Modeling non-linear dependence between risk and return in Latin markets

SERGIO GUILHERME SCHLENDER[1], MARCELO BRUTTI RIGHI[2], PAULO SERGIO CERETTA[3]
Departament of Business
Federal University of Santa Maria
Avenue Roraima, 1000, Rio Grande do Sul
BRAZIL
sergio.schlender1@gmail.com[1], marcelobrutti@hotmail.com[2], ceretta10@gmail.com[3]

Abstract - Financial theory is based on the trade-off between risk and return. However there is not a mathematical formulation of this dependence. Therefore, in this paper we estimate the dependence, through copula families, between risk and return in Latin markets, taking into consideration the U.S. market. For that, we use data from S&P500, Ibovespa, merval and IPC daily prices from January 2009 to December 2010, totaling 483 observations. In order to estimate risk we use a copula-based multivariate GARCH model. To test the copulas´ fit we use an adaptation of the Cramér-von Mises statistics. Results indicate that although linear correlation is not significant, risk and return are dependent on Latin markets. Further, there is a difference in the relationship between the studied markets.

Key-Words: Risk Analysis, Multivariate Volatility Modeling, Non-Linear Dependence, Copula-GARCH Model, Copulas Families, Latin Markets.

1 Introduction

Dependence between risk and return has always been of fundamental importance to financial economics. The study of volatility has great importance in finance, particularly in derivative pricing and risk management. Traditionally, the calculation of volatility estimates, as well as its application in determining value at risk (VaR) or hedging a portfolio relies on daily price changes [23].

Since the proposal of Generalized Auto-Regressive Conditional Heteroscedastic (GARCH) family models by [15] and [5] to account for variance heterogeneity in financial time series, a huge number of multivariate extensions of GARCH models have been introduced. The most consolidated models in literature are the Constant Conditional Correlation (CCC-GARCH) model of [6], the BEKK model of [16] and later the Dynamic Conditional Correlation (DCC-GARCH), developed by [14] and [38]. These models are based on multivariate Gaussian distributions, or a mixture of elliptical distributions, where care has to be taken to result in positive definite covariance matrices.

However, this assumption is unrealistic, as evidenced by numerous empirical studies, which show that many financial asset returns are skewed, leptokurtic, and asymmetrically dependent ([30]; [1]; [33]). Hence, these characteristics should be considered in the specifications of any effective hedging model or estimative of portfolio VaR.

These difficulties can be treated as a problem of Copulas. A copula is a function that links univariate marginals to their multivariate distribution. Since it is always possible to map any vector of random variables into a vector with uniform margins, we are able to split the margins of that vector and a summary of the dependence, which is the copula. The concept of copula was introduced by [37] and studied by many authors such as [10], [18]. The use of copulas for modeling the residual dependence between assets has recently appeared in empirical studies ([28]; [3]; [31]).

Thus, this paper aims to determine which family of copulas is best suited for the relationship between risk and return of the Latin American financial markets (Brazil, Argentina and Mexico) considering the U.S. market influence, analyzing the period after the 2007/2008 subprime crisis. [35] estimate conditional variance in these relationships and verify that Latin markets are relevant options for international diversification of investors with positions in U.S. assets. These Latin American emerging markets rank among the most mature markets within the universe of emerging countries and they actually attract a particular attention from global investors because of their great market openness ([2]). This question is investigated testing the fitting performance of copulas for S&P500, Ibovespa, Merval and IPC from January, 3, 2009 to December, 31, 2010, totaling 483 observations.

The econometric procedure by [36] estimates copula families to determine shape and magnitude of non-linear serial and cross-interdependence between returns and volatilities of financial assets. First, we estimate dynamic conditional covariance matrix through Copula-based multivariate GARCH models. This approach eliminates serial dependence, conditional heteroscedasticity and captures dynamic dependence between the Latin markets and the U.S. Secondly, we estimate the following copula families: Normal, student's *t*, Frank, Gumbel and Galambos, in order to identify which model presents the best fit for the relationship between risk and return in these Latin markets. A rank-based version of the well-known Cramér–von Mises statistic is employed to determine which family of copula has the best fit for the studied data.

The remainder of this paper is structured as follows: Section 2 briefly presents the theory of multivariate volatility modeling and copula functions; Section 3 exposes the material and methods of the study; Section 4 presents the results found and its discussion; Section 5 concludes the manuscript.

2 Theory

2.1 Multivariate Volatility Modeling

Multivariate models of volatility have attracted considerable interest during the last decade. This may be associated to the increase in the availability of financial data, the increase of the processing capacity of computers, and the fact that the financial sector began to realize the potential advantages of these models.

However, when it comes to the specification of a multivariate GARCH model, there is a dilemma. On one hand, the model should be flexible enough to be able to represent the dynamics of variance and covariance. On the other, as the number of parameters in a multivariate GARCH model often increases rapidly with the size of assets, the specification must be parsimonious enough to allow the model to be estimated easily, as well as allowing a simple interpretation of its parameters.

A feature that must be taken into account in the specification is the restriction of positivity (covariance matrices must necessarily take its determinants defined as positive). Based on this idea, we consider the model with multivariate GARCH parameterization VECH, proposed by Bollerslev, Engle and Wooldridge (1988), represented by (1).

$$vech(H_t) = A_0 + \sum_{j=1}^{q} \beta_j \, vech(H_{t-1}) + \sum_{j=1}^{p} A_j \, vech\left(\varepsilon_{t-j}, \varepsilon_{t-j}^{'}\right). \quad (1)$$

In (1), *vech* is the operator that contains the lower triangle of a symmetric matrix into a vector; H_t describes the conditional variance; the error term is $\varepsilon_t = H_t^{1/2}\eta_t, \eta_t \sim iidN(0,1)$. The disadvantage of this model is that it has a large number of parameters and in order to ensure the positivity of H_t, restrictions must be imposed.

Thus, the BEKK parameterization method emerges as an alternative, as suggested by Engle and Kroner (1995). The BEKK parameterization, which essentially takes care of the problems mentioned above about the VECH model, is defined as shown in (2).

$$H_{t+1} = C^{'}C + B^{'}H_t B + A^{'}\epsilon_t \epsilon_t^{'}. \quad (2)$$

The matrices A, B and C, which contain the coefficients for the case with two assets, are defined as.

$$A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}, \quad B = \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix}, \quad C = \begin{bmatrix} c_{11} & c_{12} \\ 0 & c_{22} \end{bmatrix}. \quad (3)$$

In (2), H_{t+1} is a conditional covariance matrix. In the bivariate case, the parameter B explains the relationship between the past conditional variances with the current ones (GARCH). The parameter A measures the extent to which conditional variances are correlated with past squared errors, i.e. it captures the effects of past shocks or volatility (ARCH). The total number of estimated parameters in bivariate occasion is eleven. In this case, the BEKK parameterization, the volatilities of the equation (2) are presented by formulations (4) and (5).

$$h_{11,t+1} = c_{11}^2 + b_{11}^2 h_{11,t} + 2b_{11}b_{12}h_{12,t} + b_{21}^2 h_{22,t} + a_{11}^2 \varepsilon_{1,t}^2 + 2a_{11}a_{12}\varepsilon_{1,t}\varepsilon_{2,t} + a_{21}^2 \varepsilon_{2,t}^2. \quad (4)$$

$$h_{22,t+1} = c_{12}^2 + c_{22}^2 + b_{12}^2 h_{11,t} + 2b_{12}b_{22}h_{12,t} + b_{22}^2 h_{22,t} + a_{12}^2 \varepsilon_{1,t}^2 + 2a_{12}a_{22}\varepsilon_{1,t}\varepsilon_{2,t} + a_{22}^2 \varepsilon_{2,t}^2. \quad (5)$$

Nevertheless, the BEKK parameterization model has as a disadvantage the difficulty to interpret its estimated parameters. Formulations (4) and (5) show that even for the case of bivariate modeling, the interpretation of the coefficients can be confusing because there are no parameters that are governed exclusively by an equation (Baur, 2006).

Thus, an approach to circumvent the problem of interpretation of the parameters is the model of conditional covariance matrix, observed indirectly through the matrix of conditional correlations. The first model of this kind was the constant conditional correlation (CCC) proposed by Bollerslev (1990) and Bollerslev and Wooldridge (1992). The conditional correlation was assumed to be constant and only the conditional branches are variable in time. The CCC model can be defined as shown by formulation (6).

$$H_t = D_t R D_t = \left(\rho_{ij}\sqrt{h_{ii,t}h_{jj,t}}\right). \quad (6)$$

In formulation (6) $D_t = diag\left(h_{11,t}^{1/2} \dots h_{NN,t}^{1/2}\right)$, where $h_{ii,t}$ is defined similarly to any univariate GARCH model; $R = \left(\rho_{ij}\right)$ is a symmetric positive definite matrix, with $\rho_{ii} = 1, \forall i$, i.e., R is the matrix containing the constant conditional correlations ρ_{ij}.

However, the assumption that the conditional correlation is constant over time is not convincing, since, in practice, the correlation between assets undergoes many changes over time. Thus, Engle and Sheppard (2001) and Tse ans Tsui (2002) introduced the model of dynamic conditional correlation (DCC). The DCC model is a two-step algorithm to estimate the parameters which makes it relatively simple to use in practice. In the first stage, the conditional variance is estimated by means of univariate GARCH model, respectively, for each asset. In the second step, the parameters for the conditional correlation, given the parameters of the first stage, are estimated. Finally, the DCC model includes conditions that make the covariance matrix positive definite at all points in time and the covariance between assets' volatility a stationary process. The DCC model is represented by formulation (7).

$$H_t = D_t R D_t. \quad (7)$$

Where,

$$R_t = diag\left(q_{11,t}^{-\frac{1}{2}} \dots q_{NN,t}^{-\frac{1}{2}}\right) Q_t diag\left(q_{11,t}^{-1/2} \dots q_{NN,t}^{-1/2}\right) \quad (8)$$

Since the square matrix of order N symmetric positive defined $Q_t = \left(q_{ij,t}\right)$ has the formulation proposed in (9).

$$Q_t = (1-\alpha-\beta)\bar{Q} + \alpha u_{t-1}u_{t-1}^{'} + \beta Q_{t-1}. \quad (9)$$

In (9), $u_{i,t} = \varepsilon_{i,t}/\sqrt{h_{i,t}}$; $\bar{Q}$ is the $N \; x \; N$ matrix composed by unconditional variance of u_t; α and β are non-negative scalar parameters satisfying $\alpha + \beta < 1$.

All of the models mentioned in the previous section are estimated under the assumption of multivariate normality. The use of a copula function, on the other hand, allows us to consider the marginal distributions and the dependence structure both separately and simultaneously (Hsu, Tseng and Wang, 2008). Therefore, the joint distribution of the asset´s return can be specified with full flexibility, which is more realistic.

In that sense, Hansen (1994) proposes a GARCH model in which the first four moments are conditional and time varying. For the conditional mean and volatility, he built on the usual GARCH model. To control higher moments, he constructed a new density, which is a generalization of the Student-t distribution while maintaining the assumption of a zero mean and unit variance, in order to model the GARCH residuals. The conditioning is obtained by defining parameters as functions of past realizations (Jondeau and Rockinger, 2006). The conditional volatility model proposed by Hensen (1994), and later discussed in Theodossiou (1998) and Jondeau and Rockinger (2003) is represented by formulation (10).

$$h_{i,t}^2 = c_i + b_i h_{i,t-1}^2 + a_i \varepsilon_{i,t-1}^2. \quad (10)$$

Where $\varepsilon_{i,t} = h_{i,t}z_{i,t}, z_{i,t} \sim skewed - t(z_i|\eta_i, \phi_i)$. The density of skewed-t distribution is represented by formulation (11).

$$d(z|\eta,\phi) = \begin{cases} bc\left[1+\dfrac{1}{\eta-2}\left(\dfrac{bz+a}{1-\phi}\right)^2\right]^{-\eta+1/2} & ,z < -\dfrac{a}{b} \\ bc\left[1+\dfrac{1}{\eta-2}\left(\dfrac{bz+a}{1-\phi}\right)^2\right]^{-\eta+1/2} & ,z > -\dfrac{a}{b} \end{cases}. \quad (11)$$

In (11), $a \equiv 4\phi c \frac{\eta-2}{\eta-1}$; $b \equiv 1 + 3\phi^2 - a^2$; $c \equiv \frac{\Gamma(\eta+1/2)}{\sqrt{\pi(\eta-2)}\Gamma(\eta-2)}$; η and ϕ are the kurtosis and asymmetry parameters, respectively. These are restricted to $4 < \eta < 30$ and $-1 < \phi < 8$.

2.2 Copula distribution functions

2.2.1 Definition and concepts

Dependence between random variables can be modeled by copulas. A copula returns the joint probability of events as a function of the marginal probabilities of each event. This makes copulas attractive, as the univariate marginal behavior of random variables can be modelled separately from their dependence (Kovadinovic and Yan, 2010).

The concept of copula was introduced by Sklar (1959). However, only recently its applications have become clear. A detailed treatment of copulas as well as their relationship to concepts of dependence is given by Joe (1997) and Nelsen (2006). A review of the applications of copulas to finance can be found in Embrechts *et al.* (2003) and in Cherubini *et al.* (2004).

For simplicity, we restrict our attention to the bivariate case. The extensions to the *n*-dimentional case are straightforward. A function $C : [0,1]^2 \to [0,1]$ is a *copula* if, for $0 \leq x \leq 1$ and $x_1 \leq x_2$, $y_1 \leq y_2$, (x_1, y_1), $(x_2, y_2) \in [0,1]^2$, it fulfills the following properties:

$$C(x,1) = C(1,x) = x,$$
$$C(x,0) = C(0,x) = 0. \quad (12)$$

$$C(x_2, y_2) - C(x_2, y_1) - C(x_1, y_2) + C(x_1, y_1) \geq 0. \quad (13)$$

Property (12) means uniformity of the margins, while (13), the *n-increasing property* means that $P(x_1 \leq X \leq x_2, y_1 \leq Y \leq y_2) \geq 0$ for (X,Y) with distribution function *C*.

In the seminal paper of Sklar (1959), it was demonstrated that a Copula is linked with a distribution function and its marginal distributions. This important theorem states that:

(i) Let *C* be a copula and F_1 and F_2 univariate distribution functions. Then (14) defines a distribution function *F* with marginals F_1 and F_2.

$$F(x,y) = C(F_1(x), F_2(y)), \ (x,y) \in R^2. \quad (14)$$

(ii) For a two-dimensional distribution function *F* with marginals F_1 and F_2, there is a copula *C* satisfying (14). This is unique if F_1 and F_2 are continuous and then, for every $(u,v) \in [0,1]^2$:

$$C(u,v) = F\left(F_1^{-1}(u), F_2^{-1}(v)\right). \quad (15)$$

In (15), F_1^{-1} and F_2^{-1} denote the generalized left continuous inverses of F_1 and F_2.

However, as Frees and Valdez (1997) note, it is not always obvious to identify the copula. Indeed, for many financial applications, the problem is not to use a given multivariate distribution but it consists in finding a convenient distribution to describe some stylized facts, for example the relationships between different asset returns.

2.2.2 Families of copulas

The most frequently used copulas are the Elliptical and the Archimedean (Yan and Kojadinovic, 2010). Among the elliptical copulas, which are characterized by the class of symmetric copulas, we highlight the Normal and Student's *t* Copulas. In the class of Archimedean copulas, which best fit the asymmetric distributions, it stands out the families: Frank and Gumbel. The Archimedean copulas, which are of one-parameter, may be constructed using a function $\phi_\alpha : [0,1]^2 \to R^{*+}$, continuous, decreasing, convex and such that $\phi(1) = 0$. Such function ϕ is called a generator and α is a real parameter. The pseudo-inverse of ϕ is defined by formula (16).

$$\phi^{[-1]}(u) = \begin{cases} \phi^{-1}(v) & 0 \leq u \leq \phi(0) \\ 0 & \phi(0) \leq u \leq +\infty \end{cases}. \quad (16)$$

This pseudo-inverse is such that, by composing it with the generator, it gives the identity, the same way as ordinary inverses do for functions with domain and range *R*. Formulation (17) defines this relation.

$$\phi^{[-1]}(\phi(u)) = u, \ \forall\, u \in [0,1]. \quad (17)$$

Given a generator and its pseudo-inverse, an Archimedean copula C^A is generated as in (18).

$$C^A(u,v) = \phi^{[-1]}(\phi(u) + \phi(v)). \quad (18)$$

These four families of copulas (Normal, Students' *t*, Frank and Gumbel) will be defined, according to Cherubini *et al.* (2004), presented below.

The Gaussian is an elliptical copula, because of its symmetry. Although this copula is the most used due to its more tangible properties, it fails to represent the reality of the observed data in financial markets.

Let $\Phi_{\rho_{XY}}$ be the joint distribution of a bi-dimensional vector, with linear correlation coefficient ρ_{XY}. The Normal Copula is defined in (19).

$$C^{Ga}(u,v) = \Phi_{\rho XY}\left(\Phi^{-1}(u), \Phi^{-1}(v)\right). \quad (19)$$

In (19), Φ^{-1} is the inverse of the standard univariate normal distribution function Φ. The Gaussian Copula generates the standard Gaussian joint distribution function, whenever the margins are standard normal.

Another elliptical family of copulas is the Student's *t*. Let ρ be the bivariate linear correlation, and υ the degrees of freedom of the student's *t* distribution function, so the Student's *t* copula is defined in (20).

$T_{\rho,v}(u,v) = t_{\rho,v}\left(t_v^{-1}(u), t_v^{-1}(v)\right)$. (20)

In (20), t_v^{-1} is the inverse of the univariate Student's t distribution function with v degrees of freedom; and $t_{\rho,v}$ is the bivariate distribution corresponding to t_v.

The Frank Copula is an Archimedean one. The generator is defined in (21), and the Copula in (22).

$$\phi_\alpha(u) = -\ln\frac{\exp(-\alpha u)-1}{\exp(-\alpha)-1}. \quad (21)$$

$$C(u,v) = -\frac{1}{\alpha}\ln\left(1+\frac{(\exp(-\alpha u)-1)(\exp(-\alpha v)-1)}{exp(-\alpha)-1}\right). \quad (22)$$

In (22) the range for α is $(-\infty, 0) \cup (0, +\infty)$.

Another copula in this class is the Gumbel, which is defined as (23).

$$C(u,v) = exp\{-[(-\ln u)^\alpha + (-\ln v)^\alpha]^{1/\alpha}\}. \quad (23)$$

In (23), the range for α is$[1, +\infty)$. The coefficient of tail dependence is given by $\lambda_U = 2 - 2^{1/\alpha}$.

However, it is often reasonable to assume that the dependence structure of a bivariate continuous distribution belongs to the class of extreme-value copulas, as it is more efficient to model financial risk with these copulas, due to the fact that it is precise in the tails of the distribution of returns, that lies the biggest challenge of diversifying a portfolio (Genest *et al*., 2011).

C is an extreme value copula when there is a function *A*: $[0,1] \rightarrow [0.5,1]$ such that for all $(u,v) \in [0,1]^2$, there is a relation as expressed in (24).

$$C(u,v) = uv^{A[\log(v)/\log(uv)]}. \quad (24)$$

It was shown by Pickands (1981) that *C* is a copula if and only if *A* is convex and $\max(t, t-1) \leq A(t) \leq 1$. By reference to this work, the function *A* is often referred to as the Pickands dependence function (Genest and Segers, 2009). Among the extreme value copulas, which are characterized by capturing the tail dependence, we highlight the families: Galambos and TEV.

A family of extreme value copulas is the Galambos. It is represented by formulation (25).

$$C(u,v) = uv\exp\left([(-\ln u)^\alpha + (-\ln v)^{-\alpha}]^{-1/\alpha}\right). \quad (25)$$

In (25), the range for α is $[0,+\infty)$. The coefficient of tail dependence is given by $\lambda_U = 2 - 2^{1/\alpha}$.

3 Material and Methods

We collected data of the daily prices of S&P500, Ibovespa, Merval and IPC, from January, 3, 2009 to December, 31, 2010, totaling 483 observations. These indices were chosen because they are commonly used in academic papers as proxies for the financial markets in these countries. Both are compounds by the stocks that are more representative in terms of liquidity and value. We considered the period after the recent financial crisis of 2007/2008, in order to avoid possible vestiges that could cause some bias in the results.

The ADF test (Augmented Dickey Fuller) was initially employed in prices and their logarithmic differences (returns), to eliminate problems of non-stationarity. The ADF test, proposed by Dickey and Fuller (1981) is represented by (26).

$$\Delta P_t = \gamma P_{t-1} + \sum_{i=1}^{n} \delta_i \Delta P_{t-i} + \varepsilon_t. \quad (26)$$

In formulation (26), ΔP_t is the price change at time t, γ and $\delta_i$are constant, and ε_t is a white noise series. If the null hypothesis cannot be rejected, the price series $\{P\}$ contains a unit root, non-stationary.

We used an autoregressive vector (VAR) to obtain the average estimate of the return and the residual series of each index. The mathematical form of the bivariate VAR model used is represented by (27).

$$VAR(L,A) = \begin{cases} \Delta L_{j,t} = \beta_0 + \sum_{i=1}^{n}\beta_i \Delta L_{j,t-i} + \sum_{j=1}^{m}\beta_j \Delta A_{t-j} + \varepsilon_{1,t} \\ \Delta A_t = \alpha_0 + \sum_{i=1}^{n}\alpha_i \Delta A_{t-i} + \sum_{j=1}^{m}\alpha_j \Delta L_{j,t-j} + \varepsilon_{2,t} \end{cases}. \quad (27)$$

In (27), ΔL_t and ΔA_t are, respectively the daily returns of Latin American *j* and U.S. markets; β_k and α_k are regression parameters; $\varepsilon_{1,t}$ and $\varepsilon_{2,t}$ are, correspondingly, the estimated residuals of returns.

Subsequently, using the residuals that were obtained through the VAR applied to the series, we used the copula-based GARCH model, represented by (10). Through this inference, the estimates of conditional variances and covariance of these markets were obtained, taking in consideration the American market, due to its influence as a benchmark. After that, we sought to identify the presence of serial correlation on the residuals of the copula-based model, by using *Q* statistic of Ljung and Box (1978), represented by (28), which tests the null hypothesis that the data are random against the alternative of non-randomness of these.

$$Q = n(n+2)\sum_{k=1}^{h}\frac{\hat{\rho}_k^2}{n-k}. \quad (28)$$

In (28), *n* is the size of sample; $\hat{\rho}_k^2$ is the autocorrelation of sample in lag *k*; *h* is the number of lags being tested; The Ljung-Box *Q* statistics follow a chi-squared (χ^2) distribution with *k* degrees of freedom. In order to do this, the data was standardized into pseudo-observations $\boldsymbol{U}_j = (U_{1j}, \ldots, U_{ij})$ through the ranks as $U_{ij} = R_{ij}/(n+$

1). The next step was, to estimate the copula's parameters, it was employed the procedure of inversion of the copula based *Kendall's Tau (τ)*, that is used to measure the monotonic dependence, which is calculated as shown in (29).

$$\tau(x, y) = 4 \int_0^1 \int_0^1 C(u, v) dC(u, v) - 1. \quad (29)$$

To determine which copula model best fits the residuals of the markets studied, we applied a rank-based version of the familiar Cramér–von Mises statistic, discussed in Genest, Rémillard and Beaudoin (2009), and extended in Genest *et. al.* (2011), which enables to check the validity of the dependence structure separately from the margins. These authors emphasize that it is a blanket test, i.e., a procedure which implementation requires neither an arbitrary categorization of the data, nor any strategic choice of smoothing method, whether it could be kernel- or wavelet-based, or any of them. The goodness-of-fit test employed is defined in (30), which tests the null hypothesis that data is fitted by $C_{\theta n}$, a copula with vector of parameters θ.

$$S_n = \int_{[0,1]^d} \mathbb{C}_n(\boldsymbol{u})^2 \, dC_n(\boldsymbol{U}). \quad (30)$$

In (30), $C_n(\boldsymbol{U}) = \frac{1}{n}\sum_{i=1}^{n} \mathbf{1}(U_{i1} \leq u_1;\, U_{i2} \leq u_2)$ is known as the empirical copula; $\boldsymbol{U}_j = (U_{1j}, \dots, U_{ij})$ are the pseudo-observations; $\boldsymbol{u} = (u_1, u_2) \in [0,1]^2$; $\mathbb{C}_n = \sqrt{n}(C_n - C_{\theta n})$ is the empirical process that assess the distance between the empirical copula and the estimation $C_{\theta n}$; n is the number of observations. In practice, the limiting distributions of S_n depend on the family of copulas under the composite null hypothesis, and on the unknown parameter value θ in particular. Thus, the probability associated with the S_n, and not its calculated value, can be used to compare distinct estimated copula families as a selection criteria, so that the larger value indicates a better fit.

4 Results and discussion

Initially, we performed the ADF test of unit root in level and first difference for all series of logarithm (daily returns). As the presence of a unit root in all series was confirmed, we calculated the daily returns by the difference of logarithms of prices. Table 2 displays the descriptive statistics of these returns, whereas Figure 1 shows the temporal evolution of these series.

The results in Table 2 confirm the fact that Brazil, Argentina and Mexico being emerging countries, should have a higher standard deviation, representing greater risk and therefore requiring higher returns, as it is verified by higher values for mean and median. The U.S. market, by contrast had lower mean and median and lower standard deviation of returns, representing a more stabilized economy. It is also noticed that all sets of returns are leptokurtic, a fact quite common, being widely recognized by financial professionals.

Figure 1 endorses these results. It gives a visual confirmation of the greater dispersion of the daily returns of the Latin American markets compared to the U.S. It is noteworthy that there is a volatility cluster at the beginning of the observations, extending for about 100 trading days. It was the vestiges of the North American financial crisis.

Subsequently, it was estimated a copula-based GARCH model to obtain the estimated variances and covariances of the bivariate relationship of the U.S. and Latin markets. Table 3 presents the results of these models.

The results in Table 3 indicate that the conditional volatility of all markets studied was significantly affected at the level of 5% by the past squared shocks, and lagged volatility. Moreover, these impacts had similar magnitudes in the models estimated for the three bivariate relationships. Nevertheless, the shape of the probability distribution of conditional volatilities estimated showed differences between the analyzed markets regarding to the number of degrees of freedom of the skewed-t function. The degrees of freedom were 7.94 for Brazil, 3.79 for Argentina and 4.69 for Mexico. These results emphasize that the return of Latin marketsare skewed and leptokurtic.

Complementing, the estimated volatilities and dynamic correlations are shown, respectively, in Figure 2 and 3, for the bivariate relationships proposed in this study. Furthermore, the Q statistics are presented in Table 4, in order to verify the serial dependence of the residues of GARCH estimates.

The results in Table 4 suggest that the estimated residuals from the copula-based GARCH model do not exhibit significant serial correlation. Therefore, the estimated models were able to fit the sample of bivariate relationships between the daily returns of the U.S. market with the Latin countries, filtering the serial dependence and the heteroscedastic dynamic behavior of data. Thus, the estimates of variance and covariance of the studied markets are valid for the computation of the optimal hedge ratio and the value of assets at risk in question.

The plots of Figure 2 emphasize that after the observation 100, the vestiges of the American financial crisis of 2007/2008 began to disappear, returning to the stability period. It is noteworthy that with the resumption of normal variability of returns in these markets, Mexico was more stable, followed by Brazil and Argentina.

Figure 3, which exposes the dynamic correlations estimated for the series of returns of the markets studied, suggests the existence of a pattern behavior. At the beginning of the analyzed period, the dependence among the markets was higher, decreasing with the path of the sample. The Argentinean market was the first to show this reduction, followed by the Brazilian and the Mexican. It is worth noting also that the correlation coefficient between the Mexican and American markets was the highest among Latin American countries, followed by Brazil and Argentina.

After this empirical analysis of the markets, as well as the estimation of their conditional volatilities, it was also estimated the parameters of the copulas Normal, Student's *t*, Frank, Gumbel, and Galambos, through inversion of Kendall's Tau. Then, it was statistically verified the goodness of fit of the estimated copulas by the S_n test exposed in the method of this study. The results of the estimated parameters, as well as values and significance of the S_n tests, for the period before crisis, are shown in Table 5.

The results in Table 5 allow us to conclude, initially, that in Latin markets the correlation between risk and return is very close to zero, emphasizing that the use of linear models to capture the dependence of this relationship is inadequate. This is reinforced by the fact that most of the estimated copulas did not reject the null hypothesis of fit to the joint distribution of probability of risk and return in Latin markets. Exceptions, at the 5% level of significance, are the Gumbel and Galambos copulas for the Brazilian market; Normal, Student's t and Frank copulas for the Mexican market.

Further, with respect to the fit, in the Brazilian market, the Normal copula had the largest value for the p-value of the S_n test. The Frank copula had a closer value for the test and its p-value. In the Argentinean and Mexican markets, the Galambos copula had the largest p-value for the S_n test. The difference is that for the Argentinean market none of the estimated copula families rejected the null hypothesis of fit the data, while for the Mexican market only extreme value copulas fitted the data. This result highlights the importance of risk management in these markets. This is because such results indicate empirically that risk and return in these markets have dependence in their probability distributions.

This relationship between risk and return in the analyzed markets is very relevant to the international portfolio diversification because over the last ten years, the volatility of Latin American financial markets has become a key determinant for explaining the risk taking behaviors of investors, especially the substitution in their portfolios between different categories of securities (Dufrenot, Mignon and Péguin-Feissolle, 2011).

5 Conclusion

In this paper we analyzed the dependence between risk and return in Latin financial markets, considering the exerted influence of the United States. To do that, we used daily data from the Brazilian, Argentinean, Mexican and U.S. markets. Initially, we estimated copula-based multivariate GARCH models, in order to estimate the conditional variance and covariance of the bivariate relationships of the U.S. market with these three Latin markets. Thus, the estimated volatilities of the markets were obtained, being used as a proxy for the risk.

After that, through a rank-based version of the familiar Cramér–von Mises statistic we determined which family of copula (Normal or Gaussian, student's *t*, Frank, Gumbel and Galambos) had the best fit to the data. The results evidenced that the linear correlation between risk and return in these countries is not significant, emphasizing the lack of linear dependence

However, this relationship in the three Latin markets studied was fitted by the estimated copulas, with exceptions of the Gumbel and Galambos copulas for the Brazilian market and Normal, Student's t and Frank copulas for the Mexican market. This result indicates that there is dependence in the joint probability distribution of risk and return in these markets, but not in a linear form. Thus, the use of linear models is unable to correctly estimate the dependence between risk and return of an asset traded in these Latin markets producing incorrect results and prompting investors to achieve optimal allocation of its portfolio ineptly.

As suggestions for future studies, we highlight the application of a similar test to estimate the copula families of the joint probability of risk and return of Asian and European markets, taking into account the influence of a greater market.

References:

[1] Ang, A., Chen, J. Asymmetric correlations of equity portfolios. *Journal of Financial Economics*, Vol. 63, 2002, pp.443–494.

[2] Arouri, M., Bellalah, M., Nguyen, D. K. The comovements in international stock markets: new evidence from Latin American emerging countries. *Applied Economics Letters*, Vol. 17, No 13, 2010, pp. 1323-1328.

[3] Ausin, M. C., Lopes, H. F. Time-varying joint distribution through copulas. *Computational Statistics & Data Analysis*, Vol. 54, No 11, 2010, pp. 2383-2399.

[4] Baur, D. A Flexible Dynamic Correlation Model. Econometric Analysis of Financial and Economics Time Series/Part A. Advances in econometrics, Elsevier, Vol.20, 2006, pp.3-31.

[5] Bollerslev, T. Generalized autoregressive conditional heteroscedasticity. Journal of Econometrics, Vol. 31, 1986, pp. 307–327.

[6] Bollerslev, T. Modeling the coherence in short-run nominal exchange rates: A multivariate generalized ARCH model. Review of Economics & Statistics, Vol. 72, No 3, 1990, pp. 498-505.

[7] Bollerslev, T., Engle, R., Wooldridge, J. M. A Capital Asset Pricing Model with Time Varying Covariances. Journal of Political Economy, Vol. 96, 1988, pp.116-131.

[8] Bollerslev, T., Wooldridge, J. M. Quasi-Maximum Likelihood Estimation and Inference in Dynamic Models with Time-Varying Covariances. Econometric Reviews, Vol. 11, 1992, pp.143-172.

[9] Cherubini, U., Luciano, E., and Vecchiato, W. *Copula Methods in Finance*. Wiley Series in Financial Engineering. Wiley, Chichester, UK, 2004.

[10] Deheuvels, P. La fonction de dépendance empirique et ses propriétés — Un test non paramétrique d'indépendance. Académie Royale de Belgique. Bulletin de la Classe des Sciences, Vol. 65, 1979, pp. 274-292.

[11] Dickey, D., Fuller, W. Likelihood ratio statistics for auto-regressive time series with a unit root. Econometrica, Vol. 49, 1981, pp.1057-1072..

[12] Dufrenot, G., Mignon, V., Peguin-Feissolle, A. The effects of the subprime crisis on the latin american financial markets: an empirical assessment. Economic Modelling, Vol. 28, No. 5, 2011, pp. 2342–2357.

[13] Embrechts P., Lindskog F., McNeil A. Modeling dependence with copulas and applications to Risk Management. Handbook of Heavy Tailed Distributions in Finance. Elsevier, 2003, pp. 329-384.

[14] Engle, R. F., Sheppard, K. Theoretical and Empirical properties of Dynamic Conditional Correlation Multivariate GARCH, NBER Working Paper, 2001, n.8554.

[15] Engle, R. Autoregressive conditional heteroscedasticity with estimates of the variance of UK inflation. Econometrica, Vol. 50, 1982, pp. 987–1007.

[16] Engle, R., Kroner, K. Multivariate simultaneous generalized ARCH. Econometric Theory, Vol. 11, 1995, pp.122–150.

[17] Frees, E., Valdez, E. Understanding relationships using copulas, North American Actuarial Journal, Vol. 2, 1998, pp.1-25.

[18] Genest, C., MacKay J. The joy of copulas: Bivariate distributions with uniform marginal. American Statistician, Vol. 40, 1986, pp. 280-283.

[19] Genest, C., Remillard, B. Validity of the parametric bootstrap for goodness-of-fit testing in semiparametric models. Annales de l'Institut Henri Poincare: Probabilites et Statistiques, Vol. 44, 2008, pp.1096-1127.

[20] Genest, C., Segers, J. Rank-based inference for bivariate extreme-value copulas. Annals of Statistics, Vol. 37, 2009, pp. 2990-3022.

[21] Genest, C., Kojadinovic, I., Nešlehová, J., Yan, J. A goodness-of-fit test for bivariate extreme-value copulas. Bernoulli, Vol. 17, 2011, pp. 253-275.

[22] Genest, C., Remillard, B., Beaudoin, D. Goodness-of-fit tests for copulas: A review and a power study. Mathematics and Economics, Vol.44, 2009, pp.199-214.

[23] Goodhart, C. A. E., O'hara, M. High frequency data in financial markets: Issues and applications. Journal of Empirical Finance, Vol. 4, 1997, pp. 73-114.

[24] Hansen, B. E. Autoregressive conditional density estimation. International Economic Review, Vol. 35, No 3, 1994, pp. 705-730.

[25] Hsu, C. C., Tseng, C. P., Wang, Y. H. Dynamic hedging with Futures: a copula-based GARCH model. The Journal of Futures Markets, Vol. 28, 2008, pp.1095–1116.

[26] Joe, H. Multivariate models and dependence concepts. Chapman Hall, 1997.

[27] Jondeau, E., Rockinger, M. Conditional volatility, skewness, and kurtosis: existence,

persistence, and comovements. Journal of Economic Dynamics and Control, Vol. 27, No 10, 2003, pp.1699-1737.

[28] Jondeau, E., Rockinger, M. The Copula-GARCH model of conditional dependencies: An international stock market application. Journal of International Money and Finance, Vol. 25, 2006, pp. 827-853.

[29] Kojadinovic, I., Yan, J.. Modeling Multivariate Distributions with Continuous Margins Using the copula R Package. Journal of Statistical Software, 34, 2010, 1-20.

[30] Longin, F., Solnik, B. Extreme correlation of international equity markets. Journal of Finance, Vol. 56, 2001, pp. 649.676.

[31] Min, A., Czado, C. Bayesian inference for multivariate copulas using pair-copula constructions. Journal of Financial Econometrics, Vol. 8, No 4, 2010, pp.511-546.

[32] Nelsen, R. An introduction to copulas. Springer, New York, second edition, 2006.

[33] Patton, A. J. Estimation of copula models for time series of possibly different lengths. Journal of Applied Econometrics, Vol. 21, No 2, 2006, pp.147-173.

[34] Pickands, J. Multivariate extreme value distributions (with a discussion). Proceedings of the 43rd Session of the International Statistical Institute. Bulletin of Institute International Statististic Vol. 49, 1981, pp. 859-878/894-902.

[35] Righi, M. B. Ceretta, P. S. Estimating value at risk and optimal hedge ratio in Latin markets: a copula-based GARCH approach, Economics Bulletin, Vol. 31,No 2, 2011, pp. 1717-1730.

[36] Righi, M. B. Ceretta, P. S. Journal of Banking & Finance, Vol. 37, No 3, 2013, pp. 837–846.

[37] Sklar A. Fonctions de Repartition á n Dimensions et leurs Marges. Publications de l'Institut de Statistique de l'Université de Paris, Vol. 8, 1959, pp. 229-231.

[38] Tse, Y. K., Tsui, A. K. C. A multivariate GARCH model with time-varying correlations. Journal of Business and Economic Statistics, American Statistical Association, Vol. 20, 2002, pp. 351-362.

[39] Yan, J., Kojadinovic, I., 2010. Package Copula. Version 0.9-7, May 28, 2010. Available in: http://cran.r-project.org/web/packages/copula/copula.pdf.

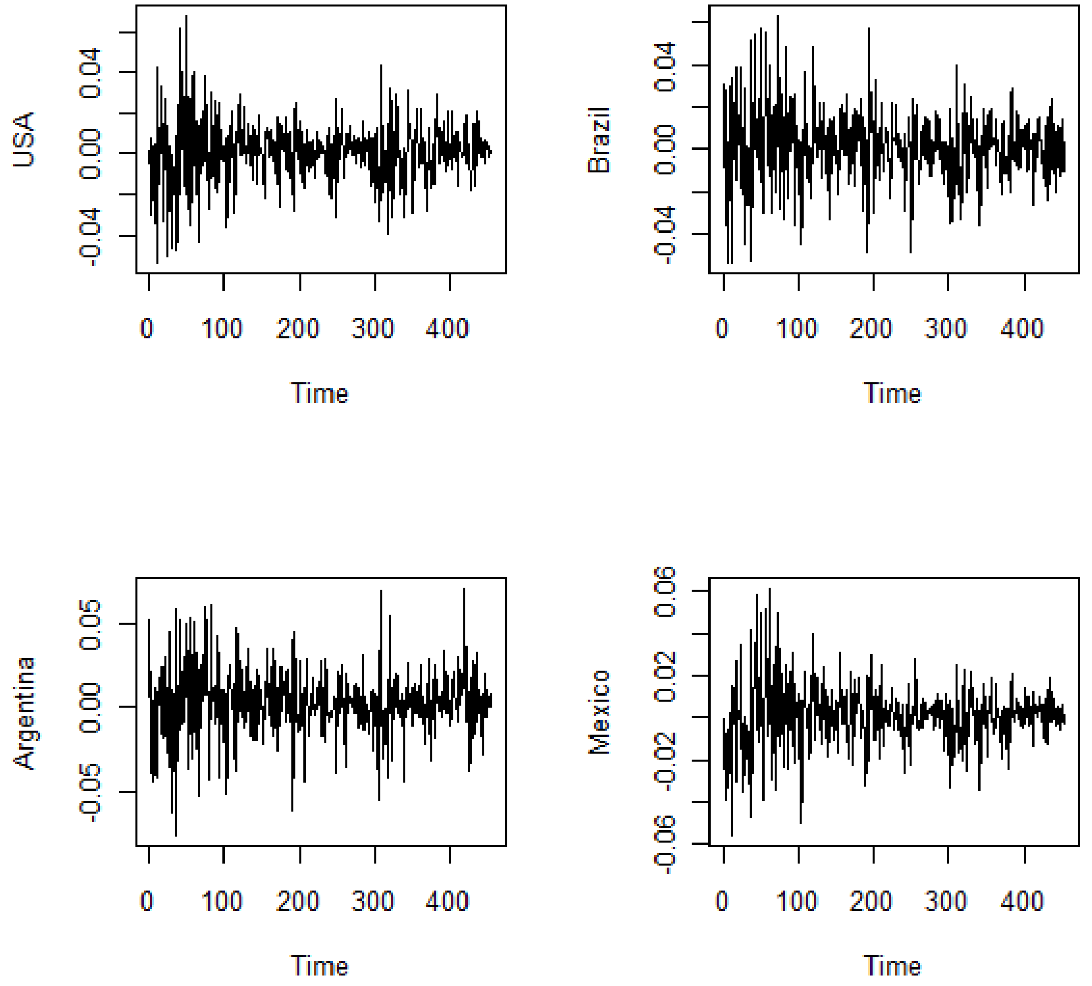

Fig.1. Time series of daily returns of S&P500 (USA), Ibovespa (Brazil), Merval (Argentina) and IPC (Mexico).

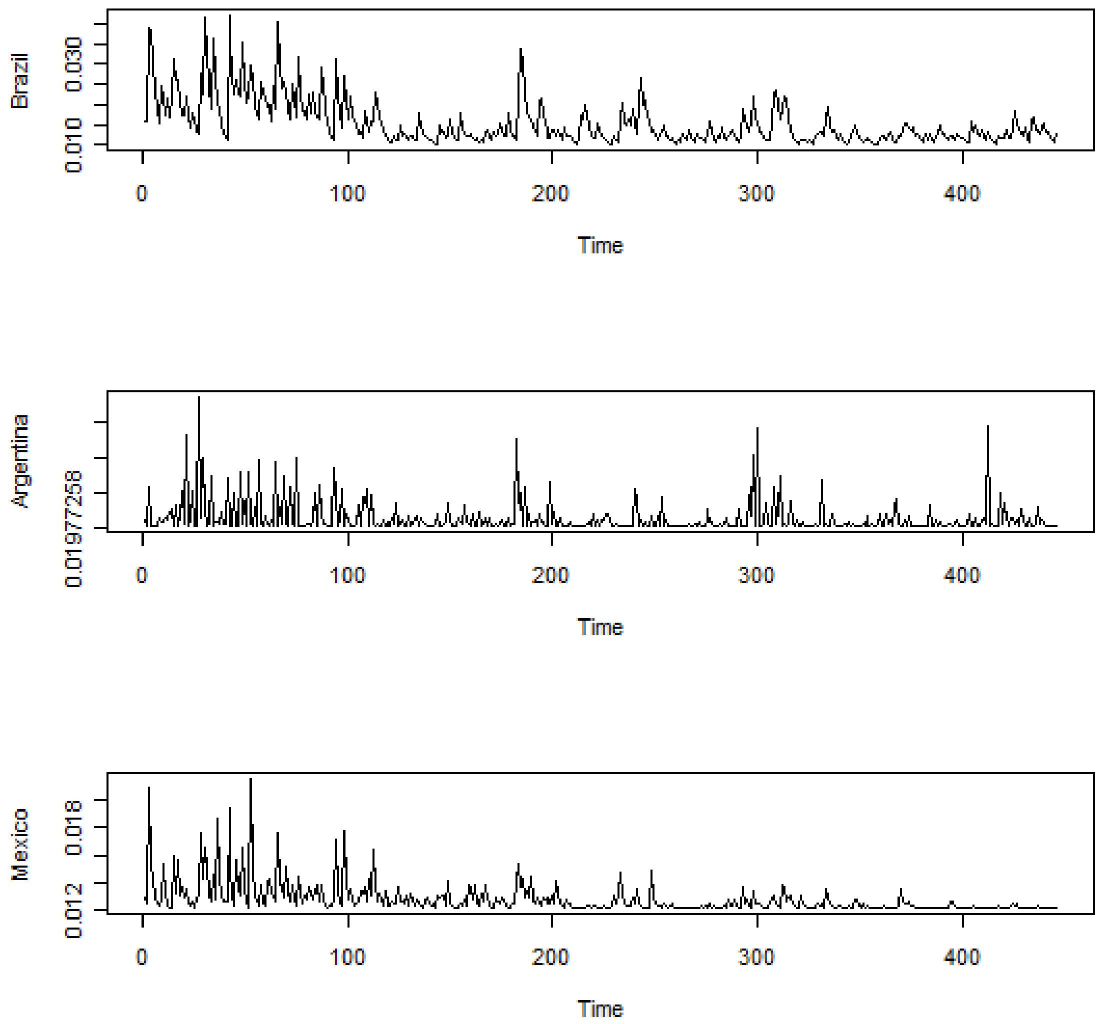

Fig.2. Estimated conditional volatilities of the bivariate relationships of daily log-returns of S&P500 with Ibovespa, Merval and IPC.

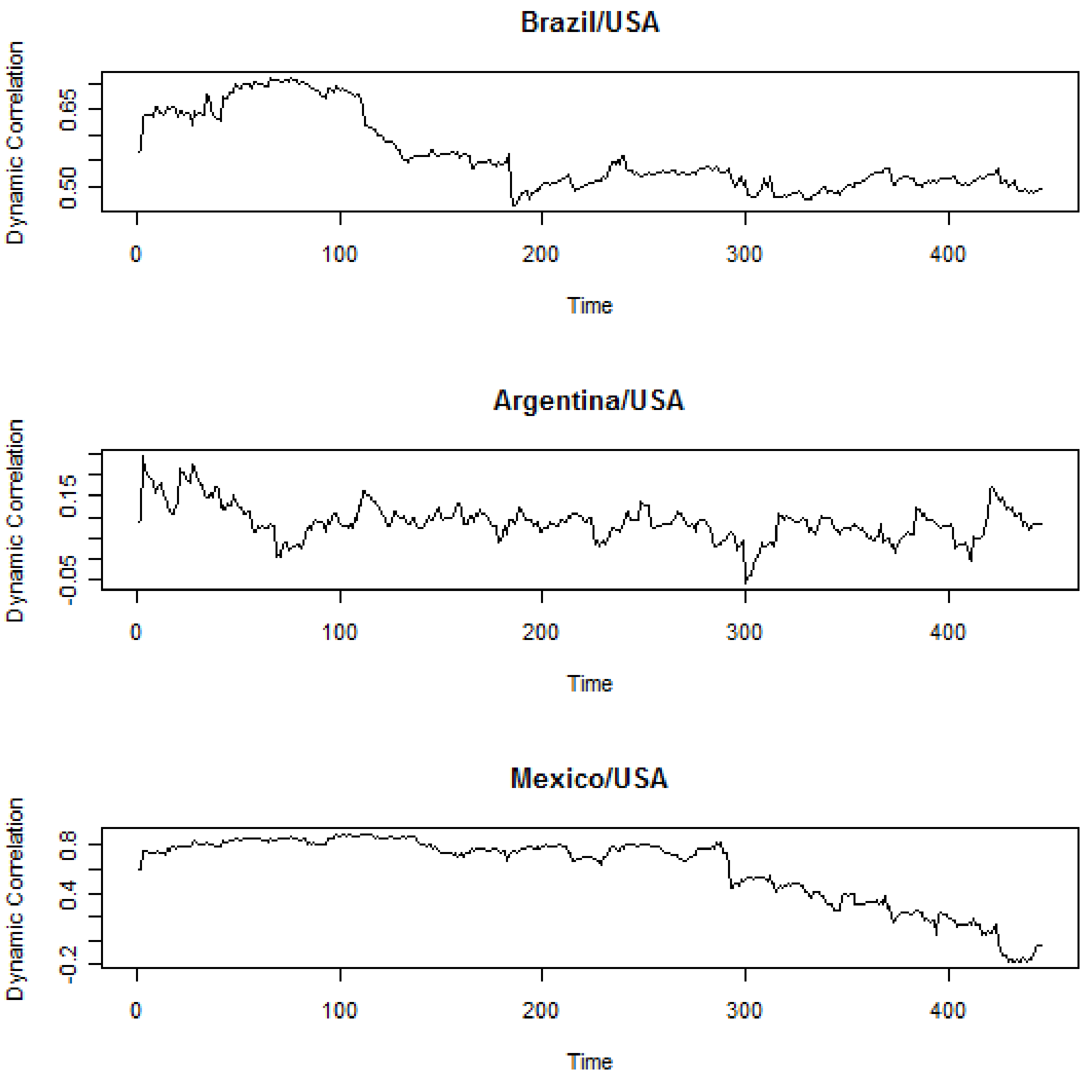

Fig.3. Estimated conditional correlation of the bivariate relationships of daily log-returns of S&P500 with Ibovespa, Merval and IPC.

Table 1 Descriptive statistics of daily returns of S&P500 (USA), Ibovespa (Brazil), Merval (Argentina) and IPC (Mexico).

Statistic	**Δln(USA)**	**Δln(Brazil)**	**Δln(Argentina)**	**Δln(Mexico)**
Mean	0.0007	0.0011	0.0025	0.0011
Median	0.0012	0.0014	0.0026	0.0020
Minimum	-0.0543	-0.0540	-0.0770	-0.0563
Maximum	0.0684	0.0638	0.0712	0.0618
Standard deviation	0.0151	0.0171	0.0201	0.0142
Skewness	-0.0734	0.0432	-0.1404	0.0428
Kurtosis	5.3436	4.6304	4.7111	6.0122

Table 2 Results of the estimated copula-based GARCH models for the bivariate relationships of daily log-returns of S&P500 (USA) with Ibovespa (Brazil), Merval (Argentina) and IPC (Mexico).

Statistic	**S&P500-Ibovespa**			**S&P500-Merval**			**S&P500-IPC**		
Variable	Coef.	Error	Prob.	Coef.	Error	Prob.	Coef.	Error	Prob.
USA a_i	0.1130	0.0437	0.0097	0.0968	0.0356	0.0066	0.1089	0.0486	0.0251
USA b_i	0.8799	0.0536	0.0000	0.8657	0.0253	0.0000	0.8825	0.0408	0.0000
USA d.f.	6.3504	2.4676	0.0101	7.9201	2.9564	0.0074	6.3718	3.5838	0.0754
Latin a_i	0.0836	0.0215	0.0001	0.0766	0.0277	0.0057	0.0555	0.0133	0.0000
Latin b_i	0.8938	0.0519	0.0000	0.8984	0.1201	0.0000	0.9435	0.0128	0.0000
Latin d.f.	7.9471	2.9435	0.0069	3.7965	1.7292	0.0281	4.6953	1.4610	0.0013
AIC			-11.8120			-11.5980			-12.6720

*All estimated parameters obtained statistical significance at 5% level.

Table 3 Ljung-Box Q statistic for residuals of daily returns of S&P500 (USA), Ibovespa (Brazil), Merval (Argentina) and IPC (Mexico) estimated by copula-based GARCH model.

	S&P500-Ibovespa		**Ibovespa**		**S&P500-Merval**		**Merval**		**S&P500-IPC**		**IPC**	
Lag	Q	prob.	Q	prob.	Q	prob.	Q	prob.	Q	prob.	Q	prob.
1	0.011	0.915	1.568	0.210	1.517	0.218	0.041	0.838	0.000	0.998	0.227	0.634
2	0.904	0.636	3.102	0.212	1.685	0.430	1.423	0.490	0.116	0.943	0.291	0.865
3	0.904	0.824	4.965	0.174	4.932	0.176	1.439	0.696	0.117	0.989	0.462	0.927
4	2.287	0.683	5.762	0.217	5.167	0.270	1.442	0.836	1.964	0.742	0.462	0.977
5	5.325	0.377	5.762	0.330	5.795	0.326	1.685	0.890	4.625	0.463	1.085	0.955
6	7.012	0.319	7.805	0.252	6.267	0.393	1.720	0.943	6.496	0.369	1.439	0.963
7	7.060	0.422	7.870	0.344	6.293	0.506	3.013	0.883	6.514	0.481	3.558	0.829
8	7.753	0.457	9.780	0.280	6.300	0.613	3.372	0.908	7.284	0.506	4.632	0.796
9	14.909	0.093	15.629	0.075	6.551	0.683	4.023	0.909	14.153	0.117	16.144	0.064
10	15.042	0.130	16.827	0.078	6.833	0.741	4.792	0.904	14.372	0.156	16.170	0.095

*All estimated parameters obtained statistical significance at 5% level.

Table 4 Estimated parameters of the copulas, values and significance of S_n tests for the relationships of daily volatility and return of the Latin markets considering the influence of U.S. market.

	Ibovespa			**Merval**			**IPC**		
Copula	Parameter	S_n test	Sig.	Parameter	S_n test	Sig.	Parameter	S_n test	Sig.
Normal	0.0375	0.0223	0.248*	-0.0149	0.0208	0.321	0.0249	0.0823	0.000
Student t	0.0375	0.0268	0.143*	-0.0149	0.0321	0.055	0.0249	0.0926	0.000
Frank	0.2148	0.0227	0.229	-0.0851	0.0209	0.304	0.1429	0.0823	0.000
Gumbel	1.0244	0.148	0.037	1.0000	0.0235	0.672	1.0161	0.1225	0.053
Galambos	0.2002	0.1449	0.042	0.0187	0.0235	0.7058	0.1791	0.1217	0.066*

*All estimated parameters obtained statistical significance at 5% level