of the Euro on the dependence structure, we split the observation period in two subperiods, before and after the introduction of the currency. Tables 7 and 8 contain the results for the respective subperiods. We find that the introduction of the Euro increases the correlation in the normal regime between the French and German markets. Before the introduction of the Euro, in the normal regime, the cross-country correlation between assets of the same type is on average 80%, against more than 96% after the introduction. The cross-asset correlations exhibit a similar pattern since all correlations increase after the introduction of the Euro. This result is consistent with those of Cappiello et al. (2006) who find that the introduction of a fixed exchange rate leads to a structural break characterized by a high correlation.  

The filtered probabilities to be in asymmetric regime are obtained as a by-product of estimation. They provide at each time period a probabilistic assessment of being in the asymmetric regime conditional on the information available at time \( t \) (Fig. S). For

### Table 5

<table>
<thead>
<tr>
<th>Normal regime</th>
<th>Asymmetric regime</th>
<th>Tail dependence coefficient</th>
<th>( \tau )</th>
<th>( TDC ((1 - \pi) \tau) )</th>
</tr>
</thead>
<tbody>
<tr>
<td>US equity–CA equity</td>
<td>0.8739</td>
<td>0.9100</td>
<td>0.7917</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.1560)</td>
<td>(0.0185)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>US bond–CA bond</td>
<td>0.3870</td>
<td>0.6234</td>
<td>0.5424</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.0831)</td>
<td>(0.0124)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1 - ( \pi )</td>
<td>0.6897</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

### Table 6

<table>
<thead>
<tr>
<th>Normal regime</th>
<th>Asymmetric regime</th>
<th>Tail dependence coefficient</th>
<th>( \tau )</th>
<th>( TDC ((1 - \pi) \tau) )</th>
</tr>
</thead>
<tbody>
<tr>
<td>FR equity–DE equity</td>
<td>0.9083</td>
<td>0.9554</td>
<td>0.7787</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.0267)</td>
<td>(0.0603)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>FR bond–DE bond</td>
<td>0.9901</td>
<td>0.8261</td>
<td>0.6733</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.058)</td>
<td>(0.027)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1 - ( \pi )</td>
<td>0.8151</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

### Table 7

<table>
<thead>
<tr>
<th>Normal regime</th>
<th>Asymmetric regime</th>
<th>Tail dependence coefficient</th>
<th>( \pi )</th>
<th>( TDC (\pi) )</th>
</tr>
</thead>
<tbody>
<tr>
<td>US bond</td>
<td>CA bond</td>
<td>0.1234</td>
<td>0.1300</td>
<td>0.0169</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.0416)</td>
<td>(0.041)</td>
<td></td>
</tr>
<tr>
<td>CA equity</td>
<td>0.4085</td>
<td>0.1385</td>
<td>0.0180</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.0207)</td>
<td>(0.0103)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.3102</td>
<td>(0.0207)</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

### Table 8

<table>
<thead>
<tr>
<th>Normal regime</th>
<th>Asymmetric regime</th>
<th>Tail dependence coefficient</th>
<th>( \pi )</th>
<th>( TDC (\pi) )</th>
</tr>
</thead>
<tbody>
<tr>
<td>FR bond</td>
<td>DE bond</td>
<td>0.1893</td>
<td>0.1209</td>
<td>0.0171</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.0170)</td>
<td>(0.0129)</td>
<td></td>
</tr>
<tr>
<td>DE equity</td>
<td>0.1175</td>
<td>0.1294</td>
<td>0.0179</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.0214)</td>
<td>(0.030)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.1849</td>
<td>(0.0294)</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Parameters of transitional probability matrix

<table>
<thead>
<tr>
<th>( P )</th>
<th>( Q )</th>
<th>( P )</th>
<th>( Q )</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.9020</td>
<td>0.9586</td>
<td>0.1838</td>
<td>0.9373</td>
</tr>
<tr>
<td>(0.0207)</td>
<td>(0.0206)</td>
<td>(0.0270)</td>
<td>(0.0373)</td>
</tr>
</tbody>
</table>
France and Germany, these probabilities show a clear break after the introduction of the Euro. Before its introduction, the dependence is more likely asymmetric and becomes more Gaussian after the event. To investigate this relationship further, we perform a logistic regression of the conditional probabilities to be in the asymmetric regime on the volatility of the exchange rate.29

For France and Germany, we have:

$$\hat{P}_t = \frac{e^{a \cdot Vol_t + b \cdot e_t + c_t}}{1 + e^{a \cdot Vol_t + b \cdot e_t + c_t}}$$

The dependent variable is $\hat{P}_t = \log(P_t/(1 - P_t))$, where $P_t$ is the conditional probability to be in the asymmetric regime given the time-$t$ available information, and $Vol_t$ is the exchange rate volatility between the two countries obtained by a M-GARCH (1,1) filter. Standard deviations are reported between parentheses. The coefficient is positive and highly significant (the R$^2$ of the regression is 0.86) suggesting a strong relationship between exchange rate volatility and asymmetric dependence.

One may be concerned that the relation between the probability of the asymmetric regime and exchange rate volatility is due to the correlation of the latter with the volatility of the equity market or the bond market of the respective countries. To address this issue, we perform an orthogonalization. We regress the exchange rate on all equity and bond return volatilities in a first step and keep residuals. Then we perform an orthogonalization. We regress the probability of asymmetric regime on these residuals in a second step. The relation remains significant.

We run the same regression for US and Canada to investigate if the relation holds when no structural change occurs. The results are similar to the European results.

$$\hat{P}_t = \frac{e^{a \cdot Vol_t + b \cdot e_t + c_t}}{1 + e^{a \cdot Vol_t + b \cdot e_t + c_t}}$$

The R$^2$ of the regression remains high at 0.75.

The fact that high exchange rate volatility is associated with asymmetric dependence appears to be consistent with the results in the literature, since asymmetric dependence was mainly found to be present in international equity markets (see Longin and Solnik, 2001). Our own results suggest the presence of asymmetric dependence in international bond markets as well.

Intuitively, the persistence of each dependence regime depends on the persistence of exchange rate volatility. A high exchange rate volatility increases extreme comovements. When bad news in a country combine with a very active currency market, transmission through the latter makes downside joint movements more likely than in a fixed exchange rate regime. This may provide an insight about the strong change in the persistence of different regimes after the introduction of the Euro. Before the exchange rate between the French Franc and the German Deutsch Mark was especially volatile and this may explain a strong persistence in the asymmetric regime when the model is estimated over this subperiod. After the introduction of the Euro, the volatility is reduced to zero and the normal regime becomes the only persistent regime. By putting the two subperiods together, both regimes appear persistent, which is consistent with our explanation.

These results are consistent with Cappiello et al. (2006) who find a structural break in the dependence structure of European markets after the introduction of the Euro. They find an increase to a near...
Table 9
Dependence structure tests. Panel A gives results for Dufour (2006) MMC tests of asymmetric dependence. LR is the likelihood ratio statistic computed from the data. The p-value is obtained from 1000 Monte Carlo repetitions with size 1043 (equal to the sample size) each. Panel B presents Chen and Fan (2005) pseudo-likelihood ratio tests for comparing results from competing multivariate copula models. We compare three models: (i) our regime-switching model with one Gaussian copula regime and one mixture of rotated Gumbel copula regime (N-G RS), (ii) the regime switching with two Gaussian regimes (Norm RS), and (iii) our four-variable mixture of rotated Gumbel copula (Gumbel). For each pair of models the best model at 1% level is indicated in the corresponding cell.

<table>
<thead>
<tr>
<th></th>
<th>US–CA</th>
<th>FR–DE</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Panel A: Dufour (2006) MMC test</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>LR</td>
<td>0.0731</td>
<td>0.7889</td>
</tr>
<tr>
<td>p-value</td>
<td>0.0090</td>
<td>0.0000</td>
</tr>
<tr>
<td>N-G RS (Decision at 1% level)</td>
<td>Norm RS</td>
<td></td>
</tr>
<tr>
<td>France–Germany</td>
<td></td>
<td></td>
</tr>
<tr>
<td>N-G RS</td>
<td>–1.3445</td>
<td>–1.3022</td>
</tr>
<tr>
<td>Gumbel</td>
<td>Norm RS</td>
<td>–0.95229</td>
</tr>
<tr>
<td>Gumbel</td>
<td>N-G RS</td>
<td></td>
</tr>
<tr>
<td>Gumbel</td>
<td>–4.7217</td>
<td>–3.2309</td>
</tr>
<tr>
<td></td>
<td>Norm RS</td>
<td></td>
</tr>
<tr>
<td></td>
<td>–4.1763</td>
<td></td>
</tr>
</tbody>
</table>

perfect correlation. However, due to the fact that the dynamic conditional correlation model they use is based on the normal distribution the correlation before the introduction of Euro is misleading since as we find, the dependence between European countries was more asymmetric and therefore the dependence was more in the tail and cannot be completely captured by correlation.

4.4. Empirical implications of the mixture copula constraints

Our mixture copula model aims at capturing extreme dependence, which is of course very useful if an investor wants to manage extreme risks.30 In Fig. 1, we show that the high exceedance correlation present in the data for the Canada–US pair of equity returns is rather well captured by the model. However, in building the asymmetric copula model to capture tail dependence, we impose some restrictions, either by construction or for parsimony reasons, that may affect the overall level of dependence between pairs of variables. These restrictions are explained in Section 3.3. Chollette et al. (2009) translate these restrictions in terms of Spearman correlation in comparing their canonical vine copulas to our mixture copulas.31

4.4.1. Spearman correlation

For two variables X and Y, the Spearman’s rho is based on concordance and discordance and is given by the formula

\[
\rho_S(X, Y) = 3\left(\text{Pr}([X_1 - X_2][Y_1 - Y_2] > 0) - \text{Pr}([X_1 - X_2][Y_1 - Y_2] < 0)\right)
\]

where \((X_1 - X_2), (X_2 - Y_2),\) and \((X_2 - Y_2)\) are three independent random vectors with the same joint distribution like \((X, Y)\), which is defined by the copula \(C\) and corresponding marginal distributions. The Spearman rho is related to the copula by the formula32

\[
\rho_S(X, Y) = \rho_S(C) = 12 \int \int C(u, v) du dv - 3
\]

For the Normal copula, the analytical expression is given by:

\[
\rho_S(C_{\text{norm}}(p)) = 1 - \sqrt{2 \pi} \text{arcsin}(\rho/2),
\]

where \(\rho\) is the correlation coefficient of a Normal distribution with copula component \(C_{\text{norm}}(p)\). For the Gumbel copula, there is no analytical expression, so we use numerical integration to compute the Spearman rho.

4.4.2. Parametric and nonparametric estimates of the Spearman correlation

Table 10 reports the estimates of the Spearman correlation obtained directly from the data (nonparametric estimate) as well as the estimates obtained from our model with a mixture copula with and without imposing a cross-country constraint of zero dependence between the bond and the equity markets (parametric estimates).33

One can say that for the two pairs of countries, Canada–US and France–Germany, the model estimates for the Spearman correlation between the equity markets and between the bond markets are roughly similar to what is observed in the data. For example, for France and Germany, the correlation between the equity markets is 0.84 in the data and 0.83 in the models with and without the constraint. For the correlation between the bond markets, the data estimate is 0.93 but 0.84 and 0.83 with and without the constraint respectively, that is a bit lower. However, where we see a big difference, is in the correlation between the equity and the bond markets, even in the same country. In France it is estimated at 0.49 with the data and at best at 0.10 with the model. This is quite a discrepancy that is due to our model construction. The effect is still present but less damaging in the Canada–US case, where such cross-market correlations are less important, except for the Canadian bond and equity markets. It is estimated at 0.46 in the data and at best at 0.22 in the model. The release of the cross-country constraint helps in getting a bit closer to the data estimates but it does not solve the major problem of underestimating the same country, cross-market correlation. One can conclude that capturing extreme dependence entails some costs and we find that this is specially the case of pairs with lower overall dependence. These costs have to be weighted with the gains of estimating rather well the dependence within markets across countries.

5. Conclusion

We propose a copula-based model of extreme dependence asymmetry that can rationalize the stylized facts put forward by Longin and Solnik. We apply it to the characterization of the extreme dependence in the equity and bond markets of two pairs of countries, the United States and Canada and France and Germany respectively. We capture the well-known strong asymmetric behavior across equity markets, but we also put forward a similar pattern in bond markets. The proposed model allows us to discover a relationship between the filtered probabilities to be in the asymmetric regime and the volatility of exchange rates. This is not possible with the extreme value approach of Longin and Solnik (2001) since only the tails of the distributions are modeled. While useful for extreme risk management, our model has limits for capturing dependence across markets, where the dependence is less strong.

30 See Tsafack (2009) for the importance of capturing extreme dependence while managing extreme risk.
31 Chollette et al. (2009) show that the restrictions become increasingly binding when the number of variables increases for mixture copulas. The canonical vine specification generalizes better to higher dimensions but not all applications lend themselves to finding a natural candidate for the first series on which everything is made conditional.
32 See Nelsen (1999) for more detail about this formulation.
33 For the results without the constraint, we had to estimate the model again by relaxing the zero dependence between the bond markets and the equity markets in different countries. For space considerations, we do not report the parameter estimates of this larger model. Estimation results are available from the authors upon request.
Since the exchange rate volatility may be a factor behind the asymmetric behavior of international equity and bond market dependence, it will be interesting to extend the model to incorporate the exchange rate in order to study the portfolio of an international investor. Moreover, the asymmetry put forward between positive and negative extreme returns suggests to investigate the behavior of an investor endowed with disappointment aversion preferences as in Ang et al. (2006).

Table 10
Unconditional Spearman rho between bonds and equity for US, Canada (CA), France (FR), and Germany (DE). Spearman rho from data is the non-parametric estimates using data, while the Spearman rho from models are estimated using ergodic probabilities from the Markov chain combined with the Spearman rho from each regime. The cross-country constraint assumes the independence between equity in one country and bond in the other for asymmetric regime.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Panel A: Spearman rho from data</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>US bond</td>
<td>0.1554</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CA equity</td>
<td>0.6631</td>
<td>0.0775</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CA bond</td>
<td>0.2002</td>
<td>0.4482</td>
<td>0.4590</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>FR equity</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.4888</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>FR bond</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.8370</td>
<td>0.4515</td>
<td></td>
<td></td>
</tr>
<tr>
<td>DE equity</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.5017</td>
<td>0.9324</td>
<td>0.4879</td>
<td></td>
</tr>
<tr>
<td>DE bond</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.5017</td>
<td>0.9324</td>
<td>0.4879</td>
<td></td>
</tr>
<tr>
<td>Panel B: Spearman rho from the model without cross-country constraint</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>US bond</td>
<td>0.0286</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CA equity</td>
<td>0.7307</td>
<td>0.0215</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CA bond</td>
<td>0.0982</td>
<td>0.4611</td>
<td>0.2207</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>FR bond</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.1025</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>DE equity</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.8292</td>
<td>0.0832</td>
<td></td>
<td></td>
</tr>
<tr>
<td>DE bond</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.1280</td>
<td>0.8440</td>
<td>0.0951</td>
<td></td>
</tr>
<tr>
<td>Panel C: Spearman rho from the model with cross-country constraint</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>US bond</td>
<td>-0.0094</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CA equity</td>
<td>0.7382</td>
<td>-0.0230</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CA bond</td>
<td>0.0350</td>
<td>0.5073</td>
<td>0.1407</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>FR bond</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.0640</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>DE equity</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.8349</td>
<td>0.0314</td>
<td></td>
<td></td>
</tr>
<tr>
<td>DE bond</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.0540</td>
<td>0.8297</td>
<td>0.0492</td>
<td></td>
</tr>
</tbody>
</table>

Fig. 5. The upper graph represents the probability to be in the asymmetric regime conditional on available information. The lower graph shows the exchange rate conditional volatility filtered with the M-GARCH (1,1) model.
Acknowledgements

The first author is a research Fellow of CIRANO and CIREQ. He gratefully acknowledges financial support from the Bank of Canada, the Fonds de la Formation de Chercheurs et l’Aide à la Recherche du Québec (FCAR), the Social Sciences and Humanities Research Council of Canada (SSHRC), and the MITACS Network of Centres of Excellence. Financial Support by CIREO and CIRANO is gratefully acknowledged by the second author. We thank a referee for comments which improve the quality of the paper, we also thank Jean Marie Dufour, Eric Girardin, Silvia Gonçalves, Eric Jacquier, Nour Meddahi, Benoit Perron, Eric Renault, Romeo Tedongap, Per Strömberg and participants at the Financial Management Association Doctoral Consortium at Chicago, the Conference on Multivariate Models for Volatility at Algarve, the Conference on Measuring Dependence in Economics and Finance at Cass Business School, the Seminar at London School of Economics, GREQAM-Marseille, Ryerson University, Suffolk University, and Stockholm School of Economics for their constructive comments and suggestions. We are grateful to Lorenzo Cappiello, Robert Engle and Kevin Sheppard for providing us with their dataset.

Appendix A. Proofs

A.1. Proof of Proposition 2.1

To prove this proposition, we need the two following lemmas:

Lemma 1. (a) Let \( f^{(1)} \) \( n \in \mathbb{N} \) be a family of symmetric multivariate density functions of \( n(\leq \infty) \) variables with same mean. The mixture \( f = \sum_{s=1}^{n} \pi_{s} f^{(1)}(\mu - \chi) \) where \( \sum_{s=1}^{n} \pi_{s} = 1 \) and \( \pi_{s} \geq 0 \) for any \( s \), is a symmetric multivariate density function. (b) Moreover for a continuum of symmetric multivariate density function \( f^{(1)}(\mu - \chi) \) with same mean, the mixture \( f = \int_{A} \pi_{s} f^{(1)}(\mu - \chi) d\sigma \) where \( \int_{A} \pi_{s} d\sigma = 1 \), is a symmetric multivariate density function.

Proof. Let \( \mu \) be the mean of all \( f^{(1)} \) (and all \( f^{(1)}(\mu - \chi) \))

\[
f(\mu - \chi) = \sum_{s=1}^{n} \pi_{s} f^{(1)}(\mu - \chi)
\]

by symmetry of all \( f^{(1)} \), we have, \( \sum_{s=1}^{n} \pi_{s} f^{(1)}(\mu - \chi) = \sum_{s=1}^{n} \pi_{s} f^{(1)}(\mu + \chi) = f(\mu + \chi) \) i.e. \( f(\mu - \chi) = f(\mu + \chi) \) and the part (a) follows. Similarly for mixture of continuum, \( f(\mu - \chi) = \int_{A} \pi_{s} f^{(1)}(\mu - \chi) d\sigma \) where \( \int_{A} \pi_{s} d\sigma = 1 \) and we have (b).

Lemma 2. Let \( \{ F^{(1)} \} \) \( n \in \mathbb{N} \) be a family of bivariate cdf with zero lower (upper) TDC. The mixture \( F = \sum_{s=1}^{n} \pi_{s} F^{(1)} \) where \( \sum_{s=1}^{n} \pi_{s} = 1 \) and \( \pi_{s} \geq 0 \) for any \( s \), is a bivariate density function with lower (upper) TDC.

Proof. We do the proof for lower tail since by “rotation” we have the same result for upper tail.

Let \( \tau_{F}^{L} \) be the lower TDC of \( F \), we have

\[
\tau_{F}^{L} = \lim_{x \to -0} \frac{F(X > x)}{F(X > y)}
\]

= \[ \lim_{x \to -0} \frac{F(X > x)}{F(Y > x)} \]

= \[ \lim_{x \to -0} \frac{F(X > x)}{F(Y > x)} \]

and since \( F = \sum_{s=1}^{n} \pi_{s} F^{(s)} \), we have

\[
\tau_{F}^{L} = \lim_{x \to -0} \frac{\sum_{s=1}^{n} \pi_{s} F^{(s)}(X > x)}{\sum_{s=1}^{n} \pi_{s} F^{(s)}(Y > x)}
\]

= \[ \lim_{x \to -0} \frac{\sum_{s=1}^{n} \pi_{s} F^{(s)}(X > x)}{\sum_{s=1}^{n} \pi_{s} F^{(s)}(Y > x)} \]

by definition \( F^{(s)}(X > x) = F^{(s)}(Y > x) \)

where \( C^{(s)} \) is the copula and \( F^{(s)}(y) \) the marginal cdf corresponding to \( F^{(s)} \), we have

\[
\tau_{F}^{L} = \lim_{x \to -0} \frac{\sum_{s=1}^{n} \pi_{s} F^{(s)}(X > x)}{\sum_{s=1}^{n} \pi_{s} F^{(s)}(Y > x)}
\]

so \( F^{(s)}(X > x) \leq \tau_{F}^{L} \) for all \( s \) and similarly \( F^{(s)}(Y > x) \leq \tau_{F}^{L} \).

\[
\tau_{F}^{L} = \lim_{x \to -0} \frac{\sum_{s=1}^{n} \pi_{s} F^{(s)}(X > x)}{\sum_{s=1}^{n} \pi_{s} F^{(s)}(Y > x)}
\]

we therefore have \( \tau_{F}^{L} = 0 \).

The part (i) and (ii) of the proposition is the straightforward application of above lemma

- For GARCH with constant mean and symmetric conditional distribution

\[
\tau_{F}^{L} = \mu + \Sigma_{L}^{1/2} \epsilon
\]

(\( + \) any GARCH dynamic equation of \( \tau_{F_{L}}^{L} \))

where \( \epsilon_{t} \) is stationary with symmetric distribution such that \( E(\epsilon_{t}) = 0 \). The unconditional distribution of \( X_{t} \) is a mixture of distribution of symmetric variable with same mean \( \mu \) but possibly different variance covariance matrix. By applying the Lemma 1, we conclude that the unconditional distribution of \( X_{t} \) is symmetric and (i) follows:

- For RS model with zero TDC

\[
\tau_{F}^{L} = \mu + \Sigma_{L}^{1/2} \epsilon
\]

where \( \tau_{t} \) takes a discrete value. Without loss of generality assume that \( X_{t} \) is bivariate and that \( \tau_{t} = x, \mu + \Sigma^{1/2} \epsilon_{t} \) is zero TDC such as in the normal case, therefore the unconditional distribution of \( X_{t} \) is a mixture of distribution with zero TDC. By applying Lemma 2, we conclude that the unconditional distribution of \( X_{t} \) has zero TDC and (ii) follows.

For (iii), with the same notations as Lemma 1, keeping marginal distribution unchanged across mixture components means that for discrete case
\[ f^{(i)}(x_1, \ldots, x_n; \delta, \rho) = c^{(i)}(u_1, \ldots, u_n; \theta) \times \prod_{i=1}^{n} f_i(x_i; \delta), \quad \text{with} \quad u_i = F(x_i; \delta), \] hence
\[ f(x_1, \ldots, x_n; \delta, \rho) = \sum_{i=1}^{n} \pi_i f^{(i)}(x_1, \ldots, x_n; \delta, \rho) \]
\[ = \sum_{i=1}^{n} \pi_i c^{(i)}(u_1, \ldots, u_n; \theta) \times \prod_{i=1}^{n} f_i(x_i; \delta) \]
\[ = c(u_1, \ldots, u_n; \theta) \times \prod_{i=1}^{n} f_i(x_i; \delta) \]
with \( c(u_1, \ldots, u_n; \theta) = \sum_{i=1}^{n} \pi_i c^{(i)}(u_1, \ldots, u_n; \theta) \) is the copula of \( f \) and we can see that \( c \) is a mixture of copula with symmetric TDC and hence is a copula with symmetric TDC. For the continuous case
\[ f(x_1, \ldots, x_n; \delta, \rho) = \int \pi \sigma f^{(i)}(x_1, \ldots, x_n; \delta, \rho) \sigma \]
\[ = \int \pi \sigma c^{(i)}(u_1, \ldots, u_n; \theta) \sigma \times \prod_{i=1}^{n} f_i(x_i; \delta) \]
\[ = c(u_1, \ldots, u_n; \theta) \times \prod_{i=1}^{n} f_i(x_i; \delta) \]
with \( c(u_1, \ldots, u_n; \theta) = \int \pi \sigma c^{(i)}(u_1, \ldots, u_n; \theta) \sigma \sigma \) which is a copula with symmetric TDC for same the reasons mentioned above. \( \square \)

A.2. Proof of Proposition 3.2
By definition of a copula, we have
\[ \eta_t = \left[ \frac{f(X_t; \theta, X_{t-1}; \delta, \rho)}{f(X_t; \theta, X_{t-1}; \delta) \times \prod_{i=1}^{n} f_i(x_i; \delta)} \right] \]
\[ = \left[ \frac{c(u_1(\delta_1), \ldots, u_k(\delta_k); \theta; |S_t = 1|) \times \prod_{i=1}^{n} f_i(x_i; \delta)}{c(u_1(\delta_1), \ldots, u_k(\delta_k); \theta; |S_t = 0|) \times \prod_{i=1}^{n} f_i(x_i; \delta)} \right] \]
with \( u_i(\delta_i) = F(x_i; \delta_i) \). By denoting \( \tilde{z}_{t-1} = (\tilde{z}_{t-1}^{(1)}, \ldots, \tilde{z}_{t-1}^{(n)}) \), the likelihood can be rewritten
\[ L(\delta, \theta, X_T) \]
\[ = \sum_{t=1}^{T} \log \left( \tilde{z}_{t-1}^{(1)} \eta_t^{(1)} \right) \]
\[ - \sum_{t=1}^{T} \log \left( \sum_{k=0}^{1} \tilde{z}_{t-1}^{(k)} c(u_1(\delta_1), \ldots, u_k(\delta_k); \theta; |S_t = k|) \times \prod_{i=1}^{n} f_i(x_i; \delta) \right) \]
\[ = \sum_{t=1}^{T} \log \left( \sum_{k=0}^{1} \prod_{i=1}^{n} f_i(x_i; \delta) \right) + \log \left( \sum_{k=0}^{1} \tilde{z}_{t-1}^{(k)} c(u_1(\delta_1), \ldots, u_k(\delta_k); \theta; |S_t = k|) \right) \]
\[ \right) \]
it follows that
\[ L(\delta, \theta, X_T) = \sum_{t=1}^{T} L(\delta, \theta, X_t) + L(\delta, \theta, X_{t-1}) \]
where
\[ L(\delta, \theta, X_T) = \sum_{t=1}^{T} \log f_i(x_i; \delta) \]
\[ L(\delta, \theta, X_{t-1}) = \sum_{t=1}^{T} \log \tilde{z}_{t-1}^{(1)} \eta_t^{(1)} \]
with
\[ \eta_t = \left[ \frac{c(u_1(\delta_1), \ldots, u_k(\delta_k); \theta; |S_t = 1|)}{c(u_1(\delta_1), \ldots, u_k(\delta_k); \theta; |S_t = 0|)} \right] \]
\[ \text{by noticing that} \quad \eta_t = \eta_{t}^{(1)} \times \prod_{i=1}^{n} f_i(x_i; \delta), \quad \text{we have that} \]
\[ \hat{\tilde{z}}_{t-1} = \left( \tilde{z}_{t-1}^{(1)} \eta_t \right) \]
Hansen, B.E., 1996. Inference when a nuisance parameter is not identified under the null hypothesis. Econometrica 64, 413–430.