
NOTES D'ÉTUDES

ET DE RECHERCHE

**THE TAIL BEHAVIOR OF STOCK RETURNS:
EMERGING VERSUS MATURE MARKETS**

Eric Jondeau and Michael Rockinger

June 1999

NER # 66



**THE TAIL BEHAVIOR OF STOCK RETURNS:
EMERGING VERSUS MATURE MARKETS**

Eric Jondeau and Michael Rockinger

June 1999

NER # 66

Les Notes d'Études et de Recherche reflètent les idées personnelles de leurs auteurs et n'expriment pas nécessairement la position de la Banque de France.

Ce document est disponible sur le site Web de la Banque de France « www.banque-france.fr ».

The Tail Behavior of Stock Returns: Emerging versus Mature Markets

Eric Jondeau*

Michael Rockinger[†]

June 1999

Abstract

For Central Banks, institutional, and individual investors it is crucial to understand the frequency and importance of drops or sudden rises in financial markets. Extreme value theory (evt) is an interesting tool providing answers to questions such as: -with what frequency do we find variations of returns beyond a given threshold? -over a given period, what type of extreme variation can be expected? -with what type of unconditional distribution of returns are the tails of returns compatible? -in a cross country setting of emerging and mature financial markets do extreme variations behave in a similar manner? -can we learn about the evolution of returns of presently developing economies from the early returns of presently mature markets? -do countries behave similarly in terms of up or down crashes for a given level of development?

In the following paper we start with a review of theoretical elements of evt. In the empirical section of this study we consider five mature markets, nine Asian, six Eastern European, and seven Latin American emerging markets. The tail-behavior of returns is found to be compatible with the existence of up to the third moment but not beyond. The estimation of the tail distribution as a Generalized Pareto Distribution shows that great care has to be taken for emerging markets where little data is available and returns' distribution is subject to violate the iid assumption. Using a subsample of countries we demonstrate the limitations of evt. We also show that little can be learned from 19th century US data about presently emerging markets' tail behavior.

Keywords: Extreme Value Theory, Generalized Pareto Distribution, Stock Market Returns.

JEL classification: C13, C22, G15, O16.

*Banque de France, Centre de recherche.

[†]HEC - School of Management; Banque de France, Centre de recherche; CEPR.

Corresponding author: Michael Rockinger, HEC-School of Management, Department of Finance, 78351 Jouy-en-Josas, France, rockinger@hec.fr. Our data has been extracted from Datastream. We would like to thank William Schwert for kindly providing his global US index. Rockinger acknowledges help from the TMR grant on *Financial market's efficiency* as well as from the Fondation HEC. We also wish to thank Jean Cordier, Pierre Sicsic, as well as Sanvi Avouyi-Dovi for their comments. Françoise Dumolard helped us by finding explanations for crashes. The usual disclaimer applies. This paper does not necessarily reflect the view of the Banque de France.

1 Introduction

Few researchers in empirical finance would argue with the affirmation that financial markets are subject to extreme variations generally due to stock market crashes. Those crashes can be brought about by political reasons or economic ones. For a Central Bank, the understanding of the expected frequency of a crash or of its magnitude may be of importance in the management of its reserves. For portfolio management the observation that returns may take extreme values, incompatible with the assumption of normality, is also of importance since it implies a break-down of traditional mean-variance analysis.¹ Furthermore, the possibility of a sudden deterioration of financial market's conditions has implications for Value-at-Risk (VaR) analysis. Similarly, the increased risk of default during a crash implies for the regulator, who has to set margin requirements in a futures market, that his settings may not be sufficient if he does not explicitly take into account the possibility of large sudden movements in the market. A thorough understanding of extreme value theory (evt) seems to be, therefore, also of great importance for empirical finance.

The literature which finds that a simple description of stock returns' distribution as a normal one is insufficient started early with Mandelbrot (1963) or Fama (1963).² Those authors suggest the use of stable distributions as an alternative to the normal one in an effort to take into account the observation that stock returns have excess skewness and kurtosis. In an alternative approach, aiming at describing patterns of variable volatility, Engle (1982) developed the ARCH model. This type of model is also able to generate skewed and leptokurtic returns.³ Whereas both types of approaches model the entire distribution of returns it is possible to focus only on the distribution of the tails of returns and, thus, to neglect the more central part. In the statistics literature, DuMouchel (1983) noticed the importance to distinguish the tails from the more central part of returns' distribution. In the finance literature, the investigation how tails of returns behave goes back to Hols and de Vries (1991), Jansen and de Vries (1991). Koedijk, Schafgans, and de Vries (1990) investigated the tail behavior of foreign exchange data. Recent research is by Danielsson and de Vries (1997), or Dacorogna, Müller, Pictet, and de Vries (1995) for exchange rates. For stock returns recent work has been made by Loretan and Phillips (1994) and Longin (1996). There is also some cross country evidence that the tail behavior of returns is leptokurtic. Longin (1996) considers a database of daily US stock between 1885 and 1990 and confirms that the tail distribution is of the Fréchet type, hence fat-tailed. Lux (1998) finds a similar behavior for German stocks sampled tick-by-tick.

The goal of this paper is to recall some fundamentals of extreme value theory and to enrich the academic literature by studying the tail behavior of stock returns in a cross-country setting. To do so we consider a database of daily stock returns for five mature markets, nine Asian markets, six Eastern European as well as seven Latin

¹The study how extreme values can be incorporated into portfolio management is left for further research.

²Non-normality is widely documented. We just mention some works with the risk of omitting many others: Blattberg and Gonedes (1974), Fama and Roll (1971), Fielitz (1976), Fielitz and Rozell (1983), Simkowitz and Beedles (1980), Kon (1984), Tucker (1992), and Mandelbrot (1997).

³In this type of model, innovations are found to remain excessively leptokurtic. This result is well known and has led to modeling innovations either with a Student-t or a Generalized Error distribution. For a study relating GARCH with extreme value theory see Stărică and Pictet (1999). Fat-tailedness has been further modeled by Baillie, Bollerslev, and Mikkelsen (1996) with the FI-GARCH model. To the authors' knowledge the tail-behavior of this model is not known presently.

American ones.⁴ A cross-country investigation of the tail-behavior of returns has not been done to our knowledge. There exists, however, a strand of literature that investigates the stock market behavior of emerging markets. Claessens, Dasgupta, and Glen (1995) investigate the existence of return anomalies and predictability for a set of 20 countries represented in the International Finance Corporation's emerging markets data base. Bekaert and Harvey (1995, 1997), using the same data, study volatility determinants for emerging markets. They also ask how well those markets are integrated with developed ones. See also Rockinger and Urga (1998) who focus on Eastern European markets. Shields (1997) compares asymmetries in volatility response to news between mature and emerging markets.

The structure of this paper is as follows. In section 2 we present theoretical elements of extreme value theory. We recall that under the assumption of independence and identical distribution of returns, asymptotically the distribution of maxima must be of one of three types which can be nested within a Generalized Extreme Value (gev) distribution. Similarly, under the same assumptions we recall that the distribution of tails must be of the Generalized Pareto Distribution (gpd). In section 3 we provide applications of evt that should be useful for the practitioner. In section 4 we implement evt on our database. In section 5 we present the results of our empirical investigation. We report various descriptive statistics and the Hill estimates. We show that for most indices under investigation up to the second moment exists. For many indices even the third one seems to exist. Next, we report the goodness of the fit of the gpd to the tails of returns. It is shown that the fit leaves space for improvements for some emerging markets. We also present the results of the maximum likelihood estimation of the generalized extreme value distribution. In section 6 we compare the tail behavior of returns for the early days of a global US index with those of presently emerging markets. We find that the behavior then has little to do with the one of currently emerging markets and, thus, that it is very difficult to make any predictions how currently emerging markets' returns tails will evolve in the future. In section 7 we conclude.

2 Extreme value theory

In this section we consider theoretical issues concerning extreme realizations. A very thorough description of the theory, at textbook level, can be found in Leadbetter, Lindgren, and Rootzén (1983) or Embrechts, Klüppelberg, and Mikosch (1997).

There are two different, yet related, approaches to modeling extreme values. A first approach studies the law followed by the maximum or minimum of returns over given time horizons. An alternative approach considers the entire tail of a set of realizations and estimates its associated distribution. We will start our theoretical description with the former approach.

2.1 Distribution of extremes

We will be concerned in this section with the distribution of the maximum or the minimum return of a stock. We will recall that such a distribution of extremes can be of three different types. Before reaching this result we need some notations. Let

⁴Datastream provided us with this data.

X_1, \dots, X_T be a sequence of random variables corresponding to stock returns.⁵ We notice that

$$-\min(-X_1, \dots, -X_T) = \max(X_1, \dots, X_T) \equiv M_T$$

which shows that without loss of generality it is enough to develop a theory for the upper tail of the distribution of returns. We next notice that if the X_t are independent and identically distributed, (iid), then, if $F_X(\cdot)$ is the cumulative distribution function (cdf) of any X_t it follows that

$$\begin{aligned} \Pr[M_T < x] &= \Pr[\max(X_1, \dots, X_T) < x] \\ &= \Pr[X_1 < x, \dots, X_T < x] \\ &= [F_X(x)]^T. \end{aligned}$$

This indicates that under the iid assumption, the law of the maximum for a finite sample can be easily obtained if $F_X(\cdot)$ is known. What if T becomes large? We notice that for a given x

$$\lim_{T \rightarrow \infty} [F_X(x)]^T = \begin{cases} 1 & \text{if } F_X(x) = 1, \\ 0 & \text{else.} \end{cases}$$

and, thus, we obtain a degenerate law. This raises the question if a scaled version of M_T converges to a finite distribution. After all, a sum of iid random variables also degenerates as the sample size increases whereas a scaled version converges, thanks to the central limit theorem, to a well defined law. For the problem at hand we have the Fisher-Tippett theorem which characterizes the limit law for maxima. It was Gnedenko (1943) who provided the first formal proof.

Theorem 1 *Let X_t be a sequence of iid random variables. If there exist norming constants $\mu_T \in \mathcal{R}$, $\psi_T > 0$, and some non-degenerate cdf H such that*

$$\frac{M_T - \mu_T}{\psi_T} \Rightarrow H,$$

where \Rightarrow designs convergence in distribution, then H belongs to one of the following three cdfs:

$$\begin{aligned} \text{Gumbel} &: \exp(-\exp(-x)), \text{ for } x \in \mathcal{R}, \\ \text{Weibull} &: \begin{cases} \exp(-(-x)^\alpha), & \text{for } x \leq 0, \\ 1, & \text{for } x > 0, \end{cases} \quad \alpha > 0, \\ \text{Fréchet} &: \begin{cases} 0, & \text{for } x \leq 0, \\ \exp(-x^{-\alpha}), & \text{for } x > 0, \end{cases} \quad \alpha > 0. \end{aligned}$$

The three distributions are called *standard extreme value distributions*. In Figure 1 we represent those various distributions.⁶ It is important to notice that the distributions cannot be flipped symmetrically around a vertical axis. This implies intuitively that to get a distribution such as the Weibull, the support of the distribution underlying the random variables must be finite. In other words, the Weibull generates thin tails. On the other hand, empirical evidence suggests that returns are

⁵Continuously compounded returns are defined by $X_t = 100 \cdot \ln(S_t/S_{t-1})$ where S_t is the closing value of a stock price at time t .

⁶In the graph we have set $\alpha=1$.

heavy tailed and, therefore, either the Gumbel or the Fréchet distribution are likely to describe the behavior of extremes of stock market returns.⁷

Sometimes it is useful to nest the three distributions. This has been originally achieved by Jenkinson and von Mises with the *generalized extreme value* distribution (gev). They define

$$H_\xi(x) = \begin{cases} \exp\left(-(1 + \xi x)^{-1/\xi}\right) & \text{if } \xi \neq 0, \\ \exp(-\exp(-x)) & \text{if } \xi = 0, \end{cases} \quad (1)$$

where $1 + \xi x > 0$. We then notice that the standard extreme value distributions can be recovered with

$$\begin{aligned} \xi = \alpha^{-1} > 0 & \quad \text{for the Fréchet distribution,} \\ \xi = 0 & \quad \text{for the Gumbel distribution,} \\ \xi = -\alpha^{-1} > 0 & \quad \text{for the Weibull distribution.} \end{aligned}$$

The parameter ξ is called the *tail index* and $1/\xi$ is called the *shape index*.⁸ Our econometric problem is to decide which is the correct distribution of extremes of returns for the data at hand and to estimate the norming constants μ_T , ψ_T , and ξ .

Having obtained the limit distribution, one can reverse the reasoning and ask what type of distribution of the cdf $F_X(\cdot)$ underlying the X_t will yield convergence to a given limit distribution. It can be shown that the normal, and lognormal laws yield maxima converging to the Gumbel. Cauchy, Pareto, or t-distributed returns will yield the Fréchet distribution. A uniform distribution would yield the Weibull.

In the empirical-finance literature researchers have tried to capture fat-tailedness through various means. Kon (1984) modeled returns as a mixture of normals. It can be shown that this mixture distribution yields a Gumbel type distribution. For ARCH processes Jansen and de Vries (1991) have derived the Fréchet distribution as the limit distribution.

To estimate the various parameters, a typical approach is to take the maxima of non-overlapping subsamples and to derive descriptive statistics as well as maximum likelihood estimates for those maxima. This approach implies, however, that the sample gets strongly reduced since subsamples have to be taken. In this case the precision of estimates reduces. An alternative approach, considered in the next section, is to consider thresholds and to focus on those realizations exceeding a given threshold. The task is then to fit a distribution to the tail of the distribution. Intuitively a link between maxima and tails of the underlying distribution must exist. This intuition is also developed in the following section. McNeil and Saladin (1997) provide an application of this method.

2.2 The tail distribution

Definition 2 Let u be a fixed real number, the threshold, in the support of X_t . The function

$$F_u(x) = \Pr[X_t - u \leq x | X_t > u], \quad x \geq 0$$

⁷Returns cannot be beyond -100%, which corresponds to bankruptcy. For this reason the left tail of the returns' support is bounded. However, as long as on average returns are far away from this boundary, fat-tails are a possibility for a given sample.

⁸In this work we follow the notations of Embrechts, Klüppelberg, and Mikosch (1997).

is called the *excess distribution function (edf)* of the random variables X_t over the threshold u . The function

$$e(u) = E[X_t | X_t > u] - u$$

is called the *mean-excess function (mef)*.

The excess distribution function measures the probability that the excess realization relative to the threshold is below a certain value, given that the realization is above the threshold. The mean-excess function averages those realizations that exceed u and considers the distance between the mean and u . We have the following mean-excess functions given a certain distribution for the tail of a distribution:

$$\begin{array}{ll} \text{Pareto} & \frac{k+u}{\alpha-1}, \alpha > 0, \\ \text{Weibull} & \frac{u^{1-\tau}}{c^\tau}, \\ \text{Exponential} & \frac{1}{\lambda}. \end{array}$$

In Figure 2 we represent the graph of those theoretical mefs associated with several distributions for X_t .

It has been shown by Balkema and De Haan (1974) and Pickands (1975) that for a certain class of distributions there exists a positive scaling function $a(u)$ such that

$$\lim_{u \rightarrow +\infty} \Pr \left[\frac{X - u}{a(u)} \leq x | X > u \right] = \begin{cases} 1 - (1 + \xi x)^{-1/\xi} & \text{if } \xi \neq 0, \\ 1 - \exp(-x) & \text{if } \xi = 0. \end{cases}$$

This means that the scaled excess function $F_u(x)$ has a limit distribution which will be called the *Generalized Pareto distribution (gpd)* written as $G_{\xi, \psi}(x)$ and defined by

$$G_{\xi, \psi}(x) = \begin{cases} 1 - (1 + \xi x / \psi)^{-1/\xi} & \text{if } \xi \neq 0, \\ 1 - \exp(-x / \psi) & \text{if } \xi = 0, \end{cases} \quad \psi > 0 \quad (2)$$

where $x \geq 0$ if $\xi \geq 0$ and $0 \leq x \leq -\psi/\xi$ if $\xi < 0$. Clearly, the tail index ξ from the gev is the same as for the gpd. The density of the gpd can be easily derived as

$$g_{\xi, \psi}(x) = \begin{cases} \psi^{1/\xi} (\psi + \xi x)^{-\frac{1}{\xi}-1} & \xi \neq 0, \\ \psi^{-1} \exp(-x/\psi) & \xi = 0. \end{cases} \quad (3)$$

An illustration of this density can be found in Figure 3 for a set of parameters that may typically arise in Finance. We notice that an increase of ξ for a constant level of the scale factor ψ increases the tail while steepening the slope at the more central part of the density. On the other hand, an increase of ψ given ξ yields a flattening for the central part of the density accompanied by an increase of the tails. Those results indicate that if one chooses a high enough threshold, then the fit of a gpd to the tail realizations will also yield an estimate of the tail index.

3 Practical applications of extreme value theory

At this stage we have recalled various theoretical results how extreme returns or tails of distributions should behave. Once parameters have been obtained it is possible to address several useful issues such as how to estimate the average waiting time between extreme realizations or how to estimate the realization of high quantiles. It is also possible to determine the number of moments that exist for the return generating distribution and which are compatible with a given tail behavior.

3.1 Waiting time between extremes

Once the tail distribution has been characterized it is possible to compute the mean waiting time between specific extreme events. This result is useful for the practitioner who wishes to estimate the average time before a given extreme value occurs. Clearly, an extreme realization is defined as the exceeding of a return of a given level. Let X_t be iid random variables with cdf G and u a threshold. The fact that $X_t > u$ or not corresponds to a Bernoulli event with success probability $p = 1 - G(u)$. The probability that X_t exceeds the threshold at time k and not before is given by the geometric distribution $p(1-p)^{k-1}$, $k = 1, 2, \dots$. As a consequence the average waiting time before u is crossed is

$$E[\min(t \geq 1 : X_t > u)] = \sum_{k=1}^{+\infty} k p(1-p)^{k-1} = 1/p.$$

Once the parameters of a gpd are estimated it becomes easy to evaluate this quantity.

3.2 Estimation of high quantiles

In the previous section we showed how to compute the average waiting time before an extreme event occurs. For the practitioner it is of similar importance to know with what type of frequency a certain extreme event will occur. For instance, how often can one expect a drop of returns beyond a certain threshold. This type of question is for instance relevant for Value-at-Risk analysis. There are investors who need to know the probability that the stock market drops by so and so much on a given day. Traditionally, one would take historical data and count the frequency of stock market drops of the desired amount. There are several problems with this approach. First, there will be very few days with exceedance of the desired level, especially if the level is large, and one will get a poor estimate of the actual probability. For a reason of statistical stability one may be better off by modeling the entire tail. Second, it is possible that one is interested in variations that never occurred before. Again, a model which explicitly models the tail can help.

A similar, but slightly more complex problem is the measurement of the largest possible realization over a certain time horizon. For instance what type of extreme realization can be expected over a horizon of 5 or 50 years. If some daily data is sampled over 200 trading days, then, under the assumption that there are 200 trading days in a year, we are looking for the largest among 1'000 respectively 10'000 realizations. As a consequence, what we have to compute are very *high quantiles*.

Formally, define p as the probability of a very rare event. We try to find an estimate x_p such that $F(x_p) = p$. To solve this issue, Rootzén and Tajvidi (1996) or McNeil (1997) suggest the following approach. They first notice that for $x > u$ we have

$$F(x) = \Pr[X \leq x] = (1 - \Pr[X \leq u])F_u(x - u) + \Pr[X \leq u].$$

Hence, it is possible to estimate the probability of being below a certain threshold, $F(x)$, with

$$\hat{F}(x) = (1 - F_T(u))G_{\xi, u, \psi}(x) + F_T(u)$$

where $F_T(u)$ is the empirical distribution function. Furthermore, if we give ourselves the probability of occurrence of a rare event and seek the associated quantile, inversion

of this formula yields

$$x_p = u + G_{\xi, u, \psi}^{-1} \left(\frac{p - F_T(u)}{1 - F_T(u)} \right) = \begin{cases} u + \frac{\psi}{\xi} \left[\left(\frac{1-p}{1-F(u)} \right)^{-\xi} - 1 \right] & \text{if } \xi \neq 0, \\ u - \psi \ln \left(\frac{1-p}{1-F(u)} \right) & \text{else.} \end{cases}$$

This formula can be easily implemented once we have estimates of the various parameters.

3.3 Existence of moments

The fact that returns of stock markets are leptokurtic has led to their modeling with fat-tailed distributions. Some of those distributions such as Mandelbrot's (1963) stable law does not allow for a finite variance. Given the tail behavior of returns, this question can again be addressed with evt.⁹ Embrechts, Klüppelberg, and Mikosch (1997, p. 165) recall that if X follows a gpd then for all integers r such that $r < 1/\xi$ the r -th moment exists with

$$E[X^r] = \frac{\psi^r}{\xi^{r+1}} \frac{\Gamma(\xi^{-1} - r)}{\Gamma(\xi^{-1} + 1)} r!$$

4 Empirical techniques

We now address the question how the tail behavior of a given sample can be characterized and how its parameters can be estimated. A first step in evt is to use exploratory methods to analyze the data.

4.1 Histograms of extrema

This is the simplest and most straightforward explanatory method to get an idea of the type of tail behavior. One constructs m -histories (sometimes called *blocks* in evt), that is non-overlapping subsamples of length m , and one considers the maximum over each m -history. This gives a sample of maxima which can be represented by an histogram. Once an histogram is obtained, a comparison with the densities displayed in Figure 1 allows an educated *guess* of the type of extreme value distribution one is dealing with. Gumbel (1958) as well as Embrechts, Klüppelberg, and Mikosch (1997) insist on the importance of such a graphical analysis.

4.2 QQ-Plots (quantile plots)

We define the generalized inverse of the cumulative distribution function F as

$$F^{\leftarrow}(t) = \inf(x \in \mathcal{R} : F(x) \geq t), \quad 0 < t < 1.$$

This function is called the quantile function of the cdf F . $F^{\leftarrow}(t)$ defines the t -th quantile of F . For the case that F is a continuous function, F^{\leftarrow} is simply the inverse function. Let x_1, \dots, x_T be the maxima over m -histories and $x_{T,T} < \dots < x_{1,T}$ the

⁹See also Longin (1996, p. 399). He concludes that the US daily returns series allows up to the third moments but not beyond.

ordered realizations, then the plot $\left\{x_{t,T}, F^{\leftarrow}\left(\frac{t}{T+1}\right)\right\}$ is referred to as the *quantile plot*.

To test if x_1, \dots, x_T follow a certain distribution, such as a Gumbel, one takes the ordered sample and plots $x_{t,T}$ against $-\ln(-\ln(\frac{t}{T+1}))$, that is the inverse of the Gumbel cdf. If the data is truly generated by a Gumbel distribution then quantiles of the theoretical and the empirical distribution should match and a roughly linear QQ-plot is expected. To check for linearity it is customary to also trace the OLS regression line of the fit of the theoretical on the empirical quantiles. As already discussed by Gumbel (1958) if the QQ-plot is (with the Gumbel distribution as reference function) concave, then the limit distribution is a Fréchet one. If the QQ-plot is convex the limit distribution is Weibull.

This method can be pushed a bit further to yield parameter estimates which can be used in further estimations. To do so, one assumes that the sample of maxima follows a gev distribution (1). In this case we obtain after taking logs twice and after introducing a random noise ε_t that

$$\ln\left(-\ln\left(\frac{t}{T+1}\right)\right) = -\frac{1}{\xi} \ln\left(1 + \xi \frac{x_{t,T} - \mu}{\psi}\right) + \varepsilon_t. \quad (4)$$

This equation can now be fitted to the sorted sample with a simple NLLS fit. The obtained parameters can be used as starting values for a further maximum likelihood estimation.

4.3 Mean-excess function plots

Estimates of the mef $e(u)$, written as $\hat{e}(u)$, can be obtained easily from a realization x_1, \dots, x_T of X_1, \dots, X_T . Indeed, if $\mathcal{I}_{\{\text{condition}\}}$ is the indicator variable taking the value 1 if the condition is true and 0 otherwise,

$$\hat{e}(u) = \frac{1}{N_u} \sum_{t=1}^T (x_t - u) \mathcal{I}_{\{x > u\}}(x_t), \quad u > 0,$$

where N_u is the number of realizations exceeding u , that is $N_u = \sum_{t=1}^T \mathcal{I}_{\{x > u\}}(x_t)$.

The *mean-excess plot* consists in the graph

$$\{(x_{t,T}, \hat{e}(x_{t,T})) : t = 1, \dots, T\}.$$

The mean-excess plot allows distinction between distributions with light- or heavy-tailed distribution. Furthermore, a simple OLS fit of a straight line gives the parameters underlying the distribution. We recall that some characteristic mean-excess functions are displayed in Figure 3.

4.4 Estimation of the tail index

As we have seen earlier the asymptotic behavior of extreme values depends on the tail index ξ . Depending on the values taken by this index we will end up with either a Fréchet, a Gumbel, or a Weibull distribution. It is therefore of great importance to be able to quickly characterize the tail index. In this section we present two simple ways to estimate the index before turning in the next section to the full estimation of all parameters with the maximum likelihood method.

4.4.1 Pickands tail index estimation

This method can be used for all $\xi \in \mathcal{R}$. This time we consider those realizations located in the tails and not only block maxima. Let $x_{t,T} < \dots < x_{1,T}$ be an ordered sample of size t . Pickands' (1975) estimator is defined by

$$\hat{\xi}_{t,T}^P \equiv \frac{1}{\ln(2)} \ln \left[\frac{x_{t,T} - x_{2t,T}}{x_{2t,T} - x_{4t,T}} \right].$$

It can be shown that the Pickands estimator is consistent as long as t is chosen in such a way that $t/T \rightarrow 0$ as $T \rightarrow +\infty$. Dekkers and De Haan (1989) have further shown that Pickands' estimator is asymptotically normal:

$$\sqrt{t} \left(\hat{\xi}_{t,T}^P - \xi \right) \Rightarrow \mathcal{N} \left(0, \nu(\xi) \right) \quad \text{where } \nu(\xi) = \frac{\xi^2 \left(2^{2\xi+1} + 1 \right)}{\left(2(2^\xi - 1) \ln(2) \right)^2}. \quad (5)$$

For empirical purposes, once the parameter ξ has been estimated one can use (5) to compute the standard error. Since Pickands' estimator is defined for all ξ it can be used to discriminate between the three extreme value distributions. There clearly remains the empirical issue on how to choose the optimal t for a given sample. We will turn to that issue later on.

4.4.2 Hill's tail index estimation

This method can only be used for the Fréchet extreme value distribution ($\xi > 0$). We consider again $x_{j,T}$ issued from the raw sample data. Hill's (1975) method estimates ξ using

$$\hat{\xi}_{t,T}^H = \frac{1}{t-1} \sum_{j=1}^{t-1} \ln(x_{j,T}/x_{t,T}). \quad (6)$$

Again if t grows such that $t/T \rightarrow 0$ as $T \rightarrow +\infty$ it can be shown that $\hat{\xi}_{t,T}^H$ is consistent. We also have asymptotic normality:

$$\sqrt{t} \left(\hat{\xi}_{t,T}^H - \xi \right) \rightarrow \mathcal{N} \left(0, \xi^2 \right).$$

For empirical purposes the variance ξ^2 can be estimated using $\hat{\xi}_{t,T}^H$. It has been shown that if the X_t are generated by a Fréchet distribution then the Hill index will yield a more efficient tail index than the Pickands' one. In the appendix we address the issue how an optimal t can be chosen using a bootstrap method.

4.5 Maximum likelihood estimation of the gev distribution

We first notice that the gev distribution of a general non-centered, non-reduced random variable is defined by

$$H_{\xi, \mu, \psi}(x) = \begin{cases} \exp \left(- \left(1 + \xi \frac{x-\mu}{\psi} \right)^{-1/\xi} \right) & \text{if } 1 + \xi \frac{x-\mu}{\psi} > 0, \xi \neq 0, \\ \exp \left(- \exp \left(- \frac{x-\mu}{\psi} \right) \right) & \text{if } \xi = 0. \end{cases}$$

Its density $h_{\xi,\mu,\psi}(x)$ is given by

$$h_{\xi,\mu,\psi}(x) = \begin{cases} \frac{1}{\psi} \left(1 + \xi \frac{x-\mu}{\psi}\right)^{-\frac{1}{\xi}-1} \exp\left(-\left(1 + \xi \frac{x-\mu}{\psi}\right)^{-\frac{1}{\xi}}\right) & \text{where } 1 + \xi \frac{x-\mu}{\psi} > 0, \xi \neq 0, \\ \frac{1}{\psi} \exp\left(-\frac{x-\mu}{\psi} - e^{-\frac{x-\mu}{\psi}}\right) & \text{if } \xi = 0. \end{cases}$$

To compute the maximum likelihood estimates one chooses, say, the T maximum returns x_1, \dots, x_T for T m -histories and fits the likelihood

$$L(\xi, \mu, \psi) = \prod_{t=1}^T h_{\xi,\mu,\psi}(x_t) \mathcal{I}_{\{1+\xi(\frac{x-\mu}{\psi})>0\}}(x_t). \quad (7)$$

Using a maximum likelihood routine the various parameters can now be easily estimated. It turns out (see Smith (1985)) that the usual asymptotic properties hold whenever $\xi > -1/2$.

Expression (7) correctly describes the likelihood if the underlying variables are iid. In empirical work the use of m -histories with m sufficiently large is likely to yield uncorrelated realizations. As we will see later on, the estimates fluctuate according to the method chosen, for financial data the assumption of an identical distribution may not hold.

4.6 Fitting ‘excesses over a threshold’

An other estimation method is based on exceedances of high thresholds. As we have recalled earlier, for large thresholds, the excess distribution function $F_u(x)$ behaves as a Generalized Pareto distribution, $G_{\xi,\psi}(u)$, where the scaling parameter depends on the threshold. The idea of this method is to chose a threshold u , to select the x_t larger than the threshold u , and to fit to all excess returns $x_t - u$ the gpd defined in (2). The density of the gpd has already been determined in equation (3). The likelihood of one observation can be written as

$$l_t = \begin{cases} -\left(\frac{1}{\xi} + 1\right) \ln(\psi + \xi x) + \frac{1}{\xi} \ln(\psi) & \text{if } \xi \neq 0, \psi + \xi x > 0, \\ -\ln(\psi) - \frac{x}{\psi} & \text{if } \xi = 0. \end{cases} \quad (8)$$

Again, it is nearly trivial to set up the numerical maximization for such a function. The advantage of this approach is that one can use all realizations in the sample exceeding u and not only the maximum over m -histories. A drawback is that one has to choose a threshold u . A very high u will again strongly reduce the amount of available data. The moral of evt is that nothing comes for free.

There remains the choice of the threshold. Whereas McNeil and Saladin (1997) use simulations to find reasonable levels of the threshold (finding that about 200 values which are actually located in the tails yield good estimates) an alternative approach is to use a graphical method based on the observation that if the random variable X follows a gpd then, as was shown by Daragahi-Noubary (1989), the mef

$$e(u) = E[X - u | X > u] = \frac{\psi + \xi u}{1 - \xi}$$

is a linear function in u .¹⁰ A threshold should, therefore, be chosen such that the relation between the obtained excesses and the mef is roughly linear. Moreover,

¹⁰It must be that $\xi < 1$, otherwise tails are so heavy that the mean does not exist.

there exists a one-to-one relation between the parameters of the distribution and the intercept and slope of the mef. A simple plot of $\widehat{\epsilon}(u)$ against u reveals the parameters ψ and ξ . Once an empirical mef has been obtained, comparison with the theoretical shapes allows identification of the type of distribution one is dealing with.

5 Empirical results

Since the aim of this study is to provide cross-country evidence for extreme events we consider a large database of global indices. We consider five indices for mature financial markets (Standard and Poor composite 500, Nikkei, Dax, CAC40 and the FTSE100). For emerging markets we consider seven Asian ones (China (Shanghai), Hong Kong, Indonesia, Malaysia, the Philippines, Singapore, South Korea, Taiwan, and Thailand), six Eastern European ones (the Czech Republic, Hungary, Poland, Russia, the Slovak Republic, and Slovenia), as well as seven Latin American ones (Argentina, Brazil, Chile, Colombia, Mexico, Peru, Venezuela).

This database has been extracted out of Datastream. For some of the series care had to be taken for the earlier part of the sample since the reported frequency was not daily.¹¹ We used data from the moment on that the frequency was daily. For some countries we had to append two indices.¹² After these corrections, all the data extracted has been used.¹³

5.1 Descriptive statistics

5.1.1 Cross-country evidence

Table 1 contains various descriptive statistics. We display for the various countries the starting date of the index as well as the number of observations. All series end on December 31st 1998. The sample size ($Nobs$) ranges between 1108 for Poland up to 7826 for mature markets. The average daily return is positive for most countries. The median often takes the value 0. An explanation for this is that prices remain constant over holidays which are not filtered out specially here. From a distributional point of view this suggests that zero returns will be over-sampled. We do not insist on this issue since in this study we focus on the tails of distributions and not on the central part.

Skewness (Sk) and its standardized version (Sk^*) are signed measures of the tail behavior of returns. For developed markets this statistics is generally found to be negative. For those markets it is crashes that introduce asymmetry. For emerging markets the picture is not so clear. There are many markets with a positive skewness. There it is sharp increases in prices that induce asymmetry of returns.

Kurtosis is a symmetric measure of the behavior of the tails of a distribution. This statistic measures the heaviness of tails as compared to the normal one. Considering kurtosis (Ku) and its standardized version (Ku^*), we notice that this statistics is too large for all indices to be reconciled with a normal distribution. Globally returns

¹¹This is the case for Indonesia before October 1989, the Czech Republic before September 1994 and Poland before October 1994. For Brazil the price figures are so small before January 1992 that price changes are meaningless.

¹²For the French index, we used the CAC Général index before July 9th, 1987; for the UK index, we used the FT all-shares before January 1st, 1980.

¹³There is only one exception. For Russia, we suppressed the data for October 16th, 1998, with a +50% return outlier.

have excessively fat-tails. In all cases the Jarque-Bera statistics is very large implying non-normality of the series. Those preliminary statistics confirm our prior that the behavior of extreme values is of great importance in understanding returns.

To glean further insight in the extremal behavior of stock returns we center and reduce all return series and consider the standardized minimum and maximum as well as various percentiles (1%, 5%, 95% and 99%). If the data was truly normal, we would expect for standardized returns that the empirical percentile corresponds to the one of the normal distribution. The 1%, 5%, 95%, and 99% critical values are then -2.3263 , -1.6449 , 1.6449 , and 2.3263 . Comparison of $q1$ and $q99$ reveals for all countries that the extreme 1 percentiles are too large to be compatible with a normal distribution. This confirms our earlier finding that returns' distributions are fat-tailed. When comparing $q5$ and $q95$ with the associated critical values we notice that at this level there are not enough realizations to be compatible with a normal distribution. It is truly the extreme values which generate non-normality in the data. We notice further that if returns were normal, the realization corresponding to the min and max would have virtually no probability to exist.

We next consider heteroskedasticity by regressing squared returns on $l = 1, \dots, 5$ lagged past squared returns. The Engle statistics $T \cdot R^2$, where T is the sample size and R^2 is the coefficient of determination, which is distributed as a χ_l^2 under the null of homoskedasticity takes very large values. We therefore conclude that there is a fair amount of heteroskedasticity in the data.

We now wish to test for autocorrelation among returns. Given the high level of heteroskedasticity we consider a version of the Box-Ljung test which corrects for heteroskedasticity (see also White, 1980). For 5 (10) lags the QW is distributed as a χ_5^2 (χ_{10}^2) whose 95% critical level is 11.07 (18.31). As a consequence the QW statistic is significant for most indices investigated. For some countries the $QW10$ statistics takes particularly large values such as for the CAC40, Thailand, the Philippines, the Czech Republic, Chile, Mexico, Peru, and Colombia. A possibility for such a large autocorrelation is stale prices. Inspection of the $AR(1)$ and $AR(2)$ coefficients suggests that the first order autocorrelation is very high for those countries. Since it is known that the evt developed earlier only holds for independent variables we consider in the empirical part of the study $AR(5)$ filtered series. To check if those series are still autocorrelated we consider again the Box-Ljung statistics. We find that there is no residual autocorrelation for all series. By using $AR(5)$ filtered series we can believe that we took care at least of some dependencies in the data.

5.1.2 Closer inspection of selected indices

To provide further insights in some of the methods, we will consider four indices in particular. Those indices are the CAC40, as well as the one of Singapore, Russia, and Mexico.¹⁴ In Figures 4a and 4b we represent the plot of returns for those indices.

In Table 2 we present for each of the four selected indices the five most extreme positive and negative returns and their associated dates. For the CAC40 we notice the large drop in the market in May 1981 which came from Mitterrand's first election. The October 1987 crash also shows up but its magnitude is not as big as the 1981 crash. We notice that the large drop, for the CAC40, on October 19th 1987 was followed during the next days by other extreme returns of both signs. This turbulent

¹⁴We choose the CAC40 since this index corresponds to our home country. The other indices were chosen randomly, one for each geographical area.

period is associated with the worldwide October crash as well as with the end of the first period of *cohabitation* (October 28th) and the beginning new electoral campaign.

For Singapore we notice the importance of the 1987 October crash which largely dominates the picture. It should be noticed that this period also coincides with the beginning of an electoral campaign as well as multi-racial conflicts. The crash on October 16th, 1989 was similarly due to very troubled elections.

The Russian index only exists for more recent dates.¹⁵ We notice that both the Russian index and the Mexican one reacted jointly during the 1997 and the more recent 1998 crash. The amplitude of returns for the Russian market is noticeable. Anecdotal evidence reveals that on October 28th, 1997 General Lebed was ousted. On September 17th, 1998 occurred a change of the Prime Minister. On June 18th, 1994 elections took place in Russia.

5.2 Hill Estimates

In this section we will discuss the results of the estimation for the Hill tail index. We focus directly on this index rather than on Pickand's since there is clear evidence of fat-tailedness of returns. We would like to emphasize that our analysis is always done on $AR(5)$ filtered series to eliminate autocorrelation.¹⁶ Given our prior concerning the extreme value distribution we follow the advice of Gumbel (1958) and of Embrechts, Klüppelberg, and Mikosch (1997) and start with a graphical analysis. In Figure 5a and 5b we trace the Hill estimates and their associated standard errors for the left and the right tails of returns. Gumbel suggests that one chooses the tail index in an area where the index is roughly linear. We notice that for the left and right tails of the CAC40 the values taken by the tail index are fairly stable once the sample contains about 100 observations. When turning to the Hill-plots associated with Singapore we notice a quickly increasing plot that stabilizes at values around 0.37 for the left tail. The right tail is first stable but increases as samples become larger than 150 observations.

The Hill-plots for Russia are, in Embrechts, Klüppelberg, and Mikosch terminology, *horror*. This indicates that it will be difficult for this country to come up with anything reasonable. Possible reasons are the very short sample or a changing distribution of returns.

Last we turn to the Mexican stock market. Here the sample is more than twice as large as the Russian market (2868 vs 1153 observations). Inspection of the left tail reveals a rather stable part up to about 80 observations. From there on the Hill estimate increases. For the right tail it is even more delicate to find a reasonable estimate. This plot illustrates well the difficulty when estimating tail indices. If one uses a large threshold, i.e. few observations, then the Hill estimate is bad because there are just too few observations. If one lowers the threshold then one uses too many observations from a region too distant from the extreme and the estimate will be biased. This illustrates, especially for small samples the importance of a method allowing a choice of optimal tail size.

To get the optimal Hill index we follow the procedure described in the appendix of the paper and perform a bootstrap estimation. Table 3 displays our results for the left and the right tails of the returns distribution. For the group of mature markets

¹⁵Notice the change in scale in Figure 4b.

¹⁶It should be noticed that if one works on unfiltered returns the estimates do not differ qualitatively from the ones reported.

we notice that on average the left and right tail index are of similar magnitude. A similar result holds for the Asian markets but there the level of ξ is higher suggesting that those markets are more prone to crashes. For the Eastern European markets the left tail is stronger. On average the opposite holds for the Latin American markets. Considering the standard deviation of the estimates of ξ within geographic groups we notice that the mature markets have very similar indices. The DAX and CAC40 exhibit strikingly similar ξ estimates of the left tail. The estimates for the right tail are close. The dispersion of ξ among Asian and Latin American markets is similar. The group where indices vary the most is the Eastern European one. For instance, considering the left tail, the Czech Republic has a very small tail index (0.23) whereas Hungary has a very large (0.41) one.

When considering the accuracy of the estimation of the tail index measured by $\text{STD}(\xi)$ we find that for developed markets the estimation is best. The accuracy falls for Asia, even more for Latin America followed by the least precise Eastern European countries. This loss of precision can be explained by the decreasing sample size. It is also because of this smaller sample size that there are less tail observations useful for the tail index estimation (measured by k^*).

5.3 Existence of moments

Now we wish to discuss the existence of moments of the return generating distribution for the various countries. Let (ξ, ξ^2) be the tail index and its variance.¹⁷ As we have seen in section 3.3 the r -th moment of a distribution exists if $r < 1/\xi$. We now describe a strategy to identify if the r -th moment exists or not.

Let n be an integer such that $n < 1/\xi \leq n + 1$. We first test $H_0^1 : n = 1/\xi$ against $H_1^2 : n > 1/\xi$. If we reject H_0^1 we will be confident (up to the usual level of significance) that at least the n -th moment exists. For the case where H_0^1 cannot be rejected we further test if we can reject the null in $H_0^2 : n - 1 = 1/\xi$ against $H_1^2 : n - 1 > 1/\xi$ (for all empirical applications this is as far as we need to go). If we reject H_0^2 we are confident of the existence of the $(n - 1)$ -th moment. For those cases where we rejected H_0^1 we further test $H_0^3 : \xi = n + 1$ against $H_1^3 : \xi \leq n + 1$. In all cases where this issue mattered we could reject H_0^3 and, thus, we did not need the $(n + 2)$ -nd moment.

In Table 4 we implement those various tests by distinguishing the left from the right tail. For the S&P index we confirm Longin's (1996) finding that the third moment exists but not the fourth one. For all mature markets we cannot reject the existence of the third moment except for the Nikkei whose right tail creates some difficulties. We can however confirm without any doubt that the second moment exists.

Further scrutiny of Table 4 confirms the existence up to a second moment for nearly all indices except for Thailand where the right tail does not allow for a second moment or more.

Concerning the existence of a third moment, the picture is fuzzier. Globally we find for the left tail that for 18 indices, and for the right tail for 14 indices, the third moment exists. It is in Eastern European and Asian markets that the existence of this moment appears to fail.

¹⁷It should be noticed that a trivial application of the δ -method indicates that if a random variable ξ has variance ξ^2 then $1/\xi$ has variance $1/\xi^2$.

Turning to the existence of a fourth moment, we see that for the Czech Republic the left tail and for Brazil the right tail appear to allow for a fourth moment. However, when considering the distribution as a whole, thus, combining both the left and right tail, for no country do we find existence of a fourth moment.

The clear indication of the existence of a second moment, maybe even of a third moment, sheds some doubts on the use of distributions such as stable laws for which the second moment does not exist. For a similar conclusion see also Loretan and Phillips (1994) and Lux (1998).

5.4 ML estimation of the Generalized Pareto distribution

To further assess the data and also to set the way for further estimates such as of high quantiles it is necessary to focus not only on the tail index but the tail distribution. Given the theoretical elements recalled above we know that this distribution should be a Generalized Pareto distribution.

Again, in a similar manner as for the Hill estimate the question of an *optimal* threshold arises. If one chooses a threshold too much in the tail one obtains a very inaccurate estimate because just too few observations are used in the estimation. On the other hand, when using too many observations one contaminates tail observations with observations from the center.

Before presenting ML estimates we provide some graphical analysis allowing for a better understanding of what goes on in the data.

5.4.1 Mean-Excess plots

To get an idea what level of the threshold one should use, we trace in Figures 6a-b the mean-excess functions for the left and right tail of returns of the four countries under closer scrutiny. We notice that for a high threshold the mef behaves more erratically than for low thresholds. On the other hand, as the threshold increases, after a flat part the mean-excess function increases linearly suggesting that it is there that one enters the true tail region.

When turning to Russia and to Mexico we notice the very erratic behavior of the mean-excess function. Those countries illustrate in a rather dramatic manner what happens if one tries to estimate tail indices on rather small samples. For such situations it would be useful to have a method for choosing the optimal sample size. Since to our knowledge, there exists no algorithm who does this for the gpd we use as threshold the value corresponding to the optimal sub-sample (t^*) obtained with the method outlined in the appendix and applied earlier in the quest for an optimal Hill index.

5.4.2 The parameter estimation

Before discussing the results of the gpd estimation reported in Table 5, we would like to mention that we started with an estimation of the gpd where ξ and ψ were estimated unrestrictedly. We do not report those estimates here.¹⁸ We found for many countries a large difference in the estimates of ξ relative to Table 3. Usually the tail index tended to be smaller than the one reported in Table 3.

¹⁸It should be noticed that for mature markets (with the exception of the Nikkei) the unrestricted ξ was always very close to the one reported here.

We experimented with various tail sizes and, yet, obtained similar results. For this reason we decided to report the estimates of ψ setting the ξ parameter equal to the value obtained in Table 3. Furthermore we use a likelihood-ratio test to see if the data rejects the gpd when ξ is restricted to the value reported in Table 3. Only very few restrictions are rejected. A noticeable case is the Nikkei. Further rejections occur for the right tail of Indonesia, both tails of Thailand, the left tail of Hungary, and the right tail of Peru.

Whereas the parameter ξ reported in Table 3 measures the tail of returns asymptotically, the parameter ψ controls the size of the tail at finite distance. See also Figure 3 to understand what happens as the parameters change. In Table 5 we notice that ψ takes the smallest value for mature markets. Both for the left and the right tails ψ takes the value 0.66. Exceptions in this group are the S&P with rather small estimates (0.47 for the left and 0.54 for the right tail) and the Nikkei with large estimates (0.90 and 0.89 for the left and right tail). Those findings are compatible with a smaller standard deviation for the S&P than for the Nikkei, yet with extreme events of equal or larger magnitude. Consideration of the other geographic areas reveals that the Asian and Latin American markets have a roughly similar dispersion of large values whereas the Eastern European ones have the largest ones. The precision of ψ , written as $\text{STD}(\psi)$, varies also across the geographic regions proportionally to the available sample size. Noticeable is again Russia whose ψ oscillates around 3.9 implying large variations in returns.

5.4.3 The actual fit of the gpd

Having obtained parameter estimates for the gpd it is possible to check how well such a distribution actually fits the tails. For this purpose we trace in Figures 7a-d the estimated gpd (as a solid line) as well as the empirical distribution function (with large dots). We notice both for the CAC40 as well as for Singapore a rather good fit of the gpd. Since the fitted gpd is above actual observations this suggests an underestimation of extreme returns. When turning to Mexico and to Russia we notice unfortunately a strong deterioration in the fit of the gpd. The gpd tends to overstate the frequency of extreme events. Those graphs, therefore, show the limited abilities of the gpd to describe the tails for certain emerging markets. For small samples, more *ad-hoc* distributions might be useful. Such constructions are left for further research.

This means that if one were to estimate mean waiting times between large events, one would find too short time intervals.

5.4.4 Estimation of High Quantiles

As an illustration of the use that can be made with those estimations we compute the largest value that is expected to occur over a 5 respectively 50 years time horizon. This extrapolation clearly assumes some form of stationarity. Given that today's emerging markets will be fully developed within the next 50 years, one can then expect a rather different behavior of those markets, more similar to today's mature markets. This observation clearly stresses the assumption of stationarity.

In Table 6 we present the result of the estimation of high quantiles for the left and right tails. We report again the sample min and max. We notice that for most indices the sample min and max are bounded by the estimates found respectively for the 5 and 50 year horizons. An example which does not obey to this rule is given by the

S&P. Whereas, according to the gpd, over a 50 year period we expect as a worse event a drop of -9.79% , we actually found on October '87 a crash of -22.83% . Had one based VaR computations entirely on evt one might have lost quite heavily. Similar finding actually hold in a lesser extent for the other developed countries. For instance for France, the worse event over a 50 year period is expected to be a -13.27% drop, whereas we found a drop of -13.91% in the sample. Those observations confirm our earlier suspicion that the gpd underestimates extreme realizations for mature markets.

Turning to the set of emerging markets we notice that on average our gpd estimates tend to bound the sample extreme values. There are some exceptions: we notice that the -40% crash of Hong Kong or the -30% crash in Singapore could not be guessed given the gpd estimation. On the other hand when turning to the Eastern European markets we tend to find overly dramatic results. For instance the gpd estimates would foresee a -62% crash (-56%) for Hungary (respectively Russia) over a 50 year horizon. With the same token, for those countries an upward crash of $+79.45\%$ (respectively $+62.57\%$) would be expected over a 50-year horizon. Those values appear too large. For emerging markets there is therefore a tendency to overestimate the probability of extreme variations.

For the Latin American countries our results seem to be rather reasonable.

5.5 Estimation of the Generalized Extreme Value distribution

5.5.1 A preliminary analysis of extremes

An alternative estimation method consists in obtaining extremes over m -histories and to estimate the resulting generalized extreme value distribution. As a preliminary step in this type of analysis we trace in Figures 8a-8b the histograms of the minimum and maximum 20-histories that is of returns grouped month by month.

We notice that the extremes of the CAC40 do not spread out very far. On the contrary, for the emerging markets the range of possible extreme values is very widespread. This preliminary analysis of extremes in conjunction with Figure 1, which displayed the density of the various limit distributions, confirms that extremes of returns are expected to follow a Fréchet distribution.

To foster this prior, we present in Figures 9a-9b the QQ-plots of the extremes. For all cases under consideration we notice concave QQ-plots. This graphical analysis confirms our prior that the distribution of extremes is of the Fréchet type in this international setting (see also Longin, 1996).

5.5.2 ML estimates of the generalized extreme value distribution

In Table 7 we present the parameter estimates for the gev while considering 20-histories. Considering ξ over the various geographic areas, we notice that the tail estimates take smaller values. We used various m -histories and tended in general to find slightly lower estimates than for the gpd estimation. For instance, the average left tail estimate for mature markets was 0.3128 with the Hill estimate, now it is only 0.2487. One possibility explaining this is that the dynamics followed by stock prices is very complicated and that the required assumptions for the various limit theorems do not apply.

It should be noticed that this finding changes nothing concerning our discussion of the existence of at least up to a second moment for returns. On average the standard

errors of the tail index are larger than for the Hill estimate. This corroborates the result that the Hill estimate tends to have good statistical properties. Within the various groups we also find higher dispersion of the parameters than before.

To sum up, the estimation of the gev confirms our prior that returns are distributed according to a Fréchet distribution. We find that the tail estimates tend to be smaller now than the Hill estimates.

6 A comparison with other data

6.1 The early days of a global US index

So far we have established that the tail parameters of various geographic areas take significantly different values. This observation suggests that financial markets could have a behavior which converges through time. If such was the case, then if we considered a presently developed market that existed already a long time ago, we could expect such a market to have behaved in its early days like a currently emerging one. To test such a conjecture we consider William Schwert's (1990) global US stock index during its early 7826 days (that means that we have a same sample length as for today's mature markets). The period covered goes from February 17th 1885 to March 2nd 1911. In Figure 10 we plot the early returns of this global US index. Casual inspection of Figure 10 does not reveal any particular behavior for the early days of the US index. Formal descriptive statistics displayed in Table 8 indicate a significant negative skewness and fat tails. The Jarque-Bera statistic takes the value 16'289.35 and is highly significant. We also report that an Engle test reveals conditional heteroskedasticity.

We also display in Table 9 the results of the optimal Hill estimate. For the left tail we find an estimate of 0.3401 and for the right tail an estimate of 0.4181. This latter estimate is quite high when compared with more recent estimates. When we compare the earlier standard deviation (0.87) with the more recent one (0.94) we notice that volatility is presently higher. Clearly, this level of volatility is much lower (nearly half) than the type of volatility met in presently emerging markets. Those observations lead us to conclude that the early years of this historical database have little to do with the tail behavior of presently emerging markets. As a consequence little can be learned from such a series about how presently emerging markets may evolve in terms of tail behavior.

6.2 The recent dynamics of mature markets

One may explain the differences in the tail behavior of mature and emerging markets by the fact that the sample used for mature markets is much larger than the sample for emerging markets. Indeed our dataset begins in the late 60s for mature markets, whereas it begins in the mid-90s for Eastern European markets and for some Latin American markets. We therefore reestimate the different parameters representing the tail behavior for the recent days of mature markets: the sample covers the period from April 2nd, 1993 to December 31st, 1998, that is 1500 observations.

As far as summary statistics, presented in Table 8, are concerned, we first note that over the recent period, the daily average return is negative for Germany and France, whereas it is positive for the other countries. Moreover the skewness is

significantly positive for both countries. Lastly it is worth noting that the excess kurtosis is far smaller for all countries than for the whole period.

Table 9 reports optimal Hill estimates. For the left tail we obtain basically the same estimates than over the whole period. Estimates of ξ are close to each other between 0.28 and 0.33. The optimal number of observations t^* is quite low (between 23 and 51). The estimates of the right tail Hill parameters are more volatile, between 0.28 and 0.36. For the Nikkei the ξ parameter is larger than over the whole period (0.36 vs 0.28). This indicates that over the recent period the 3-rd moment may exist for the right tail of the Nikkei distribution.

Gpd estimates give more ambiguous results: for the left tail of the distribution, the parameter ψ is larger than over the whole period, especially for the DAX and the CAC40. For these indices, ψ is as high as 1.1114 and 0.8541 respectively. These figures are comparable to some of the figures obtained for emerging markets. For the right tail on the contrary estimates of ψ are smaller than over the whole period. We obtain rather small estimates for the S&P and the FTSE100 (0.4136 and 0.3624 respectively).

Using a small sample for mature markets does not help to reconcile the estimates obtained for mature and emerging markets. On the one hand, Hill estimates remain very stable when the period is shortened; on the other hand, gpd parameter estimates dramatically change, but in different directions when the left and right tails are separately considered. Therefore, estimates obtained for mature financial markets over the recent period are not closer to estimates obtained for emerging markets. The smaller sample size for our developing countries, therefore, does not explain the difference in estimates.

7 Conclusion

In this paper we considered a large set of 27 countries composed by several mature markets and other emerging ones. We review important elements of extreme value theory and apply these to the database. For all indices it is shown that returns' distributions are fat-tailed. In order to be compatible with the tail behavior, the return generating process has to lay in the domain of attraction of a Fréchet distribution. This means that certain generating processes such as the mixtures of normals can be precluded. On the other hand, GARCH processes with possible jumps are compatible with the observed tails. We show that the Hill estimates of the tail indices are of roughly similar magnitude across indices. In particular, we find for nearly all indices a tail behavior compatible with the existence of a first and second moments. For many indices even a third moment seems to exist.

When turning to the estimation of the Generalized Pareto distribution we notice a certain instability in the estimation of the tail index as compared to the Hill estimate. In particular for countries with relatively small samples the estimation is difficult. Possible reasons for this difficulty is that in emerging countries the distribution of returns may change in character and its complicated evolution may invalidate the iid assumption behind the asymptotic results of evt. Since the estimation of the gpd is the basis for the estimation of high quantiles, waiting times between large given threshold exceedances, and of the extreme realization over a given time horizon, its use has to be done carefully for emerging markets. We actually show that the estimates of a gpd tend to be on the conservative side: crashes are expected too often,

exceedances of even unreasonable thresholds appear likely...

When turning to mature markets we find a rather good fit of the gpd suggesting that it could be an interesting tool for applications such as VaR, regulation of Central Bank reserves, and of Futures margin requirements. Nonetheless the user of those methods should be aware that for our sample the drop of the S&P in October '87 could not be predicted. Our tail estimate of the S&P suggests over a 50 year period as a worse possible outcome a drop of 9.79% which has to be compared with the actual drop of 22.83%!

As we mentioned, the fit of a gpd to emerging market's data is problematic. One may therefore ask if financial history can help us to improve our understanding of worse case scenarios. To partly answer this question we consider a subsample of an US daily global index covering the same sample size as our main S&P sample. We find much less volatility and extreme values than in current financial markets. This suggests that the presently available historic data may not help much in understanding how returns of currently emerging markets may evolve. Moreover, what concerns its tail behavior, the US market seems to be different from other developed markets.

From an econometric point of view several issues remain open. In this study we mostly considered the individual behavior of stock indices. For an investor interested in international diversification the study of simultaneous cross-border crashes may also be of great importance. An other issue concerns the fit of the gpd for small samples under non-standard assumptions, that is dropping the independence or the identical distribution assumption. We know, based on theoretical grounds, that if returns behaved iid, asymptotically, tails can be described by a gpd. For emerging markets the iid assumption is likely to be erroneous and clearly the sample is of small size. How can theory be improved here? A last econometric issue concerns the optimal choice of m for the estimation of the generalized extreme value distribution. Development of a bootstrap based optimal estimate of m appears valuable. Implications for the asymptotics, once the iid assumption of returns is dropped also seems relevant for the estimates of the generalized extreme value distribution.

8 Appendix

In this appendix we wish to outline how we obtained an *optimal* t^* for the Hill estimator. There is also the issue about how to obtain the level delimiting the tail over which the gpd should be estimated. Since to our knowledge, there exists no theory on how to obtain such a level, we will also use the t^* found for the Hill estimator for this purpose.

The approach that we use here goes back to work by Hall (1990), and Danielsson and de Vries (1997). Other work in this area is by Dacorogna, Muller, Pictet, and de Vries (1995), Drees and Kaufmann (1997), Beirlant, Vynckier, and Teugels (1996). The fact that we use the method by Danielsson and de Vries (1997) can be justified by the work of Lux (1998) who compares the various techniques and concludes that even though the optimal t^* varies quite a lot depending on the method chosen, the eventual tail estimate remains about the same.

The Hill index has been defined in (6) as:

$$\hat{\xi}_{t,T}^H = \frac{1}{t-1} \sum_{j=1}^{t-1} \ln(x_{j,T}/x_{t,T}).$$

where $x_{t,T} < \dots < x_{1,T}$ is the ordered sample. The index t varies from 1 to T . We are seeking the t^* minimizing the mean squared error

$$t^* \in \operatorname{argmin}_t \operatorname{MSE} \left(\xi_{t,T}^H \right) \equiv \operatorname{argmin}_t \operatorname{E} \left[\left(\xi_{t,T}^H - \xi \right)^2 \right]. \quad (9)$$

The implementation of (9) is based on the estimation of ξ , written $\tilde{\xi}$, using observations that are truly far in the tail. Danielsson and de Vries (1997, p. 246) suggest to chose an estimate of ξ obtained with 1% of the largest observations. Focusing on only 1% of the sample gave bad results for emerging markets where the entire sample is rather small anyway. For those countries we have to use more observations such as 2.5% or 5% to get a meaningful tail estimate. Next, one tries to approximate the expectation in (9) by using simulations over subsamples (a bootstrap estimation) which gets then adjusted to take into account the full sample size. More specifically, one chooses an arbitrary level $T_1 < T$ and constructs K randomly selected subsamples of size T_1 . Let $l = 1, \dots, K$ be the index of the l -th subsample. It is then possible to obtain the Hill estimate $\xi_{t_1, T_1}^{H,l}$, $t = 1, \dots, T_1$. An approximation of (9) is then

$$\frac{1}{K} \sum_{l=1}^K \left(\xi_{t_1, T_1}^{H,l} - \tilde{\xi} \right)^2.$$

An optimal t_1^* minimizing this expression has to exist since if t_1 is very small, because of the resulting small amount of observations involved, the error will be large. On the other hand, for t_1 large, we are contaminating the tail with observations from the center of the distribution which should also increase the error.

Since the choice of T_1 is arbitrary it is necessary to adjust the optimal t_1^* so that it corresponds to the full sample size T . It has been shown that the adjusted t^* is defined by

$$t^* = t_1^* \left(\frac{T}{T_1} \right)^{2\beta/(2\beta+1/\xi)} \quad (10)$$

where β is an additional parameter. To obtain β it is necessary to construct the following j -th empirical log-moments and a variable Δ defined as:

$$\begin{aligned} \tilde{m}^{(j)} &= \frac{1}{t-1} \sum_{i=1}^{t-1} [\ln(x_{i,T}/x_{t,T})]^j, \\ \Delta &= \frac{\tilde{m}^{(1)} - \tilde{m}^{(2)}/2\tilde{m}^{(1)}}{\tilde{m}^{(3)}/3\tilde{m}^{(2)} - \tilde{m}^{(4)}/4\tilde{m}^{(3)}}. \end{aligned}$$

The parameter β appearing in (10) can then be estimated with

$$\hat{\beta} = \left(\sqrt{\Delta} - 1 \right) / \hat{\xi}.$$

It is also possible that occasionally the delta is smaller than one. In such a situation β is not defined. In such a case we revert to the suggestion of Hall (1990) and select $\hat{\beta} = 1/\hat{\xi}$.

The optimal Hill estimate is then $\hat{\xi} = \xi_{t^*-1, T}^H$.

It should be noticed that in this approach there remain many arbitrary choices. First, there is the choice of the initial estimate of ξ whose estimation has already been discussed. Then there is the choice of T and T_1 . We chose for T the 10% largest observations and for T_1 half the size of T . Last, one can play with the number of bootstrap simulations to be performed. We performed always 100 bootstrap simulations.

References

- [1] Baillie, R.T, T. Bollerslev, and H.O. Mikkelsen (1996), Fractionally Integrated Generalized Autoregressive Conditional Heteroskedasticity, *Journal of Econometrics*, 74, 2-30.
- [2] Balkema, A., and L. De Haan (1974), Residual Life at Great Age, *Annals of Probability*, 2, 792-804.
- [3] Beirlant, J., P. Vynckier, and J.L. Teugels (1996a), Tail Index Estimation, Pareto Quantile Plots, and Regression Diagnostics, *Journal of the American Statistical Association*, 91, 1659-1667.
- [4] Beirlant J., P. Vynckier, and J.L. Teugels (1996b), Practical Analysis of Extreme Values. Leuven, University Press.
- [5] Bekaert, G., and C.R. Harvey (1995), Time-Varying World Market Integration, *Journal of Finance*, 50, 403-444.
- [6] Bekaert, G., and C.R. Harvey (1997), Emerging Equity Market Volatility, *Journal of Financial Economics*, 43, 29-77.
- [7] Blattberg, R., and N. Gonedes (1974), A Comparison of Stable and Student Distributions as Statistical Models for Stock Prices, *Journal of Business*, 47, 244-280.
- [8] Claessens S., S. Dasgupta, and J. Glen (1995), Return Behavior in Emerging Stock Markets, *World Bank Economic Review*, 9, 131-151.
- [9] Danielsson, J., and C.G. de Vries (1997), Tail Index and Quantile Estimation with Very High Frequency Data, *Journal of Empirical Finance*, 4, 241-257.
- [10] Daracogna M.M., U.A. Müller, O.V. Pictet, and C.G. De Vries (1995), The Distribution of Extremal Foreign Exchange Rate Returns in Extremely Large Data Sets, mimeo, Olsen & Associates, Zurich, and Erasmus University Rotterdam.
- [11] Daragahi-Noubary, G.R. (1989), On Tail Estimation: An Improved Method, *Mathematical Geology*, 21, 829-842.
- [12] De Haan, L., I.S. Resnick, H. Rootzèn, and C.G. de Vries (1989), Extremal Behavior of Solutions to a Stochastic Difference Equation with Applications to ARCH Processes, *Stochastic Processes and Their Applications*, 32, 213-224.
- [13] Dekkers A.L., and L. De Haan (1989), On the Estimation of the Extreme Value Index and Large Quantile Estimation, *Annals of Statistics*, 17, 1795-1832.
- [14] Drees, H., and E. Kaufmann (1997), Selecting the Optimal Sample Fraction in Univariate Extreme Value Estimation, mimeo, University of Cologne and University of Siegen.
- [15] DuMouchel, W. (1983), Estimating the Stable Index In Order to Measure Tail Thickness: A Critique, *Annals of Statistics*, 11, 1019-1031.
- [16] Embrechts, P., C. Klüppelberg, and T. Mikosch (1997), Modelling Extremal Events, Springer Verlag, Berlin.

- [17] Engle, R.F. (1982), Auto-Regressive Conditional Heteroskedasticity with Estimates of the Variance of United Kingdom Inflation, *Econometrica*, 50, 987-1007.
- [18] Fama, E. (1963), Mandelbrot and the Stable Paretian Hypothesis, *Journal of Business*, 36, 420-429.
- [19] Fama, E., and R. Roll (1971), Parameter Estimates for Symmetric Stable Distributions, *Journal of the American Statistical Association*, 66, 331-338.
- [20] Fielitz, B. (1976), Further Results on Asymmetric Stable Distributions of Stock Prices Changes, *Journal of Financial and Quantitative Analysis*, 11, 39-55.
- [21] Fielitz, B., and J. Rozell (1983), Stable Distributions and Mixtures of Distributions Hypotheses for Common Stock Returns, *Journal of the American Statistical Association*, 78, 28-36.
- [22] Gnedenko, B.V. (1943), Sur la Distribution Limite du Terme Maximum d'une Série Aléatoire, *Annals of Mathematics*, 44, 423-53.
- [23] Gumbel, E.J. (1958), *Statistics of Extremes*, Columbia University Press, New York.
- [24] Hall, P. (1990), Using the Bootstrap to Estimate Mean Square Error and Select Smoothing Parameter in Nonparametric Problem, *Journal of Multivariate Analysis*, 32, 177-203.
- [25] Hill, B.M. (1975), A Simple General Approach to Inference About the Tail of a Distribution, *Annals of Statistics*, 3, 1163-1173.
- [26] Hols, M.C.A.B., and C.G. de Vries (1991), The Limiting Distribution of Extremal Exchange Rate Returns, *Journal of Applied Econometrics*, 6, 287-302.
- [27] Jansen, D.W., and C.G. de Vries (1991), On the Frequency of Large Stock Returns: Putting Booms and Busts into Perspective, *Review of Economics and Statistics*, 73, 18-24.
- [28] Jenkinson, A.F. (1995), The Frequency Distribution of the Annual Maximum (or Minimum) Values of Meteorological Elements, *Quarterly Journal of the Royal Meteorology Society*, 87, 145-158.
- [29] Kon, S. (1984), Models of Stock Returns: A Comparison, *Journal of Finance*, 39, 147-165.
- [30] Kearns, P., and A. Pagan (1977), Estimating the Density Tail Index for Financial Time Series, *Review of Economics and Statistics*, 79, 171-175.
- [31] Koedijk, K.G., M.M.A. Schafgans, and C.G. de Vries (1990), The Tail Index of Exchange Rate Returns, *Journal of International Economics*, 29, 93-108.
- [32] Leadbetter, M.R., G. Lindgren, and H. Rootzèn (1983), *Extremes and Related Properties of Random Sequences and Processes*, Springer Verlag, New York.
- [33] Longin, F.M. (1996), The Asymptotic Distribution of Extreme Stock Market Returns, *Journal of Business*, 69, 383-408.

- [34] Loretan, M., and P.C.B. Phillips (1994), Testing the Covariance Stationarity of Heavy-Tailed Time Series, *Journal of Empirical Finance* 1, 211-248.
- [35] Lux, T. (1998), The Limiting Extremal Behavior of Speculative Returns: An Analysis of Intra-daily Data from the Frankfurt Stock Exchange, mimeo, University of Bonn.
- [36] Mandelbrot, B. (1963), The Variation of Certain Speculative Prices, *Journal of Business*, 35, 394-419.
- [37] Mandelbrot, B. (1997), *Fractals and Scaling in Finance: Discontinuity, Concentration, Risk*, New York, Springer Verlag.
- [38] McNeil, A. (1997), Estimating the Tails of Loss Severity Distributions Using Extreme Value Theory, *ASTIN Bulletin*, 27, 117-137.
- [39] McNeil, A., and T. Saladin (1997), The Peaks over Thresholds Method for Estimating High Quantiles of Loss Distributions, mimeo, ETH Zentrum.
- [40] Pickands, J. (1975), Statistical Inference Using Extreme Order Statistics, *Annals of Statistics*, 3, 119-131.
- [41] Rockinger, M., and G. Urga (1998), Transition Economies' Stock Markets with a Comparison with Established Stock Markets, mimeo, London Business School.
- [42] Rootzèn, H., and N. Tajvidi (1996), Extreme Value Statistics and Wind Storm Losses: A Case Study, *Scandinavian Actuarial Journal*.
- [43] Schwert, W.G. (1990), Indexes of United States Stock Prices from 1802 to 1987, *Journal of Business*, 63, 399-426.
- [44] Shields, K. (1997), Stock Return Volatility on Emerging Eastern European Markets, *The Manchester School Supplement*, 65, 18-138.
- [45] Simkowitz, M., and W. Beedles (1980), Asymmetric Stable Distributed Security Returns, *Journal of the American Statistical Association*, 75, 306-312.
- [46] Smith, R.L. (1985), Maximum Likelihood Estimation in a Class of Non-Regular Cases, *Biometrika*, 72, 67-90.
- [47] Stărică, C., and O. Pictet (1999), The Tails the Tails of GARCH(1,1) Processes Tell, mimeo, Dept. Statistics, University of Pennsylvania.
- [48] Tucker, A. (1992), A Reexamination of Finite- and Infinite- Variance Distributions as Models of Daily Stock Returns, *Journal of Business and Economic Statistics*, 10, 73-81.
- [49] White, H. (1980) A Heteroscedasticity-Consistent Covariance Matrix Estimator and a Direct Test for Heteroscedasticity, *Econometrica*, 48, 817-838.

Captions

Table 1: The first line of this Table indicates the date when our series start. All series end with December 31st 1998. *Nobs* is the number of observations in each series. *Std* is the standard deviation of returns. *Sk* (Sk^*) and *Ku* (Ku^*) represent the skewness (its standardized version) and its analogues for kurtosis. The Jarque-Bera (*JB*) statistics is defined as $(Sk^*)^2 + (Ku^*)^2$. It is distributed as a χ_2^2 under the null hypothesis of normality. *Stud. min* and *stud. max* represent the minimum and maximum of centered and reduced returns. *q1*, *q5*, *q95*, and *q99* represent the 1, 5, 95, and 99 percentile of centered and reduced returns. *TR1* (*TR5*) represents the Engle statistic for heteroskedasticity obtained by regressing squared returns on one (five) lags. AR_j , $j = 1, \dots, 5$, are the 5 coefficients of autocorrelation. *Q5*, *Q10*, and *Q20* represent the Box-Ljung statistics without correction for heteroskedasticity. Those with correction for heteroskedasticity are noted *QW5*, *QW10*, and *QW20*. *Q5* and *QW5* on *filt. data* are the Box-Ljung tests without and with correction for heteroskedasticity after the returns have been filtered by an $AR(5)$ autoregression.

Table 2: For four selected indices we represent the five most extreme positive and negative returns with their associated dates.

Table 3: We present the optimal Hill estimates both for the left and the right tail of the returns' distribution. The optimal sample size is determined by a bootstrap search. t^* represents the optimal number of observations that should be included in the tail. The threshold, that is the value taken by the t^* -th observation is also represented. We present ξ and its standard deviation $STD(\xi)$. For each group of countries we present the mean and standard deviation of the various estimates.

Table 4: We test if a given moment exists for a distribution. Let $r = 1/\xi$ and n be an integer such that $n < r \leq n + 1$. If we can accept the alternative hypothesis $r > n$ then we believe that the n -th moment exists. In this case we have a YES in the table. If $r = n$ cannot be rejected we further test if $r = n - 1$ can be rejected. If this test can be rejected we have an A in the Table. It means that the hypothesis of the existence of an n -th moment cannot be rejected.

Table 5: Maximum likelihood estimates of the parameter ψ of the Generalized Pareto distribution given the estimate of ξ from Table 3. As tail-sample we use the same one as for the Hill estimation with optimal sample size. *Lik.* represents the log-likelihood of the sample. *LRT* is a likelihood ratio statistic to check if the use of ξ from Table 3 rather than an unconstrained version is acceptable.

Table 6: Here we present the high quantile estimates of the various countries. The estimate x_{999} (resp. x_{9999}) correspond to the largest realization expected over a 5 (resp. 50) year horizon. To render those estimates comparable with actual data we recall the min/max obtained from actual data displayed initially in Table 1.

Table 7: This Table presents the parameter estimates for the gev distribution. We distinguish the left from the right tail as well as various geographic areas.

Table 8: Here we present descriptive statistics for the early days of a global US index as well as for subsamples of presently mature markets.

Table 9: For both tails of the distribution we present the results of the optimal Hill estimates and the parameters of the gpd estimation. The series considered are the early days of a global US index as well as subsamples of presently mature markets.

Figure 1: Representation of the Gumbel, Fréchet, and Weibull densities.

Figure 2: Representation of various mean-excess functions conditional on the distribution of the underlying random variables. We get for Pareto random variables

$(k + u)/(1/\xi - 1)$. In the graph we take $k = 2$, and $\xi = 0.5$. For the Weibull we obtain a mef of $u(1 - \tau)/\tau$. We choose $\tau = 0.75$ and $\tau = 1.25$. The exponential yields a mef of $1/\lambda$. Here we take $\lambda = 3$.

Figure 3: This graph traces the density of the Generalized Pareto Distribution for various parameters. As the tail index ξ increases the tail becomes fatter, meaning that less moments will exist. If for a given tail index the scale parameter increases, the tail fattens for *small* tail values. Asymptotically the tail is the same.

Figures 4a and 4b: Those graphs trace the evolution of returns for selected countries over time. Notice the smaller samples in Figure 4b.

Figures 5a and 5b: Those graphs represent the Hill plots for the left and right tails of return distributions. For all graphs the central line (continuous line) corresponds to the Hill estimates given a certain amount of observations in the tail. The dashed lines correspond to 95% confidence intervals. The Hill estimation is always computed with at least 15 tail observations.

Figures 6a-6b: Those Figures represent the graphs of the mean-excess function for various thresholds.

Figures 7a-7d: Here we trace the estimated gpd (continuous line) against the empirical distribution function (dotted line). The upper graph is always for the left tail of returns whereas the lower graph corresponds to the right tail.

Figure 8: In those Figures we represent the histogram of the maximum and minimum of blocks of 20 returns.

Figures 9a-9b: Here we represent the QQ-plots. For a given ordered sample $\{x_{T,T}, \dots, x_{1,T}\}$ we trace the $x_{t,T}$ on the horizontal axis and on the vertical axis $F^{\leftarrow}(t/T + 1)$ that is the inverse cumulative distribution function of the Gumbel. The straight line represents an OLS fit of a line through the set $\{(x_{t,T}, F^{\leftarrow}(t/T + 1)), t = 1, \dots, T\}$. Concavity of this set implies that extremes follow a Fréchet distribution.

Figure 10: This graph represents the first 7826 returns of a large US index. The period starts on February 17th 1885 and ends on March 2nd 1911.

Figure 11: Here we represent similarly to Figures 7 the estimated and the actual distribution function.

	S&P	Nikkei	DAX	CAC40	FTSE100	China	Hong Kong	Indonesia	Malaysia	Philippines
Beg. date	19690102	19690102	19690102	19690102	19690102	19920522	19750102	19891016	19800103	19860103
Nobs	7826	7826	7826	7826	7826	1725	6261	2404	4956	3390
Mean	0,032	0,025	0,028	0,035	0,035	-0,006	0,065	-0,006	0,021	0,080
Median	0,010	0,003	0,000	0,000	0,025	0,000	0,000	0,000	0,000	0,000
Std	0,935	1,073	1,059	1,039	1,034	3,174	1,812	1,472	1,621	1,935
Min	-22,833	-16,135	-13,710	-13,910	-13,029	-17,905	-40,542	-12,732	-24,153	-15,786
Max	8,709	12,430	8,872	8,225	8,943	28,860	17,247	13,126	20,817	15,657
Sk	-1,975	-0,295	-0,622	-0,651	-0,322	1,369	-2,339	0,415	-0,345	0,083
Sk*	-71,336	-10,655	-22,452	-23,518	-11,630	23,212	-75,553	8,304	-9,910	1,975
Ku	50,702	15,126	10,742	11,169	10,161	13,763	52,440	15,257	33,135	8,961
Ku*	915,571	273,142	193,974	201,679	183,479	116,678	846,988	152,700	476,160	106,498
JB	843358,5	74720,3	38130,0	41227,6	33799,6	14152,5	723096,3	23386,2	226826,8	11345,8
stud. min	-24,459	-15,066	-12,968	-13,422	-12,636	-5,638	-22,412	-8,646	-14,909	-8,201
q1	-2,541	-3,085	-2,662	-2,744	-2,683	-2,782	-2,927	-3,107	-2,776	-2,842
q5	-1,520	-1,549	-1,531	-1,552	-1,513	-1,508	-1,408	-1,342	-1,380	-1,463
q95	1,518	1,411	1,468	1,495	1,403	1,401	1,367	1,330	1,296	1,538
q99	2,512	2,884	2,417	2,457	2,530	3,391	2,589	3,340	2,451	2,900
stud. max	9,282	11,565	8,349	7,884	8,616	9,093	9,483	8,923	12,826	8,051
TR1	102,844	398,705	321,137	74,957	1938,754	15,488	29,042	85,971	1083,602	216,015
TR5	383,826	571,282	680,060	597,742	2039,764	136,932	101,443	202,643	1191,379	349,652
AR1	0,084	0,013	0,050	0,121	0,154	0,024	0,014	0,245	0,095	0,172
AR2	-0,018	-0,046	-0,055	-0,006	0,002	0,012	-0,017	0,067	0,029	0,002
AR3	-0,016	0,020	-0,003	-0,009	0,018	0,061	0,089	-0,015	0,010	0,000
AR4	-0,023	0,022	0,019	0,035	0,038	0,046	0,013	-0,038	-0,049	0,037
AR5	0,024	-0,013	0,017	0,011	0,007	0,017	-0,006	0,024	0,044	-0,005
Q5	13,739	5,236	9,681	25,132	40,217	2,351	10,860	32,267	14,195	21,026
Q10	7,789	4,361	7,135	15,512	24,741	2,712	6,533	21,045	7,568	12,398
Q20	4,482	3,770	4,468	9,038	14,914	2,704	4,162	13,723	5,493	9,569
QW5	10,019	8,158	15,652	56,260	29,971	3,179	13,178	37,255	9,542	34,989
QW10	20,364	15,862	25,577	63,272	40,852	14,112	15,167	48,261	13,571	44,933
QW20	25,917	32,550	35,282	74,902	51,316	34,200	21,006	60,937	29,113	67,993
Q5 on filt. data	0,006	0,001	0,021	0,008	0,006	0,012	0,007	0,059	0,015	0,004
QW5 on filt. data	0,005	0,003	0,042	0,013	0,010	0,026	0,015	0,104	0,012	0,008

Table 1.a: Descriptive Statistics

	Singapore	South Korea	Taiwan	Thailand	Czech Rep.	Hungary	Poland	Russia	Slovak Rep.	Slovenia
Beg. date	19750102	19750102	19750102	19750502	19940920	19910103	19941004	19940801	19930915	19940104
Nobs	6261	6261	6261	6175	1118	2086	1108	1153	1382	1303
Mean	0,030	0,034	0,056	0,021	-0,055	0,088	0,027	-0,063	-0,038	0,040
Median	0,000	0,000	0,000	0,000	0,000	0,021	0,000	0,000	0,000	0,000
Std	1,379	1,407	1,847	1,400	1,037	1,728	1,994	3,647	2,040	1,567
Min	-30,042	-17,370	-19,656	-10,028	-7,077	-18,034	-10,286	-26,245	-20,573	-9,853
Max	15,867	10,024	19,914	11,350	4,739	13,616	7,893	23,510	29,022	7,465
Sk	-1,740	-0,150	0,106	0,107	-0,426	-1,249	-0,109	-0,137	2,448	-0,396
Sk*	-56,201	-4,853	3,422	3,427	-5,819	-23,287	-1,484	-1,898	37,155	-5,836
Ku	47,726	11,051	9,646	9,742	4,460	19,166	2,509	9,607	49,067	5,344
Ku*	770,859	178,492	155,794	156,268	30,441	178,684	17,047	66,585	372,341	39,378
JB	597382,6	31883,0	24283,5	24431,5	960,5	32470,4	292,8	4437,2	140018,2	1584,7
stud. min	-21,811	-12,372	-10,674	-7,176	-6,769	-10,487	-5,172	-7,179	-10,065	-6,311
q1	-2,608	-2,893	-3,030	-2,983	-3,123	-3,545	-2,709	-3,067	-2,582	-3,249
q5	-1,431	-1,420	-1,536	-1,490	-1,691	-1,230	-1,571	-1,437	-1,241	-1,566
q95	1,431	1,541	1,501	1,429	1,574	1,461	1,640	1,362	1,263	1,512
q99	2,646	2,895	2,769	3,165	2,700	2,546	2,818	3,124	2,642	2,836
stud. max	11,487	7,101	10,753	8,091	4,621	7,828	3,945	6,463	14,244	4,737
TR1	444,812	186,143	1211,394	753,781	83,634	239,224	135,444	154,111	422,541	186,412
TR5	790,634	560,055	1442,228	1033,881	173,877	287,879	166,657	159,672	504,456	213,875
AR1	0,137	0,074	0,019	0,192	0,311	0,113	0,198	0,075	-0,232	0,300
AR2	-0,028	0,006	0,057	0,031	0,159	0,090	0,016	0,107	0,056	-0,040
AR3	0,018	-0,007	0,061	0,042	0,044	-0,028	0,009	0,055	0,255	-0,020
AR4	0,047	-0,018	0,019	0,030	0,010	0,022	0,012	0,005	-0,072	0,009
AR5	0,000	-0,028	0,014	0,011	-0,052	-0,019	-0,044	0,018	0,212	0,017
Q5	27,643	8,438	9,853	50,193	28,516	9,408	9,274	4,761	47,742	24,269
Q10	14,439	5,617	6,816	28,104	17,809	7,843	5,892	5,177	28,218	15,298
Q20	8,089	4,894	5,531	16,671	10,475	7,097	3,565	4,499	17,776	8,365
QW5	15,255	12,612	15,044	49,538	45,292	7,050	19,393	9,131	11,700	40,399
QW10	18,466	15,969	23,041	53,540	58,133	16,597	24,134	24,797	22,851	46,912
QW20	26,291	26,666	35,664	64,541	70,911	25,298	31,292	33,883	34,888	53,472
Q5 on filt. data	0,001	0,008	0,001	0,006	0,138	0,007	0,047	0,056	0,455	0,009
QW5 on filt. data	0,002	0,012	0,003	0,012	0,232	0,008	0,157	0,168	0,240	0,031

Table 1.b: Descriptive Statistics

	Argentina	Brazil	Chile	Colombia	Mexico	Peru	Venezuela
Beg. date	19930803	19920102	19870105	19920103	19880105	19910103	19940103
Nobs	1413	1826	3129	1825	2868	2086	1304
Mean	0,004	0,510	0,083	0,062	0,128	0,190	0,120
Median	0,042	0,307	0,000	0,000	0,029	0,000	0,000
Std	2,284	3,433	0,997	1,140	1,753	1,680	2,174
Min	-14,765	-17,229	-12,304	-5,289	-14,314	-8,796	-10,805
Max	12,072	22,813	6,471	9,918	12,154	8,908	20,062
Sk	-0,547	0,097	-0,363	1,094	0,002	0,315	0,875
Sk*	-8,393	1,700	-8,280	19,071	0,036	5,868	12,904
Ku	4,820	3,057	11,436	9,955	7,925	3,802	12,319
Ku*	36,987	26,669	130,580	86,810	86,633	35,444	90,806
JB	1438,5	714,1	17119,8	7899,7	7505,3	1290,7	8412,2
stud. min	-6,467	-5,167	-12,421	-4,692	-8,239	-5,347	-5,024
q1	-2,801	-2,825	-2,629	-2,473	-2,900	-2,742	-2,800
q5	-1,745	-1,587	-1,434	-1,419	-1,485	-1,396	-1,479
q95	1,404	1,703	1,548	1,612	1,453	1,725	1,378
q99	2,593	2,694	2,860	3,261	2,663	3,010	3,039
stud. max	5,284	6,496	6,405	8,645	6,861	5,188	9,171
TR1	101,939	62,803	48,746	34,223	230,266	234,384	309,788
TR5	196,359	216,283	104,777	79,423	303,194	256,103	314,215
AR1	0,095	0,071	0,291	0,352	0,200	0,383	0,238
AR2	-0,053	0,023	0,108	0,199	-0,009	0,041	0,028
AR3	-0,014	0,004	0,039	0,108	0,009	0,007	-0,036
AR4	0,034	0,017	0,058	0,059	0,069	0,052	0,048
AR5	0,015	0,024	0,070	-0,032	0,033	0,029	0,078
Q5	3,789	2,349	66,726	65,911	26,468	63,480	17,608
Q10	4,456	4,026	37,833	33,376	14,391	37,194	10,706
Q20	3,177	3,618	21,599	17,675	8,212	22,196	5,824
QW5	8,854	5,284	105,429	92,036	40,579	112,915	12,351
QW10	15,961	23,422	118,799	94,149	45,545	124,926	22,726
QW20	24,602	36,775	130,638	99,817	54,431	130,132	36,084
Q5 on filt. data	0,002	0,022	0,095	0,012	0,041	0,046	0,083
QW5 on filt. data	0,006	0,062	0,231	0,028	0,100	0,190	0,271

Table 1.c: Descriptive Statistics

CAC40		Russia	
19810513	-13,91	19971028	-26,245
19871019	-10,14	19980827	-21,338
19871028	-9,11	19980917	-19,199
19871026	-8,45	19980826	-17,679
19871110	-7,61	19971003	-16,041
19871029	6,39	19980715	15,848
19690811	6,48	19940815	16,605
19910117	6,81	19960618	16,783
19780313	7,40	19971029	21,874
19871112	8,23	19971006	23,510

Singapore		Mexico	
19871020	-30,04	19971027	-14,314
19871019	-15,15	19880106	-10,526
19871023	-14,07	19880309	-10,448
19851205	-11,11	19980910	-10,341
19891016	-9,73	19880315	-10,120
19750131	8,30	19950131	9,779
19980113	8,77	19971028	11,056
19750214	10,62	19880229	11,690
19750128	10,93	19880118	11,709
19871022	15,87	19980915	12,154

Table 2: Most extreme returns and date of occurrence

	Left tail of distribution				Right tail of distribution			
	ξ	STD	t*	threshold	ξ	STD	t*	threshold
S&P	0,3115	0,0180	301	-1,54	0,2768	0,0233	141	2,00
Nikkei	0,2805	0,0276	103	-3,00	0,3576	0,0308	135	2,43
DAX	0,3388	0,0183	342	-1,70	0,3070	0,0230	178	1,99
CAC40	0,3337	0,0214	244	-1,93	0,2770	0,0227	149	2,19
FTSE100	0,2994	0,0241	154	-2,18	0,3333	0,0284	138	2,12
MEAN	0,3128	0,0219	229	-2,07	0,3103	0,0256	148	2,14
STD	0,0242	0,0041	100	0,57	0,0354	0,0037	17	0,18
China	0,3062	0,0497	38	-6,48	0,4379	0,0762	33	7,92
Hong Kong	0,3775	0,0306	152	-3,61	0,3625	0,0286	161	3,23
Indonesia	0,3641	0,0607	36	-3,75	0,4682	0,0690	46	3,34
Malaisia	0,4100	0,0386	113	-3,11	0,4063	0,0251	263	2,06
Philippines	0,3928	0,0369	113	-3,30	0,3504	0,0331	112	3,52
Singapore	0,3716	0,0255	212	-2,27	0,3393	0,0333	104	2,95
South Korea	0,3742	0,0322	135	-2,92	0,3611	0,0266	184	2,73
Taiwan	0,2662	0,0242	121	-4,43	0,2971	0,0293	103	4,43
Thailand	0,3405	0,0313	118	-3,31	0,3392	0,0354	92	3,72
MEAN	0,3559	0,0366	115	-3,69	0,3736	0,0396	122	3,77
STD	0,0449	0,0118	54	1,20	0,0538	0,0191	71	1,69
Czech Rep.	0,2310	0,0530	19	-2,68	0,2233	0,0499	20	2,33
Hungary	0,4065	0,0742	30	-4,84	0,3442	0,0420	67	3,01
Poland	0,2746	0,0614	20	-4,52	0,2946	0,0529	31	3,91
Russia	0,3067	0,0970	10	-12,50	0,3326	0,0807	17	10,39
Slovak Rep.	0,4013	0,0518	60	-2,67	0,3928	0,0655	36	3,52
Slovenia	0,3521	0,0622	32	-3,14	0,1943	0,0486	16	3,95
MEAN	0,3287	0,0666	29	-5,06	0,2970	0,0566	31	4,52
STD	0,0705	0,0169	17	3,76	0,0757	0,0141	19	2,94
Argentina	0,2657	0,0415	41	-4,92	0,2797	0,0583	23	5,44
Brasil	0,2541	0,0464	30	-7,96	0,2827	0,0336	71	6,22
Chile	0,2644	0,0333	63	-2,10	0,3015	0,0455	44	2,43
Colombia	0,3487	0,0532	43	-2,00	0,4259	0,0615	48	2,21
Mexico	0,3271	0,0449	53	-3,81	0,3228	0,0471	47	4,01
Peru	0,2203	0,0519	18	-4,63	0,2856	0,0457	39	3,79
Venezuela	0,3033	0,0536	32	-4,31	0,3763	0,0753	25	4,65
MEAN	0,2834	0,0464	40	-4,25	0,3249	0,052	42	4,11
STD	0,0449	0,0074	15	2,01	0,0559	0,014	16	1,48

Table 3: Optimal Hill estimates

Moment	Left tail				Right tail			
	1st	2nd	3rd	4th	1st	2nd	3rd	4th
S&P	YES	YES	YES	NO	YES	YES	YES	NO
Nikkei	YES	YES	A	NO	YES	YES	NO	NO
DAX	YES	YES	A	NO	YES	YES	YES	NO
CAC40	YES	YES	A	NO	YES	YES	YES	NO
FTSE100	YES	YES	YES	NO	YES	YES	YES	NO
China	YES	YES	YES	NO	YES	YES	A	NO
Hong Kong	YES	YES	NO	NO	YES	YES	A	NO
Indonesia	YES	YES	NO	NO	YES	A	NO	NO
Malaisia	YES	YES	NO	NO	YES	YES	NO	NO
Philippines	YES	YES	NO	NO	YES	YES	NO	NO
Singapore	YES	YES	NO	NO	YES	YES	NO	NO
South Korea	YES	YES	NO	NO	YES	YES	NO	NO
Taiwan	YES	YES	YES	NO	YES	YES	NO	NO
Thailand	YES	YES	NO	NO	YES	NO	NO	NO
Czech Rep.	YES	YES	YES	YES	YES	YES	YES	NO
Hungary	YES	A	NO	NO	YES	YES	NO	NO
Poland	YES	YES	YES	NO	YES	YES	A	NO
Russia	YES	YES	A	NO	YES	YES	NO	NO
Slovak Rep.	YES	YES	NO	NO	YES	YES	NO	NO
Slovenia	YES	YES	A	NO	YES	YES	YES	NO
Argentina	YES	YES	YES	NO	YES	YES	YES	NO
Chile	YES	YES	YES	NO	YES	YES	YES	NO
Mexico	YES	YES	A	NO	YES	YES	YES	NO
Peru	YES	YES	YES	NO	YES	YES	YES	NO
Venezuela	YES	YES	YES	NO	YES	YES	NO	NO
Brasil	YES	YES	YES	NO	YES	YES	YES	YES
Colombia	YES	YES	NO	NO	YES	YES	NO	NO

Table 4: Existence of left and right integral for moment computation

	Left tail of returns				Right tail of returns			
	ψ	STD	Lik.	LRT	ψ	STD	Lik.	LRT
S&P	0,4710	0,0338	-174,01	0,06	0,5481	0,0569	-97,63	0,04
Nikkei	0,9090	0,0771	-169,70	6,95	0,8943	0,0890	-152,98	3,23
DAX	0,5983	0,0471	-239,69	0,24	0,6084	0,0567	-144,48	0,08
CAC40	0,6686	0,0564	-206,82	0,18	0,5968	0,0649	-116,85	0,39
FTSE100	0,6538	0,0694	-132,45	0,11	0,6783	0,0721	-143,39	0,10
MEAN	0,6601	0,0568	-184,53	1,51	0,6652	0,0679	-131,06	0,77
STD	0,1594	0,0173	40,57	3,04	0,1363	0,0134	23,09	1,38
China	2,2348	0,3507	-69,25	3,27	2,7475	0,7785	-66,38	0,36
Hong Kong	1,4007	0,1357	-270,95	0,94	1,2864	0,1417	-176,44	1,41
Indonesia	1,4711	0,2370	-59,59	3,32	1,6602	0,2525	-104,56	6,81
Malaisia	1,1936	0,1526	-203,50	0,29	0,8807	0,0777	-313,72	0,01
Philippines	1,3315	0,1551	-172,98	2,17	1,2187	0,1485	-176,61	0,19
Singapore	0,8874	0,0814	-230,96	0,07	0,9673	0,1198	-141,16	0,08
South Korea	1,1878	0,1245	-186,66	3,61	1,0362	0,0954	-227,34	3,61
Taiwan	1,2134	0,1133	-188,64	3,07	1,3179	0,1602	-157,58	0,08
Thailand	1,3073	0,1353	-158,76	9,94	1,3430	0,1455	-146,50	8,70
MEAN	1,3586	0,1651	-171,26	2,96	1,3842	0,2133	-167,81	2,36
STD	0,3683	0,0814	69,02	2,95	0,5619	0,2176	71,19	3,30
Czech Rep.	0,5832	0,2092	-10,53	0,19	0,6245	0,1090	-15,90	2,69
Hungary	2,2200	0,3554	-88,75	4,99	0,9589	0,1694	-92,21	0,32
Poland	1,2195	0,2912	-30,13	0,98	1,2173	0,2140	-47,87	2,84
Russia	3,9507	1,0971	-34,23	1,38	3,8757	0,8504	-47,19	2,12
Slovak Rep.	0,8938	0,1203	-124,24	0,33	1,3305	0,4306	-57,46	2,53
Slovenia	1,1359	0,2712	-36,81	0,79	0,9558	0,1756	-22,81	2,45
MEAN	1,6672	0,3907	-54,11	1,44	1,4938	0,3248	-47,24	2,16
STD	1,2471	0,3551	43,10	1,79	1,1923	0,2803	27,22	0,93
Argentina	1,3174	0,2594	-61,54	0,01	1,7929	0,2560	-33,26	3,17
Brasil	2,0153	0,3250	-70,82	2,06	1,7445	0,3265	-71,33	0,02
Chile	0,5598	0,0831	-45,22	0,00	0,7633	0,1325	-43,36	1,95
Colombia	0,7324	0,1395	-41,91	1,10	1,0060	0,1846	-55,28	1,30
Mexico	1,2962	0,2099	-80,49	1,06	1,3124	0,2665	-73,49	0,01
Peru	1,0897	0,2032	-38,91	2,27	1,1966	0,1696	-63,48	4,89
Venezuela	1,1270	0,2807	-51,99	0,69	1,6622	0,4547	-48,79	0,02
MEAN	1,1625	0,2144	-55,84	1,03	1,3540	0,2558	-55,57	1,62
STD	0,4697	0,0833	15,67	0,90	0,3951	0,1098	14,86	1,87

Table 5: GPD on filtered returns with optimal t^* (from Hill-bootstrap)

	Left tail of returns			Right tail of returns		
	x999	x9999	Sample min	x999	x9999	Sample max
S&P	-4,78	-9,80	-22,83	4,23	8,21	8,71
Nikkei	-7,03	-15,29	-16,14	6,69	14,64	12,43
DAX	-5,99	-12,89	-13,71	5,29	11,18	8,87
CAC40	-6,13	-13,27	-13,91	5,56	11,82	8,23
FTSE100	-5,30	-10,55	-13,03	5,20	10,50	8,94
MEAN	-5,85	-12,36	-15,92	5,40	11,27	9,44
STD	0,86	2,21	4,03	0,88	2,32	1,70
China	-17,72	-36,98	-17,90	19,08	37,20	28,86
Hong Kong	-12,54	-30,64	-40,54	11,24	27,12	17,25
Indonesia	-10,40	-24,17	-12,73	11,31	27,34	13,13
Malaisia	-11,01	-28,01	-24,15	13,42	35,51	20,82
Philippines	-12,81	-30,58	-15,79	13,25	31,50	15,66
Singapore	-8,38	-19,62	-30,04	7,45	16,71	15,87
South Korea	-9,55	-22,84	-17,37	10,56	25,44	10,02
Taiwan	-10,07	-19,04	-19,66	9,49	17,82	19,91
Thailand	-9,70	-22,17	-10,03	9,65	21,73	11,35
MEAN	-11,35	-26,00	-20,91	11,72	26,71	16,98
STD	2,77	5,96	9,45	3,33	7,22	5,73
Czech Rep.	-4,93	-8,06	-7,08	4,63	8,04	4,74
Hungary	-19,09	-62,08	-18,03	24,02	79,45	13,62
Poland	-10,01	-18,99	-10,29	10,69	20,89	7,89
Russia	-25,22	-56,83	-26,25	27,25	62,57	23,51
Slovak Rep.	-12,54	-32,65	-20,57	9,36	22,05	29,02
Slovenia	-9,21	-19,58	-9,85	8,56	18,07	7,46
MEAN	-13,50	-33,03	-15,34	14,09	35,18	14,37
STD	6,75	20,05	6,80	8,42	26,19	8,96
Argentina	-12,02	-22,20	-14,76	10,51	18,86	12,07
Brasil	-16,69	-30,95	-17,23	17,25	31,81	22,81
Chile	-4,75	-9,05	-12,30	4,63	8,50	6,47
Colombia	-6,18	-13,91	-5,29	6,48	14,15	9,92
Mexico	-10,08	-21,57	-14,31	9,88	20,92	12,15
Peru	-7,99	-14,98	-8,80	8,70	16,47	8,91
Venezuela	-10,53	-20,17	-10,80	9,97	18,75	20,06
MEAN	-9,75	-18,98	-11,93	9,63	18,49	13,20
STD	3,98	7,10	4,02	3,98	7,15	6,00

Table 6: High quantiles estimated

	Left tail							Right tail						
	μ	STD	ψ	STD	ξ	STD	Lik.	μ	STD	ψ	STD	ξ	STD	Lik.
S&P	1,1746	0,0285	0,5111	0,0215	0,1892	0,0387	-404,57	1,2468	0,0317	0,5435	0,0233	0,1517	0,0399	-412,43
Nikkei	1,1299	0,0432	0,7222	0,0376	0,3272	0,0411	-563,21	1,1632	0,0393	0,6434	0,0324	0,3703	0,0426	-527,16
DAX	1,2681	0,0341	0,5820	0,0261	0,2748	0,0450	-466,54	1,3147	0,0318	0,5603	0,0261	0,2233	0,0416	-440,51
CAC40	1,2444	0,0380	0,6369	0,0287	0,2542	0,0457	-497,14	1,3324	0,0358	0,6544	0,0279	0,1136	0,0359	-475,74
FTSE100	1,3047	0,0290	0,5451	0,0267	0,1983	0,0396	-424,06	1,2767	0,0250	0,4556	0,0227	0,2710	0,0413	-370,88
MEAN	1,2243	0,0346	0,5995	0,0281	0,2487	0,0420	-471,11	1,2668	0,0327	0,5714	0,0265	0,2260	0,0403	-445,35
STD	0,0710	0,0062	0,0830	0,0059	0,0569	0,0032	62,89	0,0668	0,0053	0,0812	0,0039	0,1013	0,0026	59,73
China	3,5008	0,2737	1,9383	0,2029	0,2749	0,0909	-206,14	3,2827	0,2847	2,0503	0,2491	0,5211	0,1004	-222,57
Hong Kong	1,8625	0,0717	1,0071	0,0575	0,3885	0,0517	-570,67	2,0535	0,0624	0,9801	0,0529	0,2938	0,0447	-546,40
Indonesia	1,1933	0,0892	0,8423	0,0797	0,4396	0,0716	-197,31	1,3051	0,0907	0,8827	0,0828	0,4653	0,0720	-204,84
Malaisia	1,5192	0,0668	0,9306	0,0570	0,3382	0,0561	-420,02	1,7084	0,0603	0,8664	0,0520	0,3187	0,0545	-399,70
Philippines	2,0614	0,0964	1,1383	0,0860	0,3026	0,0560	-318,23	2,2731	0,1123	1,2857	0,0891	0,2637	0,0594	-334,76
Singapore	1,5243	0,0524	0,7640	0,0391	0,2850	0,0457	-466,95	1,6464	0,0477	0,7729	0,0389	0,2455	0,0425	-456,00
South Korea	1,3959	0,0585	0,8560	0,0461	0,3338	0,0495	-503,24	1,7141	0,0652	0,9499	0,0452	0,2053	0,0460	-512,93
Taiwan	2,1648	0,0793	1,1658	0,0603	0,2092	0,0476	-577,78	2,2774	0,0749	1,1969	0,0582	0,1879	0,0420	-581,93
Thailand	1,1769	0,0622	0,8397	0,0534	0,4711	0,0489	-513,56	1,3333	0,0740	1,0035	0,0555	0,3220	0,0473	-543,49
MEAN	1,8221	0,0945	1,0536	0,0758	0,3381	0,0576	-419,32	1,9549	0,0969	1,1098	0,0804	0,3137	0,0565	-422,51
STD	0,7218	0,0687	0,3594	0,0499	0,0831	0,0147	146,27	0,6124	0,0729	0,3875	0,0654	0,1125	0,0191	141,54

Table 7.a: Parameter estimates of the gev distribution (20-histories)

	Left tail							Right tail						
	μ	STD	ψ	STD	ξ	STD	Lik.	μ	STD	ψ	STD	ξ	STD	Lik.
Czech Rep.	1,2007	0,1206	0,7445	0,0798	0,1416	0,0976	-75,08	1,2930	0,1235	0,7057	0,0717	0,0065	0,1234	-67,98
Hungary	1,6541	0,1089	1,0314	0,1120	0,4487	0,0803	-194,56	1,5905	0,1208	1,0057	0,0914	0,3295	0,0908	-184,19
Poland	2,4589	0,2014	1,3056	0,1490	0,1298	0,0873	-105,65	2,6681	0,1931	1,2200	0,1428	0,1468	0,0927	-102,47
Russia	3,5729	0,3834	2,5851	0,3327	0,3329	0,0850	-155,06	4,0370	0,4550	2,7855	0,3766	0,3755	0,0941	-160,51
Slovak Rep.	2,0056	0,1437	0,9973	0,1129	0,3632	0,1036	-121,22	1,9025	0,1322	0,9645	0,0976	0,2617	0,1160	-114,84
Slovenia	2,1934	0,1431	1,1029	0,1131	0,0752	0,0680	-109,89	1,9895	0,1558	1,0219	0,1216	0,0768	0,0968	-105,51
MEAN	2,1809	0,1835	1,2945	0,1499	0,2486	0,0870	-126,91	2,2468	0,1967	1,2839	0,1503	0,1995	0,1023	-122,58
STD	0,8104	0,1030	0,6577	0,0922	0,1523	0,0126	42,02	0,9907	0,1294	0,7538	0,1136	0,1462	0,0138	42,36
Argentina	2,9254	0,2515	1,7467	0,1706	0,1615	0,0888	-156,02	2,5518	0,1465	1,0514	0,1314	0,3555	0,0875	-128,27
Brasil	4,1765	0,2843	2,1256	0,1924	0,1717	0,0835	-221,30	4,1307	0,3544	2,4530	0,2223	0,1115	0,1133	-231,14
Chile	1,1352	0,0518	0,5770	0,0435	0,2364	0,0703	-181,47	1,1663	0,0640	0,6355	0,0484	0,2687	0,0663	-199,48
Colombia	1,2858	0,0786	0,6224	0,0581	0,2330	0,0797	-112,66	1,3993	0,1016	0,8460	0,0772	0,2768	0,0756	-142,85
Mexico	1,9800	0,1094	1,0153	0,0805	0,2949	0,0780	-251,81	2,1945	0,0937	1,0411	0,0891	0,2518	0,0668	-252,35
Peru	1,9487	0,1381	1,0301	0,1014	0,1940	0,0901	-178,97	2,0600	0,1544	1,1561	0,1126	0,2031	0,0857	-191,50
Venezuela	2,3301	0,2464	1,3645	0,1688	0,2367	0,1273	-129,55	2,4183	0,2066	1,3668	0,1563	0,2333	0,0969	-129,60
MEAN	2,2545	0,1657	1,2117	0,1165	0,2183	0,0882	-175,97	2,2744	0,1602	1,2214	0,1196	0,2430	0,0846	-182,17
STD	1,0416	0,0935	0,5725	0,0601	0,0461	0,0185	49,00	0,9652	0,0976	0,5899	0,0576	0,0747	0,0169	49,88

Table 7.b: Parameter estimates of the gev distribution (20-histories)

	Early US index	Recent S&P	Recent Nikkei	Recent DAX	Recent CAC40	Recent FTSE100
Beg. date	18850217	19930402	19930402	19930402	19930402	19930402
Ending date	19110302	19981231	19981231	19981231	19981231	19981231
Nobs	7826	1500	1500	1500	1500	1500
Mean	0,0135	0,0670	0,0655	-0,0202	-0,0215	0,0731
Median	0,0380	0,0469	0,0534	0,0000	0,0000	0,0890
Std	0,8704	0,8534	0,8702	1,1304	1,3782	1,2114
min	-8,5161	-7,1127	-7,4549	-5,3389	-5,9571	-8,3822
max	6,6109	4,9887	4,8605	6,5993	7,6605	6,1057
Sk	-0,3004	-0,6576	-0,7409	0,2219	0,1717	-0,7129
Sk*	-10,8502	-10,3979	-11,7148	3,5087	2,7153	-11,2726
Ku	7,0423	9,4405	8,4680	4,2370	3,1920	4,6815
Ku*	127,1677	74,6340	66,9453	33,4962	25,2351	37,0101
JB	16289,35	5678,34	4618,91	1134,30	644,18	1496,82
TR1	111,23	100,44	91,06	30,21	29,32	154,73
TR5	1276,44	124,81	128,13	101,56	80,51	215,66

Table 8: Summary statistics for the early days of a US index and recent mature markets indices

	ξ	Optimal Hill estimation			gpd estimation	
		STD	t*	threshold	ψ	STD
Left tail of distribution						
Early US index	0,3401	0,0182	351	-0,0265	0,4975	0,0306
Recent S&P	0,3318	0,0489	46	-1,6551	0,5624	0,1061
Recent Nikkei	0,2851	0,0399	51	-2,5356	0,7653	0,1240
Recent DAX	0,3242	0,0725	20	-3,4210	1,1114	0,2433
Recent CAC40	0,2787	0,0581	23	-2,8126	0,8541	0,1580
Recent FTSE100	0,2446	0,0471	27	-2,0318	0,5595	0,0855
Right tail of distribution						
Early US index	0,4181	0,0188	494	0,0204	0,4136	0,0218
Recent S&P	0,3039	0,0493	38	1,6185	0,4972	0,1064
Recent Nikkei	0,2763	0,0532	27	3,1945	0,8944	0,2135
Recent DAX	0,2855	0,0404	50	1,9540	0,5889	0,0932
Recent CAC40	0,2499	0,0361	48	2,1798	0,5197	0,0952
Recent FTSE100	0,3162	0,0294	116	1,1162	0,3624	0,0417

Table 9: Results for the early days of a US index and recent mature markets indices

Figure 1: Plot of the Gumbel, Frechet, and Weibull densities

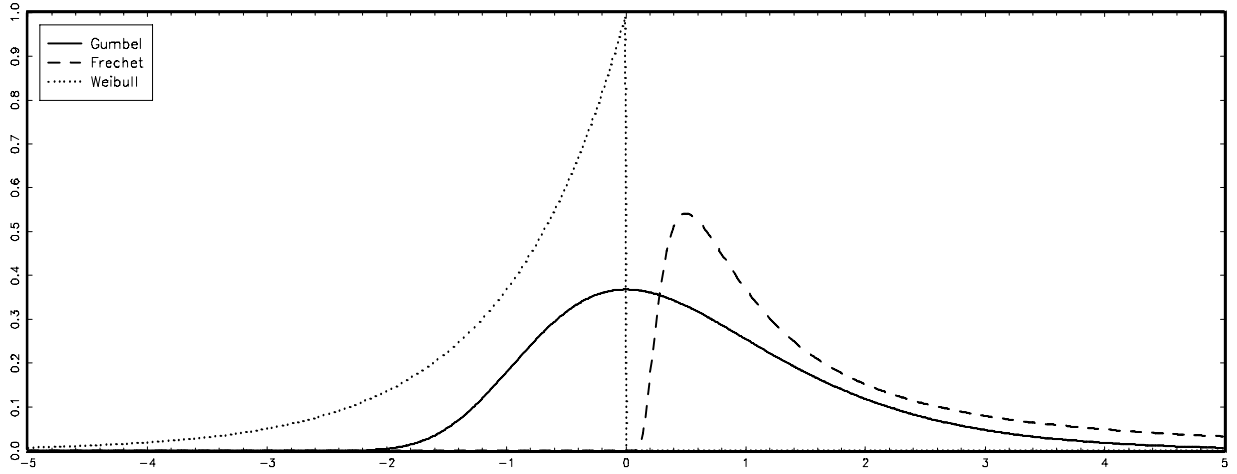


Figure 2: Plot of various Mean-Excess Functions

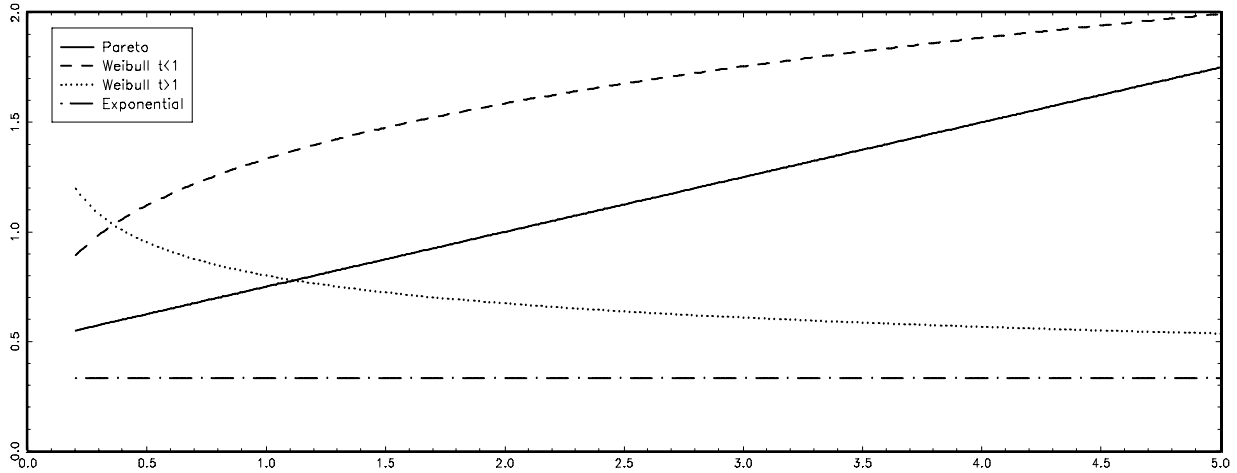
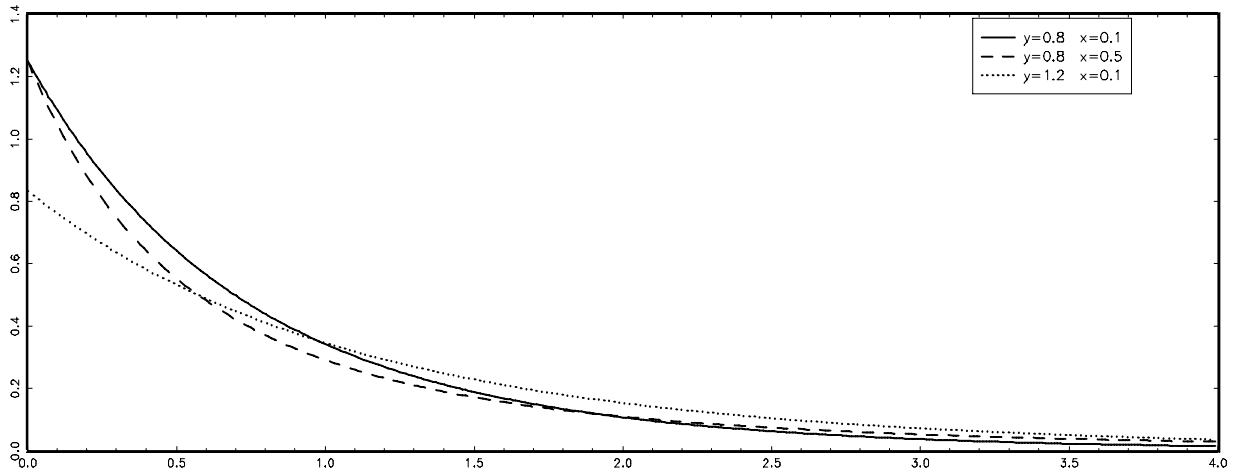
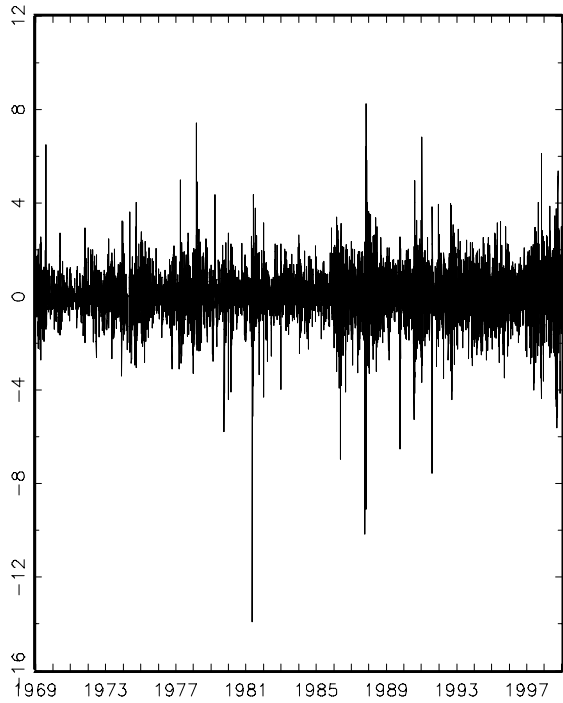


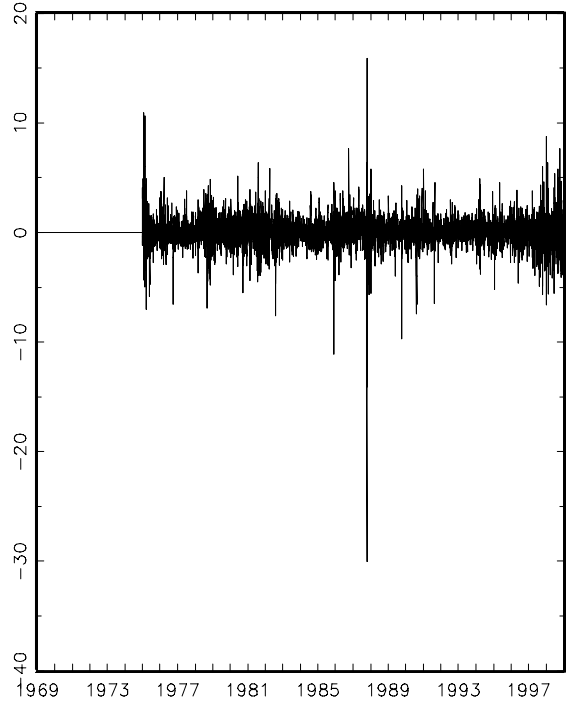
Figure 3: Plot of various Generalized Pareto Distributions



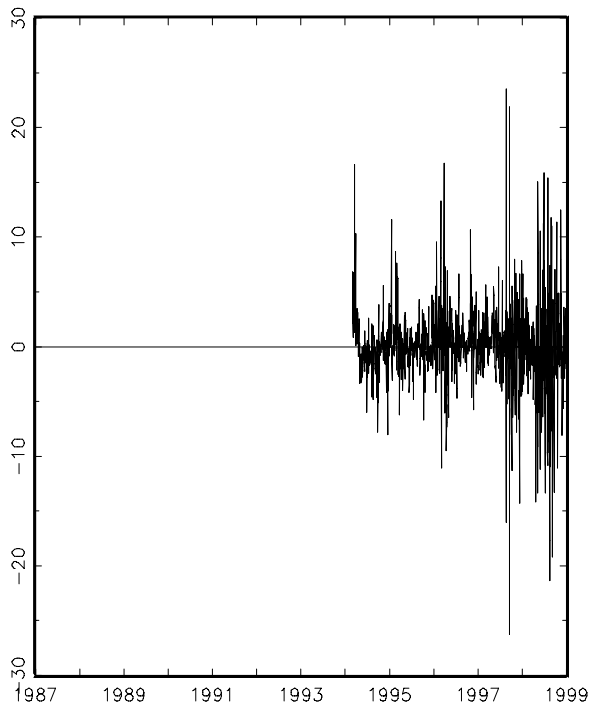
Plot of CAC40 returns



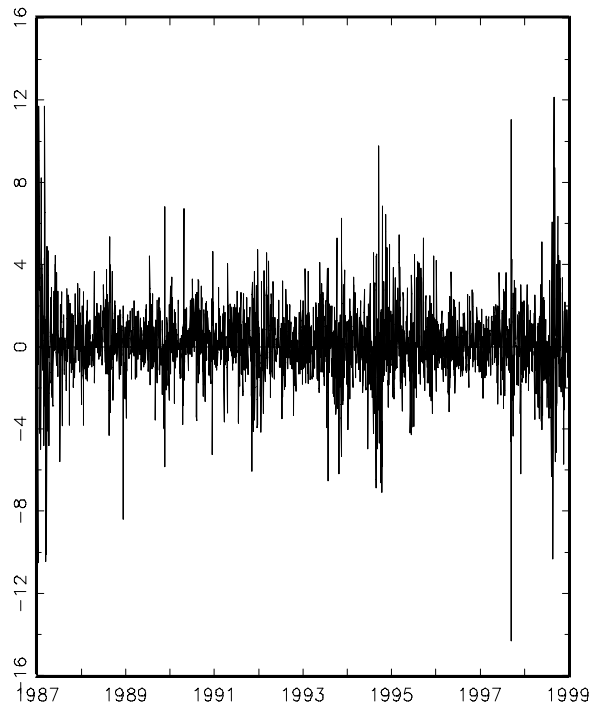
Plot of Singapore returns



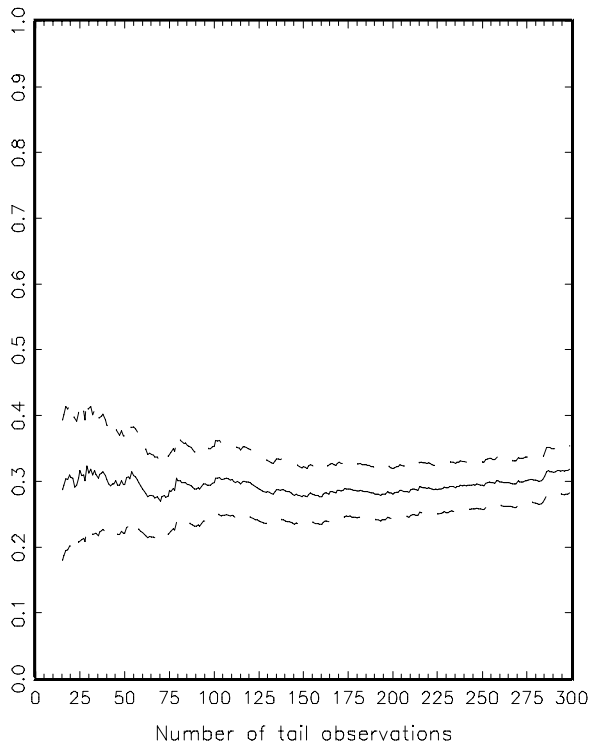
Plot of Russian returns



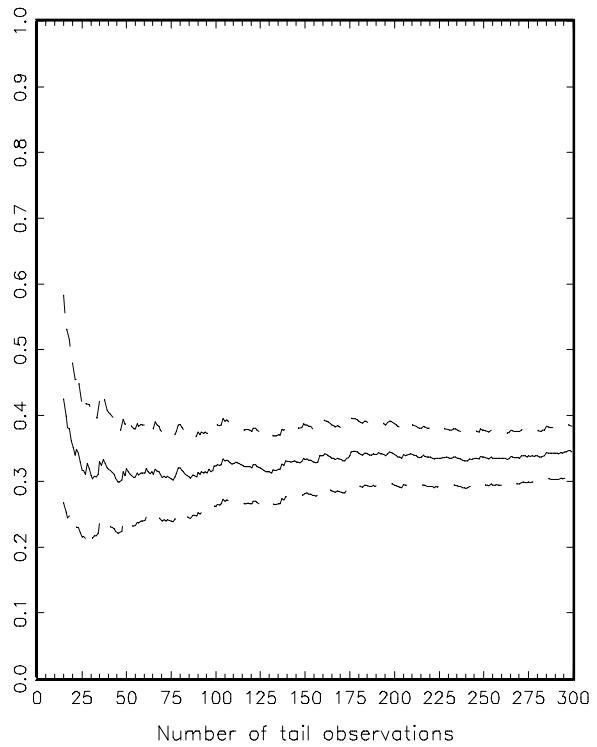
Plot of Mexican returns



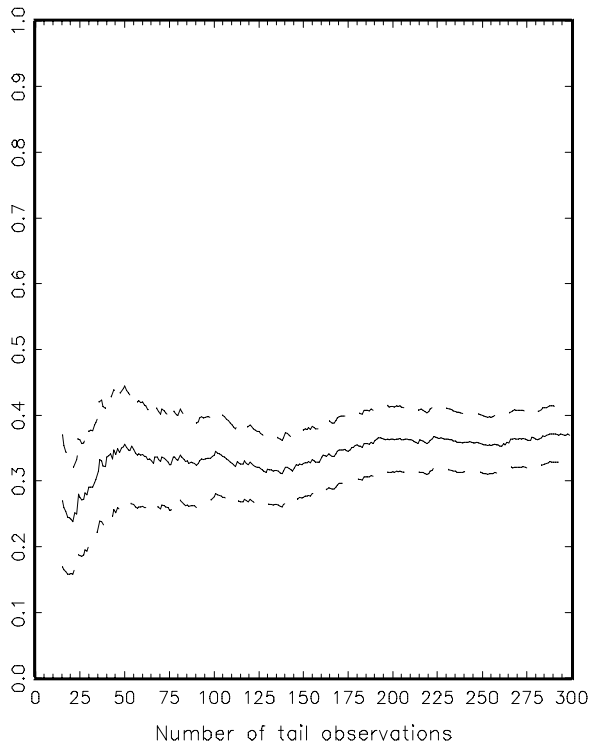
Hill-plot for left tail of CAC40 returns



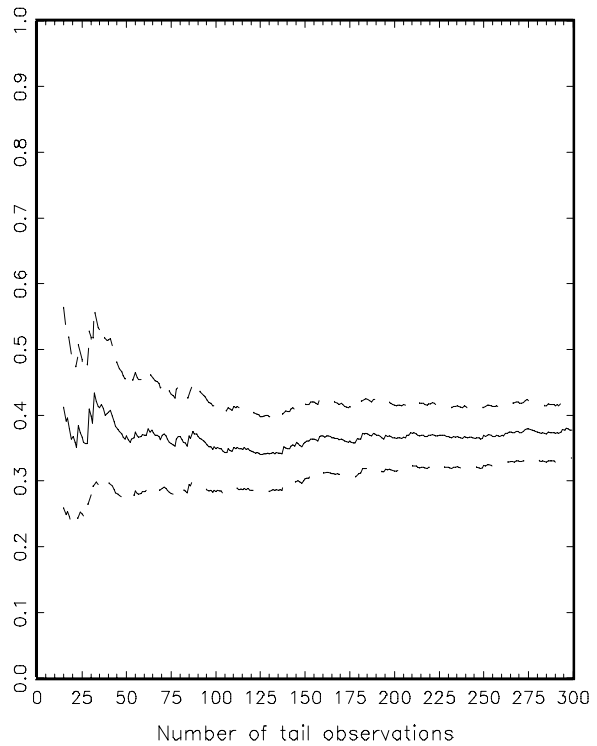
Hill-plot for right tail of CAC40 returns



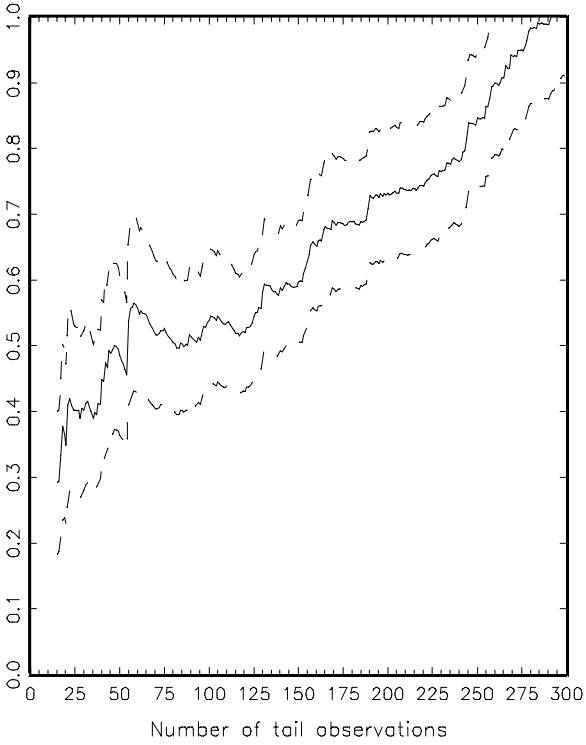
Hill-plot for left tail of returns of Singapore



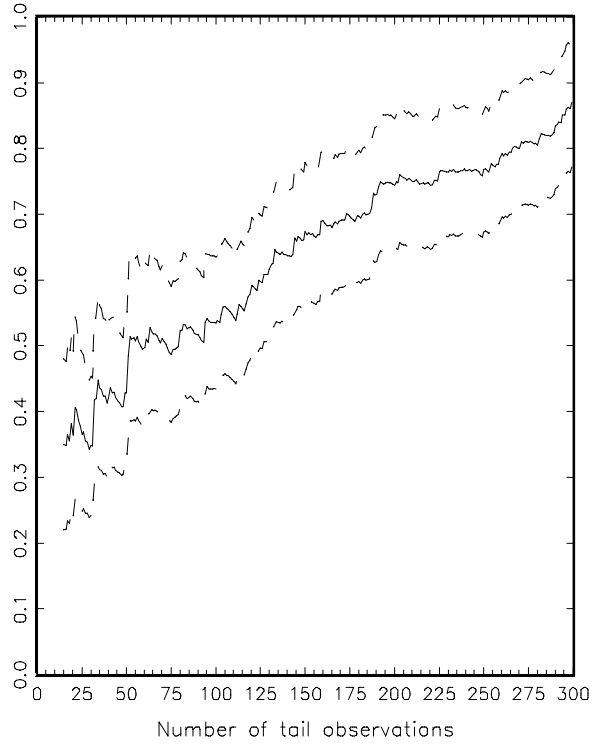
Hill-plot for right tail of returns of Singapore



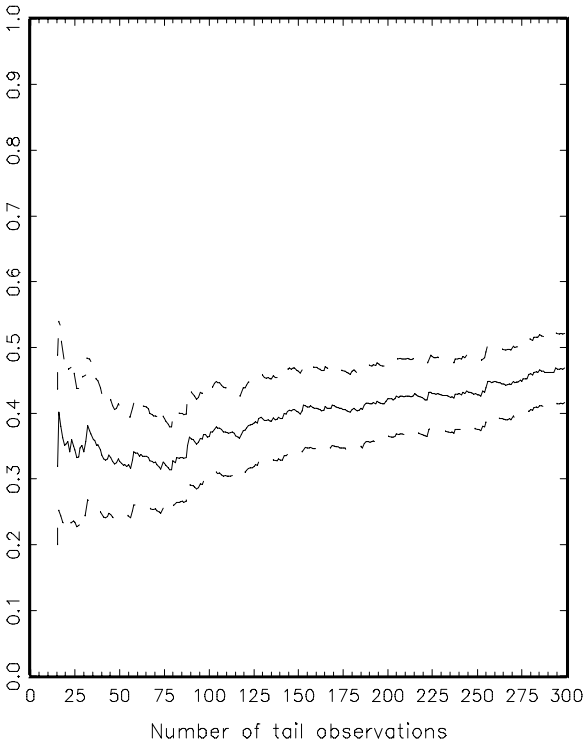
Hill-plot for left tail of Russian returns



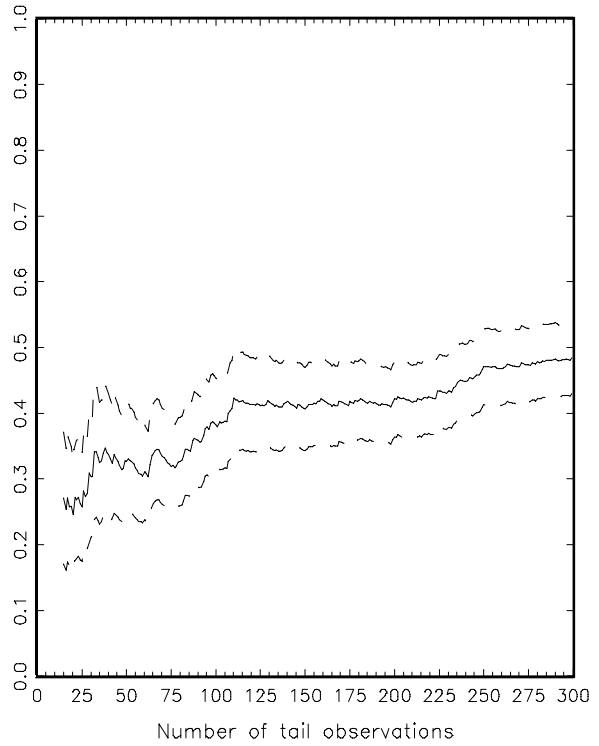
Hill-plot for right tail of Russian returns



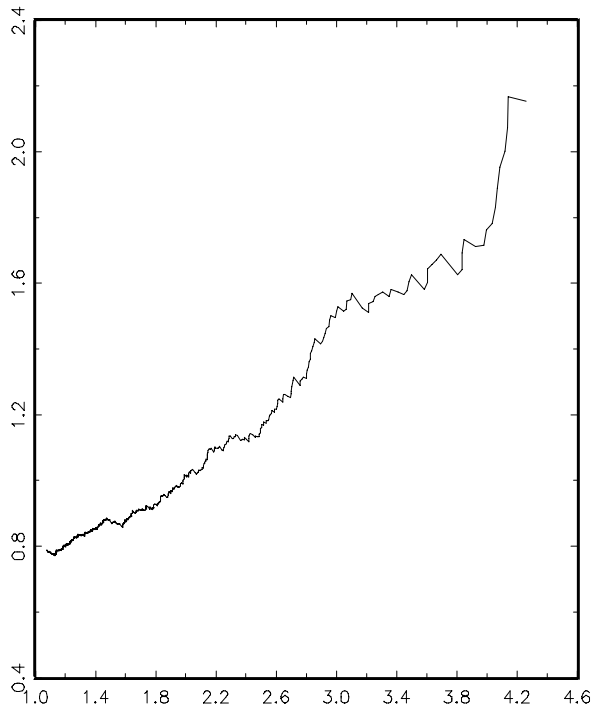
Hill-plot for left tail of Mexican returns



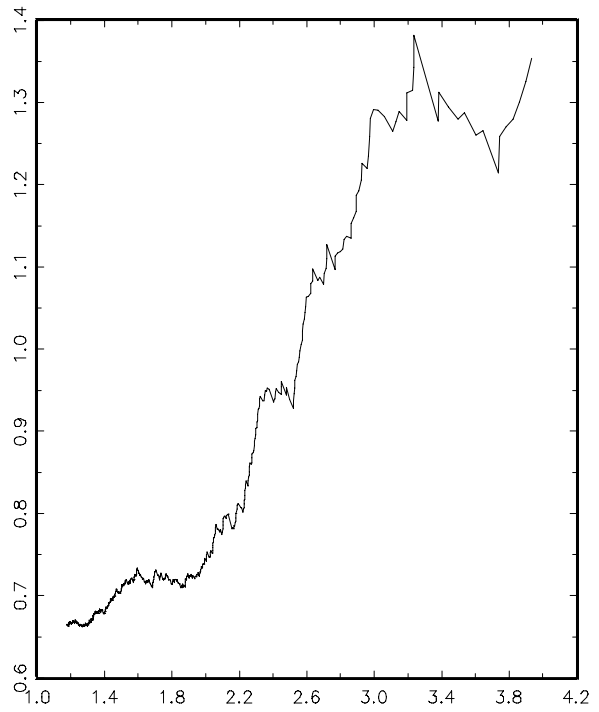
Hill-plot for right tail of Mexican returns



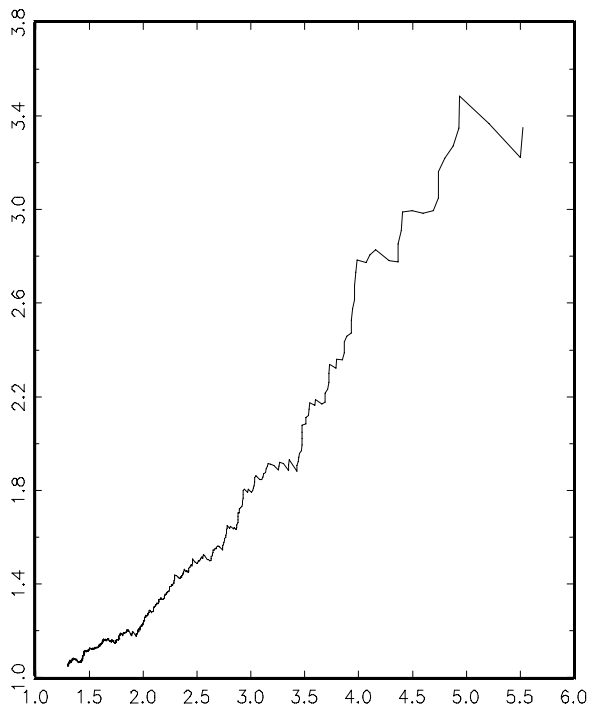
Mean-excess function for 10% of absolute values of smallest returns CAC40



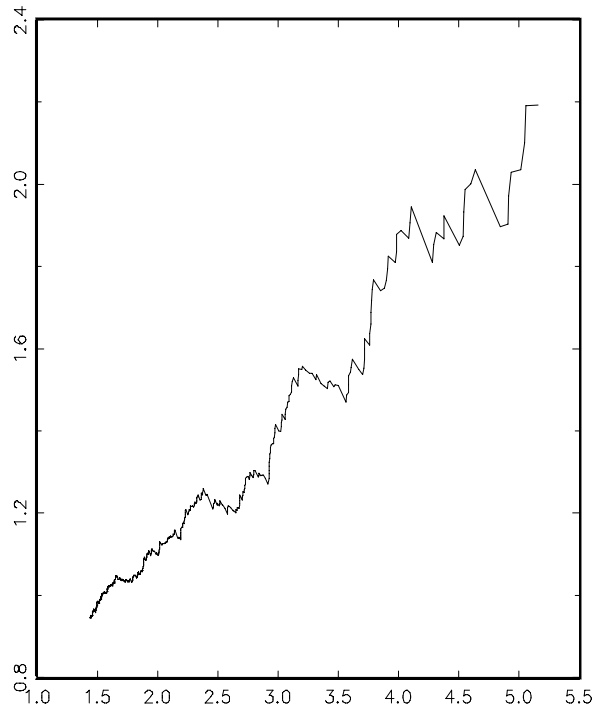
Mean-excess function for 10% of largest returns CAC40



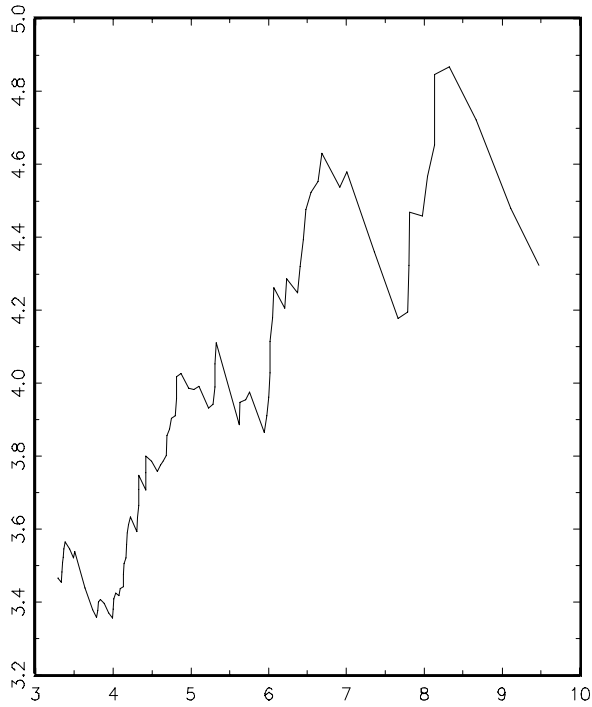
Mean-excess function for 10% of absolute values of smallest returns Singapore



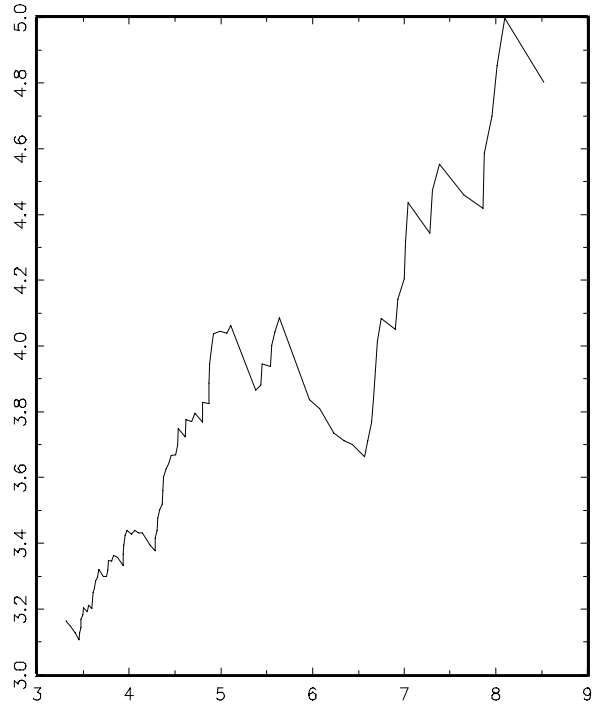
Mean-excess function for 10% of largest returns Singapore



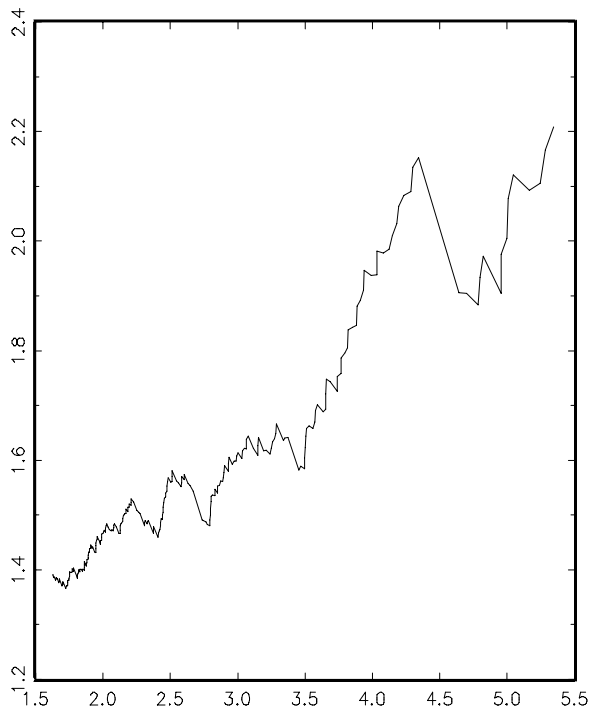
Mean-excess function for 10% of absolute values of smallest returns Russia



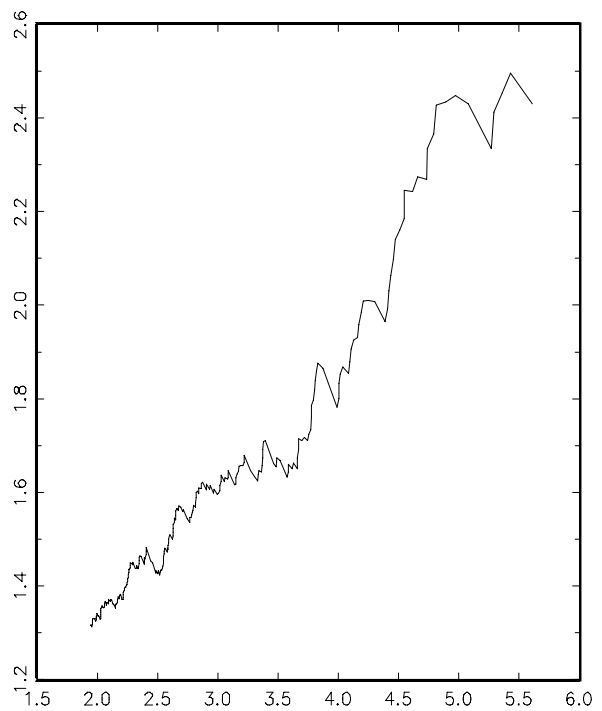
Mean-excess function for 10% of largest returns Russia



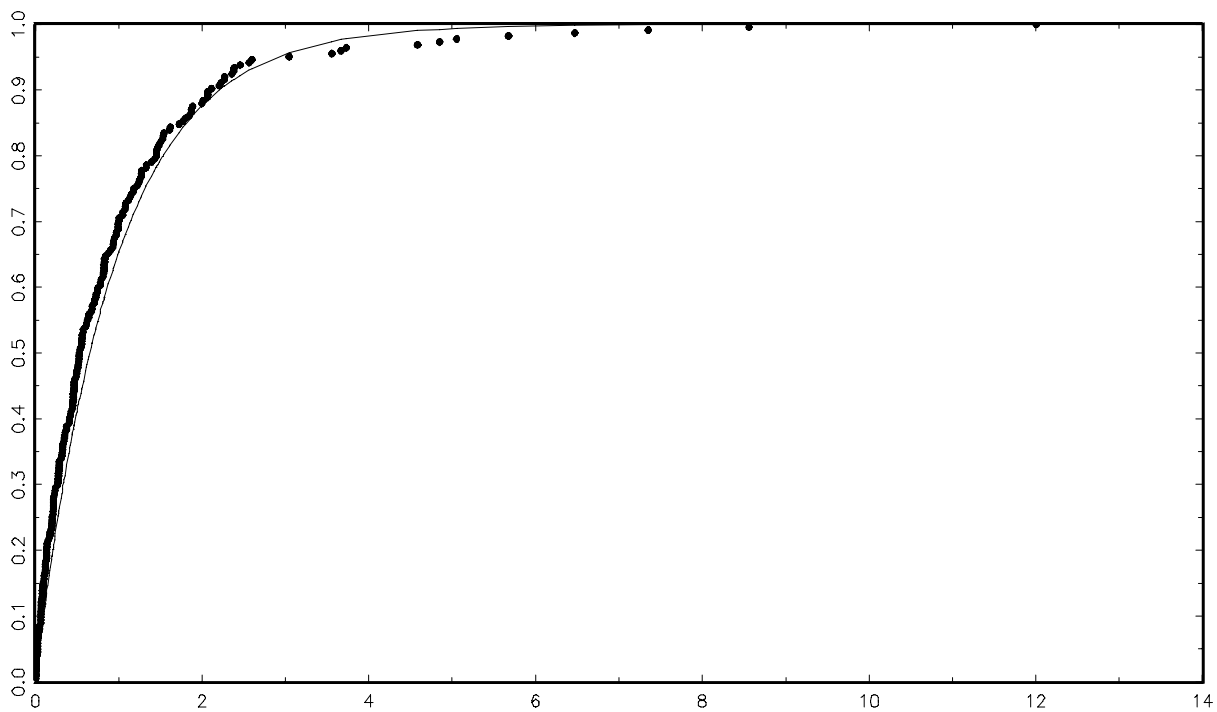
Mean-excess function for 10% of absolute values of smallest returns Mexico



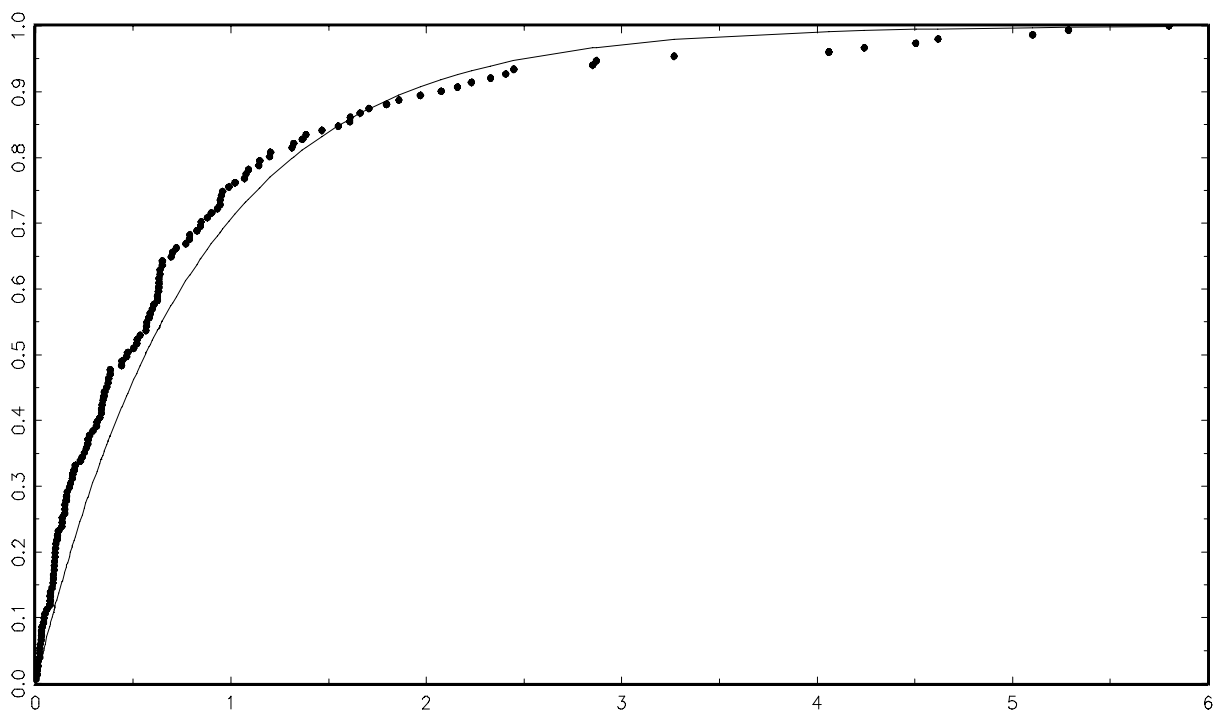
Mean-excess function for 10% of largest returns Mexico



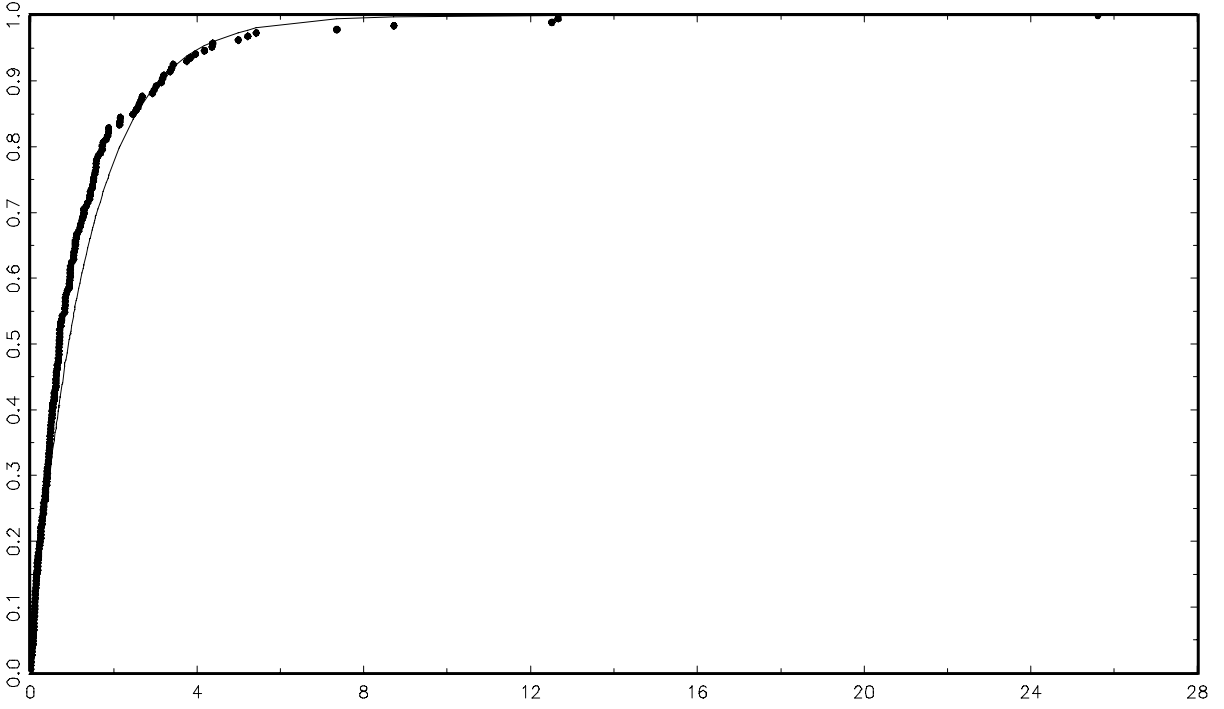
Comparison of tail observations and fitted GPD
Left tail for CAC40 returns



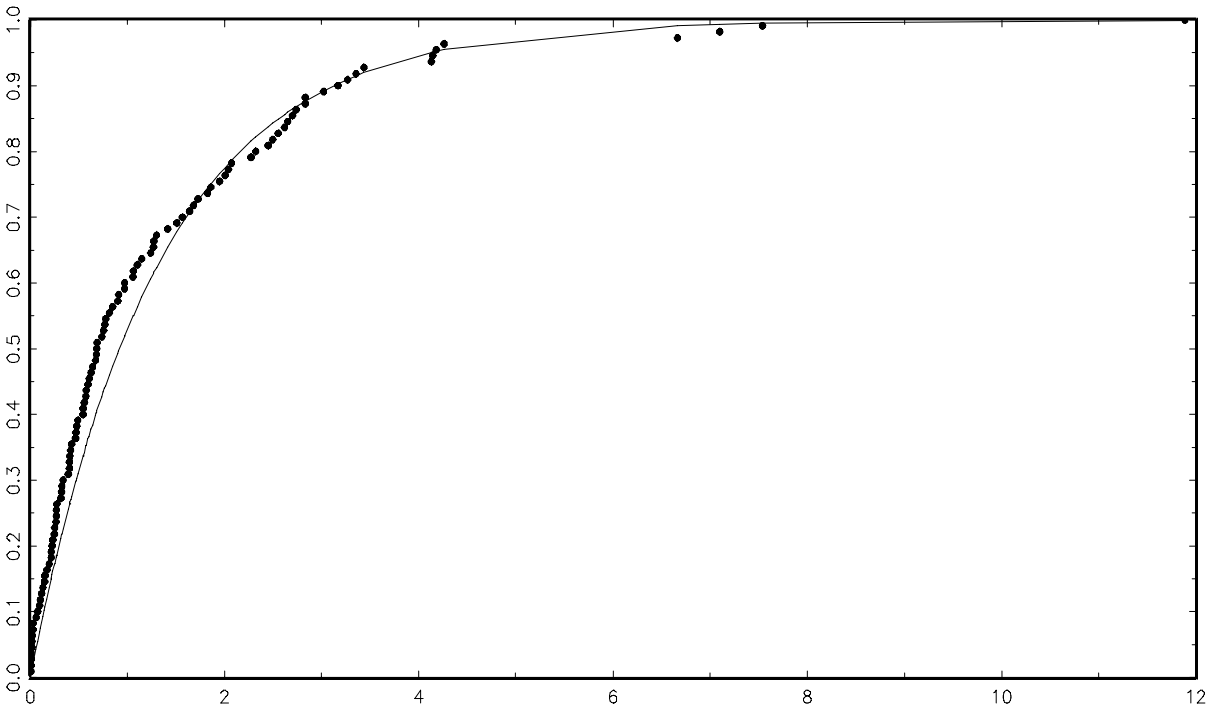
Comparison of tail observations and fitted GPD
Right tail for CAC40 returns



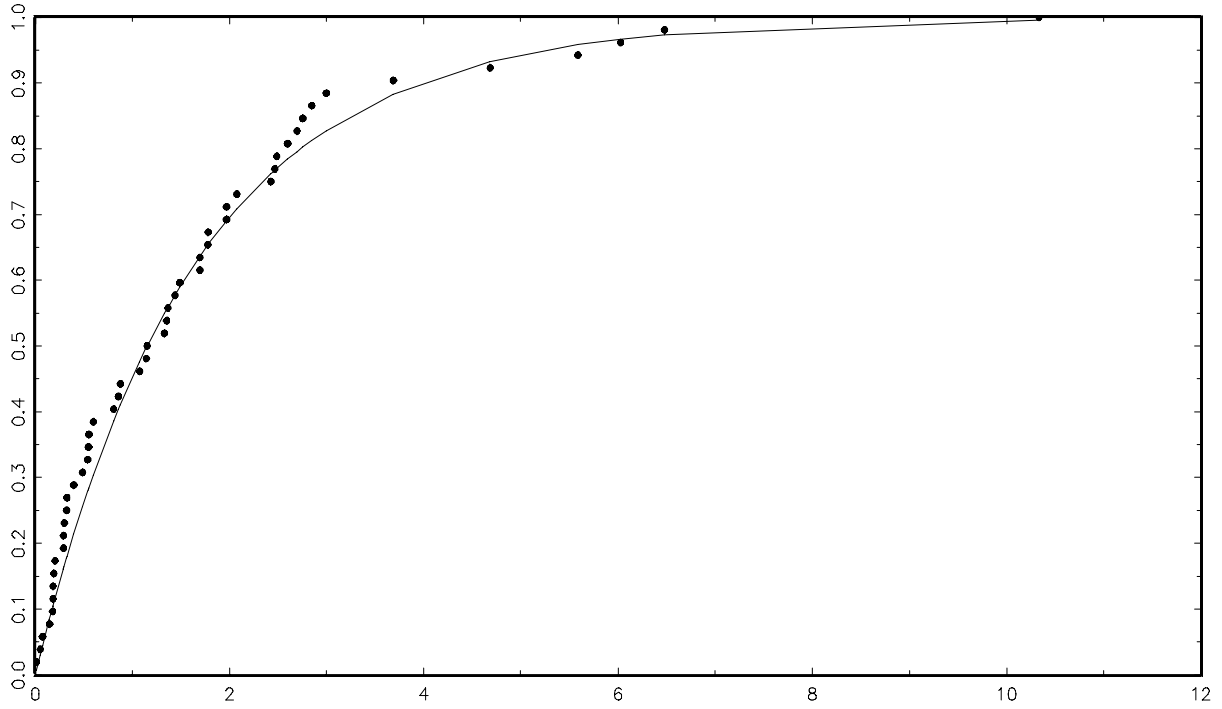
Comparison of tail observations and fitted GPD
Left tail for Singapore returns



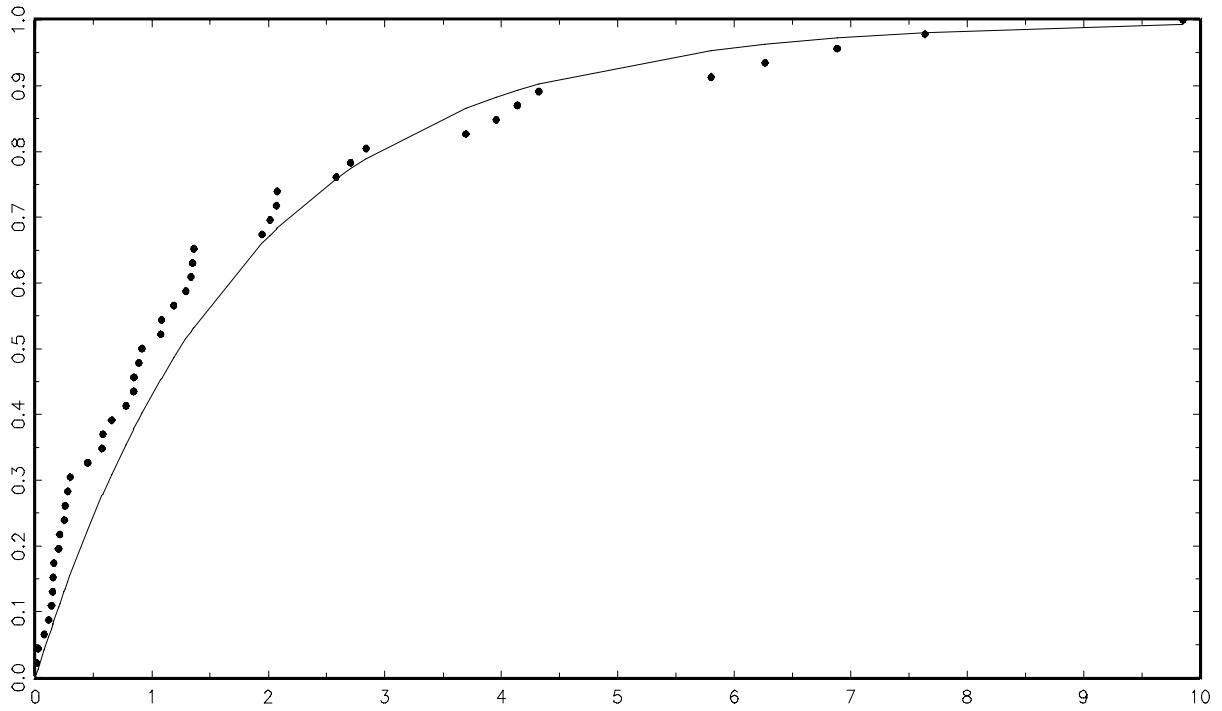
Comparison of tail observations and fitted GPD
Right tail for Singapore returns



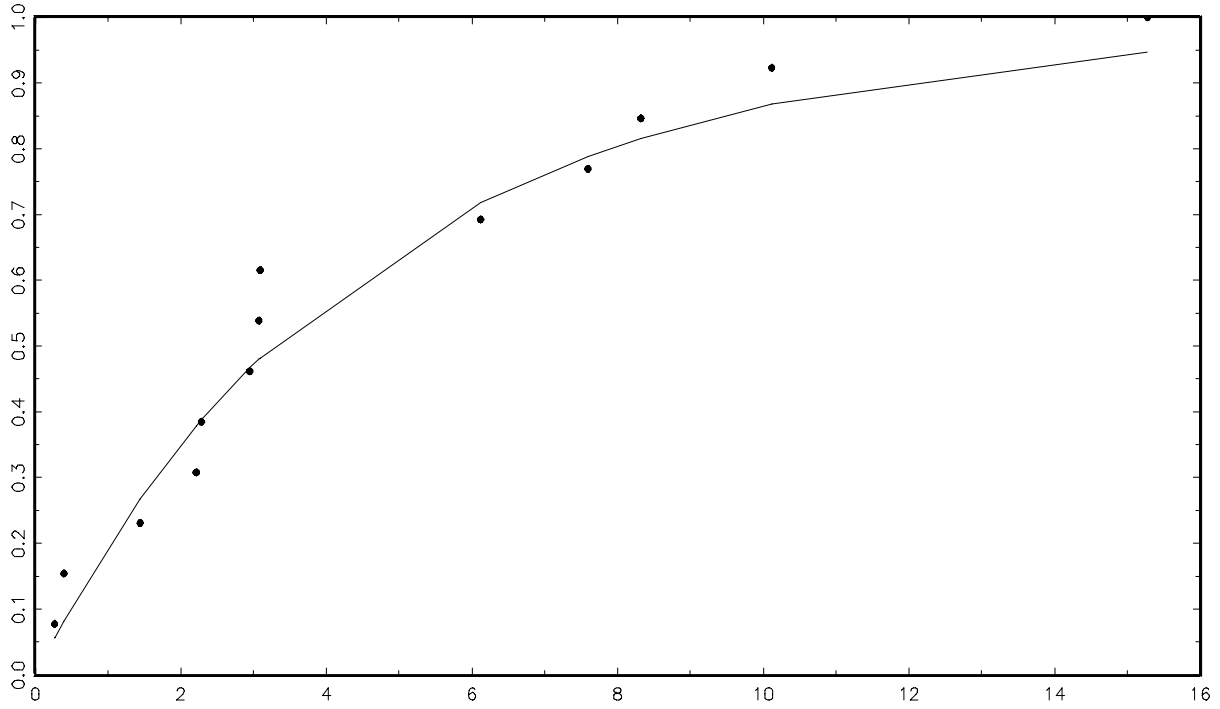
Comparison of tail observations and fitted GPD
Left tail for Mexican returns



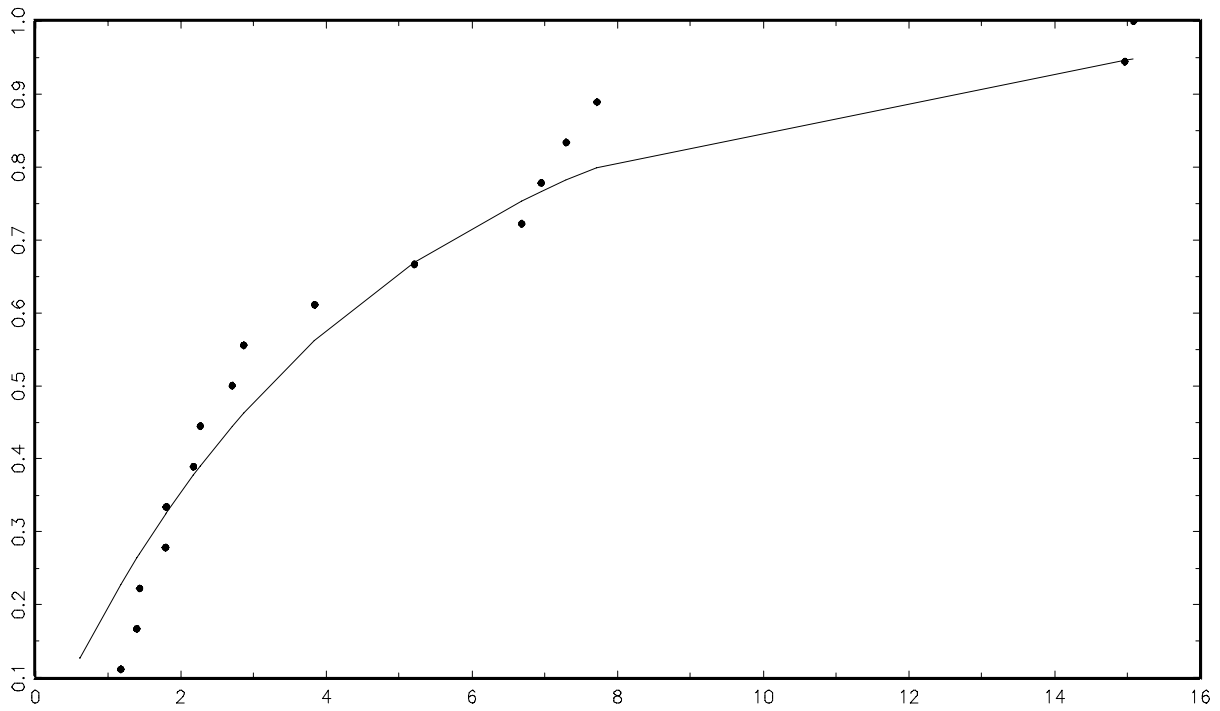
Comparison of tail observations and fitted GPD
Right tail for Mexican returns



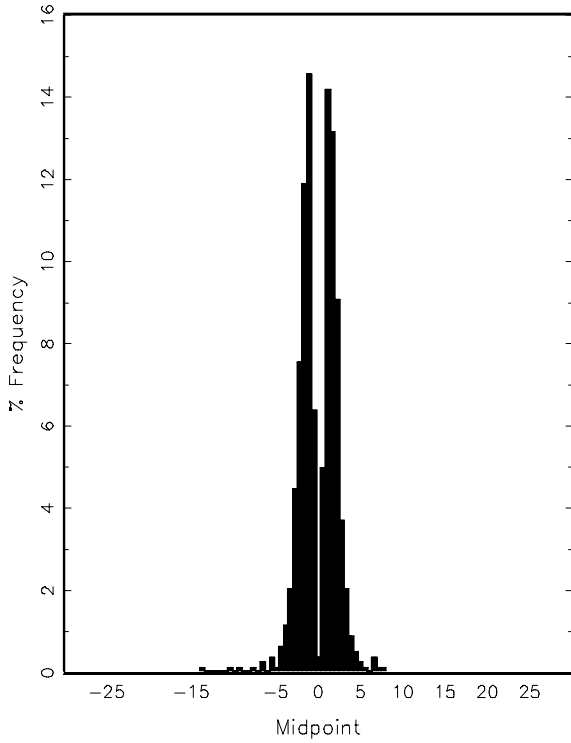
Comparison of tail observations and fitted GPD
Left tail for Russian returns



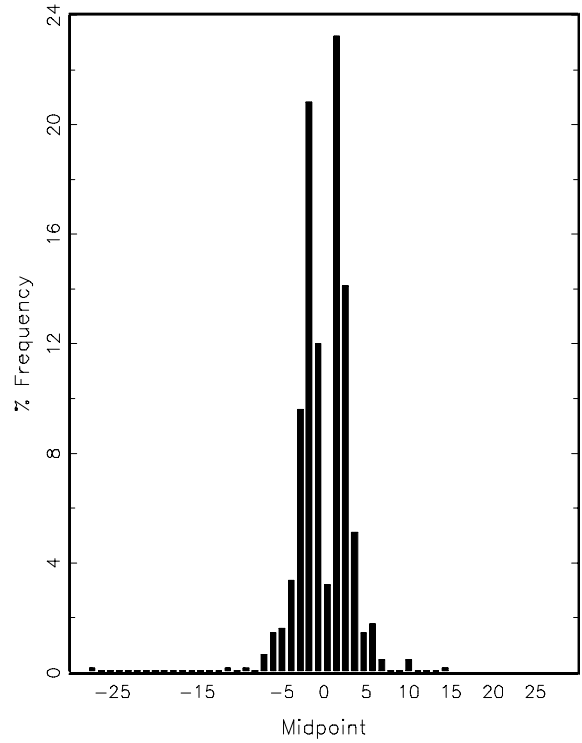
Comparison of tail observations and fitted GPD
Right tail for Russian returns



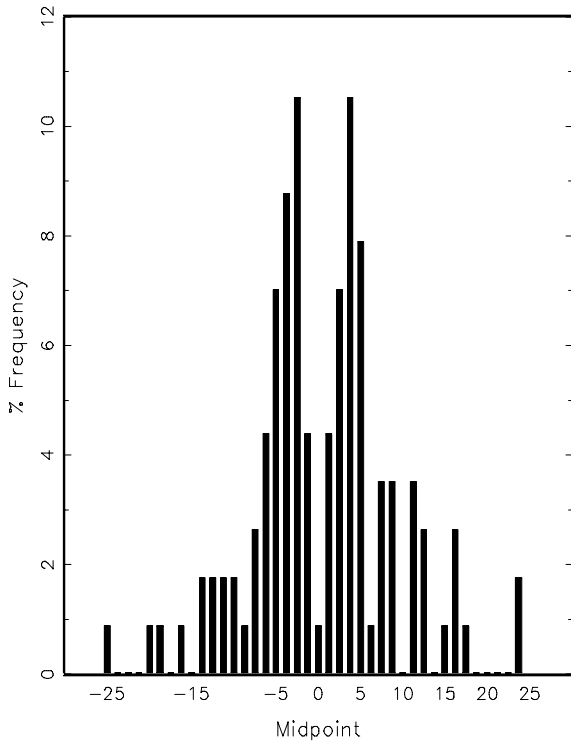
Histogram for min (max) of 20-histories for left (right) tail of CAC40 returns



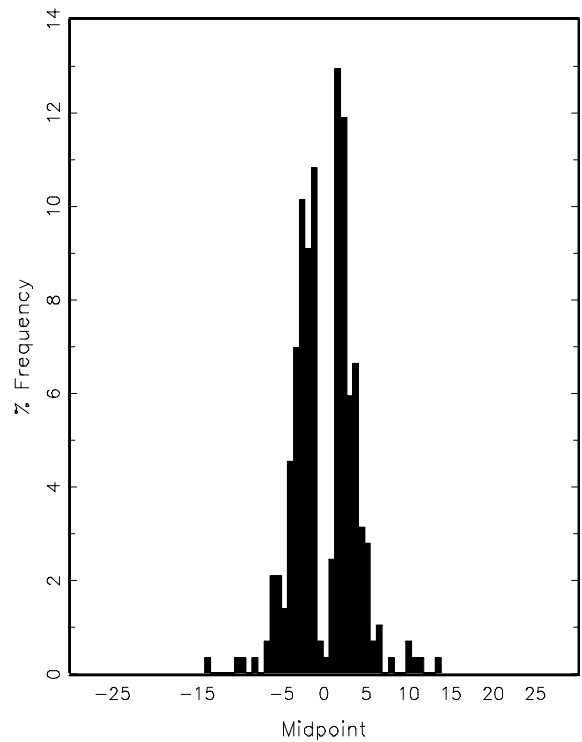
Histogram for min (max) of 20-histories for left (right) tail of returns of Singapore index



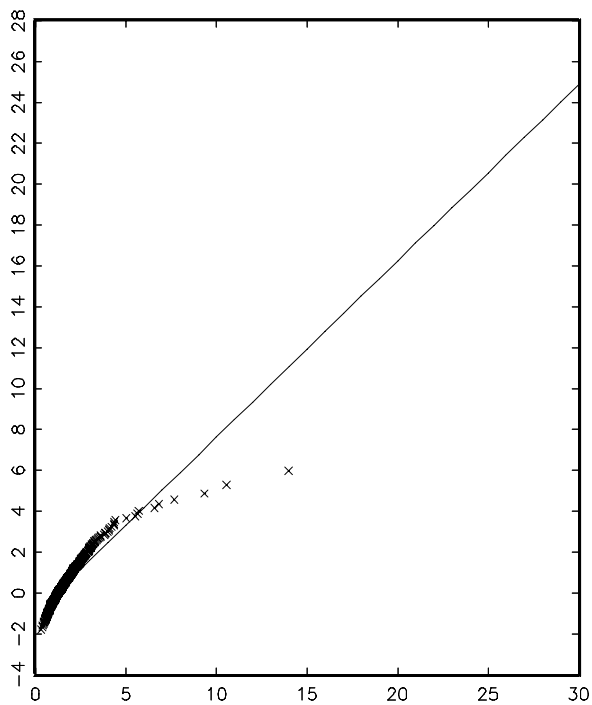
Histogram for min (max) of 20-histories for left (right) tail of returns of Russian index



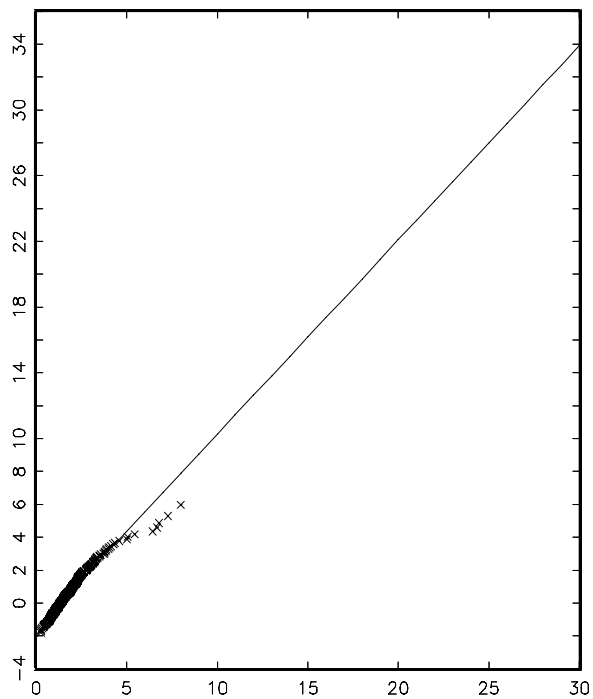
Histogram for min (max) of 20-histories for left (right) tail of returns of Mexican index



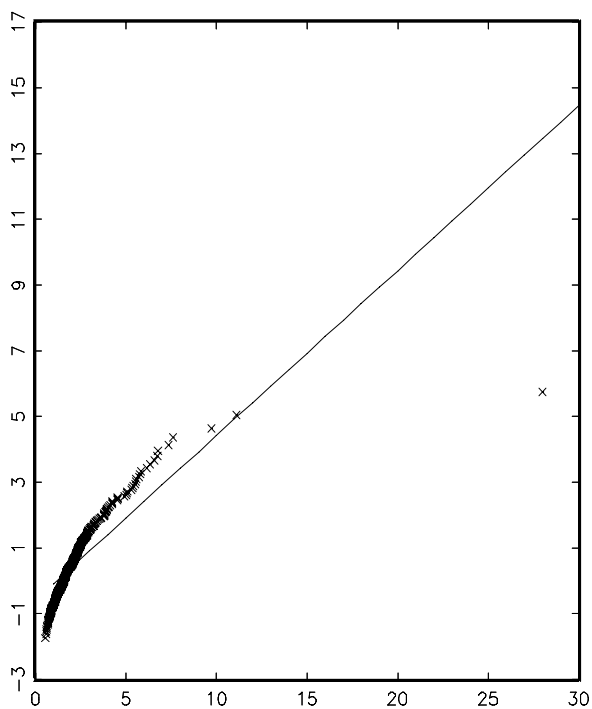
QQ-plot for min of 20-histories
left tail of CAC40 returns



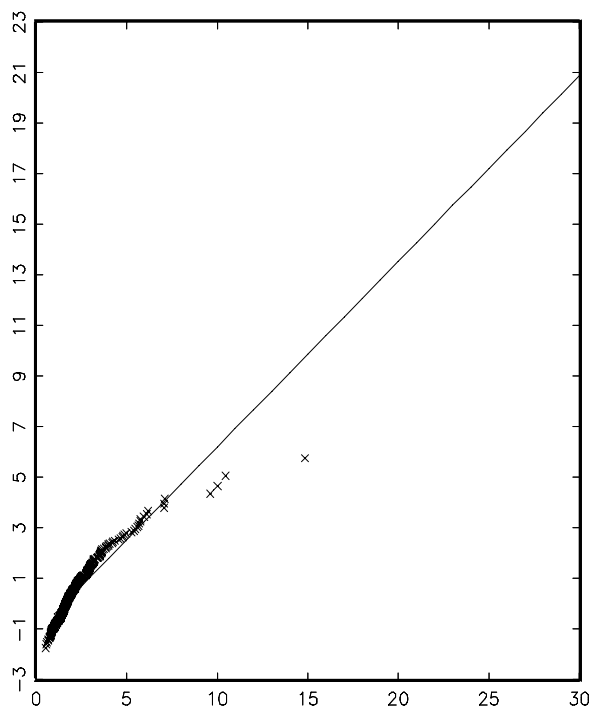
qq-plot for max of 20-histories
right tail of CAC40 returns



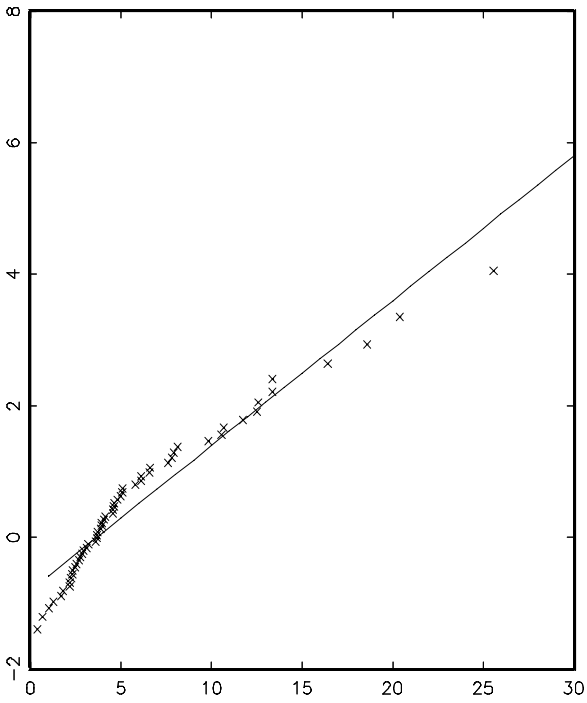
QQ-plot for min of 20-histories
left tail of returns of Singapore index



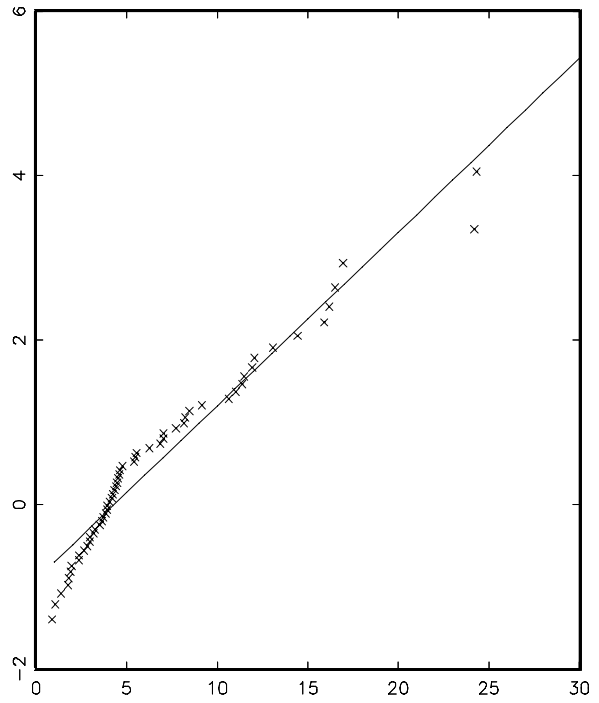
qq-plot for max of 20-histories
right tail of returns of Singapore index



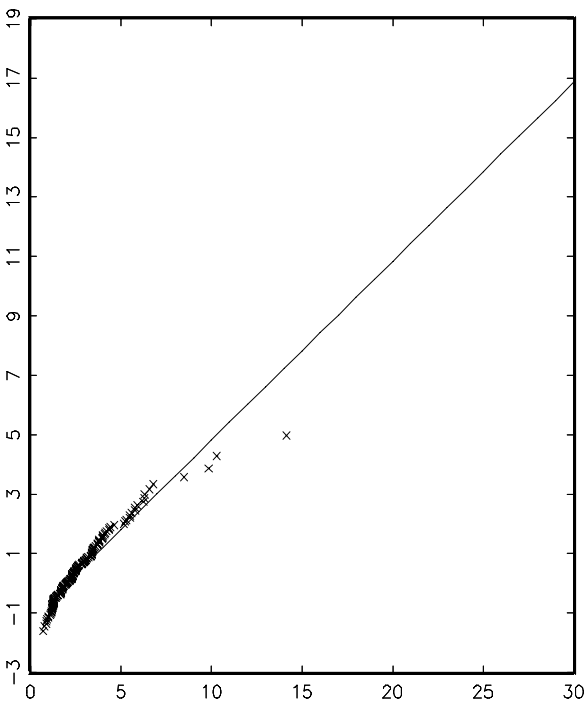
QQ-plot for min of 20-histories
left tail of returns of Russian index



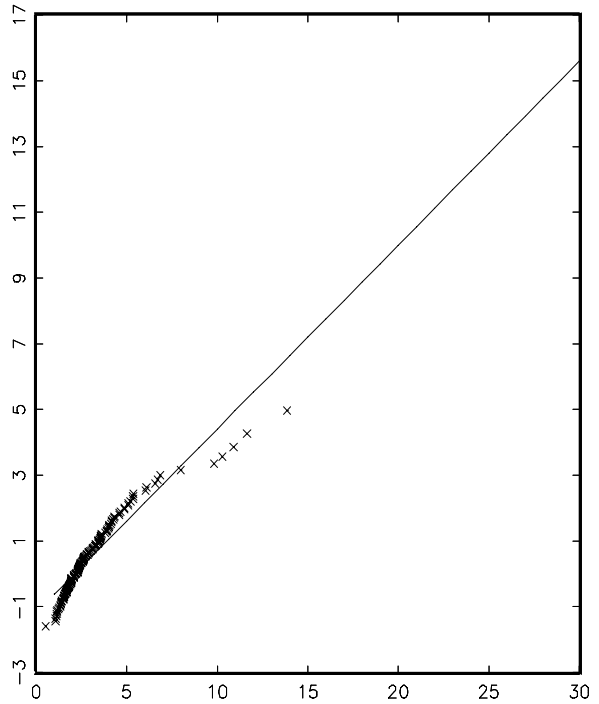
qq-plot for max of 20-histories
right tail of returns of Russian index



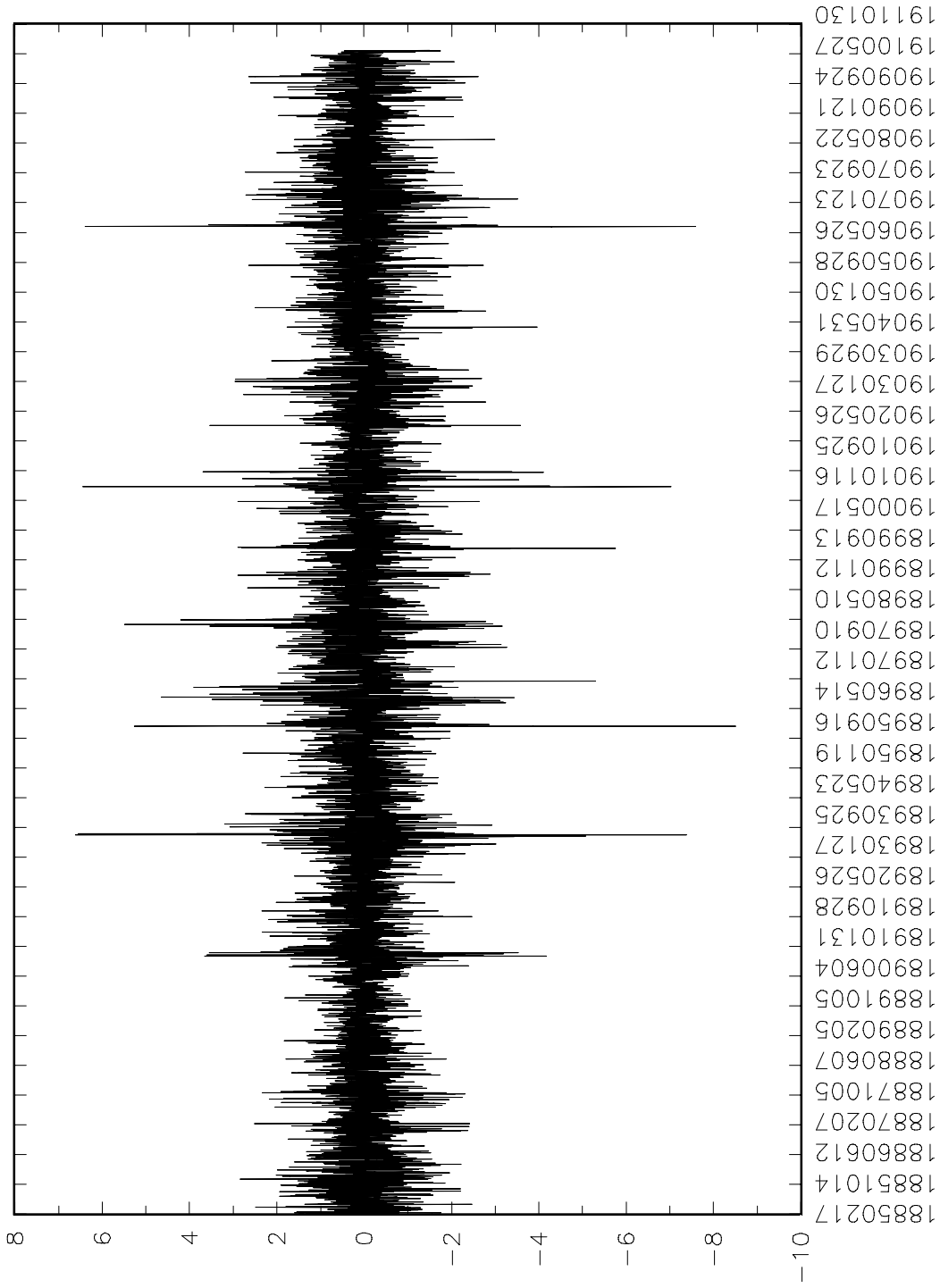
QQ-plot for min of 20-histories
left tail of returns of Mexican index



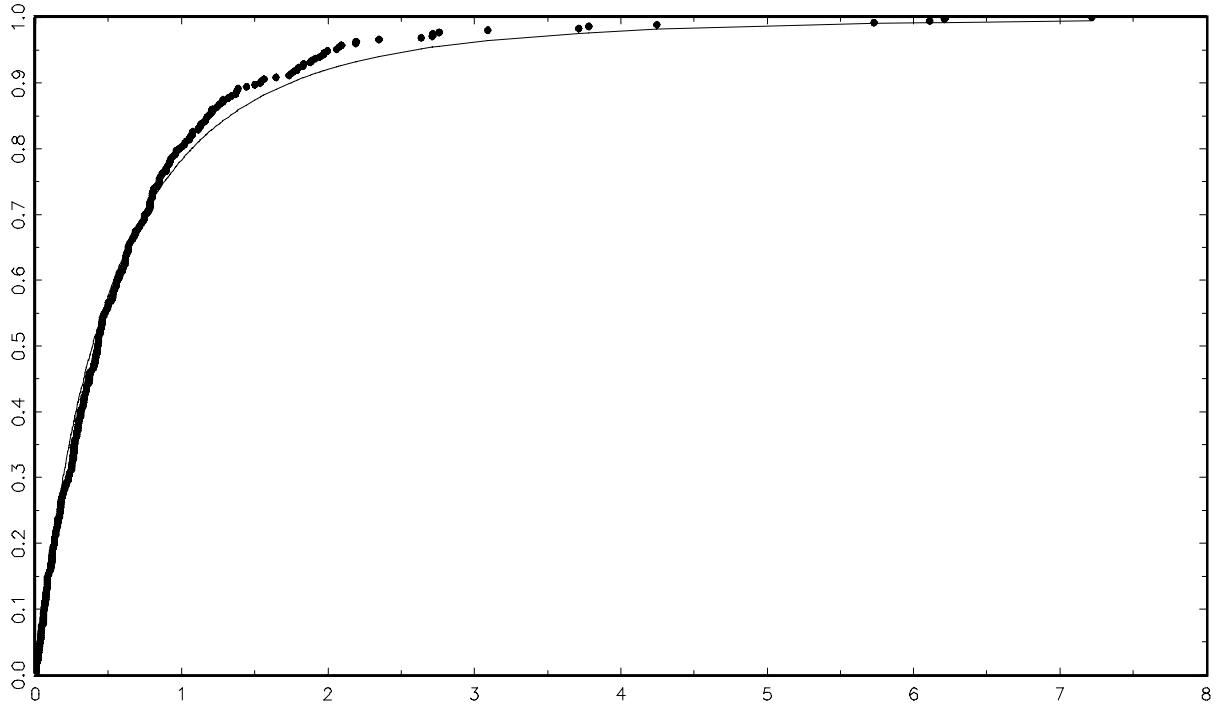
qq-plot for max of 20-histories
right tail of returns of Mexican index



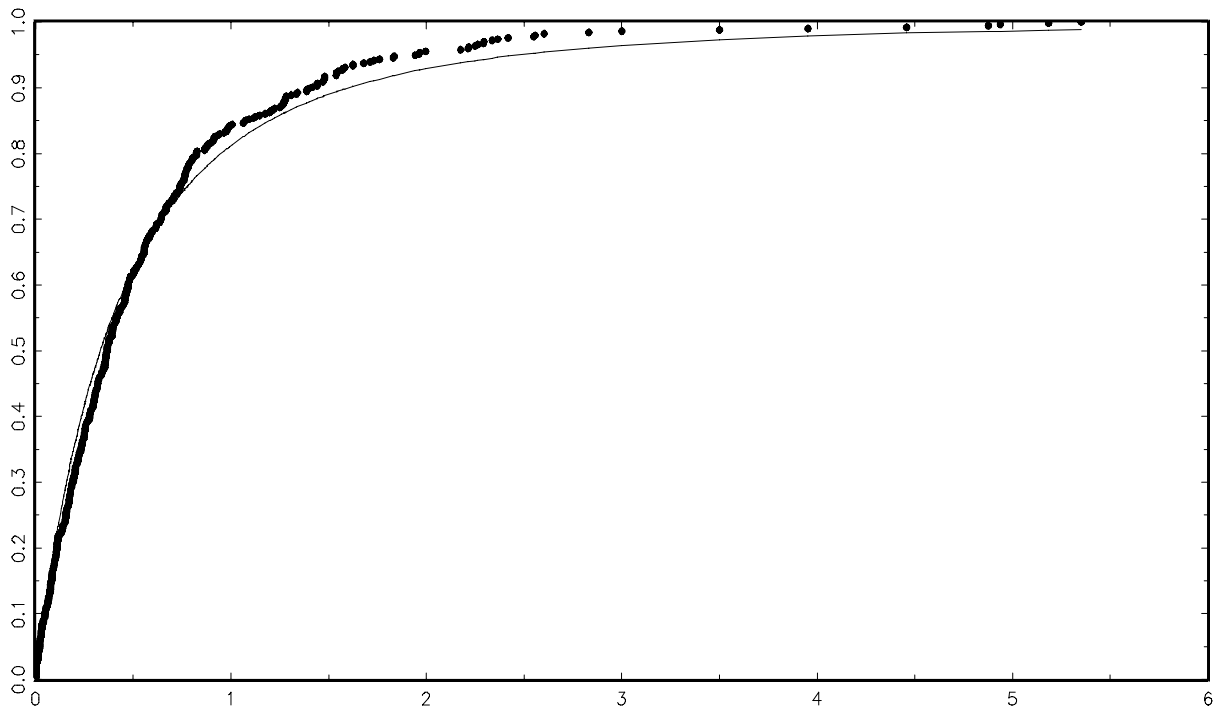
Global US index
February 17th 1885 – March 2nd 1911



Comparison of tail observations and fitted GPD
Left tail for Global US index returns (early years)



Comparison of tail observations and fitted GPD
Right tail for Global US index returns (early years)



Notes d'Études et de Recherche

1. C. Huang and H. Pagès, "Optimal Consumption and Portfolio Policies with an Infinite Horizon: Existence and Convergence," May 1990.
2. C. Bordes, « Variabilité de la vitesse et volatilité de la croissance monétaire : le cas français », février 1989.
3. C. Bordes, M. Driscoll and A. Sauviat, "Interpreting the Money-Output Correlation: Money-Real or Real-Real?," May 1989.
4. C. Bordes, D. Goyeau et A. Sauviat, « Taux d'intérêt, marge et rentabilité bancaires : le cas des pays de l'OCDE », mai 1989.
5. B. Bensaid, S. Federbusch et R. Gary-Bobo, « Sur quelques propriétés stratégiques de l'intéressement des salariés dans l'industrie », juin 1989.
6. O. De Bandt, « L'identification des chocs monétaires et financiers en France : une étude empirique », juin 1990.
7. M. Boutillier et S. Dérangère, « Le taux de crédit accordé aux entreprises françaises : coûts opératoires des banques et prime de risque de défaut », juin 1990.
8. M. Boutillier and B. Cabrillac, "Foreign Exchange Markets: Efficiency and Hierarchy," October 1990.
9. O. De Bandt et P. Jacquinot, « Les choix de financement des entreprises en France : une modélisation économétrique », octobre 1990 (English version also available on request).
10. B. Bensaid and R. Gary-Bobo, "On Renegotiation of Profit-Sharing Contracts in Industry," July 1989 (English version of NER n° 5).
11. P. G. Garella and Y. Richelle, "Cartel Formation and the Selection of Firms," December 1990.
12. H. Pagès and H. He, "Consumption and Portfolio Decisions with Labor Income and Borrowing Constraints," August 1990.
13. P. Sicsic, « Le franc Poincaré a-t-il été délibérément sous-évalué ? », octobre 1991.
14. B. Bensaid and R. Gary-Bobo, "On the Commitment Value of Contracts under Renegotiation Constraints," January 1990 revised November 1990.
15. B. Bensaid, J.-P. Lesne, H. Pagès and J. Scheinkman, "Derivative Asset Pricing with Transaction Costs," May 1991 revised November 1991.
16. C. Monticelli and M.-O. Strauss-Kahn, "European Integration and the Demand for Broad Money," December 1991.
17. J. Henry and M. Phelipot, "The High and Low-Risk Asset Demand of French Households: A Multivariate Analysis," November 1991 revised June 1992.
18. B. Bensaid and P. Garella, "Financing Takeovers under Asymmetric Information," September 1992.

19. A. de Palma and M. Uctum, "Financial Intermediation under Financial Integration and Deregulation," September 1992.
20. A. de Palma, L. Leruth and P. Régibeau, "Partial Compatibility with Network Externalities and Double Purchase," August 1992.
21. A. Frachot, D. Janci and V. Lacoste, "Factor Analysis of the Term Structure: a Probabilistic Approach," November 1992.
22. P. Sicsic et B. Villeneuve, « L'Afflux d'or en France de 1928 à 1934 », janvier 1993.
23. M. Jeanblanc-Picqué and R. Avesani, "Impulse Control Method and Exchange Rate," September 1993.
24. A. Frachot and J.-P. Lesne, "Expectations Hypothesis and Stochastic Volatilities," July 1993 revised September 1993.
25. B. Bensaïd and A. de Palma, "Spatial Multiproduct Oligopoly," February 1993 revised October 1994.
26. A. de Palma and R. Gary-Bobo, "Credit Contraction in a Model of the Banking Industry," October 1994.
27. P. Jacquinet et F. Mihoubi, « Dynamique et hétérogénéité de l'emploi en déséquilibre », septembre 1995.
28. G. Salmat, « Le retournement conjoncturel de 1992 et 1993 en France : une modélisation V.A.R. », octobre 1994.
29. J. Henry and J. Weidmann, "Asymmetry in the EMS Revisited: Evidence from the Causality Analysis of Daily Eurorates," February 1994 revised October 1994.
30. O. De Bandt, "Competition Among Financial Intermediaries and the Risk of Contagious Failures," September 1994 revised January 1995.
31. B. Bensaïd et A. de Palma, « Politique monétaire et concurrence bancaire », janvier 1994 révisé en septembre 1995.
32. F. Rosenwald, « Coût du crédit et montant des prêts : une interprétation en terme de canal large du crédit », septembre 1995.
33. G. Cette et S. Mahfouz, « Le partage primaire du revenu : constat descriptif sur longue période », décembre 1995.
34. H. Pagès, "Is there a Premium for Currencies Correlated with Volatility ? Some Evidence from Risk Reversals," January 1996.
35. E. Jondeau and R. Ricart, "The Expectations Theory : Tests on French, German and American Euro-rates," June 1996.
36. B. Bensaïd et O. De Bandt, « Les stratégies "stop-loss" : théorie et application au Contrat Notionnel du Matif », juin 1996.

37. C. Martin et F. Rosenwald, « Le marché des certificats de dépôts. Écarts de taux à l'émission : l'influence de la relation émetteurs-souscripteurs initiaux », avril 1996.
38. Banque de France - CEPREMAP - Direction de la Prévision - Erasme - INSEE - OFCE, « Structures et propriétés de cinq modèles macroéconomiques français », juin 1996.
39. F. Rosenwald, « L'influence des montants émis sur le taux des certificats de dépôts », octobre 1996.
40. L. Baumel, « Les crédits mis en place par les banques AFB de 1978 à 1992 : une évaluation des montants et des durées initiales », novembre 1996.
41. G. Cette et E. Kremp, « Le passage à une assiette valeur ajoutée pour les cotisations sociales : Une caractérisation des entreprises non financières "gagnantes" et "perdantes" », novembre 1996.
42. S. Avouyi-Dovi, E. Jondeau et C. Lai Tong, « Effets "volume", volatilité et transmissions internationales sur les marchés boursiers dans le G5 », avril 1997.
43. E. Jondeau et R. Ricart, « Le contenu en information de la pente des taux : Application au cas des titres publics français », juin 1997.
44. B. Bensaid et M. Boutillier, « Le contrat notionnel : Efficience et efficacité », juillet 1997.
45. E. Jondeau et R. Ricart, « La théorie des anticipations de la structure par terme : test à partir des titres publics français », septembre 1997.
46. E. Jondeau, « Représentation VAR et test de la théorie des anticipations de la structure par terme », septembre 1997.
47. E. Jondeau et M. Rockinger, « Estimation et interprétation des densités neutres au risque : Une comparaison de méthodes », octobre 1997.
48. L. Baumel et P. Sevestre, « La relation entre le taux de crédits et le coût des ressources bancaires. Modélisation et estimation sur données individuelles de banques », octobre 1997.
49. P. Sevestre, "On the Use of Banks Balance Sheet Data in Loan Market Studies : A Note," October 1997.
50. P.-C. Hautcoeur et P. Sicsic, "Threat of a Capital Levy, Expected Devaluation and Interest Rates in France during the Interwar Period," January 1998.
51. P. Jacquinot, « L'inflation sous-jacente à partir d'une approche structurelle des VAR : une application à la France, à l'Allemagne et au Royaume-Uni », janvier 1998.
52. C. Bruneau et O. De Bandt, « La modélisation VAR structurel : application à la politique monétaire en France », janvier 1998.
53. C. Bruneau et E. Jondeau, "Long-Run Causality, with an Application to International Links between Long-Term Interest Rates," June 1998.
54. S. Coutant, E. Jondeau et M. Rockinger, "Reading Interest Rate and Bond Futures Options' Smiles: How PIBOR and Notional Operators Appreciated the 1997 French Snap Election," June 1998.

55. E. Jondeau et F. Sédillot, « La prévision des taux longs français et allemands à partir d'un modèle à anticipations rationnelles », juin 1998.
56. E. Jondeau et M. Rockinger, "Estimating Gram-Charlier Expansions with Positivity Constraints," January 1999.
57. S. Avouyi-Dovi et E. Jondeau, "Interest Rate Transmission and Volatility Transmission along the Yield Curve," January 1999.
58. S. Avouyi-Dovi et E. Jondeau, « La modélisation de la volatilité des bourses asiatiques », janvier 1999.
59. E. Jondeau, « La mesure du ratio rendement-risque à partir du marché des euro-devises », janvier 1999.
60. C. Bruneau and O. De Bandt, "Fiscal policy in the transition to monetary union: a structural VAR model," January 1999.
61. E. Jondeau and R. Ricart, "The Information Content of the French and German Government Bond Yield Curves: Why Such Differences?" February 1999.
62. J.-B. Chatelain et P. Sevestre, « Coûts et bénéfices du passage d'une faible inflation à la stabilité des prix », février 1999.
63. D. Irac et P. Jacquinot, « L'investissement en France depuis le début des années 1980 », Avril 1999.
64. F. Mihoubi, « Le partage de la valeur ajoutée en France et en Allemagne », mars 1999.
65. S. Avouyi-Dovi and E. Jondeau, "Modelling the French Swap Spread," April 1999.
66. E. Jondeau and M. Rockinger, "The Tail Behavior of Stock Returns: Emerging Versus Mature Markets," June 1999.

Pour tous commentaires ou demandes sur les Notes d'Études et de Recherche, contacter la bibliothèque du Centre de recherche à l'adresse suivante :

For any comment or enquiries on the Notes d'Études et de Recherche, contact the library of the Centre de recherche at the following address:

BANQUE DE FRANCE
41.1391 - Centre de recherche
75 049 Paris CEDEX
tél : 01 42 92 49 55
fax : 01 42 92 62 92
email : thierry.demoulin@banque-france.fr