RT-MAE 2006-14

AN IMPLICIT MEASURE OF TAIL DEPENDENCE AND ITS APPLICATION FOR IDENTIFICATION OF THE CHANGE OF DEPENDENCE BETWEEN ASSETS IN MARKET BOOMS AND CRASHES

by

Adonírio Panzieri, Vladimir Belitsky,
Paulo Ricardo Magalhães Rocha
and
Fernando Pigeard de Almeida Prado

Palavras-Chave: Dependence between assets, extreme dependence, tail dependence, implicit dependence measure, extreme value theory, value at risk, risk management, extreme market conditions, market booms and crashes.
Classificação AMS: C13, C16, C53, F36, G10, G15.

- Agosto de 2006 -

An implicit measure of tail dependence and its application for identification of the change of dependence between assets in market booms and crashes

Adonírio Panzieri, *Banco do Brasil S.A, Brazil,*
Vladimir Belitsky[1], *University of São Paulo, Brazil,*
Paulo Ricardo Magalhães Rocha, *University of São Paulo, Brazil,*
Fernando Pigeard de Almeida Prado, *University of São Paulo, Brazil.*

Abstract

We suggest a new empirical measure of dependence between extremely small or extremely large values of two random variables – such measures are usually called tail dependence measures –, and show how it may be used to identify the change of dependence between two risk factors in extreme market conditions.

Our results and ideas are useful for building portfolios that would be stable in market booms and crashes.

The construction of our empirical measure of dependence employs theoretical results and statistical procedures from the Extreme Value Theory and a particular formula (constructed in Longin (2000)) that suggests how the value-at-risk of a portfolio may be aggregated from the values-at-risk of its components, when the components' returns do not obey a Gaussian distribution. The tools employed in the construction have been chosen to provide an adequate inference at the distribution of exceedences of the random variables which tail dependence is of the interest.

Our empirical dependence measure is the implicit coefficient of correlation of a specific bi-variate extreme value distribution. We note that implicit correlations have been used by practitioners, but mainly on intuitive bases. Thus, an independent contribution of our work is the exposition of how an implicitly defined dependence measure may be studied and applied.

JEL classification: C13, C16, C53, F36, G10, G15.

Key words and phrases: Dependence between assets, extreme dependence, tail dependence, implicit dependence measure, extreme value theory, value at risk, risk management, extreme market conditions, market booms and crashes.

---

[1]The corresponding author: Institute of Mathematics and Statistics, University of São Paulo, Rua do Matão, 1010, CEP 05508-090, São Paulo, SP, Brazil; +55-11-30916129, Fax +55-11-38144135; e-mail belitsky@ime.usp.br

This work was supported by Banco do Brasil S.A., by the Brazilian federal foundations CNPq and CAPES, and by the São Paulo state foundation FAPESP.

1 Introduction

1.1 The first goal of our paper is to suggest a new approach to the following question: [2]

> Given a sample $(x_1, y_1), \ldots, (x_n, y_n)$ of observations of n daily returns of two assets, deduce whether their dependence weakens or strengthens in a market boom. (1)

The suggested approach is based on a new empirical measure of the upper tail dependence of two random variables. This measure is the implicit coefficient of linear correlation of a specific bi-variate extreme value distribution related to these variables. We note that implicit dependence measures are being used by practitioners but the theoretical study of their properties is rather poor. This fact justifies the second goal of our paper: to expose how the properties of an implicit dependence measure may deduces from the theoretical results and the empirical procedures that have been used in its definition.

1.2 In this sub-section, we shall explain what motivated us to introduce a new approach to the question (1).

We start our explanation with the observation that a (1)-like question arises quite commonly. For example, one would face it, when one's objective is to measure financial risks and to manage them via building optimal portfolios. From the importance and the commonness of the question (1) a reader might conclude that the question has been studied by now. Indeed, a brief search in the academic literature reveals that there exists a number of different approaches to the question (1). Some of them use specific coefficients, like the coefficient of the lower tail dependence (see Joe (1997)), some others are grounded on a specific structure of the bivariate return distribution, like, the normality or the ellipticity (see Genton (2004)), and there are others whose advantage over the above mentioned approaches is in that they concern with choosing the distribution that fits best the sample (as the copula approach, for example, exposed in Cherubini *et al.* (2004) or Kolev *et al.* (2005)). Actually, there is a vast literature that shows that one approach may work better than another one depending on the structure of the sample and on final objective of the study of the question (1).

The variety of the existing and actually employed approaches to the question (1) may be explained as a consequence of the variety of possible answers to the following two questions that are embedded in (1):

Which of the sample points are typical for a market boom? (2)

What measure quantifies adequately the dependence in a boom? (3)

The difficulty in finding the unique correct answer to the question (2) is easy to illustrate. For example, consider an observation (x, y) composed of a large return value of one asset and of a moderate positive return value of another one (for concreteness, let $x > 0$ and $|x|$ be relatively large while $y > 0$ and $|y|$ be relatively close to 0). Since a boom is interpreted as the market condition in which asset returns are likely to assume

[2]Certainly, the one day period is taken in (1) just for the sake of argument.

A similar question may be posed with a market crash instead of a market boom. A simple modification, to be presented in Section 3.3, makes our approach applicable for market crashes.

their largest values then the large value of x indicates that the observation (x, y) was observed in a boom. On the other hand, by the same argument, the moderate value of y contradicts this indication. This gives an example of the uncertainty in determining the border between the boom and the non-boom observations. Thus arises the non-uniqueness of the answer to (2).

Turning now our attention to the question (3), we note that the difficulty in finding the uniquely correct answer to it stems essentially from the following two facts: first, that the proportion of the sample points that are observed in booms is relatively small, and second, that they are significantly sparse. The fewness of the boom sample points is due to the fact that booms are typically rare. Their sparseness is due to the fact that they come from the right tail of a non-limited distribution. This, in turn, is a consequence of the fact that an asset return is measured by the logarithmic difference of price, and consequently, its values are assumed to be unlimited. Both the fewness and the sparseness of the boom observations affect the precision of the statistical inference at the dependence structure of two assets in booms. In order to make the inference procedure reliable, practitioners usually resort to simple numerical measures of dependence, like, for example, the coefficient of linear correlation. However, the linear correlation coefficient is usually unable to express adequately a complex dependence structure (for examples of this disability, see Embrechts *et al.* (2004) and Boyer *et al.* (1997)). This trade-off between the adequateness of a measure for representing the dependence and our ability to estimate it precisely from the available data causes the non-uniqueness of the answer to (3).

The above presented arguments imply that one cannot expect to find the unique approach to the question (1) that would be more effective than any other one for all possible structures of the data set $(x_1, y_1), \dots, (x_n, y_n)$. Thus, the existence of different approaches to (1) should be viewed as reasonable and even desirable.

The last conclusion and the importance of the question (1) justify the introduction of new approaches to (1). Ours is formulated and analyzed in the present paper. Its construction is the contents of Section 3.1. The results of its application to several real cases are reported in Section 3.3. These results are then analysed in Section 4; the analysis' conclusions suggest the usefulness and the efficiency of our approach.

1.3 Our approach to the question (1) is based on a new measure of dependence between two random variables that we construct and study in the present paper. In this subsection, we shall depict the construction of this measure and explain how it is employed by our approach.

Our measure of dependence between two random variables is constructed from a sample of their observations and expresses the cointegration between the variables in the upper tail of their distribution function. The measure has a free parameter $\alpha \in [0.95; 1)$ that reflects how far this upper tail is. Given the value of α, the construction steps are as follows: (1) we employ two distinct methods to get two estimates of the α-quantile of the linear combinations of the random variables, (2) we find the value of a particular quantity involved in these methods so that the methods' estimations coincide; the found value is declared to be the dependence measure. The quantile estimation methods employed in our construction use tools from the Extreme Value Theory (the classical reference for the EVT is Leadbetter *et al.* (1983); recent developments and applications are exposed in

Embrechts *et al.* (1997)). This allows us to measure the dependence between the random variables in the upper tail of their bi-variate distribution function, and, at the same time, to contend with the problem of the fewness and sparseness of the sample points that came from this tail (this problem has been described above). The price paid for this advantage is the implicit character of the construction of our dependence measure. The point here is that the results of implicit constructions do not admit an easy analysis.[3] There are diverse reasons for this. In our case, the main difficulty of the analysis of our implicit dependence measure stems from the fact that our construction employs a statistical procedure that has no analytic formulae for the estimators that it produces. Nevertheless, using empirical facts about the functioning of this procedure, we succeed to prove that (a) the constructed measure indeed expresses a dependence between the random variables, and that the closer the measure's value to zero, the weaker the dependence (Statement 2 in Section 2.6); and that (b) the constructed measure reflects mainly the structure in the upper tail of the distribution function of the variables, and that a farer tail corresponds to a larger value of the parameter α (Statement 4 in Section 2.6). The properties (a), (b) just depicted suggest that our implicit dependence measure may be used to answer the question (1) in the following way: calculate from the sample $(x_1, y_1), \ldots, (x_n, y_n)$ the implicit dependence measures for a sequence of values of α that tend to 1; if the measure's value increases as $\alpha \nearrow 1$, conclude that the dependence between the assets strengthens along with the strengthening of a market boom, if the measure's value decreases conclude that the dependence weakens. This is the essence of our approach to the question (1).

2 Implicit measure of tail dependence

2.1 Preliminary notations and definitions

In this section, we shall construct a specific measure of dependence between two random variables. These variables will be denoted generically by X and Y. In the later section, the constructed measure will be applied for answering the question (1). Because of this application, we shall interpret X and Y as the daily returns of two market assets already in the construction arguments. We shall abuse the notations and for the sake of brevity, call the assets by the names of their returns: "the X and the Y assets".

For $p \in [0; 1]$, we introduce

$$Z_p = pX + (1-p)Y \tag{4}$$

It follows from (4) and from the interpretation of X and Y that Z_p expresses the daily return of the portfolio composed of the X and the Y assets with the respective weights p and $1-p$. We shall refer to this portfolio as "the Z_p portfolio". It will play an important role in our constructions.

For a random variable $\mathcal{X}$, we shall denote by $F_{\mathcal{X}}$ its distribution function. For an arbitrary $\alpha \in (0; 1)$, we shall denote by $\mathrm{VaR}_\alpha(\mathcal{X})$ the α-quantile of $F_{\mathcal{X}}$ and call it *the*

[3] As a corroborating example, think of what the implied volatility coefficient measures when we extract it from the Black-Scholes option valuation formula applied to an asset whose return does not follow a Brownian motion.

value at risk at level α of the asset $\mathcal{X}$. These name and notation belong to the field in which an α-quantile of $\mathcal{X}$ is interpreted as a measure of risk of a risk factor (an asset, for example) whose return has the distribution $F_{\mathcal{X}}$ (see Jorion (1997)). [4] Since we shall employ results and techniques that have been developed within this field, thus we accede to these name and notation.

2.2 Two methods for estimating high quantiles of a portfolio

What we shall construct and denominate *implicit measure of dependence between random variables X and Y* is a quantity that stems from a comparison between two estimates of a high quantile of the random variable Z_p. These estimates are obtained by two different methods from a sample $(x_1, y_1), \ldots, (x_n, y_n)$ of observations of (X, Y). These methods are described in the present section. From now one, they will be denominated *the first* and *the second methods for estimating portfolio's risk* (note: "a portfolio's risk" is the interpretation of "a quantile of Z_p" in accordance to our explanation in Section 2.1). These names will facilitate the referring to the methods in our arguments.

2.2.1 *The first method* employs essentially the fact that x_j and y_j from the sample are simultaneous[5] observations of X and Y. This fact together with the definition (4) of the random variable Z_p ensure us that if for an arbitrarily fixed $p \in [0; 1]$ we define

$$z_1(p) := px_1 + (1-p)y_1, \ldots, z_n(p) := px_n + (1-p)y_n \tag{5}$$

then we can regard the sequence $z_1(p), \ldots, z_n(p)$ as a sample of observations of Z_p. The first method estimates $\mathrm{VaR}_\alpha(Z_p)$ from this sample using a specific statistical procedure called Peaks over Threshold method (to be abbreviated further as the POT method), that has been developed within the Extreme Value Theory (the EVT, for short).

We postpone the exposition of the POT method to Section 2.5, and now we shall only explain what obliges us to employ it. The cause is the necessity to reconstruct the right tail of F_{Z_p} from the sample (5) of observations of Z_p (we note that we need the far right tail of F_{Z_p} because our aim is to estimate high quantiles of Z_p). However, due to the reasons explained in the Introduction, the sample (5) would contain quite few observation from the tail of F_{Z_p} and these observations would be significantly sparse. This fewness and sparseness of observations have an adverse effect on the precision and the confidence of the estimation of the right tail of F_{Z_p} that one would obtain by any non-parametric statistical procedure. A way to enhance the estimation precision is to use the Parametric Statistics. Among the parametric estimation methods applicable to the case, the POT method is preferable by two reasons: first, it is frequently and successfully used in contemporary studies of financial data (for example, Longin (2005) and Tsay (2002) and the references therein); second, it is intimately related to a so-called Annual Maxima method from the EVT; Remark 1 will specify this relation and explain its importance in our case.

[4] Banks, investors and financial institutions would be typically interested in the losses lower than the 5%-quantile. Accordingly, α is usually either smaller than 0.05 or larger than 0.95.

[5] That is, observed at the same moment. This is an assumption of our mathematical framework; it is implicitly assumed when we say "(x, y) is an observation of (X, Y)". Note that this assumption may fail in applications, in particular, in financial ones. For example, it fails if X and Y mean the daily equity index returns of two countries which stock exchange markets close at significantly different times.

2.2.2 To present and to justify *the second method*, we need the notion of the extreme random variable from the EVT. We recall it below.

Let $\{(X_i, Y_i), i \in \mathbb{N}\}$ be a sequence of independent copies of (X, Y). Let

$$M_m := \max\{X_1, \ldots, X_m\}, \qquad L_m := \max\{Y_1, \ldots, Y_m\} \text{ for all }, \ m \in \mathbb{N} \tag{6}$$

Let us then assume that F_X and F_Y belong to the maximum domain of attraction of the Extreme Value distributions[6]. Then, the Central Theorem of the EVT (see Theorem 1.4.2 in page 11 of Leadbetter *et al.* (1983)) guarantees that there exist appropriate numeric sequences $\{a_m\}$, $\{b_m\}$, $\{a'_m\}$ and $\{b'_m\}$ such that

$$M := \lim_{m\to\infty} (a_m(M_m - b_m)), \text{ and } L := \lim_{m\to\infty} (a'_m(L_m - b'_m)) \tag{7}$$

are two non-generate and well defined random variables. Those are called *the extreme random variables corresponding to the distribution functions F_X and F_Y*, respectively. Let

$$\rho = (\mathrm{Var}(M) \cdot \mathrm{Var}(L))^{-1/2} \times \mathrm{E}\left[(M - \mathrm{E}(M))(L - \mathrm{E}(L))\right] \tag{8}$$

denote *the coefficient of linear correlation between the extreme random variable corresponding to F_X and F_Y* (in (8), Var means the variance).

Now, we can formulate the second method: it gets $\mathrm{VaR}_\alpha(X)$ and $\mathrm{VaR}_\alpha(Y)$ from the respective samples $x_1, \ldots, x_n$ and $y_1, \ldots, y_n$ using the POT method, it gets then ρ from the sample $(x_1, y_1), \ldots, (x_n, y_n)$,[7] and finally, it estimates $\mathrm{VaR}_\alpha(Z_p)$ by the value of the expression

$$\left[(p\mathrm{VaR}_\alpha(X))^2 + 2\rho p(1-p)\mathrm{VaR}_\alpha(X)\mathrm{VaR}_\alpha(Y) + ((1-p)\mathrm{VaR}_\alpha(Y))^2\right]^{1/2} \tag{9}$$

The second portfolio risk estimation method is based on the following statement which will be referred to as Longin's result, after the name of its author:

Statement 1 (Longin (2000)) *Let α be from* $[0.95; 1)$ *(this range of α has been explained in the footnote 4). Let F_X and F_Y belong to the maximum domain of attraction of the Extreme Value distributions. Let $\mathrm{VaR}_\alpha(X)$, $\mathrm{VaR}_\alpha(Y)$ and ρ be estimated by appropriate methods from the Extreme Value Theory* (these methods will be specified in Remarks 1 and 2 below). *Then there are empirical evidence of and theoretical indication to the fact that for all $p \in [0; 1]$, the expression (9) provides an estimate of the true value of $\mathrm{VaR}_\alpha(Z_p)$ that is more precise and more reliable than the estimates provided by traditional methods* (the traditional methods tested by Longin are the historical method, the method based on the assumption that (X, Y) is bi-variate normal, and the methods based on fitting the GARCH and the EWMA models to the data).

Remark 1. Longin (2000) employs the Maxima in Blocks method (known also under the name the Annual maxima method) to calculate the VaR's in his expression (9). The description of this method may be found in Reiss and Thomas (1997) and in

[6] This is not a restrictive assumption, as we shall explain in Remark 3.

[7] How one can estimate ρ from this sample will be discussed in Section 2.3 after the eq. 14. This aspect however, is not important for us, since in our constructions, the value of ρ will be determined implicitly.

Embrechts *et al.* (1997). These references present also theoretical results that ensure that the Annual Maxima method and the POT method produce the same estimate of $\text{VaR}_\alpha(\mathcal{X})$ of a random variable $\mathcal{X}$ from an infinite sample of its observations. In practice, however, samples are finite, and empirical studies indicate that if in a finite sample case both the POT method and the Annual maxima method apply then the former is more effective than the latter (Këllegi and Gilli (2000), Panzieri (2002), Smith (2000), Tsay (2002)). By this reason, our second method employs the POT method for calculating VaR 's in the expression (9).

Longin's result is recent. The rest of this section is devoted to the exposition of our understanding of its importance and usefulness.

Longin's result resembles the formula

$$\begin{aligned}&\text{VaR}_\alpha(Z_p) =\\&= \left[(p\text{VaR}_\alpha(X))^2 + 2\rho p(1-p)\text{VaR}_\alpha(X)\text{VaR}_\alpha(Y) + ((1-p)\text{VaR}_\alpha(Y))^2\right]^{1/2}\end{aligned} \tag{10}$$

that is known to be exact for any α and any p, if (X, Y) is a Gaussian (that is, bi-variate normal) pair of random variables and if ρ is the coefficient of linear correlation between X and Y (see Jorion (1997) for details). However, asset returns are usually non-Gaussian and the equality (10) is known to fail for non-Gaussian cases. Thus, the formula (10) is of no help, if one wants to use its right hand side as an estimate of $\text{VaR}_\alpha(Z_p)$.

What Longin shows in his paper (2000) is that if (X, Y) are not Gaussian, then the closeness between the true value of $\text{VaR}_\alpha(Z_p)$ and the right hand side of (10) is significantly enhanced when the coefficient of linear correlations between X and Y is substituted by the coefficient of linear correlations between the extreme random variables corresponding to X and Y. To see the usefulness of Longin's discovery, recall from Section 2.1 that when α is close to 1 then $\text{VaR}_\alpha(Z_p)$ is a measure of the risk of the short position composed of the X and the Y assets with the respective weights p and $1-p$. Thus, Longin's result is useful in practice when one wishes to estimate the risk of this portfolio but does not have a sample of its values. This situation happens, for example, if one possesses solely the largest daily returns of the X and the Y assets over the week period during several weeks, but needs to estimate the daily value at risk of the portfolio composed of these assets. In this case, the Longin approach[8] would be possibly the only mean to get an estimate of $\text{VaR}_\alpha(Z_p)$. The usefulness of the Longin's result just exposed compensates the absence of its rigorous analytic proof (as Longin states on page 1109, the structure of (9) was inspired by (10) on an ad hoc basis).

2.3 The intuition

Assume we are given a sample $(x_1, y_1), \ldots, (x_n, y_n)$ of observations of (X, Y). Let us fix two real numbers to be denoted by α and p; the first one must be close to 1, say larger than 0.95, the second is in $(0; 1)$. Let us then define $\varrho_{X,Y}(\alpha; p)$ as the solution of the

[8] By "Longin's approach" we mean the formula (9) and the methods used for estimating the quantities involved in it.

following equation:

$$\mathrm{VaR}_\alpha(Z_p) = \\ = \left[(p\mathrm{VaR}_\alpha(X))^2 + 2\varrho_{X,Y}(\alpha;p)p(1-p)\mathrm{VaR}_\alpha(X)\mathrm{VaR}_\alpha(Y) + ((1-p)\mathrm{VaR}_\alpha(Y))^2\right]^{1/2} \tag{11}$$

where $\mathrm{VaR}_\alpha(Z_p)$ is calculated from $(x_1, y_1), \ldots, (x_n, y_n)$ by the first portfolio risk estimation method (described in Section 2.2.1), and where $\mathrm{VaR}_\alpha(X)$ and $\mathrm{VaR}_\alpha(Y)$ are gotten from the respective samples $(x_1, \ldots, x_n)$ and $(y_1, \ldots, y_n)$ by the POT method.

The construction of $\varrho_{X,Y}(\alpha;p)$ is simpler than, but nevertheless intrinsically equivalent to the construction of the implicit dependence measure that will be defined in the next sub-section and that is the principal subject of our study. In order to provide a reader with an intuition that will help him/her to understand the construction and properties of the "true" implicit dependence measure, we shall analyse now its "toy model" $\varrho_{X,Y}(\alpha;p)$ that will be called for simplicity by the same name, i.e., implicit dependence measure. The principal conclusions of this analysis are emphasized in (12), (13) and (14).

Let us observe that $\varrho_{X,Y}(\alpha;p)$ presents the value of the coefficient ρ (defined in (8)) that makes the first and the second portfolio quantile estimation methods produce the same value of $\mathrm{VaR}_\alpha(Z_p)$. Let us then observe that the second method uses a unique quantity that bears the information about the dependence between X and Y; it is the coefficient ρ. Combining these two observations, we conclude that

what we have called the implicit dependence measure
indeed expresses a dependence between X and Y (12)

We note also that the first of the above observations explains the word "implicit" in the name of $\varrho_{X,Y}(\alpha;p)$.

What kind of dependence is it? To answer this question, note that the sample $(x_1, y_1), \ldots, (x_n, y_n)$ "influences" the value of $\varrho_{X,Y}(\alpha;p)$ through the values of $\mathrm{VaR}_\alpha(X)$, $\mathrm{VaR}_\alpha(Y)$ and $\mathrm{VaR}_\alpha(Z_p)$. Note next that when α is close to 1 then these three VaR's depend on relatively large values of the respective samples $x_1, \ldots, x_n$, $y_1, \ldots, y_n$, and $z_1(p), \ldots, z_n(p)$. Note finally that an observation $z_i(p)$ from the latter sample cannot be large, if both x_i and y_i are small. Thus,

the expressed dependence measures mainly the relation between large values of X and Y (13)

We note that we shall formalize the meaning of the term "dependence between large values" in Sections 2.6. For this, we shall need to present a detailed analysis of the functioning of the POT method.

Let us finally explain why

we expect advantages from the implicit construction of the dependence measure (14)

We start our explanation with the observation that it is impossible to construct a single observation of the pair (M, L) (recall M and L have been defined in (7)) from the sample $(x_1, y_1), \ldots, (x_n, y_n)$; this is because the sample is finite while both M and L are limit random variables defined through infinite sequences of X's and Y's. Since ρ is the coefficient of linear correlation between M and L, then this impossibility becomes the principal obstacle that one would face, when one's objective is to estimate ρ directly from the sample $(x_1, y_1), \dots, (x_n, y_n)$. A typical approach that achieves this objective and avoids the obstacle consists of the following steps: *(a)* to fix an integer m; *(b)* to divide the sample in $k = n/m$ blocks with m observations in each block and to define

$$\begin{array}{l} x_1^{(m)} := \max\{x_1, \dots, x_m\}, \dots, x_k^{(m)} := \max\{x_{n-m+1}, \dots, x_n\} \\ y_1^{(m)} := \max\{y_1, \dots, y_m\}, \dots, y_\ell^{(m)} := \max\{y_{n-m+1}, \dots, y_n\} \end{array} \tag{15}$$

(c) to use then the sequence $(x_1^{(m)}, y_1^{(m)}), \dots, (x_k^{(m)}, y_k^{(m)})$ for estimating $\rho_{(M_m, L_m)}$ (which is possible because this sequence is a sample of k observations of (M_m, L_m)), and to accept the constructed estimate of $\rho_{(M_m L_m)}$ as an estimate of ρ. Note that the last step is justified by the convergence $\rho_{(M_m, L_m)} \to \rho$, as $m \to \infty$, that follows from the definitions of M, L, M_m and L_m.

Remark 2. An estimate of $\rho_{(M_m L_m)}$ may be calculated as the sample correlation coefficient of the sequence $\{(x_i^{(m)}, y_i^{(m)})\}$, but may be also obtained employing more elaborated techniques. One of such techniques may be built on the basis of the result of Tiago de Oliveira (1973) about the dependence structure between extreme random variables. This technique has been developed and applied by Longin (2000). This is one of the contributions of Longin (2000).

The arguments presented above make it clear that a direct estimation of ρ requires establishing the value of m. This is a rather difficult problem, because if m is taken small then the distribution of (M_m, L_m) is essentially different from that of (M, L), and consequently, $\rho_{(M_m, L_m)}$ is a bad approximation of ρ, while if m is taken large then the size of the sample of $\{(x_i^{(m)}, y_i^{(m)})\}$ becomes small and consequently, the precision of estimation of $\rho_{(M_m, L_m)}$ worsens. Calculating ρ in an implicit way frees a statistician from the necessity of looking for the optimal value of m. This is the first principal reason for us to affirm (14).

The second principal reason for us to affirm (14) is our belief that the construction of our implicit dependence measure allows the data to express adequately the information they contain in respect to the dependence between X and Y. This belief has two principal grounds. They are described below.

(a) The first ground is the combinations of the following two facts: (1) that the POT method is an effective tool for obtaining $\text{VaR}_\alpha(\mathcal{X})$ from a sample of $\mathcal{X}$ when α is close to 1 (we have mentioned this property of the POT method in Section 2.2.1); and (2) that we use the POT method to calculate all the quantities of (11) that, due to our construction, must be determined directly from the data.

(b) The second ground is the combination of the following two facts: (1) that among the methods designed for estimating $\text{VaR}_\alpha(Z_p)$ from $(x_1, y_1), \dots, (x_n, y_n)$ for large α, the first portfolio risk estimation method is probably the most efficient; and (2) that our implicit dependence measure is the value of a coefficient from the second portfolio quantile estimation method that forces it to produce the same $\text{VaR}_\alpha(Z_p)$ as the first method. We note that our statement about the efficiency of the first portfolio risk estimation method stems from the fact that it gets $\text{VaR}_\alpha(Z_p)$ by the POT method from the sample

$z_1(p), \ldots, z_n(p)$ that bears all the information about the dependence between X and Y that is contained in the sample $(x_1, y_1), \ldots, (x_n, y_n)$ and that is relevant for estimating a quantile of Z_p; this is because each $z_i(p)$ is constructed by the expression $px_i + (1-p)y_i$ which is equivalent to the expression (4) that defines Z_p as $pX + (1-p)Y$.

2.4 The definition

Given a sample $(x_1, y_1), \ldots, (x_n, y_n)$ of independent observations of a pair of random variables (X, Y),[9] and given a real number α close to 1 (how close α must be to 1 and why will be explained at the end of this section), we denote by $\varrho_{X,Y}(\alpha)$ the value of ρ that minimizes the value of the expression

$$\sum_{p\in\mathcal{G}} \Big\{ \left[(p\mathrm{VaR}_\alpha(X))^2 + 2\rho p(1-p)\mathrm{VaR}_\alpha(X)\mathrm{VaR}_\alpha(Y) + ((1-p)\mathrm{VaR}_\alpha(Y))^2\right]^{0.5} - [\mathrm{VaR}_\alpha(Z_p)] \Big\}^2 \quad (16)$$

where

$$\mathcal{G} = \{0.01, 0.02, \ldots, 0.99\} \quad (17)$$

and where $\mathrm{VaR}_\alpha(X)$ and $\mathrm{VaR}_\alpha(Y)$ are obtained by the POT method from the respective data sets $x_1, \ldots, x_n$ and $y_1, \ldots, y_n$, and where for each p from the grid (17), $\mathrm{VaR}_\alpha(Z_p)$ is obtained by the POT method from the data set $z_1(p), \ldots, z_n(p)$ defined in (5).

We call the quantity $\varrho_{X,Y}(\alpha)$ thus defined *implicit measure of the upper tail dependence at the level α between the variables X and Y*, or, for short within the present paper, *implicit dependence measure.*

Note that the first line of the expression in the figure brackets of (16) is the estimate of the α-quantile of Z_p provided by *the first method* (described in Section 2.5.2), while the second line of this expression is the estimate of the same quantile provided by *the method method* (described in Section 2.5.1). Thus, $\varrho_{X,Y}(\alpha)$ may be interpreted as the value that reconciles uniformly in p, the estimates of $\mathrm{VaR}_\alpha(Z_p)$ provided by these two methods.

The above given interpretation of $\varrho_{X,Y}(\alpha)$ exposes clearly the difference between it and its "toy analogue" $\varrho_{X,Y}(\alpha; p)$ (that has been constructed in Section 2.3): the latter depends on p while the former does not. Thus, a reader may inquire why we focus our interest at the former, or, in other words, why we suppress the dependence on p of our implicit dependence measure. We have had two reasons for this. The first reason is that we have an empirical evidence of this independence: for several real samples $(x_1, y_1), \ldots, (x_n, y_n)$ that we have analysed (these samples are described in Section 3.3), we calculated $\varrho_{X,Y}(\alpha; p)$ by the formula (11) and verified that its value changes insignificantly with the variation of p within $[0, 1]$. The second reason stems from the combination of the following two facts: that our implicit dependence measure is an implied value of the coefficient ρ from the Longin's formula (9), and that ρ does not depend on p by the very definition.

[9] Our arguments hold true also for dependent observations, provided the results from the Extreme Value Theory used in our arguments apply for the dependence structure between the observations. Such dependence structures may be described following the tools and ideas presented in Leadbetter *et al.* (1983), but this description is not an objective of our work.

We note that we purge the dependence on p of our implicit dependence measure by averaging with respect to p. This is why $\sum_{p \in \mathcal{G}}$ has appeared in (16). In our applications (see Section 3.3), this summation runs over the grid (17). Our choice of the grid was ad hoc, but proved to be satisfactory; a more or a less refined partition of $[0; 1]$ may be used.

We note that an optimization procedure is required for calculating $\varrho_{X,Y}(\alpha)$ from a real data set since it is defined as an *argmin* of the expression (16). In our application (see Section 3.3), we optimized over the grid $-1, -0.99, \ldots, 0.99, 1$. Certainly, other optimization procedures may be employed, if desired.

Finally, let us explain and quantify the closeness of α to 1 that has been assumed in the definition of $\varrho_{X,Y}(\alpha)$. We observe that our implicit dependence measure is designed to get the dependence between large values of X and Y. This is the intrinsic reason why the measure construction requires for estimating high quantiles of X, of Y and of $Z_p = pX + (1-p)Y$. The POT method that we employ for this estimation shows to be more effective than other methods for estimating α-quantiles with $\alpha \geq 0.8$ (certainly, this lower bound is not exact and depends ultimately on the underlying distribution). However, besides the POT method, we employ the Longin's result that has been established for $\alpha \geq 0.95$ (recall the Statement 1). Thus, the constraint $\alpha \in [0.95; 1)$.

2.5 How the POT method estimates VaR

Estimates of VaR of the variables X, Y and Z_p play the central role in our definition of the implicit dependence measure. We recall that these estimates are calculated from samples of these variables by the POT method. Certain details of this method will be essentially used in our further arguments. In order to expose these details, we shall show below how the POT method would estimate $\mathrm{VaR}_\alpha(X)$ from the sample $x_1, \ldots, x_n$ of observations of X.

The theoretical background of the POT method is the result of Pickands (1975)[10] that states that given the distribution of X, there may be found a real number ξ and a function $\beta(u), u \geq 0$, such that (below we assume that F_X has an infinite support and write accordingly $u \to \infty$)

$$\lim_{u \to \infty} \left| \mathrm{P}\left[X - u \leq x \mid X > u\right] - G_\xi(x/\beta(u)) \right| = 0 \text{ uniformly in } x \geq 0 \tag{18}$$

where G_ξ is the so-called Generalized Pareto distribution (GPD) with the parameter value ξ (the form of the GPD may be found in Embrechts et al. (1997), for example). It is worth noting that the convergence (18) does not occur for any F_X, as some authors state. However, Pickands showed that it does occur for a wide class of distribution functions which includes "most "textbook" continuous distribution functions" (the citation is from Pickands (1975)). Thus, when the Pickands' result is employed, it is implicitly assumed that the approximated F_X belongs to this class. (Actually, in the real cases to which we apply our results, we shall get an indication that this assumption is valid; see Section 4(e) for details in respect.)

[10]This result is stated in Thm. 3.4.13(b) of Embrechts *et al.* (1997); the latter reference and Reis and Thomas (1997) provide the details of the POT method, that are beyond those presented by us.

Remark 3. In accordance to the above statement, we shall assume that F_X and F_Y, the distribution functions of X and Y, admit the approximation by a GPD in the sense of (18) (this implies automatically that F_{Z_p} also admits such approximation). The important fact that we want to emphasize here is that this assumption implies that F_X and F_Y belong to the maximum domain of attraction of the Extreme Value distributions. This implied property of F_Y and F_X is necessary for us to be able to employ *the second portfolio quantile estimation method* (described in Section2.2.2). This implication is a well known fact. A reader can find the details in Reiss and Thomas (1997) or in Embrechts *et al.* (1997) in the chapters that present the theoretical relation between the POT and the Annual Maxima methods.

The "POT method" is the name for the following procedure:

(a) Establish the value of u, starting from which the approximation of $\mathrm{P}\left[X-u \le x \mid X>u\right]$ by $G_\xi(x/\beta(u))$ is sufficiently good (the criterion for goodness depends on diverse factors that will be discussed in Section 3.2);

(b) Estimate ξ, $\beta(u)$ and $\mathrm{P}[X>u]$ from the data set $x_1,\ldots,x_n$ (we shall denote the estimates by $\hat{\xi}$, $\hat{\beta}(u)$ and $\hat{\mathrm{P}}[X>u]$).

The POT methods produces the function (the expression in the figure brackets in (19) below is determined by the form of Generalized Pareto distribution)

$$\hat{F}_X(x) = 1-\hat{\mathrm{P}}[X>u]\left\{\left(1+\hat{\xi}\cdot(x-u)/\hat{\beta}(u)\right)^{-1/\hat{\xi}}\right\},\ x\ge u;\ \hat{F}_X(x)=0,\ x<u \quad (19)$$

that is believed, due to the Pickands' result, to be a good approximation for $F_X(x)$ in the region $x \ge u$ (called *the tail of F_X beyond the threshold u*). Accordingly, $\hat{F}_X^{-1}(\alpha)$ is taken as the estimate of $\mathrm{VaR}_\alpha(X)$ produced by the POT method (provided $u<\hat{F}_X^{-1}(\alpha)$, which however, will be always the case in our study). Note that the estimate of the true $\mathrm{VaR}_\alpha(X)$ will be denoted by $\mathrm{VaR}_\alpha(X)$ either; this will not cause any confusion because the true value is unknown in the framework of our study (since F_X is assumed to be unknown).

The POT method leaves a statistician with a choice of how exactly he/she would establish the value of u and find the estimates $\hat{\xi}$, $\hat{\beta}(u)$ and $\hat{\mathrm{P}}[X>u]$. In fact, there exist various statistical procedures for this. In our case, we got $\hat{\xi}$ and $\hat{\beta}(u)$ via the Maximum Likelihood procedure that has been encoded by Reiss and Thomas and distributed together with their book (1997). We calculated $\hat{\mathrm{P}}[X>u]$ as the fraction of the sample points that lie to the right of u. As for the way we used to establish u, it will be discussed in Section 3.2 below.

2.6 The principal properties

What we call *the principal properties* of our implicit dependence measure are the properties that make it applicable for solving the question (1). We have stated them on a heuristic level in (12) and (13). These properties will be formalized and proved in the Statements 2 and 4 below.

Statement 2. *The implicit measure of tail dependence between X and Y is an estimate of the coefficient of linear correlation between two extreme random variables that are related to X and Y. This relation is such that the stronger is the linear correlation the stronger is the dependence between X and Y.*

Proof. Recall from (8) that ρ denotes the coefficient of linear correlation between M and L and that these M and L are two extreme random variables that correspond to X and Y. Recall that the implicit coefficient, $\varrho_{X,Y}(\alpha)$, is the value of ρ implied by the relation (16) whose parameters are determined by the sample $(x_1, y_1), \ldots, (x_n, y_n)$ of observations of (X, Y). Thus, $\varrho_{X,Y}(\alpha)$ is an estimate of ρ determined by the sample. These facts altogether justify the assertion of the first sentence of Statement 2.

Recall the definitions (6) of the random variables M_m and L_m. Recall next that the coefficient of their linear correlation has been denoted by $\rho_{(M_m,L_m)}$. Recall finally from our arguments presented right after (14) that $\varrho_{X,Y}(\alpha)$ actually estimates $\rho_{(M_m,L_m)}$ for some large but finite m. (This fact does not contradict the conclusion that $\varrho_{X,Y}((\alpha)$ estimates ρ because ρ and $\rho_{(M_m,L_m)}$ must be close for large m.) Now, since for a finite m, it holds that

$$\mathrm{P}\left[M_m \leq x,\, L_m \leq y\right] = \left(\mathrm{P}\left[X \leq x,\, Y \leq y\right]\right)^m, \qquad \forall x, y \tag{20}$$

then the assertion of the second sentence of Statement 2 follows.

Remark 4. Once our dependence measure between X and Y is the implicit linear correlation coefficient between M_m and L_m, thus the measuring of dependence between X and Y by our measure has the same advantages and disadvantages as the measuring of dependence between M_m and L_m by the coefficient of their linear correlation. In particular:

(a) Given the value of the linear correlation between M_m and L_m, we cannot infer[11] exactly at the structure of their dependence, and, in particular, at the values of $\mathrm{P}[M_m \leq x,\, L_m \leq y]$ for all $(x, y) \in \mathbb{R}^2$. Accordingly, the relation (20) does not help to infer at the dependence structure between X and Y from the value of the value of our implicit dependence measure. The role of this relation in our arguments is that it implies the second assertion of the Statement 2.

(b) The second assertion of the Statement 2 and the usual interpretation of the value of the linear correlation coefficient justify altogether the following rule: the closer to 0 the value of our implicit dependence measure, the weaker the dependence between X and Y. This rule will be essentially used in our approach to the question (1).

The second property of our implicit coefficient, to be formulated in Statement 4, is based on the fact that higher order statistics contribute more than lower order statistics to the estimate of VaR_α obtained by the POT method, when α is close to 1. There are actually two causes of this fact. We shall start our presentation below with the one that is easer to be exposed and that makes clear the mathematical meaning of the concept "the contribution of a sample point to an estimate". Next, we shall formalize this concept and use this formalism to present the second cause of the cited fact. Then, we shall formulate

[11]Unless the linear correlation coefficient is either 1 or −1.

this fact in Statement 3, and finally, we shall use it to deduce Statement 4. We recall that Statement 4 formalizes in what sense $\varrho_{X,Y}(\alpha)$ measures the tail dependence between X and Y.

Let us recall from Section 2.5 certain steps of the implementation of the POT method and of its use in getting $\mathrm{VaR}_\alpha(X)$ from a sample $\{x_i\}$ of n observations of X. Recall first that given the value of the threshold u, estimates of the parameters ξ and $\beta(u)$ are determined from a log-likelihood function that depends on those points from $\{x_i\}$ that lie to the right of the threshold u,[12] while the estimate of $\mathrm{P}[X > u]$ is determined as the proportion of the sample points that lie to the right of u. Recall next the threshold u and these estimates determine the function $\hat{F}_X$ via (19). Recall finally, that we set $\mathrm{VaR}_\alpha(X)$ as $\hat{F}_X^{-1}(\alpha)$. All the facts listed above imply that if we take arbitrarily an observation x that lies to the left of the threshold u and change a little its position, assuring that it still does not exceed u, then the value of $\mathrm{VaR}_\alpha(X)$ will not change, while if we take an observation x that lies to the right of u and change a little its position, assuring that it stays still to the right of u, then the value of $\mathrm{VaR}_\alpha(X)$ will change.

The arguments of the above paragraph justify that our intuitive concept of "the influence of an observation x on an estimate" may be given the mathematical form in the following way. Let $\delta_{x_1}, \ldots, \delta_{x_n}$ denote non intersecting intervals with centers at $x_1, \ldots, x_n$ respectively, and such that for each i, $\mathbb{P}[X \in \delta_{x_i}] = \varepsilon$ for a fixed very small $\varepsilon > 0$ (to avoid unnecessary complications, we take ε such that none of δ_x contains the threshold u). Let $\hat{\theta}$ denote the estimate obtained from $(x_1, \ldots, x_n)$ of some quantity θ.[13] By the *influence of an observation x on $\hat{\theta}$* we shall mean the absolute value of the amplitude of the change in $\hat{\theta}$ caused by the change of the position of x within the interval δ_x, when all other observations are kept fixed. By "relative influence of an observation x" we then mean its influence normalized by the sum of the influences of all the observations.

We can now reformulate the conclusion of the last but one paragraph in the following way: the observations that are larger than the threshold u influence $\mathrm{VaR}_\alpha(X)$, while those that are smaller than u do not. This is the first cause for the Statement 3 to hold. As we have said, there is one more cause. It will be presented below.

We start noting that the POT method has been designed to infer at the behavior of a tail of a distribution function. Our concept of "the tail of a distribution" suggests then that the larger the value of an observation from a sample $(x_1, \ldots, x_n)$, the more significant is its influence on the inference at the right tail of the distribution function from which the sample has been extracted. Now, since the POT method assumes that the right tail lies to the right of the threshold u, then one expects that the order statistics $x_{(i+1)}$ influences the estimate of the right tail more than $x_{(i)}$ does, provided both are larger than u. This is true and may be indeed observed when the POT method is being applied. Unfortunately, we cannot prove this fact analytically because there do not exist analytic expressions for the estimates $\hat{\xi}$ and $\hat{\beta}(u)$ that are used by the POT method to determine the estimate of the right tail of a distribution function. In order to expose what we observe empirically when we apply the POT method, let $\hat{F}_X$ denote the estimate of F_X produced by the POT

[12]The exact expression of the log-likelihood function is not important for our arguments here.

[13]In our case, the quantity in interest is $\mathrm{VaR}_\alpha(X)$. Note that in our text, VaR_α actually means $\widehat{\mathrm{VaR}}_\alpha$. We do not use $\widehat{\mathrm{VaR}}_\alpha$ because the text makes it clear that "VaR_α is calculated by the POT method from a sample".

method from a sample $(x_1, \dots, x_n)$ and let $\hat{F}_X^{\text{new},i}$ denote the estimate produced by the same method after we have perturbed the position of the order statistics $x_{(i)}$ of this sample (provided $x_{(i)} > u$). What we typically observe is the following fact: $|\hat{F}_X(x) - \hat{F}_X^{\text{new},i}(x)|$ grows with the growth of x for each fixed i, and grows with the growth of i for each fixed x. The combination of this fact with the definition $\text{VaR}_\alpha(X) := \hat{F}_X^{-1}(\alpha)$ yields that the influence of an observation $x_{(i)}$ on $\text{VaR}_\alpha(X)$ increases with α when i is kept fixed, and increases with i when α is kept fixed. This fact is the second cause of the following property.

Statement 3. *Let $(x_{(1)}, \dots, x_{(n)})$ be the ascending order statistics of the sample $(x_1, \dots, x_n)$ of observations of X and let $\text{VaR}_\alpha(X)$ be obtained from this sample by the POT method. Then, $x_{(i+1)}$ influences $\text{VaR}_\alpha(X)$ more than $x_{(i)}$, for all i. Moreover, as α increases to 1, the distribution of the relative influences of $x_{(1)}, \dots, x_{(n)}$ on $\text{VaR}_\alpha(X)$ becomes more concentrated on the highest order statistics.*

Certainly, Statement 3 holds for Y and its sample $(y_1, \dots, y_n)$, as well as for Z_p and its sample $(z_1(p), \dots, z_n(p))$, for each p. Combining these statements with the construction of $\varrho_{X,Y}(\alpha)$ we deduce the following result.

Statement 4. *The value of $\varrho_{X,Y}(\alpha)$ is influenced by the observations of the samples*

$$\begin{aligned} & x_1, \dots, x_n; \\ & y_1, \dots, y_n; \\ & \textit{and } px_1 + (1-p)y_1, \dots, px_n + (1-p)y_n, \textit{ for each } p \textit{ from the grid } (17) \end{aligned} \tag{21}$$

However, within each one of the samples, the influence differs from observation to observation according to the following rules:
(a) the influence increases with the increase of the observation value;
(b) as α increases to 1, the distribution of the relative influences tends to be more concentrated on the highest order statistics.

Remark 5. The existing upper tail dependence measures (see Joe (1997)) typically define the tail cut-off (for example, $x + y = \theta$ for some large θ) and calculate the dependence between the occurrences of (X, Y) that lie above the cut-off. Our implicit dependence measure acts in a similar way. However, the essential difference is the absence of the sharp cut-off in our case: the weight with which an occurrences of (X, Y) is taken by our measure may be any number from $[0; 1]$. The Statement 4 specifies that the weight of (x, y) increases with the increase of either x or y, and that the actual weight depends not only on how large solely x or solely y is, but also on how larger $px + (1-p)y$ is for diverse values of p. Such weight attribution has the following economic interpretation: the dependence between the asset returns X and Y is determined by the risks of the portfolios composed out of the X and the Y assets with different weights. This interpretation reveals a novel – to the best of our knowledge – approach to the measurement of the dependence between two asset returns. We intend to study it thoroughly in the future. At the present, we can only state that: (a) this is an economically meaningful approach, since when one infers at the dependence between two assets, one usually aims at the optimization of the risk of a portfolio composed out of these assets; and (b) this approach provides meaningful results, as will be shown in Section 4.

3 Identification of the dependence change

3.1 The approach

Let the question (1) be posed for two assets. We suggest it may be solves with the help of our implicit dependence measure in the way described in **(1-2)** below. In this description, the random variables X and Y denote the returns of the assets for which the question has been posed, and the data base for calculating $\varrho_{X,Y}(\alpha)$ is the sample of observations of the returns in the question.

(1) Fix a finite sequence $\{\alpha_j\}$ such that $\alpha_j \nearrow 1$ as j increases (for example, the sequence we choose in the applications presented in Section 3.3 is 0.95, 0.96, 0.97, 0.98, 0.99, 0.995, 0.999), and get $\varrho_{X,Y}(\alpha_j)$ for each α_j from the sequence.
(2) If $\varrho_{X,Y}(\alpha_j)$ keeps the same sign and $|\varrho_{X,Y}(\alpha_j)|$ decreases (respectively, increases) as $\alpha_j \nearrow 1$, conclude then that the dependence between the assets' returns X and Y weakens (respectively, strengthens) as a market boom strengthens.

That **(1-2)** is a meaningful approach to the question (1) may be justified on the basis of the properties of our implicit dependence measure and of the economic interpretation of a market boom. This justification is as follows.

Although a market boom is a cumbersome event, we shall interpret it here simply as the market condition in which large returns of assets occur more frequently than their medium or small returns. Thus, if an asset return is distributed according to a distribution functions F_X, then conditioned on the ocurrence of a market boom, the distribution of the return is not F_X any more: the conditional distribution, when compared with F_X, attributes larger weights to larger return occurrences. Combining this conclusion with the Statement 4, we conclude that if X and Y mean the returns of two assets and if α is close to 1 then $\varrho_{X,Y}(\alpha)$ is determined by the conditional distribution of (X, Y) conditioned on a boom ocurrence, and, moreover, the larger the value of α the stronger the boom. Consequently, if $\{\alpha_j\}$ is an arbitrary sequence increasing to 1 then the sequence $\{\varrho_{X,Y}(\alpha_j)\}$ may indicate how the distribution of (X, Y) changes along with the strengthening of a boom. The principal question is then what property of a distribution is represented by $\varrho_{X,Y}(\alpha_j)$. The Statement 2 answers this question: $\varrho_{X,Y}(\alpha_j)$ expresses the dependence between X and Y in the boom of the "strength" α_j. Moreover, the same statement tells us that the strength of the measured dependence relates to the closeness of $|\varrho_{X,Y}(\alpha_j)|$ to 1. Altogether, the arguments of the present paragraph justify the approach rule **(1-2)**.

Remark 6. Since the value of $\varrho_{X,Y}(\alpha)$ is related to the value of a particular coefficient of linear correlation (as specified in Statement 2), then
(a) we would not be able to make any reliable conclusion when $|\varrho_{X,Y}(\alpha_j)|$ decreases or increases but $\varrho_{X,Y}(\alpha_j)$ alternates its sign as $\alpha_j \nearrow 1$;
(b) the convergence of $\varrho_{X,Y}(\alpha_j)$ to 1 indicates that X and Y tend to a complete dependence as a boom propagates, while the convergence of $\varrho_{X,Y}(\alpha_j)$ to 0 may occur because X and Y tend to be independent, but does necessarily implies this tendency.

3.2 Our modification of the approach and its cause

There is a problem that arises in the course of the execution of the approach suggested above. This problem admits different solutions. The one adapted by us will be presented here.

Let us first present the problem. For this, we recall to a reader that the construction of the implicit dependence measure uses values-at-risk of various random variables. Then, we recall that these values are obtained by the POT method (as has been explained in the Section 2.5). And finally, we recall that the execution of the POT method requires establishing the value of the threshold u (as explained in the text right above eq. (19)). We now note that for the POT method to give a reliable result, it is extremely important to find the "optimal" value of u. The point here is that taking u too small worsens the precision of estimation of the right tail of a distribution function (because $\mathrm{P}[X - u \leq x \,|\, X > u]$ and $G_\xi(x/\beta(u))$ from formula (18) are significantly different for $x > u$ when u is small), while, on the other hand, taking u too large diminishes the number of the sample points used for calculating the estimates $\hat{\xi}$ and $\hat{\beta}(u)$ (the points used are those that lie to the right of u) and thus, also worsens the precision of estimation of the right tail of a distribution function. Establishing the optimal value of u is the problem that we intend to discuss in the present section, together with the approach suggested by us for its solution.

The problem exposed above does not have an analytic solution, but may be solved in principal by one of several visual methods that have been developed within the Extreme Value Theory. However, an employment of one of these methods for the execution of the approach **(1-2)** has the following disadvantage: it creates a visual analysis step in the middle of other steps that may be executed by a computer. To eliminate this disadvantage and at the same time, to avoid the problem of finding the most optimal u, we modify **(1-2)** to **(1′-2′)** described below:

(1′) Fix a finite sequence $\{\alpha_j\}$ such that $\alpha_j \nearrow 1$ as j increases. Fix an integer N (in our applications, $N = 50$) and for each α_j, realize N values of $\varrho_{X,Y}(\alpha_j)$, choosing, for each realization, u uniformly from the interval delimited by the 0.85 and the 0.95 quantiles of each sample.

(2′) For each α_j, represent the obtained values of $\varrho_{X,Y}(\alpha_j)$ by a box-plot, and analyze then whether the sequence of box-plots increases, decreases or is stable as $\alpha_j \nearrow 1$; the tendency identified for the box-plot sequence is then declared to be the tendency of $\varrho_{X,Y}(\alpha_j)$ as $\alpha_j \nearrow 1$.

Let us give details in respect to the step **(1′)**. It uses the empirical evidence that the optimal value of u should be close to the 0.9-quantile of the sample (i.e., approximately 10% of the sample points should lie to the right of the optimal u). To calculate N values of $\varrho_{X,Y}(\alpha)$ we repeat N times the following procedure: we calculate $\mathrm{VaR}_\alpha(X)$ by the POT method with the value of u that we choose at random and according to the uniform distribution between the 0.85 and the 0.95 quantiles of the sample of X; then, in the similar way and independently, we calculate $\mathrm{VaR}_\alpha(Y)$ and $\mathrm{VaR}_\alpha(Z_p)$ from the respective samples of Y and of Z_p; finally, we plug the obtained VaR's in (16) and get the corresponding value of $\varrho_{X,Y}(\alpha)$.

Let us finally specify how we decide in respect to the "tendency" of box-plots, when we execute the step (2′). Our decision rule is as follows. If the sequences of the maxima, of the minima, and of the first, of the second and of the third quantiles all grow with the growth of α_j, then we identify the increasing tendency in the box-plot sequence; similarly, if all they decrease then we identify the decreasing tendency. When neither of these tendencies can be identified and when each of the sequences is not very much irregular, we identify stability. Otherwise, we say that our approach cannot provide a trustworthy conclusion. (A notable fact however, allowed us to identify a tendency or a stability in each real case to which in has been applied; see Section 3.3 below.) For example, we would conclude that $\varrho_{X,Y}(\alpha)$ definitely decreases with $\alpha \to 1$ from the analysis of the box-plot sequence showed in the "IBOVESPA-NASDAQ (short-short)" panel of Figure 2, while we would suggest that $\varrho_{X,Y}(\alpha)$ probably decreases with $\alpha \to 1$ from the analysis of the box-plot sequence showed in the "IBOVESPA-NASDAQ (long-long)" panel of Figure 2; in the latter case, if we were asked for a more definite conclusion, we then would say that the box-plot sequence identifies certainly that $\varrho_{X,Y}(\alpha)$ does not increase as $\alpha \to 1$.

3.3 Results for twenty real cases

We took the following data sets (the data source was Bloomberg L.P.):

NASDAQ (US equity market index) 7339 observations of daily returns from the period Feb. 08, 1971 - Feb. 16, 2001;

IBOVESPA (Brazilian equity market index) 7790 observations of daily returns from the period Jan. 02, 1968 - Feb. 16, 2001;

MERVAL (Argentinean equity market index) 2663 observations of daily returns from the period Jan. 09, 1990 - Feb. 16, 2001;

EI (a representative security from Brazilian sovereign debt) 1597 observations of daily returns from the period Nov. 08, 1994 - Feb. 16, 2001;

FRB (a representative security from Argentinean sovereign debt) 1382 observations of daily return from the period Jan. 27, 1995 - Feb. 16, 2001.

and by matching the sets in pairs over the maximal common period, we constructed 10 samples of simultaneous daily return observations. The scatterplots of all ten samples are presented in Figure 1.

For each sample, we posed the question (1) twice: first, regarding the change of dependence between the corresponding assets in a market boom, and second, regarding the change of dependence in a market crash.[14] To answer the questions about booms, we applied the approach (1′-2′) described in Section 3.2. To answer the questions about crashes, we applied the same approach but to the sample $(-x_1, -y_1), \ldots, (-x_n, -y_n)$ in the place of $(x_1, y_1), \ldots, (x_n, y_n)$. The value of the approach's parameter N was 50 in each application. We recall that the approach's conclusion stems from analysis of box-plots of values of $\varrho_{X,Y}(\alpha)$. These box-plots are showed in Figure 2. In Table 1, we present the answers that our approach gave to the 20 studied cases.

[14]It is intuitively clear that the dependences in booms and in crashes should be considered separately. Longin and Solnik (2001) provide pairs of indices for which these two dependences are indeed different.

4 For the conclusion: Testing the efficiency of our constructions

Here we shall comment the results obtained in Section 3.3 and presented in Table 1 and argue that they indicate the efficiency of the approach that we have constructed in Section 3. Since this approach is based on the properties of our implicit dependence measure, then the obtained results confirm empirically the properties of this measure that we have derived in Section 2.6.

(a) It is reasonable that the NASDAQ index shows the decrease of dependence both in market booms and crashes, when taken in a pair with any one of the Latin America indices. An explanation is that crises and booms in the Latin America do not necessarily propagate to the rest of the world, causing essential losses or gains in the principal markets.

It is interesting to note that the box-plots from the "IBOVESPA-NASDAQ (short-short)" panel of Figure 2 indicates that $\varrho_{X,Y}(\alpha) \downarrow 0$, which could be interpreted (with a certain proviso explained in Remark 6(b)) as that IBOVESPA and NASDAQ are asymptotically independent in a boom.

(b) We recall to a reader that shortly before the 2001 Argentinean crisis, certain analysts affirmed that the Brazil-Argentina economic and financial links were not strong enough and predicted thus that this crisis would not affect significantly the Brazilian economy. This affirmation may be accepted as true since in fact the 2001 Argentinean crisis did not ruined the Brazilian economy. We now note that this affirmation is in agreement with the lower tail relation between EI and FRB that is identified by our approach (see Table 1). It is important to stress that the data we have used to analyze this relation do not include the referred crisis' period, so that it is not the case that our approach predicts an event that is already present in the sample used for the prediction.

(c) It is reasonable that Brazilian equity index is related to the Brazilian sovereign debt in the way such that in a market crash their dependence increases. The same is true for Argentina.

(d) Our approach seems to be applicable for a wide class of dependence structures. We ground this conclusion on the fact that the approach worked well for all ten data sets that we have studied, although the scatter-plots of these data sets form quite different patterns. We note that we conclude that it "worked well" since its produced a smooth sequence of box-plots in each one of the 20 cases studied.

(e) The use of box-plots not only enhance our confidence in determining the tendency of $\varrho_{X,Y}(\alpha)$ as $\alpha \nearrow 1$, but can also support the assumption made in Section 2.5 that the distribution functions F_X and of F_Y belong to the class for which the Pickands' result holds. Namely, if this assumption is true, then the approximating function (defined in (19)) should not vary very much with the variation of u, and consequently, neither should $\varrho_{X,Y}(\alpha)$. In accordance to our constructions this would result in the narrowness of the corresponding box-plot. Thus, the narrowness of box-plots, which one observes in Figure 2, may be viewed as a confirmation of the cited assumption.

When analysing panels in Figure 2 from the point of view of the narrowness of the box-plots, one easily sees that in almost each one of the 20 cases presented there, the box-plot of the values of $\varrho_{X,Y}(.999)$ is much wider than the box-plot of any $\varrho_{X,Y}(\alpha)$ with $\alpha < .999$. Moreover, in some cases, $\varrho_{X,Y}(.999)$ has the sign different from that of all other $\varrho_{X,Y}(\alpha)$, and, in some other cases, $\varrho_{X,Y}(.999) > 1$, that means that the true minimizer of (16) is > 1 and that contradicts the definition that obliges the minimizer to pertain to $[-1;1]$. Nevertheless, we do not take these facts as a disapproval of our arguments. We attribute them to the badness-of-fit of the POT approximation $\hat{F}$ (given in (19)) to the remote right tail of the true distribution function F (see the discussion on page 364 of Embrechts et al. (1997) in respect); this badness-of-fit is usually caused or aggravated by the discrepancy between u and the optimal threshold value. This fact together with the fact that our 50 estimates correspond to different values of u explain the irregularities of the box-plot of estimates of $\varrho_{X,Y}(.999)$.

(f) Our approach can reveal properties that are not totally obvious or expected from some general considerations. For example, it shows that the Brazilian and Argentinean equity markets weaken their dependence both in market booms and in market crashes. This behavior is different from that observed for the pairs composed of the index of the US equity market and of European countries equity market indices: those strengthen their dependence in a crisis but weaken it in a boom, as has been shown by Longin and Solnik (2001).

References

Boyer, B.H., Gibson, M.S., Loretan, M. (1997) Pitfalls in tests for change in correlations. *International Finance Discussion Paper Nr. 597* of the Board of Governors of the Federal Reserve System.

Embrechts, P., Kluppelberg, C. and Mikosch, T. (1997) *Modeling Extremal Events for Insurance and Finance.* Springer, Berlin.

Cherubini, U., Luciano, E., Vecchiato, W., (2004) *Copula Methods in Finance.* Wiley Finance: Chichester.

Embrechts, P., McNeil, A.J., Straumann, D.(2002) Correlation and dependency in risk management: properties and pitfalls. In *Risk Management: Value at Risk and Beyond* Edited by M. A. H. Dempster, Cambridge University Press.

Genton, M. G., editor (2004) *Skew-elliptical distributions and their applications. A journey beyond normality.* Chapman&Hall/CRC.

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

Jorion, P. (1997) *Value at Risk.* McGraw-Hill, Chicago.

Këllegi, E., Gilli, M. (2000) Extreme value theory for tail-related risk measures. Preprint, University of Geneva. Available also at `http://citeseer.ist.psu.edu/kellezi00extreme.html`.

Kolev, N., Mendes, B. V. de M., Anjos, U. (2005) Copulas: a Review and Recent Developments. To appear in *Stochastic Models.*

Leadbetter, M.R., Lindgren, G. and Rootzén, H. (1983) *Extremes and Related Properties of Random Sequences and Processes.* Springer-Verlag, Berlin.

Longin, F.M. (2000) From value at risk to stress testing: the extreme value approach, *Journal of Banking & Finance* 24, 1097-1130.

Longin, F. M., (2005) The choice of the distribution of asset returns: How extreme value theory can help? *Journal of Banking & Finance* 29, 1017-1035.

Longin, F. M., Solnik, B. (2001) Extreme correlation in international equity markets. *The Journal of Finance*, LVI, Nr. 2, 649-676.

Panzieri Fl., A. (2001) *Teoria de valores extremos aplicada a finanas: dois ensaios.* Ph.D. thesis. EAESP/FGV, São Paulo, Brazil.

Pickands, J. (1975) Statistical inference using extreme order statistics. *Ann. Statist.*, **3**, 119-131.

Reiss, R.-D., Thomas, M. (1997) *Statistical analysis of extreme values.* Birkhäuser Verlag, Basel - Boston - Berlin.

Smith, R. L. (2000) Measuring Risk with Extreme Value Theory. In: P. Embrechts (editor), *Extremes and Integrated Risk Management.* Risk Books, London.

Tiago de Oliveira, J. (1973) *Statistical Extremes - a survey.* Center of Applied Mathematics, Faculty of Sciences, Lisbon.

Tsay, R. S. (2002) *Analysis of Financial Time Series.* John Wiley & Sons. Inc.

The dependence between the returns as a market crash deepens	Pair of assets	The dependence between the returns as a market boom strengthens
Increases	IBOVESPA-EI	Increases
Increases	MERVAL-EI	Stable (around 0.4)
Increases	MERVAL-FRB	Decreases
Increases	IBOVESPA-FRB	Decreases
Stable (around 0.26)	IBOVESPA-MERVAL	Decreases
Decreases	IBOVESPA-NASDAQ	Decreases (to 0)
Decreases	EI-FRB	Decreases
Decreases	NASDAQ-MERVAL	Decreases
Decreases	NASDAQ-FRB	Decreases (to 0)
Decreases (to 0)	NASDAQ-EI	Decreases

Table 1.

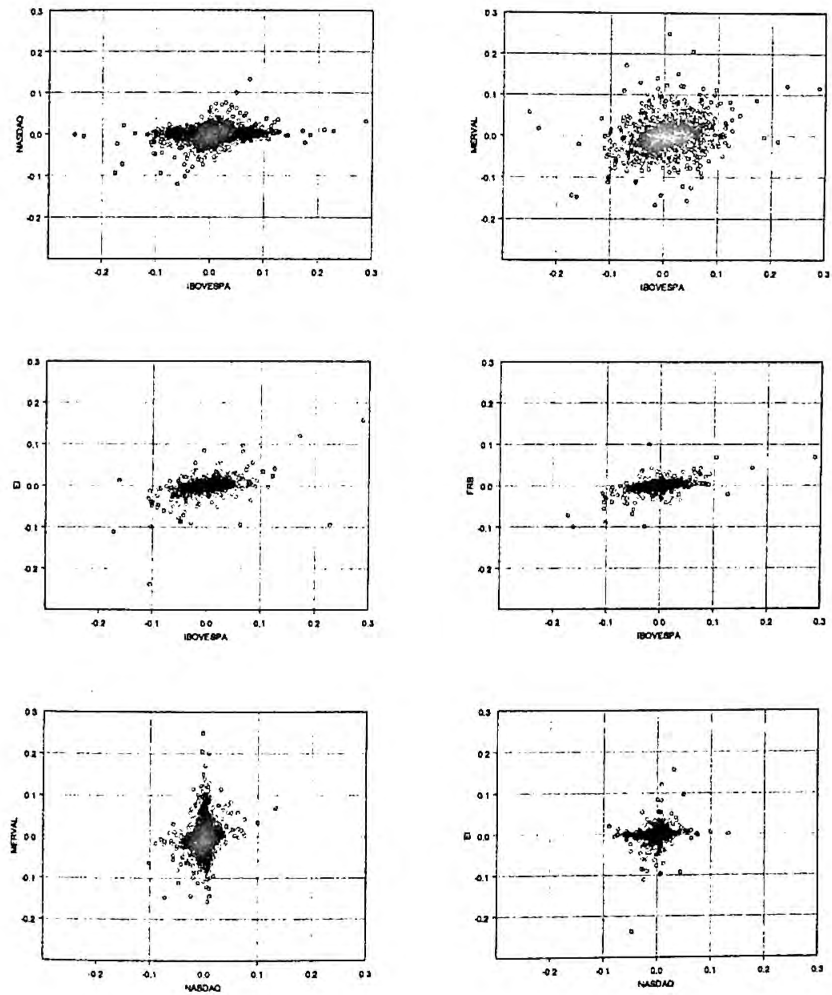
NASDAQ
IBOVESPA
MERVAL
IBOVESPA
EI
IBOVESPA
FRB
IBOVESPA
MERVAL
NASDAQ
EI
NASDAQ

Figure 1.

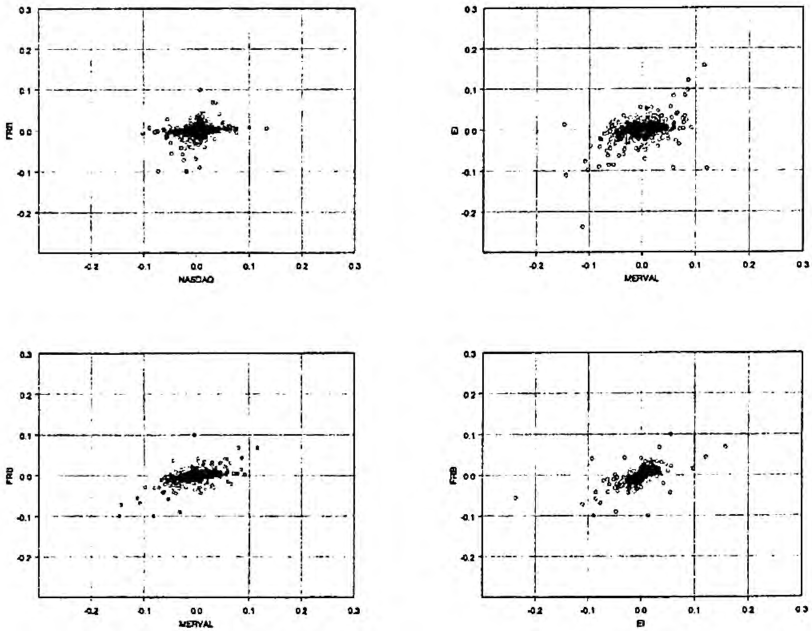

Figure 1, continuation.

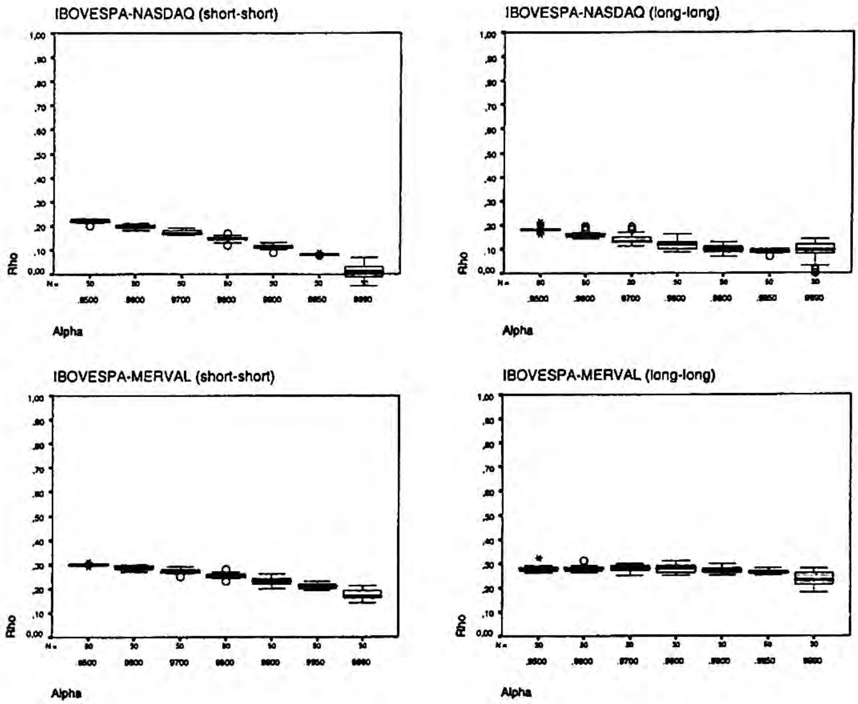
IBOVESPA-NASDAQ (short-short)
IBOVESPA-NASDAQ (long-long)
IBOVESPA-MERVAL (short-short)
IBOVESPA-MERVAL (long-long)
Rho
Alpha

Figure 2.

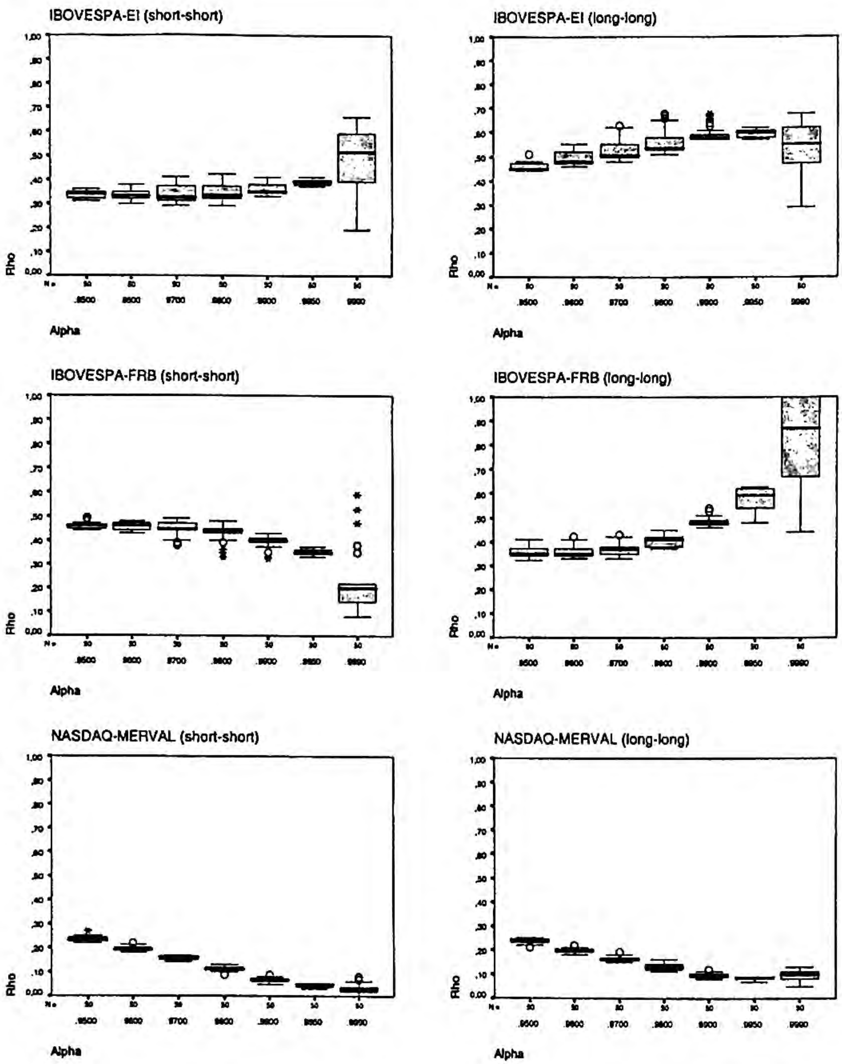

Figure 2, continuation.

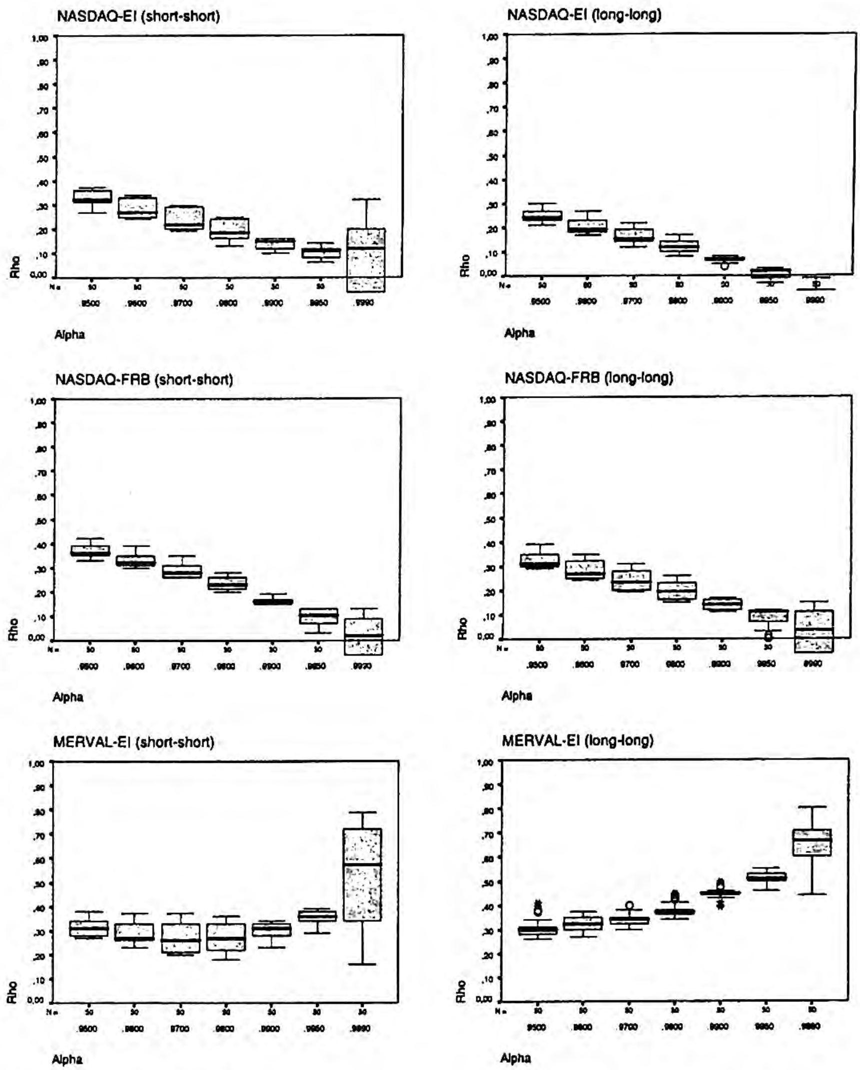

Figure 2, continuation.

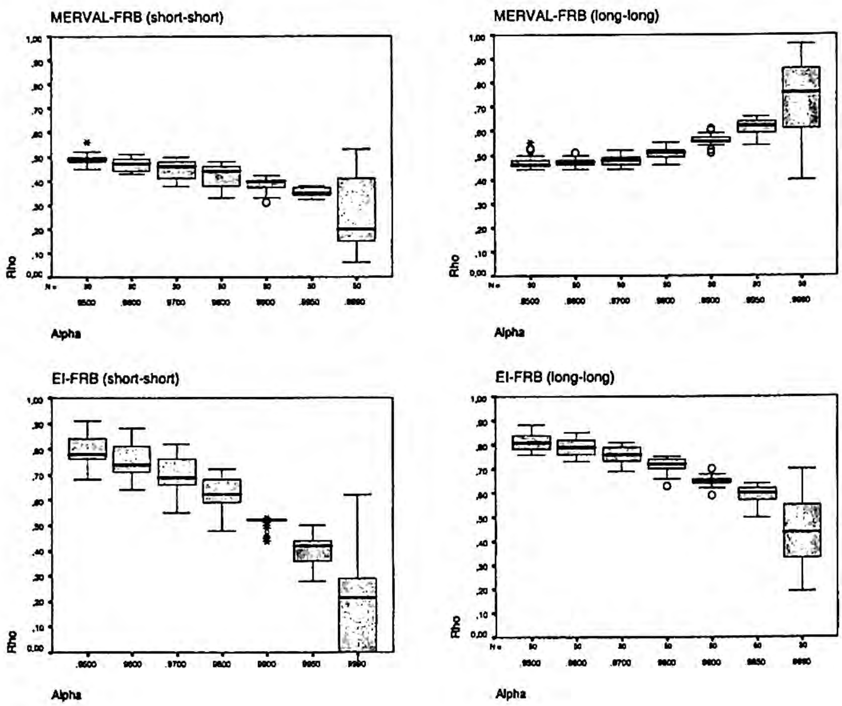

Figure 2, continuation.

Captions for the figures and tables

Table 1: The dependence change between pairs of asset returns as a market crash deepens and as a market boom strengthens.

Here we present our conclusions in respect to the change in dependence between 10 pairs of assets in market booms and crashes. The assets involved are described in Section 3.3. The observations of pairs of their daily returns are presented in Figure 1. The conclusions were derived from these observations via the approach presented in Sections 3.1 and 3.2.

Figure 1: Scatter-plots of pairs composed of observations of daily returns of NASDAQ, IBOVESPA, MERVAL, EI and FRB.

The description of these equity market indices and sovereign debt securities is given in Section 3.3.

Figure 2: The box-plot sequences that were used to identify the tendency of the implicit dependence measure $\varrho_{X,Y}(\alpha)$ as $\alpha \searrow 0$ and $\alpha \nearrow 1$.

The "long-long" panels (resp. "short-short" panels) determine the tendency of $\varrho_{X,Y}(\alpha)$ as $\alpha \searrow 0$ (resp., $\alpha \nearrow 1$), which, in turn, determines the change of the dependence in a market crash (resp., boom).

ÚLTIMOS RELATÓRIOS TÉCNICOS PUBLICADOS

2006-01 - SILVA, M.F. DA, FERRARI, S.L.P., CRIBARI-NETO, F. Improved likelihood inference for the shape parameter in Weibull regression. 2006. 24p. (RT-MAE 2006-01)

2006-02 - LACHOS, V.H., BOLFARINE, H., ARELLANO-VALLE, R. B., MONTENEGRO, L. C. Inference in multivariate skew-normal regression models. 2006. 21p. (RT-MAE 2006-02)

2006-03 - BUENO, V. DA C. Redundancy allocation for a coherent system of dependent components. 2006. 16p. (RT-MAE 2006-03)

2006-04 - ESPINHEIRA, P.L., FERRARI, S. L. P., CRIBARI-NETO, F. On Beta regression residuals. 2006. 17p. (RT-MAE 2006-04)

2006-05 - BAZÁN, J. L., BOLFARINE, H., BRANCO, M. D. A generalized skew probit class link for binary regression. 2006. 22p. (RT-MAE 2006-05)

2006-06 - MIRANDA, J. C. S. DE; MORETTIN, P. A. On the estimation of the intensity of point processes via wavelets. 2006. 45p. (RT-MAE 2006-06)

2006-07 - DUARTE, D.; GALVES, J.; GARCIA, N. L. Markov approximation and consistent estimation of unbounded probabilistic suffix trees. 2006. 12p. (RT-MAE 2006-06)

2006-08 - RIFO, L. R.; WECHSLER, S. Regularly varying densities for St. Petersburg envelopes. 2006. 8p. (RT-MAE 2006-08)

2006-09 - FERRARI, S. L. P.; CYSNEIROS, A. H. M. A. Skovgaad's adjustment to likelihood ratio tests in exponential family nonlinear models. 2006. 15p. (RT-MAE 2006-09)

2006-10 - POLPO, A., PEREIRA, C. A. B., Parallel and series systems: rates, distributions and sub-distributions. 2006. 13p. (RT-MAE 2006-10)

2006-11 - LACHOS, V. H., BOLFARINE, H., GENTON, M.G. Bayesian inference for measurement error models with flexible skew-symmetric distributions. 2006. 29p. (RT-MAE 2006-11)

2006-12 - BUENO, VANDERLEI DA COSTA, A Martingale Estimator for a Structural Importance Measure of a Coherent System's Component Through its Critical Level.

2006-13 - BUENO,VANDERLEI DA COSTA, Estimating system´s component reliability importance under dependence conditions. 2006. 09p. (RT-MAE 2006-13)

The complete list of "Relatórios do Departamento de Estatística", IME-USP, will be sent upon request.

Departamento de Estatística
IME-USP
Caixa Postal 66.281
05311-970 - São Paulo, Brasil