Some empirical evidence concerning stable distributions

José Miguel Pinto DosSantos[a)]

1. Introduction

The most frequent assumption made in the theoretical and empirical analysis of financial markets is that the distribution of price changes or price returns of financial assets is approximately normal[1)]. Nor is this view limited to financial economics: all areas that use econometric tools make frequent use of the normality assumption. An alternative view holds that the distribution of price changes or returns belongs to a family of distributions designated as stable distributions[2)]. Which view best reflects reality is not without interest.

The fixed and floating exchange rate regimes provide a good starting

a) I am grateful to Professors J. Fukuchi, T. Kitaoka, S. Komura and K. Maekawa for their comments and to S. Yoshimura for her research assistance. Remaining errors are my own.

1) This assumption is called the normal hypothesis, Gaussian hypothesis (Fama, 1963), or also the Gibrat law or hypothesis (Steindl, 1965). The distribution that it refers to is the normal or Gaussian distribution.

2) This view or assumption is also known as Pareto-Levy law (Mandelbrot, 1960), stable Paretian hypothesis (Mandelbrot 1961), Pareto law (Steindl, 1965), stable law (Monrad and Stout, 1989) or the Mandelbrot hypothesis (Fama, 1965b). The distributions that it refers are known as stable distributions, stable non-Gaussian distributions, sub-gaussian distributions, or stable Paretian distributions.

example of the economic implications of the two views. Which of the following two regimes implies more risk to economic agents that produce goods for exportation or consume foreign goods? An exchange rate regime that only allows fluctuations of 1% around a central rate but requires realignments when certain unpredictable crises occur, or an exchange rate regime that allows the free flotation of two currencies? The unaware observer might answer that the first is less risky than the second. However the answer is more subtle than that. Careful observation of Figure 1 will give an idea of why this answer is probably wrong. In this figure, the probability distribution of the rate of variation (or return) implied by the first, fixed-rate, regime may be represented by distribution α_1 and the implied by the second, floating-rate, regime by distribution α_2. The fixed-rate regime does indeed imply an higher probability, translated in practice by an higher frequency, of very small changes than the floating-rate regime: on this account it seems less risky. However, the fixed-rate regime also implies an higher probability of very large changes: on this account it is more risky. As will be referred later, the proposition that a certain distribution has more weight in the tails than another is equivalent to the proposition that the first distribution is more risky than the sec-

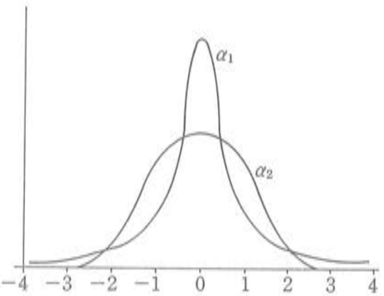

Fig. 1 Risk depends on the shape of the probability distribution

ond. Thus we arrive to the counter-intuitive result that the fixed exchange rate regime may be considered more risky than the floating-rate one (for risk averse agents).

This paper has several objectives. The first is to present in an unified manner the existing statistical theory, main estimation methods, and more important empirical results concerning stable distributions in some economic fields, with a special emphasis to those related to financial economics. A lot of work exists in this area but there is not up to now a survey that presents in a consistent way the results dispersed in a large and growing literature. Thus we present the main properties of stable functions in section 2, the several estimation methods in section 3, and the main empirical findings in section 5.

A second objective is to show the importance of this topic to some fields of economic theory and practice. Either because the relevance of this topic is thought to be self evident and deserving no further comment, or because it is thought to be of no consequence, there is up to now no explicit elaboration on this subject. This is dealt with in section 4 where the implications of stable distributions are analyzed for the Efficient Market Hypothesis, risk theory and econometrics. We also propose in this section a new test to the semi-strong form of the Efficient Market Hypothesis.

The third objective is to add another piece of evidence: in section 6 the parameters that determine the shape of the distribution of the returns of the Nikkei 225 price index are estimated using two different methods and the results of the two methods are compared.

2. Definition and basic properties

Let two random variables U and V have the same distribution:

$$U \overset{d}{=} V,$$

where $\overset{d}{=}$ indicates that the distribution is the same.

Thus if

$$U \overset{d}{=} aV+b,$$

then U and V differ only by the location and scale parameters.

Let X_1, X_2, ... and X be independent identically distributed (iid) random variables with common distribution L. Let also S_n be the sum of n of those random variables: $S_n = X_1 + \dots + X_n$. Then if the distribution L has non-zero variance and if for each n there exist scaling and centering constants $a_n > 0$ and b_n such that

$$S_n \overset{d}{=} a_n X + b_n$$

then the distribution L is stable. If this condition holds with $b_n = 0$ for all n then the distribution is said to be stable in the strict sense. If it holds for values of b_n other than zero then the distribution is said to be stable in the broad sense.

Some of the basic properties of stable distributions are the following[3]:

Property 1. (Existence of characteristic exponent) Every stable distribution has norming constants of the form $a_n = n^{1/\alpha}$ with $0 < \alpha \leqq 2$. α is called characteristic exponent or characteristic index of the stable distribution L.

Property 2. (Irrelevance of centering constant) If a distribution is stable with exponent $\alpha \neq 1$ the centering constant may be chosen so

3) For formal proofs for these properties see either Feller (1968), Gnedenko and Kolmogorov (1948) or Levy (1925); later in the text, intuitive proofs and justifications will be presented for some of these properties.

that the distribution is strictly stable.

Thus we are free to center the distribution L in an arbitrary way, and can chose the most convenient centering, $b_n=0$, when the need arises.

Property 3. (Asymptotic law of Pareto) The tails of stable distributions for values of $\alpha<2$ follow the weak or asymptotic law of Pareto, eg.:

$$\lim_{x\to\infty} \Pr(X>x)\to(x/Z_1)^{-\alpha} \tag{1}$$

and

$$\lim_{x\to\infty} \Pr(X<-x)\to(|x|/Z_2)^{-\alpha} \tag{2}$$

where X is a random variable and Z_1 and Z_2 are constants defined as:

$$\beta=\frac{Z_1^\alpha-Z_2^\alpha}{Z_1^\alpha+Z_2^\alpha}.$$

β, of which more will be said afterwards, is a parameter for skewness.

Property 4. (Stability or invariance under addition) If L is strictly stable with exponent α then all linear combinations of any two or more iid random variables with distribution L will also have distribution L.

Property 5. (Limiting distributions) A distribution L possesses a domain of attraction if and only if it is stable.

A distribution F of iid random variables is said to belong to the domain of attraction[4)] of a distribution L if there exists norming constants a_n and b_n such that

$$\frac{(S_n-b_n)}{a_n}\xrightarrow{d}Y,$$

4) For more on domain of attraction see Stout (1989).

where Y is a random variable with a distribution L, S_n is the sum of n of those iid random variables and $\xrightarrow{d}$ denotes convergence in distribution.

As Feller (1968) points, the importance of the normal distribution is due largely to the central limit theorem. If X_1, X_2, ... are iid with a common distribution F having zero expectation and finite variance, then according to the central limit theorem the distribution of S_n is asymptotically normal[5].

However, if iid random variables do not have finite variance but their sums follow a limiting distribution, then the limiting distribution must be stable with $0<\alpha<2$. All stable distributions and no others may occur as such limits. For a stable distribution with characteristic exponent $0<\alpha<2$ to be a limiting distribution of the sums of such random variables the Doeblin-Gnedenko conditions give the necessary and sufficient conditions:

$$\lim_{x\to\infty} \frac{F(-x)}{1-F(x)} \to \frac{C_1}{C_2}$$

and for every constant $k>0$,

$$\lim_{x\to\infty} \frac{1-F(x)+F(-x)}{1-F(kx)+F(-kx)} \to k^{\alpha}$$

where $F(\bullet)$ is the cumulative distribution function of the random variable X ($F(x) \equiv Pr(X \leqq x)$) and C_1 and C_2 are constants. All random variables that follow the asymptotic law of Pareto will satisfy the Doeblin-Gnedenko conditions and thus belong to the domain of attrac-

5) As will become clear later, if F has zero expectation and unit variance, then distribution of $S_n/n^{1/2}$ is asymptotically normal with zero expectation and unit variance.

tion of a stable distribution, regardless of their having a stable distribution or not. For example, for any random variable X that is not stable but is asymptotically Paretian:

$$\lim_{x\to\infty} \frac{F(-x)}{1-F(x)} \to \left[\frac{(|-x|/Z_2)}{(x/Z_1)}\right]^{-\alpha} = \frac{Z_2^\alpha}{Z_1^\alpha}$$

and

$$\lim_{x\to\infty} \frac{1-F(x)+F(-x)}{1-F(kx)+F(-kx)} \to \frac{(x/Z_1)^{-\alpha}+(|-x|/Z_2)^{-\alpha}}{(kx/Z_1)^{-\alpha}+(|-kx|/Z_2)^{-\alpha}} = k^\alpha$$

The characteristic function of stable distributions is[6]:

$$\phi(t) = \int_{-\infty}^{+\infty} \exp(ixt)\,dF(x) = \exp[i\delta t - \gamma|t|^\alpha\{1 - i\beta(t/|t|)\omega(\alpha, t)\}] \quad (3)$$

with

$$\omega(\alpha, t) = \begin{cases} \tan(\alpha\pi/2), & \text{if } \alpha \neq 1 \\ -2\log|t|/\pi, & \text{if } \alpha = 1 \end{cases}$$

where α is the characteristic exponent, β is a measure of skewness $(-1<\beta<1)$, $\gamma(\gamma=c^\alpha)$ is the scale (that is, the unit of measurement) parameter and δ is the location parameter, t belongs to the set of real numbers and i is the imaginary unit[7].

The characteristic exponent determines the leptokurtosis of a distribution, in other words, the thickness of the tails of that distribution.

---

6) We adopt closely the definitions given by Zolotarev (1957, p. 441) and McCulloch (1986). However, it should be noted that other parameterizations exist which are considered more convenient for analytical work (Zolotarev (1966), Chambers, Mallows and Stuck (1976) and Feuerverger and McDunnough (1981)).

7) Laha (1989) and Kendal and Stuart (1977, ch. 4) present a complete treatment of characteristic functions.

The thicker these are the larger the probability of occurrence of outliers. If the tails of a certain distribution are thicker than the tails of normal distribution, then the distribution is said to display excess kurtosis, or to be leptokurtic. On the other hand, if the tails are thinner, a case rarely found in practice, then the distribution is said to be platykurtic. When $\alpha=2$ the distribution is Gaussian, or mesokurtic. As α moves away from 2 towards 0 the thickness of the tails or leptokurtosis increases[8]. A consequence of this increasing leptokurtosis is that finite variance exists only in the extreme case of $\alpha=2$ and the mean only when $\alpha>1$. Figure 2 shows[9] the distribution function of 3 stable distributions with different α's: $\alpha_1<\alpha_2<\alpha_3$.

The symmetry of a distribution can be judged according to its β. β is the limiting value of the ratio of the difference of the tail probabilities to the sum of the tail probabilities:

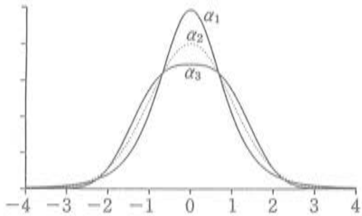

Fig. 2 Probability distributions with different characteristic exponents: $\alpha_1<\alpha_2<\alpha_3$

8) As Kendall and Stuart (1977) and Davidson and Mackinnon (1993) point out the terms leptokurtic, mesokurtic and platokurtic were first used to designate not the tails but the central part of the distribution; thus leptokurtic distributions own their name not to their thick tails but to their relatively thin central part. However, in almost all instances, it is the tails that are referred by these terms.

9) This Figure is based on Figure 1 of Fama and Roll (1968); the standardization is made using: $u=(x-\delta)/c$.

$$\beta = \lim_{x \to \infty} \frac{1-F(x)-F(-x)}{1-F(x)+F(-x)}, \text{ when } \alpha \neq 2,$$

where F(•) is the cumulative distribution function of the random variable X. As should have become apparent by now, in the above equations 1 and 2, Z_1^α and Z_2^α are the probability contained respectively in the right and left tails of a distribution. When $\beta=0$ the distribution is symmetric ($Z_1=Z_2$); when $\beta>0$ the distribution is skewed to the right[10], in other words, the right tail is longer and thicker than the left one ($Z_1>Z_2$), and the degree of right skewness increases as β approaches 1; and conversely when $\beta<0$. It should be noted that in equation (3) as α approaches 2 (and $\omega(\alpha, t)$ approaches 0), β loses its effect and the distribution approaches the symmetrical normal distribution regardless of the value of β.

Using the characteristic function presented above, a more rigorous presentation of property 3 may now be given[11]. If n iid stable random variables having equal values for the parameters α, β, γ and δ are summed, the expression for the logarithm of the characteristic function of this sum is:

$$n \log[\phi(t)] = i(\delta n)t - (\gamma n)|t|^\alpha\{1 - i\beta(t/|t|)\omega(\alpha, t)\}$$

where $\log[\phi(t)]$ is the logarithm of the characteristic function of the individual random variables. As can be seen, the distribution of the

10) It is also said to be positively skewed. As pointed by McCulloch (1986) there is some confusion in the literature, also characterized by Hall (1981) as a "comedy of errors": following Gnedenko and Kolmogorov (1948) the characteristic function (3) is usually written with a positive sign for $\{i\,\beta(t/|t|)\omega(\alpha, t)\}$. Then when $\beta>0$ the distribution is negatively skewed and it is positively skewed when $\beta<0$.

11) This passage is based on Fama (1963).

sums is, except for location δ and scale γ, exactly the same as the distribution for the individual random variables: α and β remain constant under addition. Even if the location and scale parameters are not the same for each individual variable in the sum, the property of stability under addition still holds. The expression for the logarithm of the characteristic function of the sum of n such variables is:

$$\sum_{j=1}^{n} \log[\phi_j(t)] = i(\sum_{j=1}^{n} \delta_j)t - (\sum_{j=1}^{n} \gamma_j)|t|^{\alpha}\{1 - i\,\beta(t/|t|)\omega(\alpha, t)\}$$

where δ_j and γ_j are the location and scale parameters of each different random variable. Thus the sum of random variables with stable distributions having the same values for α and β but having different values for γ and δ, is still a random variable with stable distribution having the same values for α and β and whose location and scale parameters are the sums of location and scale parameters of the distribution of the individual random variables.

3. Estimation

To determine whether a random variable is normally distributed or has a stable distribution other than the normal it suffices to estimate the value of the characteristic exponent α. However, as explicit expressions for the densities of stable distributions are known for only the Gaussian ($\alpha=2$) and the Cauchy ($\alpha=1$, $\beta=0$), there is no general sampling theory available. Thus usually it is necessary to use numerical methods to estimate α. Below we present some methods of estimation that have been proposed.

3.1 Double log graphing

According to property 3 of stable distributions presented above, when $\alpha<2$ these distributions follow the weak or asymptotic law of

Pareto. Taking logarithms on both sides of expressions (1) and (2) we have:

$$\lim_{x\to\infty} \log \Pr(X>x) \to -\alpha(\log x - \log Z_1)$$

and

$$\lim_{x\to\infty} \log \Pr(X<-x) \to -\alpha(\log|x| - \log Z_2)$$

The above expressions imply that, if $\Pr(X>x)$ and $\Pr(X<-x)$ are plotted against $|x|$ on double log paper, the two lines should become asymptotically straight as $|x|$ approaches infinity.

However, as pointed by Mandelbrot (1963c) this technique is weak when the characteristic exponent is close to 2. If $\alpha=2$, $\Pr(X>x)$ decreases faster than $|x|$ increases, and the slope of $\log \Pr(X>x)$ against $\log|x|$ will approach $-\infty$. For normal distributions the law of Pareto does not hold even asymptotically. When α is between 1.5 and 2, although the law of Pareto holds asymptotically, the absolute value of the slope in the middle of the double-log graph will be greater than the true asymptotic slope, which is only reached at the extreme bottom of the graph. For example, when $\alpha=1.5$ the asymptotic slope is attained only when $\Pr(X>x) \leqq 0.015$, and when $\alpha=1.9$ only when $\Pr(X>x) \leqq 0.0005$ (Fama, 1965, Table 6). Thus to be able to observe the slope of the graph start to converge to the distribution's true α we need to have a number of observations of more than $n=1/\Pr(X>x)$. For a stable distribution with $\alpha=1.5$ we need to have at least $n=67$, and when $\alpha=1.9$ the minimum is $n=2000$. And as the expected number of extreme values which will exhibit the true asymptotic slope is $n \Pr(X>x)$, if $\alpha=1.9$ and we have a sample size of $n=4000$ we can expect that the asymptotic slope will be observable only for the largest two observations in each tail.

Given what was said above, it is only natural that this technique was applied only when very large cross-sectional and time series samples were available. Examples of the actual use of this technique for estimating α are Mandelbrot (1963a) for the daily returns of cotton prices, Fama (1965b) for 30 different stocks, and Steindl (1965) for the distribution of wealth in Sweden, for the distribution of US corporations according to assets, for the distribution of US firms according to number of employees, and for the distribution of firms according to turnover in the Federal Republic of Germany, from among several examples. Figure 3 presents an example of this technique for the distribution of wealth in Sweden in 1959, using data from Steindl (1965, p. 188).

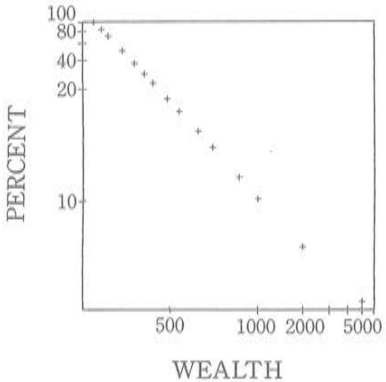

Fig. 3 Distribution of wealth in Sweden: percentage of owners with property equal to or exceeding the amount of property shown in the horizontal axis (thousands of swedish crowns)

3.2 Range Analysis[12)]

According to Property 4 of stable distributions presented in section 2, the distribution of the sum of n stable iid random variables S_n is also stable with the same characteristic exponent as the original distribution L. However, the process of summing causes the scale of the distribution to change. To keep the scale parameter of the distribution of the sums the same as the scale parameter of the original distribution we have to weight each random variable in the sum by a constant $1/a_n$ such that:

$$\gamma\, n|(1/a_n)t|^{\alpha}=\gamma|\, t\,|^{\alpha}.$$

Solving this expression for a_n we get that $a_n=n^{1/\alpha}$. Thus either the summands are divided by $n^{1/\alpha}$ or the scale of the distribution of the unweighted sums is $n^{1/\alpha}$ times the scale of the original distribution. It follows that, for example, the interquartile range of the distribution S_n will be $n^{1/\alpha}$ times that of the distribution L. Range analysis is based on this property.

Let the difference between two symmetric percentiles define an interpercentile range of a distribution and denote the interpercentile range of S_n of by R_n. The interpercentile range of the distribution of the sum of n stable iid random variables as a function of the same interpercentile range of the distribution L is given by:

$$R_n=n^{1/\alpha}\, R_1.$$

Solving for α we have that:

$$\alpha=\frac{\log n}{\log R_n-\log R_1}.$$

12) This and the next subsection are based on Fama (1965b).

Thus, for a given set of data we can easily have many different estimates of α by applying the above formula for several values of n and using different interpercentile ranges.

However, this procedure will produce biased estimates of α if the sample price changes are correlated. For example, if there is positive serial correlation in the first differences, the interpercentile range of the distribution of the sums will be more than $n^{1/\alpha}$ the interpercentile range of the original distribution and the estimated α will be downward biased. In a similar way, negative serial correlation will result in upward biased estimates of α.

Fama (1965b) used this technique to estimate the characteristic exponent of 30 different stocks. The results obtained do not differ much from those obtained with double log graphing. And as 21 out of 30 of the estimated α's had values of less than 2, Fama considered this to be conclusive evidence in favor of the Mandelbrot hypothesis.

3.3 Sequential Variance

Although the population variance of a stable distribution with $\alpha<2$ is infinite, the sample variance will always be finite. Thus, as the size of a sample drawn from a population with a stable distribution is increased the sample variance should increase. To see why, let X be a stable iid random variable with characteristic, skewness, scale and location parameters $\alpha<2$, β, γ, and δ, and define a new variable $Y=X-\delta$. Y has exactly the same distribution as X except that the location parameter is 0. Y^2, however, has no negative tail, and its positive tail is related to the tails of the distribution of Y:

$$\Pr(Y^2>y)=\Pr(Y>y^{1/2})+\Pr(Y<-y^{1/2}).$$

Since Property 3 of stable distributions applies to Y it follows that:

$$\lim_{y\to\infty} \Pr(Y^2>y)\to(y^{1/2}/Z_1)^{-\alpha}+(y^{1/2}/Z_2)^{-\alpha}=(Z_1^{-\alpha}+Z_2^{-\alpha})y^{-\alpha/2}.$$

Thus Y^2 has a stable distribution with characteristic parameter $\alpha'=\alpha/2$. For the distribution of sums of n stable iid random variables to have the same distribution as the original random variable, the sums must be weighted by $n^{-1/\alpha'}=n^{-2/\alpha}$. In this way the distributions of Y^2 and $n^{-2/\alpha}\,\Sigma\,Y^2$ are identical. The variance of a sample with size n is:

$$S^2=n^{-1}\sum_{i=1}^{n} Y_i^2$$

This can be multiplied by $n^{-2/\alpha+2/\alpha}$ so that:

$$S^2=n^{-1+2/\alpha}(n^{-2/\alpha}\sum_{i=1}^{n} Y_i^2)$$

As we know that the distribution of $n^{-2/\alpha}\,\Sigma_{i=1}\,Y_i^2$ is stable and independent of n, we have that the variance of two samples with sizes n_1 and n_2 will be related in the following way:

$$S_2^2=S_1^2(n_2/n_1)^{-1+2/\alpha}$$

Solving this equation for α we get:

$$\alpha=\frac{2(\log n_2-\log n_1)}{2\log S_2-2\log S_1+\log n_2-\log n_1}$$

The above equation can be used for different sample sizes to obtain several estimates of α. For example, Fama (1965b) using this method estimated α for 30 different stocks by averaging 56 different estimates for each stock.

This procedure has several problems: the estimates are very sensitive to the values of n_1 and n_2; if the sample variance declines as the sample size increases the estimated α will be more than 2; and in general it produces estimates, for a same time series, that are so erratic as to be of quite doubtful value. These problems make sequential variance an inferior procedure to range-analysis.

3.4 The Fama-Roll method

Much of the empirical work concerning stable distributions that has been done in the past 25 years has been based on the method developed by Fama and Roll (1968, 1971).

Assume that a random variable X is symmetric ($\beta=0$) with parameters α, $\gamma=c^{\alpha}$ ($c=a_n$), and δ. It follows that the standardized variable $U=(X-\delta)/c$ is also stable with parameters α, $\gamma=c=1$, and $\delta=0$. The logarithm of the characteristic function for U is:

$$\ln \phi(t) = -|t|^{\alpha}$$

Except for the normal ($\alpha=2$) and Cauchy ($\alpha=1$) distributions, expressions for the density and cumulative distribution functions are not known. But by using series expansions presented in Bergstrom (1952), Fama and Roll (1968) developed numerical approximations for the cumulative distribution functions which they tabulated for several values of the characteristic exponent.

To estimate α the following procedure is used: first, $c=\gamma^{1/\alpha}$, the scale parameter, is estimated using:

$$\hat{c} = (1/1.654)(x_{0.72} - x_{0.28})$$

where $x_{0.72}$ and $x_{0.28}$ are the 72nd and 28th percentiles of the data distribution. This estimator for c is based on the fortuitous finding that $(x_{0.72}-x_{0.28})/\hat{c}$ lies within 0.4% of 1.654 for all values of α in the interval [1.0, 2.0] when $\beta=0$.

Then, an inter-percentile range z_f (for a large value of f, with $0<f<1$) is calculated with the aid of the following expression:

$$\hat{z}_f = (x_f - x_{1-f}) / 2\hat{c}$$

where x_f is the $f(n+1)$th order statistic used to estimate the f th percentile of the data distribution with size n. Which is the best value for f cannot be determined analytically[13] but Fama and Roll (1971) using Monte Carlo simulation conclude that generally values of f between 0.95 and 0.97 are the best for estimating α. Perhaps because $f=0.95$ is the value which reduces the sampling error of the inter-percentile range, most empirical work uses this value.

Finally, this $\hat{z}_f$ is referred to a table of percentiles of standardized symmetric stable distributions (Table 2 in Fama and Roll, 1968) to obtain the value of α whose fth percentile best matches $\hat{z}_f$.

To judge the stability of a distribution we can use the fact that if a sample of observations is drawn from a stable distribution then every non-overlapping sum of observations of that sample will have the same characteristic exponent. On the other hand, if the distribution is not stable but has finite second moments, the estimates of α should increase towards 2 as the number of observations in each sum increases (Fama, 1963).

3.5 The McCulloch method

The McCulloch (1986) method generalizes the Fama-Roll method, providing consistent estimators for all four parameters. Its advantages over the Fama-Roll method are that it relaxes the restrictions on α (that can now be in the range [0.6, 2.0]) and β (that can be within its

13) If on one hand there is an incentive to choose large values of f because the inter-percentile range increases and the characteristic of a distribution is better seen in the extreme tails, on the other hand the sampling dispersion of the inter-percentile range increases, a problem that will be more serious the smaller the sample.

full permissible range [−1, 1]), and that it eliminates the small asymptotic bias in the Fama-Roll estimators.

Define

$$\upsilon_\alpha=\frac{x_{.95}-x_{.05}}{x_{.75}-x_{.25}} \text{ and } \upsilon_\beta=\frac{x_{.95}+x_{.05}-2x_{.5}}{x_{.95}-x_{.05}}.$$

These indexes are independent of γ and δ. Let $\hat{\upsilon}_\alpha$ and $\hat{\upsilon}_\beta$ be the corresponding sample values:

$$\hat{\upsilon}_\alpha=\frac{\hat{x}_{.95}-\hat{x}_{.05}}{\hat{x}_{.75}-\hat{x}_{.25}} \text{ and } \hat{\upsilon}_\beta=\frac{\hat{x}_{.95}+\hat{x}_{.05}-2\hat{x}_{.5}}{\hat{x}_{.95}-\hat{x}_{.05}}.$$

The statistics $\hat{\upsilon}_\alpha$ and $\hat{\upsilon}_\beta$ are consistent estimators of υ_α and υ_β. The parameters α and β may be consistently estimated by:

$$\hat{\alpha}=\psi_1(\hat{\upsilon}_\alpha, \hat{\upsilon}_\beta) \text{ and } \hat{\beta}=\psi_2(\hat{\upsilon}_\alpha, \hat{\upsilon}_\beta),$$

functions that are tabulated in McCulloch (1986, Tables III and IV).

Estimation of the scale parameter c can be done by computing the following consistent estimator $\hat{c}$:

$$\hat{c}=\frac{x_{.75}-x_{.25}}{\phi_3(\hat{\alpha}, \hat{\beta})},$$

where the function $\psi_3(\hat{\alpha}, \hat{\beta})$ is tabulated in McCulloch (1986, Table V).

Other estimation methods besides those presented above exist (for example, Paulson, Holcomb and Leitch, 1975), but they are computationally more complex and do not seem to give more exact estimates. For this reason they have seldom be used in practice and are not presented here.

4. Implications for economic theory and practice

Whether an economic variable is distributed according to the normal or according to a stable distribution other than the normal can have implications in several fields of economic theory[14]. We will see now what are its implications for the Efficient Markets Hypothesis, risk theory and econometrics.

4.1 Efficient Markets Hypothesis

According to the Efficient Markets Hypothesis, price changes between two points in time reflect the influx of new information into the market during that time period. If a piece of news that arrived to the market during that time period did not cause an appropriate price movement when the next transaction occurred (or immediately if the market is so liquid that the time between transactions is very small) but only after a certain number of transactions (or some time) had elapsed, then that market would not be informationally efficient. By appropriate we mean that it reflects the impact on the economic fundamentals of the fact that gave rise to the news.

Informational efficiency is not dependent on the shape of the distribution of the effect of new information on price changes or returns. A market can be informationally efficient both when the effect of news on prices is Gaussian or stable Paretian. It can also be inefficient under both distributions.

If a market is informationally efficient, the distribution of the effect that each individual bit of information has on price changes may be

14) It has been sugested also that the study of the distribution of economic variables can be useful for governments when they are formulating their economic policies, namely in their foreign exchange management policies (Calderon-Rossell and Ben-Horim, 1982).

stable with constant parameters α and β but with different values for δ and γ. If the influence that each piece of information has on the price is combined in a simple additive manner, then the price changes from transaction to transaction will also be a stable distribution with the same values for α and β. And since price changes for any interval of time, be it one day or one year, are simple sums of the price changes that occurred during that interval, then the price changes for any interval will also have a stable distribution with parameters α and β.

Even if the distribution of the effects that information has on price changes is not stable, as long as it is asymptotically Paretian, and if those effects combine in a simple additive way and there are enough pieces of information, then the distribution of price changes between two transactions will also be stable. And if there are not enough pieces of information between transactions to insure that limiting stable distribution is achieved by the distribution of price changes that accompany each transaction, then as long as there are enough pieces of information during a day, a week or a month, then the price changes for these intervals will exhibit stable distributions.

On the other hand, a market can also informationally inefficient under the stable Paretian hypothesis. As Mandelbrot (1963b) has shown, the distribution of price changes or returns can be asymptotically stable under various types of aggregation of information that are not informationally efficient. Price changes between transactions may depend on a subset of the information arrived to the market during the time between transactions (for example, investors may pay attention solely to the piece of news what they regard as the most important, or with largest impact, and disregard the rest). Under this case, if the effects of individual pieces of information are asymptotically Paretian with characteristic exponent α, the distribution of the largest effect will also be asymptotically Paretian with the same characteristic exponent α. It follows that the Doeblin-Gnedenko conditions will be satisfied and that the distribution of price changes will be stable with the same value of the characteristic exponent α, provided that are a sufficiently large number of transactions in the interval of time considered (hour, day, etc.). And if the Doeblin-Gnedenko conditions are satisfied for that period, the price for longer intervals will be stable with the same vale of α.

The question of whether an economic variable has a stable distribution because the process that generates news is itself a stable distribution or because the behavior of economic agents makes it so does not seem to have an easy theoretical or empirical solution. Assume that, for example, stock prices p_t follow a simple random model:

$$p_t = p_{t-1} + z_t$$

where z_t is an independently distributed random variable with $E(z_t) = 0$. z_t can be either normally distributed ($\alpha=2$) or have a stable distribution ($\alpha<2$), but in either case it is not evident whether this is due to the process of stock trading or simply to the shadowing of the underlying process that generates the news. An example will make clear the distinction: stock prices depend on economic fundamentals that are influenced by natural (floods, etc.) and social (wars, discoveries, etc.) factors. The natural and social phenomena may have a generating process that has a normal or a stable or other distribution. If the trading of stocks is neutral in relation to the underlying distribution, z_t will have the same distribution as the natural and social processes. However, the process of stock trading may alter that distribution. If as some argue (Shiller 1981, 1989) stock prices are more volatile than the underlying fundamentals then that excess volatility is due to the process of trading: if the mentioned natural and social processes are normally distributed ($\alpha=2$) the trade induced volatility may make z_t have a stable distribution with $\alpha<2$. In more formal terms:

$$z_t=T(w_t)$$

where w_t is the distribution of the natural and social phenomena assumed to be iid with $\alpha=\alpha'$, $T(\bullet)$ a function that reflects the process of stock trading and z_t is as before but distributed with $\alpha=\alpha''<\alpha'$.

Concerning natural phenomena, although some are found to be normally distributed many others are found to be stable (Monrad and Stout, 1989). In what concerns social phenomena other than economics no recent literature was found on the subject, but Zipf (1941, 1949) presents evidence that most are distributed with fat tails. Concerning economic variables (other than stock prices and exchange rates), while some of those surveyed by Steinld (1965) seem to be best approximated by the normal, most do not seem so.

A possible test of the informational efficiency of the stock market would be to compare the distributional characteristics of the natural and social processes that might affect economic fundamentals, v_t, with the distributional characteristics of stock prices changes, z_t[15]. However, as v_t is not easily observable this test can be difficult to implement. Another possibility is to assume that the information producing industry is both efficient and rational, giving proportionally more coverage to events that have more impact on economic fundamentals, in this way allowing stock market participants to have an undistorted picture of what affects fundamentals. If thus the news are assumed to

15) This type of test would in fact be a test of the informational efficiency of the stock market and the information industry.

have the same distributional characteristics as the phenomena that affect fundamentals, we could test if those characteristics are similar to the distributional characteristics of prices changes: if they are not it would be apparent that stock trading was not being rational or efficient. The same test would also allow us to form an idea of whether stock trading was adding or subtracting instability to the volatility of economic fundamentals. The problem of the difficulty of observation of the distributional characteristics of news remains but it should not be as hard to tackle as the observation of the distributional characteristics of the underlying natural and social phenomena, as psychologists have already developed reliable and valid means of analyzing and scoring the content of texts (McClelland 1962, Hampton, Summer and Webber 1987).

4.2 Risk theory[16)]

The importance of the concept of risk in economic theory and its relevance in such areas as choice of output level by a firm, saving and investment behavior, and option pricing, just to mention a few, is so evident that no further comment is needed here.

However, due to the traditional association of risk with variance (that holds only under very strict conditions), the relationship between risk and the shape of the distribution of a random variable seems to be forgotten. It is the objective of this section to stress that choice under uncertainty can be conceived as the choice of one out of several random variables with different shapes, a point stressed already half a century ago by Alchian (1950).

Following Knight (1921) it is usual to distinguish risk from uncertain-

16) This section borrows from Machina and Rothschild (1992) and Rothschild and Stiglitz (1970, 1971).

ty. A situation is said to involve uncertainty if an economic agent cannot, or does not, assign a numerical probability to the different possible occurrences. On the other hand, a situation where the randomness facing the agent can be expressed in terms of concrete numerical probabilities is said to involve risk.

An agent facing a decision under risk can be thought of as facing a choice of alternative univariate probability distributions (of returns, for example). It is usual to assume that such an agent can rank all possible distributions that he faces, and that such ranking is complete, transitive and, in an appropriate sense, continuous. Let V(•) be a real-valued preference function over the set of cumulative distribution functions F(•). An usual specification for V(•) is the Riemann-Stieltjes integral:

$$V(F) \equiv \int U(x)\,dF(x) \qquad (4)$$

for a utility function U(•). Equation (4) is an expected utility model of preferences over possible random states. The shape of U(•) reflects the attitudes of the agent toward risk, and is assumed to be an increasing and concave function of x. Thus, for a given shape of the distribution, the agent will chose the one with higher expected value. And when comparing distributions with a given expected value, he will prefer the distribution which is more concentrated around the expected value if he or she is risk averse. Thus, when facing a choice between different stable distributions with the same expected value, a risk averse agent will chose the one with thinner tails ($\alpha=2$).

The usual measure of the riskiness of a random variable is the variance, or alternatively, the standard deviation. The easiness of their use and interpretation has led to the widespread use in finance of mean-variance analysis and to the development of modern portfolio theory by Markowitz (1952, 1959) and Tobin (1958), and the capital asset pricing model of Sharpe (1965) and Litner (1966), based on it.

However the mean-variance approach has theoretical as well as empirical weaknesses. The most important theoretical objection is that an expected utility maximizer would evaluate all distributions only on the basis of their means and variances if and only if his utility function took the form $U(x) \equiv ax + bx^2$, where b indicates the degree of risk aversion. However the assumption that the utility function takes a quadratic form is disputable because if the agent is risk averse ($b<0$) his utility will decrease as x increases beyond a certain point ($a/2b$), and because the agent will be more averse to constant additive risks in high levels of x than in low levels (what is in contrast, for example, with the empirical observation that those with greater wealth take greater risks).

The traditional view held that when comparing two distributions with the same mean, the following four definitions of risk are equivalent:

1. Y is equal to V plus noise If three random variables Y, V and Z are related in the following way:

$$Y \overset{d}{=} V+Z$$

with $E(Z|V)=0$ for all V, and $\overset{d}{=}$ indicates the same distribution, then Y is more variable or riskier than V.

2. Every risk averter prefers V to Y If Y and V have the same mean but any risk averter prefers V to Y in such a way that:

$$\int U(x)\,dF_y(x) \leqq \int U(x)\,dF_v(x)$$

where $F_y(\bullet)$ and $F_v(\bullet)$ are the cumulative distribution functions of Y and V, then it is reasonable to say that Y is more risky than V.

3. Y has more weight in the tails than V If the density function of Y is obtained by transferring some mass or probability weight from around the center of the density function of V towards its tails in such a way as to leave the mean unchanged, then Y is more risky than V.

4. Y has a greater variance than V If the variance of Y is larger than the variance of V, then Y is more risky than V.

Rothschild and Stiglitz (1970) show that while the first three definitions of risk are in fact equivalent, they are different from the fourth. Thus, as the tails of a stable distribution get fatter (as α decreases from 2 towards 0) risk increases. But although riskier distributions imply higher variance, higher variance does not imply riskier distributions.

The empirical weakness is the observation that most variables in economics and especially in finance seem to be distributed not according to the normal but according to other stable distributions. As stable distributions have not defined second and higher moments, the sample variance and the standard deviation do not give an accurate measure of dispersion. This, of course, does not mean that that concepts such as diversification are meaningless in a securities market where returns have stable distributions with $\alpha<2$. Fama (1965a, 1971), using concepts of variability other than the variance, has shown that it is possible to develop a model for portfolio analysis for such markets[17].

As a result of problems that the use of the variance as a measure of risk has, other measures of risk have been proposed. These include,

17) It can be argued that although the population variances are infinite, since sample variances of returns are finite they can be used in Markowitz-type portfolio analysis. However, as this model is highly sensitive to the estimates of the variances that are used, the estimates thus produced are highly unreliable.

among others, the mean absolute deviation and the interquartile range.

4.3 Econometrics

The implications for statistical and econometric work of stable distributions lies in the non-existence of finite variance. In practical terms this means that the sample variance of a stable distribution will show a extremely erratic behavior even for very large sample sizes. Because of its very erratic behavior the sample variance is not a meaningful measure of the variability of a stable distribution other than the normal. Unfortunately stable distributions are not the only source of potential trouble to econometricians: when there is any non-normality in the data, even if it arises in a distribution with finite variance, it is possible to find estimators that are more efficient than least square estimators (Judge, Hill, Griffiths, Lutkepohl and Lee, 1988).

Thus an elementary precaution that should be taken any time OLS is employed is to test the normality of the residuals. The literature on testing for normality is vast and is beyond the scope of this work. The interested reader is referred to, for example, White and MacDonald (1980), where several well-known and easily computable statistics for testing normality and a selected bibliography of this field are presented. However, the Jarque-Bera test has recently gained popularity. Bera and Jarque (1981) propose the following statistic that, under the null hypothesis that residuals are normally distributed, has an asymptotic $\chi^2_{(2)}$ distribution and is given by:

$$\lambda = (T - k/6)[S^2 + (1/4)(K-3)^2],$$

where T is the number of observations, k is the number of regressors (zero for a non-regressed series), and S and K are respectively skewness and kurtosis.

If non-normality is found, further tests should be made to ascertain whether it is due to a stable distribution[18], or to other distributions. If it is judged that the later is the case, most econometric textbooks argue that, since the OLS estimator remains unbiased minimum variance from within the class of linear unbiased estimators and consistent, and since the conventional tests are still asymptotically justified, then least squares estimators can still be used under conditions of non-normality. However, Koenker (1982) argues that these reasons are not very compelling.

If a data set is judged to have a stable distribution, methods that assume finite variance are liable to produce misleading results. Thus when modeling data whose distribution is known to belong to this class, it does not make much sense to use a regression technique that has as its criterion the minimization of the sum of squared residuals from the estimated regression line, since the expectation of that sum will be infinite. In particular the OLS estimators which assume normal and homoskedastic residuals are inefficient and the confidence intervals associated with these estimates are incorrect and can be very misleading. In this case it makes more sense to use other methods of linear regression, such as the maximum likelihood (McCulloch, 1979). Also superior to OLS in the presence of stable distributions are robust regression methods, but since the distribution of estimators is generally not well known hypothesis testing is difficult. Among the robust regression methods superior to OLS we can refer absolute-value regression (Wagner 1959, 1962, Blattberg and Sargent, 1971, Dielman and Pfaffenberger, 1982), which do not make use of second or higher order moments statistics, and Trimmed Least Squares (Koenker and Bassett,

18) See for example the tests presented in sub-sections 3.4 and 3.5.

1978, 1982, and Ruppert and Carroll, 1980).

5. Stable distributions in economic literature

There is a vast and vigorous literature on using stable distributions to model random economic phenomena. In economics, the justification for the use of a stable non-normal instead of a normal distribution depends heavily on data-based evidence, namely that the tails of the distributions generating the data have a certain non-normal shape. Unlike the physical sciences there is no strong theoretical reason to some phenomena being distributed according to a stable instead of the normal distribution (or vice-versa). If on one hand there is ample evidence that economic phenomena are not normally distributed, on the other hand tradition, convenience and mathematical elegance (to what can be added the analytical intractability of stable distributions) have made common the assumption that economic variables are normally distributed. In what follows we present the results of a part of the literature concerning stable distributions in four areas of economics: income and wealth distribution, foreign exchange, currency futures and stock prices. However, stable distributions are also useful in other areas like population economics, urban and regional economics, etc.

5.1 Income and wealth distribution

This is an area of special significance to stable distribution theory because it was work by V. Pareto in income distribution that lead to the discovery of the empirical relationship, later denominated "Pareto's law"[19], presented above as Property 3 in Section 2. It was also work

19) It was, however, Levy (1924) who iniciated the general theory of stable distributions by finding the Fourier transforms of all stictly stable distributions. General interest in stable distributions was later stimulated by work of

（次頁へ続く）

in this area by Mandelbrot (1960, 1961, 1962) that later rekindled economists interest in this family of distributions.

Pareto found that the distribution of income is not strongly influenced either by the socio-economic structure of the society under analysis or by the definition chosen for income, and that this observation is truer the more one restricts attention to the higher range of values of income. Socio-economic structure and definition of income can at most influence the values taken by certain parameters of an apparently universal distribution law. Pareto thought that $\alpha=1.5$ was a good general approximation that described well most distributions of income. Later, stable distributions were found empirically to fit well the upper portion of income distribution of populations of so different socio-economic backgrounds and sizes as the "burghers" of certain Renaissance city-states, numbering only a few hundred, and the taxpayers in the USA, numbering about a hundred million (Mandelbrot, 1960).

As α increases and as the goodness of the fit of income distribution extends from the upper tail to median incomes, it can be considered that the distribution of income becomes more egalitarian. One famous example pertains not to income but to wealth distribution: it was found that in Sweden α increased from 1.5 to 1.7 between 1931 and 1959, and that the range of the straight-line distribution[20], which included incomes above 200,000 Swedish crowns in 1931, had extended to the left and that in 1959 the distribution of incomes above 100,000 Swedish crowns (including over 70% of the wealth owning population) had

W. Doblin on domains of attraction in 1939. In the forties Zipf (1941, 1949) presented many cases where the tails of of phenomena studied in the social sciences are well fitted by stable distributions. In the sixties Mandelbrot reintroduced and developed their use in economics.

20) See sub-section 3.1 on double log graphing for explanation.

become a straight line (Steindl, 1965). Figure 3 presents the double log graph of wealth distribution in Sweden in 1959.

5.2 Foreign exchange

Westerfield (1977) examined weekly foreign exchange rates, including spot rates and one, two and three month forward rates, for five currencies (those of the Netherlands, Switzerland, United Kingdom, Germany and Canada) against the US dollar, encompassing both the fixed exchange rate regime in the 1960's and the early floating regime of the 1970's. After finding that the weekly returns of those exchange rates were symmetric but highly leptokurtic, Westerfield compared the distributions of the actual exchange rate returns with the theoretical distributions of several stable distributions (for α=1.2, 1.3, ..., 2.0) using Chi-squared tests, and concluded that the normal distribution provided an inadequate description of the observed distributions when compared with the stable distributions for all cases analyzed. This observation is confirmed by the values of the estimated characteristic exponent presented in Table 3 (Westerfield, 1977, pp 191-192), were all the estimated α's for all the spot and forward rates have values lower than 1.6 for the two exchange rate regimes. These results seem to be stable as can be observed from the estimated values of α for the sums of two, five and ten non-overlapping adjacent observations (presented in Table 4, Westerfield, 1977, p. 194): although the estimated values of α show a tendency to increase with the number of summands, that tendency is weak.

Similar results for spot and forward exchange rates for other sample periods and exchange rates are reported by: Farber, Roll and Solnik (1977), who used weekly and monthly data for the period from 1957 to 1975 (divided in two subperiods by the end of March 1971) for the exchange rates of 17 different currencies against the US dollar, and found that all had distributions with $\alpha<2$, except for Brazil and Spain during the floating rate period; Rana (1981); McFarland, Pettit and Sung (1982), who used daily data for the period from 2 January 1975 to 29 June 1979 for the spot and forward of 7 different currencies against the US dollar, and found that that all had distributions with $\alpha<1.5$; and Calderon-Rossell and Ben-Horin (1982) who tested the hypothesis that the distribution of the exchange rates of 14 different currencies against the US dollar had characteristic exponents $\alpha=1$, $\alpha=1.5$ and $\alpha=2$ (against the alternatives $\alpha\neq1$, $\alpha\neq1.5$ and $\alpha\neq2$) using the Kolmogorov-Smirnov test for data from 1 July 1974 to 29 July 1977, could not reject the hypothesis that $\alpha=1.5$ for 9 of the exchange rates at the usual levels of significance; for the remaining 5 all three hypothesis could be rejected, what in conjunction with the fact that these distributions seemed to have a high degree of asymmetry led the authors to hypothesize that they could belong to the stable (Paretian) asymmetric family.

Coppes (1995) using daily data for the cross exchange rates of the US dollar, Japanese yen, German mark and British pound during the 1980-1992 period found also that all series suffered from leptokurtosis (however, he does not report testing for symmetry). But in contrast with results obtained by Westerfield (1977), although the daily rates of return were found to have stable distributions (with α between 1.50 and 1.60), the distributions of non-overlapping sums of adjacent observations tended to the normal. Two exceptions were found to this tendency: the characteristic exponent of the Japanese yen / British pound rate showed a slow increase as more observations were aggregated, while the characteristic exponent of the German mark / British pound rate remained stable around 1.50. Coppes attributed this later result to the Exchange Rate Mechanism of the European Monetary System because, he argued, this mechanism limited the changes of the German mark British pound exchange rate during a considerable part of the observed period[21), what may have induced dependence between daily price changes. To test the stability of the characteristic exponent, Coppes used the procedure proposed by Hall, Brorsen and Irwin (1989) of drawing randomly the daily rates of return from the original series and only then summing, and obtained results that show that the characteristic exponent of the distributions of the rates of return of the Japanese yen / British pound and the German mark / British pound exchange rates tend to two when eight daily returns are summed. This result led him to conclude that there was indeed lack of independence of consecutive price changes in the mark / pound exchange rate that might be attributable to the participation of the British pound in the Exchange Rate Mechanism during a small period in the sample under examination.

Against this interpretation DosSantos (1996) tested the independence of the daily returns of thirteen bilateral exchange rates using the Ljung-Box (1978) Q-statistics. The results of this test showed that the daily returns of all exchange rate series for the period from March 1, 1973 to July 30, 1993 suffered from autocorrelation, and thus of lack of independence. Since there is lack of independence in the daily returns of all exchange rate series, it follows that this cannot be the reason for the stability of the characteristic exponent for the mark / pound exchange rate found by Coppes, because this argument would imply that the characteristic exponent of other exchange rates would

21) In fact, as it entered in early October 1990 and left it in middle September 1992, the British pound belonged to the Exchange Rate Mechanism for less than 24 months out of the 156 months in Coppes data set, a mere 15% of the total.

also be stable under addition, what does not happen. This conclusion is reinforced by the fact that the only exception to the referred lack of independence of daily returns found for all exchange rates is precisely the pound / mark exchange rate for the period when the pound belonged to the Exchange Rate Mechanism (lack of independence was found also for other subperiods when the pound did not participate in this Mechanism).

A further result presented in DosSantos (1996) is that where exists an explicit or implicit arrangement to link one currency to another, the distribution of daily returns of that distribution seems to be stable. This result was found for all exchange rates involving the German mark against currencies belonging to the Exchange Rate Mechanism and also for the US dollar / Canadian dollar rate. But where such an arrangement does not exist, the characteristic exponent tends to two, what may indicate that the distribution of those exchange rates are a mixture of distributions with finite variance (perhaps a mixture of normals). The exchange rates of the US dollar against the European currencies and the Japanese yen, as well as the yen / mark exchange rate where found to be in this case.

These results are in line with the view that holds as a misconception the widespread idea that fixed exchange rate regimes such as the Bretton Woods system and the Exchange Rate Mechanism of the European Monetary System made possible a relatively inflation-free growth because they brought a high degree of certainty into international transactions which fostered trade and capital flows among the nations involved. As the distributions of exchange rate returns of currencies that are pegged seem to have fatter tails than those that are not, then fixed exchange regimes may imply more, not less, risk. A possible explanation to this counter intuitive result may lay in the necessity of realignments, or change of the fixed exchange rates, that arises from time to time when different countries pursue diverse macroeconomic policies. For example, between the inception of the Bretton Woods system at the end of World War II and its collapse in August 15, 1971, there were no less than 74 exchange rate changes involving a total of 45 countries; and from March 1979, when it started, to May 1993 there were 17 realignments in the Exchange Rate Mechanism of the European Monetary System, not including the ejection of the British pound and Italian lira in September 1992. As Fritz Machlup is said to have once remarked, fixed-rate systems look like being actually jumping-rate systems (Malabre, 1994). As the probability of large absolute returns (jumps) increases in fixed exchange rate regimes, so increases the associated exchange rate risk. Only a system that can absolutely guaranty the inexistence of future realignments can be said to reduce risk. In this light it is understandable that the European Union seeks this guaranty by eliminating the national currencies for a single currency and by imposing imperative limits on the values certain macroeconomic variables can attain in the different countries of the Union.

5.3 Currency Futures

Although there is a vast literature on the shape of the distribution of spot and forward exchange rates[22], relatively few studies examine the statistical properties of the distribution of futures price changes.

So (1987), using the McCulloch method, estimated α and β for the distribution of daily returns of future prices for the British pound, Canadian dollar, the Swiss franc, the German mark and the Japanese yen. He found that the characteristic exponent for all contracts (March,

22) See the previous sub-section for references.

June, September and December) for a certain currency have similar values with $\alpha<2$ and β around 0. He then proceeded to test whether the observed stable distributions can be explained by the inverse relationship between maturity and variability proposed by Samuelson (1965) and Mandelbrot (1966): since the characteristic exponent (for a certain scale parameter) determines the probability of extreme results, its decrease reveals an increasing probability of large and sudden changes. Thus So estimated α and β for each of twelve months to maturity for the several contracts and currencies and found that generally α decreased with maturity. This result did not change after correcting for the small variability found in the estimated scale parameter, allowing to accept the Samuelson-Mandelbrot hypothesis.

5.4 Stock prices

The usual assumption that the distribution of price changes is normal or approximately so is due to Bachelier (1900) and Osbone (1959), who used arguments based on the central-limit theorem to support it. However, it is well established that the normal distribution fails to represent the distribution of stock price returns properly. The first works to argue that stable distributions might fit better stock price returns than the normal distribution are Mandelbrot (1963a) and Fama (1963). Fama (1965) who, using the first three methods presented in Section 3, was the first to report actual estimates for the characteristic exponent of 30 US stocks, concluded for their non-normality.

Teichmoeller (1971) applied the Fama-Roll method to 30 alphabetically chosen US stocks to the original series and to the sum of non-overlapping sums of 2, 5 and 10 observations. These series were corrected for stock dividends and stock splits, and only calendar day returns were used (thus excluding the returns between Friday and

Monday). The estimated characteristic exponents were found to have a mean, that slightly increased with the sums, in the 1.6 to 1.7 range.

However, Fielitz and Smith (1972) using the daily returns of 200 US stocks, adjusted for stock splits, stock dividends and cash dividends, compared for various class intervals the number of observations of the empirical stock distributions with the number of observations expected to occur under a Gaussian hypothesis. They concluded that the asymmetric stable distributions are more appropriate than either the normal or stable distributions and that the Fama-Roll method (1968, 1971) should not be applied in this instance.

Leitch and Paulson (1975) report the estimates of the characteristic exponent for 20 US stocks using the graphical, Fama and Roll and another method that does not restrain the distribution to be symmetric (Paulson, Holcomb, and Leitch, 1975). They conclude that the agreement of estimates of the characteristic exponent are very good when using the three methods, when the restriction $\beta=0$ is made in the third method. Even when this restriction is not made, the effect of skewness (including those cases when the estimated β's are very much different from zero) on the estimates of α is not large (generally the estimates of the Paulson, Holcomb, and Leitch (1975) method when $\beta\neq0$ are even nearer of those of the Fama-Roll method than when the restriction $\beta=0$ is made).

From the results above we might conclude that stock price returns are better represented by stable distributions, either symmetric or non-asymmetric, than by the normal distribution; and that asymmetry does not seem to cause much of a problem when estimating the characteristic exponent with the Fama-Roll method.

6. Is the Nikkei 225 stable Paretian?

From the FOREX data base (made available by the *Nihon Keizai Shinposha*) the daily values of the Nikkei 225 average were collected for the period from October 31, 1978 to July 30, 1993, with a total of 4001 observations. This series is presented in Figure 4. From this data the daily returns were computed. Figure 5 presents their evolution. Their mean, standard deviation, skewness, kurtosis, Ljung-Box

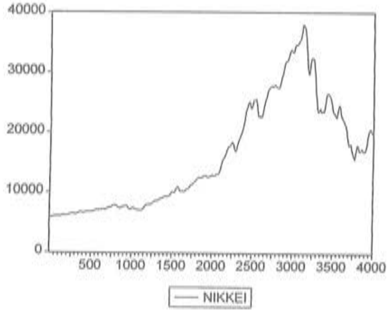

Fig. 4 Evolution of the Nikkei 225 between October 31, 1978 and July 30, 1993

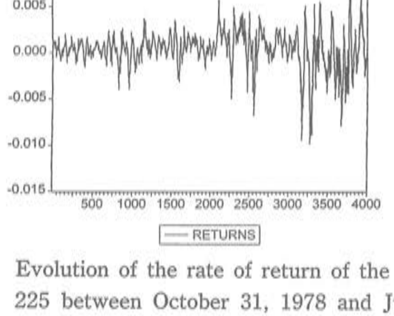

Fig. 5 Evolution of the rate of return of the Nikkei 225 between October 31, 1978 and July 30, 1993

Q-statistics Q(12) and Q(24)[23] were computed and are presented in Table 1.

The mean is approximately zero as expected. But the distribution shows every sign of not being normal: it is skewed to the left[24]; it is heavy tailed; and the p-value of the Jarque-Bera statistic is 0. There is also evidence of autocorrelation, as the p-values of the estimated Q-values are zero in both cases. A careful look at the histogram of the daily returns in Figure 6 will confirm some of these results: the distribution seems slightly skewed to the left (the left tail is fatter and somewhat longer than the right tail), and to be fat tailed (it seems more with distribution α_1 than with distribution α_2 in Figure 1).

While in principle the Fama-Roll method should be applied only to

Table 1 Descriptive statistics

Mean	3.1 E-04
Std. Deviation	1.9 E-03
Skewness	−0.962
Kurtosis	6.461
Q(12)	31023
Q(24)	39755
Jarque-Bera	2613.24

23) The choice of the span is arbitrary: for large spans some power of the test is lost, but for small values highly significant correlations at relatively high lags is not captured: thus our choice of a relatively short span (12) and a relatively large one (24).

24) To confirm the lack of symmetry of this distribution we also tested the null hypothesis that half of the observations are below the mean using the statistic $s=[(n^-/n)-0.5]\sqrt{n}$, where n^- is the number of observations below the mean and n is the total number of observations, with $s \sim N(0, 1)$; as $n^-=1819$, we have that $s=-2.9$, so the null can be rejected for the usual significance levels.

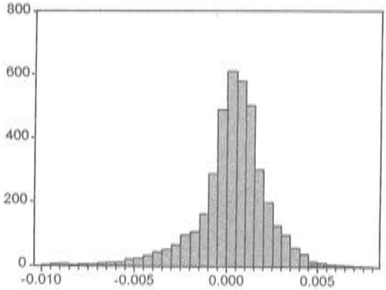

Fig. 6 Histogram of the rate of return of the Nikkei 225 between October 31, 1978 and July 30, 1993

Table 2 Estimates of α and β for the sums several consecutive non-overlapping observations

	Fama-Roll method	McCulloch method	
Sums of	$\hat{\alpha}$	$\hat{\alpha}$	$\hat{\beta}$
1	1.38	1.36	−0.24
2	1.37	1.36	−0.22
4	1.39	1.38	−0.22
5	1.36	1.36	−0.21
8	1.38	1.38	−0.22
10	1.33	1.37	−0.23
16	1.31	1.31	−0.26
20	1.32	1.31	−0.33

symmetrical distributions, we used it as well as the McCulloch method to estimate the characteristic exponent of the distribution of daily returns of the Nikkei 225 presented in Table 2.

It can be noticed first that both methods yield very similar results for α, what may suggest that some amount of skewness, as long as it is not large, does not invalidate the results obtained with the Fama-Roll method, as already noted in the previous section. Then we notice that the estimates of α are quite stable for the diverse sums, in the range between 1.3 and 1.4. The estimates for β obtained with the McCulloch method show that the distribution is in fact negatively skewed, as expected.

7. Conclusions

This paper presented the existing theory, estimation methods and principal empirical results concerning stable distributions in some economic fields.

After analyzing the results reported in the previous sections it is apparent that stable distributions seem to fit better several types of economic data than the normal distribution, and that the evidence is especially abundant and strong for foreign exchange and stock prices time series. But, since there are no explicit expressions for the densities of stable distributions (with the exception of the Cauchy and Gaussian distributions) no attempt has been made to use them in theoretical modeling of economic behavior. As this situation does not seem likely to be altered in the near future, the normal distribution of economic variables will continue to be a standard assumption in most economic areas.

However, as a certain number of observers has noted (Leamer, 1983, Rosenberg, 1992), to be able to attain its objective of explaining human behavior, and to be useful in advising policy, economics must become more and more an empirically progressive discipline. This implies that it has to conform its assumptions more realistically to the observed facts and to go beyond generic predictions to achieve an increasing number of specific predictions, and these predictions must be made with more and more precision.

References

Alchian, A. A., 1950, Uncertainty, evolution, and economic theory, Journal of Political Economy, June 1950, 211-221.

Bachelier, L., 1900, Theory of speculation, in Cootner, P. H. (ed.) The Random Character of Stock Market Prices. The MIT Press (1964).

Bera, A. K., and C. M. Jarque, 1981, An efficient large-sample test for normality of observations and regression residuals, Australian National University Working Papers in Econometrics, No. 40.

Bergstrom, H., 1952, On some expansions of stable distributions, Arkiv for Matematik, 375-378.

Blattberg, R. C., and T. Sargent, 1971, Regression with non-gaussian stable disturbances: some sampling results, Econometrica 39, 501-510.

Calderon-Rossel, J. R., and M. Ben-Horim, 1982, The behavior of foreign exchange rates, Journal of International Business Studies, 99-111.

Chambers, J. M., C. L. Mallows and B. W. Stuck, 1976, A method for simulating stable random variables, Journal of the American Statistical Association 71, 340—344.

Coppes, R. C., 1995, Are exchange rate changes normally distributed?, Economic Letters 47, 117-121.

Davidson, R. and J. G. MacKinnon, 1993, Estimation and Inference in Econometrics, (Oxford University Press, Oxford).

Dielman, T. and R. Pfaffenberger, 1982, LAV (Least Absolute Value) estimation in linear regression: a review, TIMS Studies in the Management Sciences 19, 31-52.

Dos Santos, J. M. P., 1996, Which exchange rates are stable Paretian?, Mimeo, Hiroshima University.

Fama, E. F., 1963, Mandelbrot and the stable Paretian hypothesis, Journal of Business 36, 420-429.

Fama, E. F., 1965a, Portfolio analysis in a stable paretian market, Management Science.

Fama, E. F., 1965b, The behavior of stock market prices, Journal of Business 38, 34-105.

Fama, E. F., 1971, Risk, return and equilibrium, Journal of Political Economy 79, 30-55.

Fama, E. F. and R. Roll, 1968, Some properties of symmetric stable distributions, Journal of the American Statistical Association 63, 817-836.

Fama, E. F. and R. Roll, 1971, Parameter estimates for symmetric stable distributions, Journal of the American Statistical Association 66, 331-338.

Farber, A., R. Roll, and B. Solnik, 1977, An empirical study of risk under fixed and flexible exchange, Journal of Monetary Economics, 235-265.

Feller, W., 1968, An introduction to probability theory and its applications, Vol. 1 (Wiley, New York).

Feuerverger, A., and P. McDunnough, 1981, On efficient inference in symmetric stable laws and processes, in Statistics and Related Topics, edited by M. Csörgö, D. A. Dawson, J. Rao, and A. Saleh, (North-Holland).

Fielitz, B. D. and E. W. Smith, 1972, Asymmetric stable distributions of stock price changes, Journal of the Americal Statistical Association, 67, 813-814.

Gnedenko, B. V. and A. N. Kolmogorov, 1948, Limit distributions for sums of independent random variables (Addison-Wesley, 1954).

Hall, J. A., W. Brorsen, and S. H. Irwin, 1989, The distribution of futures prices: A test of the stable Paretian and mixture of normals hypothesis, Journal of Financial and Quantitative Analysis 24, 105-116.

Hall, P., 1981, A comedy of errors: the canonical form for a stable characteristic function, Bulletin of the London Mathematical Society, 13, 23-27.

Hampton, D. R., C. E. Summer, and R. A. Webber, 1987, Organizational Behavior and the Practice of Management, 5th Ed. (Scott, Foresman and Company).

Judge, G. G., R. C. Hill, W. E. Griffiths, H. Lütkepohl, T.-C. Lee, 1988, Introduction to the Theory and Practice of Econometrics, 2nd. Ed., (John Wiley & Sons).

Kendal, M. G. and A. Stuart, 1977, The Advanced Theory of Statistics, Vol. 1, 4th ed (Charles Griffin, London).

Koenker, R. W., 1982, Robust methods in econometrics, Econometric Reviews 1, 213-290.

Koenker, R. W., and G. W. Basset, 1978, Regression quantiles, Econometrica 46, 33-50.

Koenker, R. W., and G. W. Basset, 1982, Robust tests for heteroscedasticity based on regression quantiles, Econometrica 50, 43-62.

Knight, F., 1921, Risk, Uncertainty and Profit, (Houghton Mifflin Co.).

Laha, R. G., 1989, Characteristic functions, Encyclopedia of Statistical Sciences, edited by Samuel Kotz, Norman L. Johnson and associate editor, Campbell

B. Read, (Wiley, New York).

Leamer, E. E., 1983, Let's take the con out of econometrics, American Economic Review 73, 31-43.

Leitch, R. A. and A. S. Paulson, 1975, Estimation of stable law parameters: stock price behavior application, Journal of the American Statistical Association 70, 690-697.

Levy, P., 1925, Calcul des probabilités, (Gauthier Villars).

Litner, J., 1965, The valuation of risk assets and the selection of risky investments in stock portfolios and capital budgets, Review of Economics and Statistics 47, 13-37.

Ljung, G. M., and G. E. P. Box, 1978, On a measure of lack of fit in time series models, Biometrika 65, 297-303.

Machina, M. J. and M. Rothschild, 1992, Risk, The New Palgrave Dictionary of Money and Finance (MacMillan, London).

Malabre Jr., A. L., 1994, Lost prophets, (Harvard Business School Press).

Mandelbrot, B. 1960, The Pareto-Lévi law and the distribution of income, International Economic Review 1, 79-106.

Mandelbrot, B., 1961, Stable Paretian random functions and the multiplicative variation of income, Econometrica 29, 517-543.

Mandelbrot, B., 1962, Paretian distributions and income maximization, Quarterly Journal of Economics 76, 57-85.

Mandelbrot, B., 1963a, The variation of certain speculative prices, The Journal of Business 36, 394-419.

Mandelbrot, B., 1963b, New methods in statistical economics, Journal of Political Economy 71, 421-440.

Mandelbrot, B., 1963c, The stable Paretian income distribution when the apparent exponent is near two, International Economic Review 4, 111-114.

Mandelbrot, B., 1966, Forecasts of future prices, unbiased markets, and martingale models, Journal of Business 39, 242-265.

Markowitz, H., 1952, Portfolio selection, Journal of Finance 7, 77-91.

Markowitz, H., 1959, Portfolio Selection: Efficient Diversification of Investment. (Yale University Press).

McClelland, D. C., 1962, Business Drive and National Achievement, Harvard Business Review 40(4), 99-112.

McCulloch, J. H., 1979, Linear regression with symmetric stable disturbances, Ohio State University Working Paper, No. 63.

McCulloch, J. H., 1986, Simple consistent estimators of stable distribution parameters, Commun. Statist.-Simula. 15, 1109-1136.

McFarland, J., R. Pettit, and Sam Sung, 1982, The distribution of foreign exchange prices; Trading effects and risk measurement, Journal of Finance, 693-715.

Monrad, D., W. Stout, 1989, Stable distributions, Encyclopedia of Statistical Sciences, edited by Samuel Kotz, Norman L. Johnson and associate editor, Campbell B. Read, (Wiley, New York).

Paulson, A. S., E. W. Holcomb, and R. A. Leitch, 1975, The estimation of the stable laws, Biometrika 62, 162-169.

Rana, P., 1981, Exchange rate risk under generalized floating Eight Asian countries, Journal of International Economics 11, 459-466.

Rosenberg, A., 1992, Economics-Mathematical Politics or Science of Diminishing Returns?, (University of Chicago Press).

Rothschild, M. and J. Stiglitz, 1970, Increasing risk: I. A definition, Journal of Economic Theory 2, 225-243.

Rothschild, M. and J. Stiglitz, 1971, Increasing risk: II. Its economic consequences, Journal of Economic Theory 3, 66-84.

Ruppert, D. and J. Carroll, 1980, Trimmed least squares estimation in the linear model, Journal of the American Statistical Association 75, 828-838.

Samuelson, P., 1965, Proof that properly anticipated prices fluctuate randomly, Industial Management Review (Spring 1965), 41-49.

Sharpe, W., 1964 Capital asset prices: a theory of market equilibrium under conditions of risk, Journal of Finance 19, 425-442.

So, J. C., 1987, The sub-Gaussian distribution of currency futures: stable Paretian or nonstationary?, The Review of Economics and Statistics 69, 100-107.

Steindl, J., 1965, Random processes and the growth of firms: a study of the Pareto law, (Griffin, London).

Stout, W., 1989, Domain of attraction, Encyclopedia of Statistical Sciences, edited by Samuel Kotz, Norman L. Johnson and associate editor, Campbell B. Read, (Wiley, New York).

Stuart, A. and J. K. Ord, 1987, Kendall's advanced theory of statistics, Vol. 1, Distribution theory, 5th ed. (Charles Griffin, London).

Shiller, R. J., Do stock prices move too much to be justified by subsequent changes in dividends?, American Economic Review 71, 421-435.

Shiller, R. J., 1989, Market Volatility, (The MIT Press).

Teichmoeller, J., 1971, A note on the distribution of stock price changes, Journal of the American Statistical Association 66, 282-284.

Tobin J., 1958, Liquidity preference as behavior towards risk, Review of Economic Studies 25, 65-86.

Wagner, H. M., 1959, Linear programming techniques for regression analysis, Journal of the American Statistical Association, 54, 206-212.

Wagner, H. M., 1962, Non-linear regression with minimal assumptions, Journal of the American Statistical Association, 57, 572-578.

Westerfield, J. M., 1977, An examination of foreign exchange risk under fixed and floating rate regimes, Journal of International Economics 7, 181-200.

White, H. and G. M. MacDonald, 1980, Some large-sample tests for non-normality in the linear regression model, Journal of the American Statistical Association 75, 16-28.

Zipf, G. K., 1941, National Unity and Disunity, (Principia Press).

Zipf, G. K., 1949, Human Behavior and the Principle of Least-Effort, (Addison-Wesley Press).

Zolotarev, V. M., 1957, Mellin-Stieljes transforms in probability theory, Theory of Probability and Its Application 2, 433-460.

Zolotarev, V. M., 1966, On representation of stable laws by integrals, in Selected Translations in Mathematical Statistics and Probability, Vol. 6, American Mathematical Society.