STOCK RETURN DISTRIBUTION AND MARKET CAPITALISATION

PÉTER LUKÁCS[1]

The paper focuses on the relationship between the distribution of stock returns and the market capitalisation of stocks. The analysis is based on the returns of 21 stocks listed on the Budapest Stock Exchange (BSE). First, these stocks are ranked according to their market capitalisation, and then different moments of distribution as well as normalised moments such as skewness and kurtosis are calculated. Results are evaluated both by charts and rank-correlation. A significant relationship is demonstrated between the distribution of returns and the market capitalisation.

KEYWORS: Distribution; Skewness; Kurtosis.

The basic issue of financial modelling, and specifically the modelling of stock prices, is how to approach the uncertainty characterising the prices of different stocks, indices and derivatives. The treatment of uncertainty may result in difficulties both in theory and in methodology. Analysing the distributions of returns and setting up different autoregressive volatility models there are two very popular methods to treat the previous uncertainty. The current paper focuses on the first i.e. the distribution-based approach of uncertainty.

Bachelier derived the first basic model of the distribution of stock returns (*Bachelier*; 1900). The lognormal model has a long and illustrious history. For other reasons the lognormal model has become the workhorse of the financial asset pricing literature.

Doubts in connection with the normality of stock returns appeared relatively early in scientific literature. Empirical research revealed extreme kurtosis and, consequently, extremely fat tails in most stock returns. Stable Pareto-Lèvy or stable Paretian distributions (*Lévy*; 1925) offered an excellent opportunity to model these phenomena and have been very popular to model fat tail problems ever since. The stable distributions are natural generalizations of the normal one in that, as their name suggests, they are stable under addition, i. e., a sum of stable random variables is also a stable variable. However nonnormal stable distributions have more probability mass in the tail areas than the normal. In fact, the nonnormal stable distributions are so fat-tailed that their variance and all higher moments are infinite. Sample estimates of variance or kurtosis for random variables with

[1] PhD student of the University of Sciences Pécs.

these distributions will not converge as the sample size increases, but will tend to increase indefinitely.

Closed-form expressions for the density functions of stable random variables are available for only three special cases: the normal, the Cauchy and the Bernoulli cases. *Lèvy* derived the following explicit expression for the logarithm of the characteristic function $\varphi(t)$ of any stable random variable X:

$$\log \varphi(t) \equiv \log \mathrm{E}\left[e^{itX}\right] = i\delta t - \gamma \left| t \right|^{\alpha} \left[1 - i\beta \,\mathrm{sgn}\,(t) \tan\left(\alpha\pi/2\right)\right],$$

where

$(\alpha,\beta,\gamma,\delta)$ – are the four parameters that characterise each stable distribution,
$\alpha \in (0,2]$ – is the exponent index,
$\beta \in (-\infty,\infty)$ – is the skewness index,
$\gamma \in (0,\infty)$ – is the scale parameter, and
$\delta \in (-\infty,\infty)$ – is said to be the location parameter.

When $\alpha = 2$, the stable distribution reduces to normal. As α decreases from 2 to 0, the tail areas of the stable distribution become increasingly 'fatter' than the normal. When $\alpha \in (1,2)$, the stable distribution has a finite mean given by δ, but when $\alpha \in (0,1]$, even the mean is infinite. The parameter β measures the symmetry of the stable distribution; when $\beta = 0$ the distribution is symmetric, and when $\beta > 0$ (or $\beta < 0$) the distribution is skewed to the right (or left). When $\alpha = 1$ és $\beta = 0$ we have the Cauchy distribution, and when $\alpha = 1/2,\ \beta = 1,\ \gamma = 1$ and $\delta = 0$ we have the Bernoulli distribution.

A very good evaluation of the application of Pareto-Lévy distributions to model stock returns can be found in the papers of *Varga* (1999, 2001) which also contain the results of empirical research. The results of empirical research on the Hungarian stock market are summarised in the work of *Rappai* and *Varga* (1997).

Research focuses on estimating the parameter α out of the four parameters of Pareto-Lévy distributions. This parameter characterises the 'peakedness' of the central part of the distribution and consequently the fatness of tails. The Hill method – to be detailed later – leads to the consistent and the most efficient estimation of the reciprocal value of the α parameter. This procedure allows us to model the phenomenon of 'peakedness' without having presume the normality of the theoretical distribution. In addition to the Hill method, plenty of procedures can be applied to model the fat tail problem; *t*-distributions with different degrees of freedom, mixture of normal distributions, etc. The literature of modelling fat tail problem does not have a long history; (*Koedjik–Schafgans–de Vries*; 1990, *Koedjik–Stork–de Vries*; 1992), (*Kähler*; 1993), (*Koedjik–Kool*; 1993).

This paper investigates the relationship between capitalisation and the previously detailed kurtosis problem using stocks listed on Budapest Stock Exchange (BSE). While the existence of extra kurtosis in the case of stock and index return distributions is widely accepted by researchers, the problem of asymmetry divides them significantly. This paper also tests the relationship between capitalisation and asymmetry. Additionally the paper also explores a third issue. How does the risk of a risk avoiding investor change as he/she rearranges his/her portfolio towards less capitalised stocks with respect to dispersion, asymmetry and peakedness.

DATA AND DEFINITION OF STOCK RETURNS

The research reported in the paper involves the closing prices of 21 stocks listed on the BSE. Stock returns are calculated according to the following formula:

$$r_t = \ln (P_t / P_{t-1}) = \ln P_t - \ln P_{t-1}, \qquad /1/$$

where

r_t – is the daily return in time t, and
P_t – is the stock price in time t.

According to equation /1/ daily returns (later returns) are calculated for all the analysed 21 stocks listed on the BSE. The distribution-characteristics of the returns are compared to the rank-position of market capitalisation. The first daily closing price is as of 1st April 1997 if the given stock had already been listed at that time. In all other cases the first daily closing price was as of the first trade-day of the given stock. The last closing prices are as of 9th May 2001. Consequently the number of returns is 1023 in most cases, and the minimum number of returns is 843 (in the case of Rába Magyar Vagon Rt., the latest listed stock).

The main consideration in stock selection was to involve the six 'market-leader' stocks of the BSE (Matáv, MOL, OTP Bank, Richter Gedeon, TVK, BorsodChem). Other stocks were selected randomly in order to represent all the capitalisation segments of the BSE.

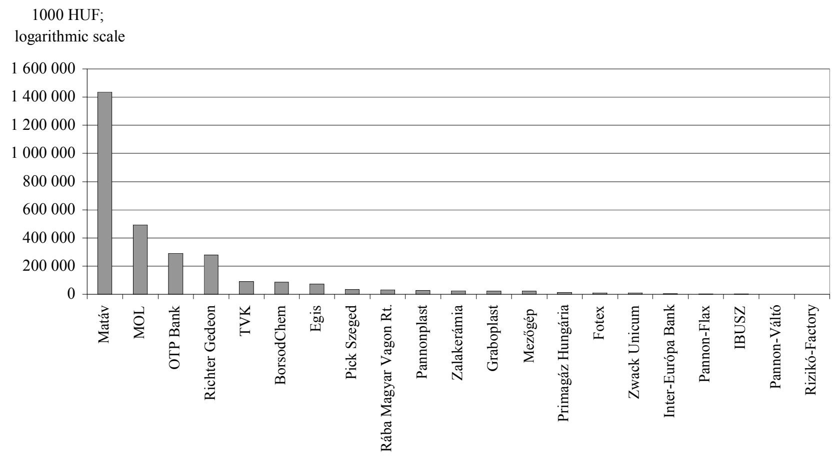
Multiplying the simple arithmetic average of prices performed the calculation of market capitalisation for a given stock and the volume introduced to the BSE on 9th May, 2001.

The following columnar composition shows the 21 analysed stocks ranked by their market capitalisation.

Stocks ranked by their market capitalisation

Stock	Capitalisation (HUF)	Stock	Capitalisation (HUF)
Matáv	1 437 621 758 438	Graboplast	23 114 041 039
MOL	492 286 344 000	Mezőgép	22 988 926 080
OTP Bank	291 126 226 500	Primagáz Hungária	13 931 460 000
Richter Gedeon	279 475 607 681	Fotex	13 326 457 370
TVK	90 677 577 007	Zwack Unicum	11 339 800 000
BorsodChem	85 602 570 878	Inter-Európa Bank	7 199 055 825
Egis	72 675 600 953	Pannon-Flax	2 406 904 923
Pick Szeged	35 439 942 545	IBUSZ	1 129 562 885
Rába Magyar Vagon Rt.	33 003 309 913	Pannon-Váltó	1 017 275 000
Pannonplast	27 227 522 892	Rizikó-Factory	392 411 244
Zalakerámia	23 727 071 258		

Figure 1. Stocks ranked by their market capitalisation



The columnar composition and Figure 1 clearly show that capitalisation decreases at an increasing pace, e.g., the capitalisation of the first, the most capitalised stock, exceeds the capitalisation of the next ten stocks. (In Figure 1 we would like to illustrate the tendency only. For a more accurate analysis different measures of concentration can be used.)

RISK AND CAPITALISATION

Asset price theories describe the risk of an asset by using the second and higher central moments of the return distribution (*Bodie–Kane–Marcus*; 1996). In the case of even moments (second, fourth…), increasing values imply increasing risk. In the case of odd moments (third, fifth…), the plus or minus sign of the values indicates whether extra-risk arises from the asymmetry of the distribution. The three-moment based portfolio selection model developed by *Gamba* and *Rossi* (1998) suggests adding a third component to the existing two components of the basic CAPM model (Capital Asset Pricing Model) in order to represent the favourable–unfavourable effect of the asymmetry of distributions. Positive skewness i.e., left asymmetry is favourable for a risk-avoiding investor, as the probability of realising huge negative returns is less. In case we accept the normality of the stock return distribution, risk can be interpreted as the second central moment, i.e., variance. The relationship between the range of dispersion and capitalisation is demonstrated in Figure 2, and the relationship between the standard deviation and capitalisation is shown in Figure 3.

Figures 2 and 3 show that both the ranges of dispersion and the standard deviation tends to increase as capitalisation decreases. Regarding the increasing pace of decreasing capitalisation values the linear trend fitted to the data naturally does not imply a linear relationship between the two indicators and the capitalisation values. A Bartlett test performed on the basis of the variances of the examined stock returns unquestionably proves that standard deviations differ from each other significantly. (Test value: 3155.4, critical value: 31.41).

Figure 2. The range of dispersion with respect to capitalisation
(percent)

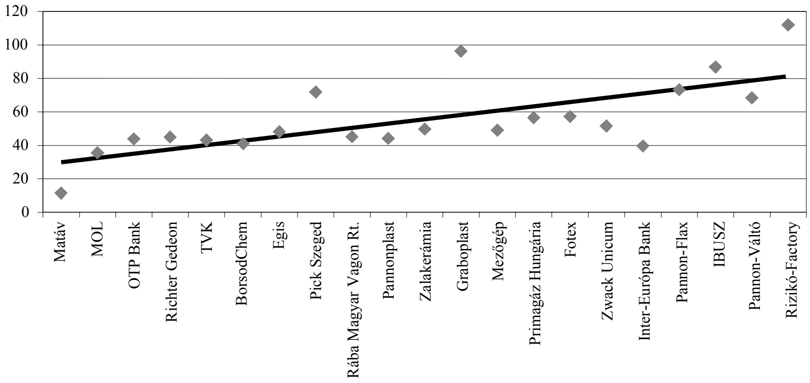
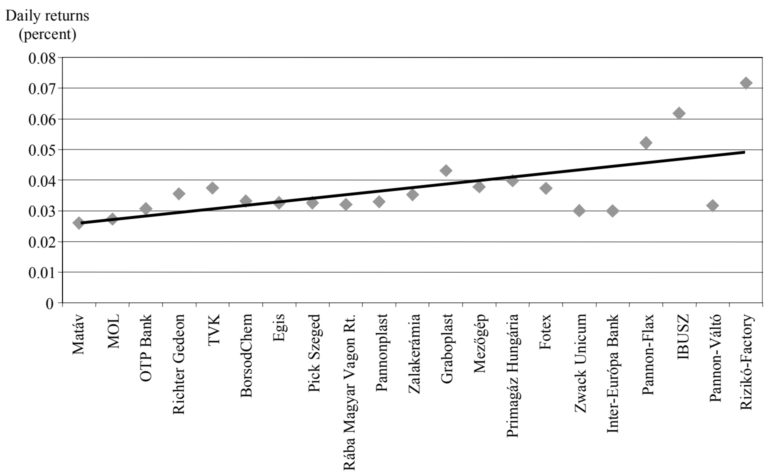


Figure 3. Standard deviation with respect to capitalisation



PRESUMING NORMALITY

If the theoretical distribution of returns is presumed normal, the difference between the theoretical and the empirical distribution can be detected by estimating the third and fourth moments of the distribution from the empirical data. More precise methods are available to test normality (e.g., Chi Squared tests) but analysing the third and fourth moments presents a clear picture. Skewness is tested by the third, and kurtosis is tested by the normalised fourth moment. Equations /2/, and /3/ define the applied formula to estimate the third and fourth normalised moments respectively.

$$\hat{S} = 1/(n\hat{\sigma}^3)\sum_{t=1}^{n}(r_t - \hat{\mu})^3, \qquad /2/$$

$$\hat{K} = 1/(n\hat{\sigma}^4)\sum_{t=1}^{n}(r_t - \hat{\mu})^4, \qquad /3/$$

where

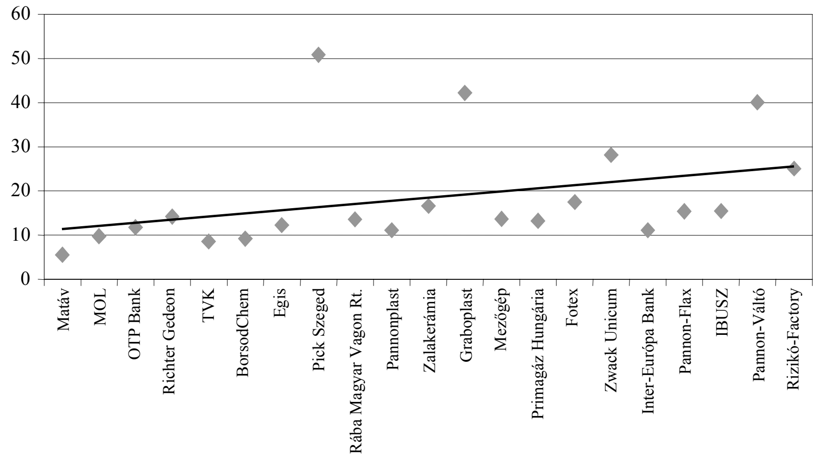
r_t – is the daily return in time t,
n – is the number of returns,
t – is the time period,
$\hat{\mu}$ – is the sample mean of returns, and
$\hat{\sigma}^2$ – is the sample variance of returns.

Presuming the normality of the theoretical distribution a confidence interval can be determined to the estimated values of skewness and kurtosis. Standard deviations can be calculated for skewness and kurtosis by the formulas $\sqrt{6/n}$, and $\sqrt{24/n}$, respectively.

It is worth mentioning that adding a confidence interval to estimate skewness and kurtosis values raises a very complex methodological problem (*Shiang et al.*; 1989).

Results are shown in Figures 4 and 5.

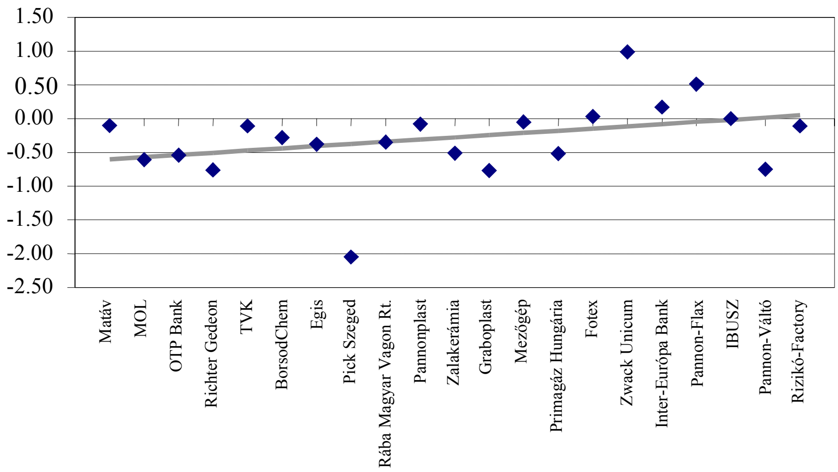
Figure 4. Kurtosis with respect to capitalisation



The hypothesis of normality of the distributions must be rejected because the estimated kurtosis values are significantly above the value of 3 of the theoretical normal distribution. Figure 4 shows that kurtosis values tend to grow as capitalisation decreases. (The fitted linear trends do not imply a linear relationship.)

Analysing skewness values the picture is more complicated. In 8 cases out of 21, the hypothesis of asymmetry must be rejected on a 99 percent confidence level. In addition, results demonstrate an interesting relationship between asymmetry and capitalisation. As far as capitalisation decreases, negative skewness values tend to zero, further turn into positive i.e., less capitalisation value means a more positive skewness measure. Conse quently, from the point of view of a risk avoiding investor, the risk decreases as the decision-maker restructures its portfolio to less capitalised papers.

Figure 5. Skewness with respect to capitalisation



Analysing kurtosis, normality had to be rejected in all cases, so more sophisticated methods seem to be necessary to model the distribution of returns.

KURTOSIS TESTS WITH THE HILL-METHOD

A high peaked distribution in our case means that more values belong to the central and tail parts of the distribution and fewer values belong to the medium parts as compared to the normal distribution. Consequently, if kurtosis values are high, fat tails are revealed.

This effect can be modelled by applying, on the one hand, Pareto-Levy stable distributions (*Palágyi*; 1999), and on the other hand distribution-free methods, e.g., the Hill-method (*Lux–Varga*; 1996), (*Varga*; 1998, 1999).

With the Hill-method (*Hill*; 1975), two indices characterise the fatness of tails. Specifying the range that contains tail data is a basic dilemma when using this method. It is important that the Hill-indices show an approximate stability when changing the tail ranges. The Hill-index values were calculated for 5, 10 and 25 percent tail range both for positive and negative values.

The Hill-indices are given in equations /4/ and /5/.

$$\gamma_{H+} = 1/\alpha_{H+} = (1/m)\sum_{i=1}^{m}\left[\log X_{(i)} - \log X_{(m)}\right], \qquad X(1) \geq X(2) \geq \ldots \geq X(m), \qquad /4/$$

$$\gamma_{H-} = 1/\alpha_{H-} = 1/n\sum_{j=1}^{n}\left[\log\left|Y(j) - Y(n)\right|\right], \qquad \left|Y(1)\right| \geq \left|Y(2)\right| \geq \ldots \geq \left|Y(n)\right|. \qquad /5/$$

In equations /4/ and /5/:

α_{H+} – is the positive tail index,
α_{H-} – is the negative tail index,
m – is the number of returns belonging to positive tail,
n – is the number of returns belonging to negative tail,
$X(i)$ – are return values belonging to positive tail, and
$Y(j)$ – are return values belonging to negative tail.

Given /4/ and /5/, the more peaked the distribution is, the fatter the tails that it has, and the smaller the value of Hill tail-index that is calculated. Figure 6 shows the Hill-index values when 25 percent of the data belongs to the upper and lower tails.

Figure 6. Hill-index values
(tail range: 25 percent)

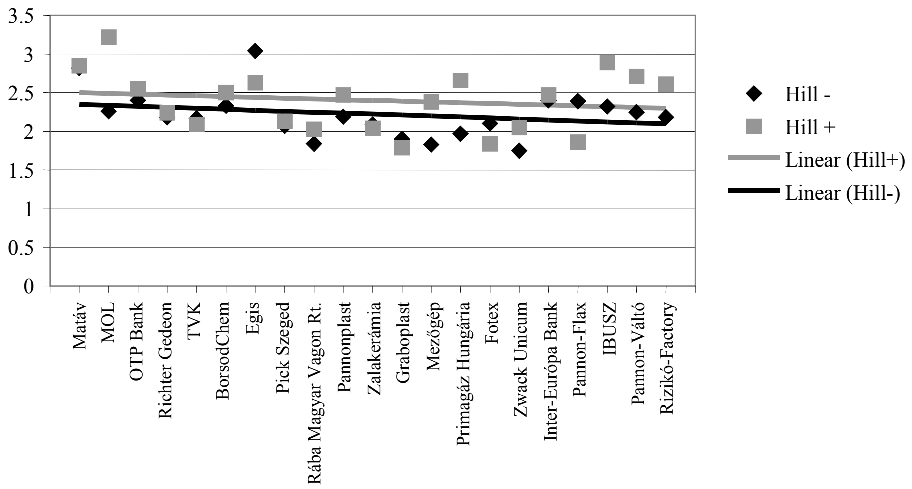


Figure 6 demonstrates that the Hill-index values do not change significantly as long as capitalisation decreases. While, presuming normality, kurtosis tends to grow as capitalisation decreases, kurtosis implied by the Hill-index values does not seem to change as capitalisation decreases. A more precise test can be conducted to test whether there is a significant difference between the Hill-index values of stock returns. Formulas for the positive and negative ranges respectively are demonstrated in equations /6/ and /7/.

$$Q^{+} = \sum_{i=1}^{21} (\alpha^{+} / \alpha_i - 1)^2 \cdot m\,, \qquad /6/$$

$$Q^{-} = \sum_{i=1}^{21} (\alpha^{-} / \alpha_i - 1)^2 \cdot n\,. \qquad /7/$$

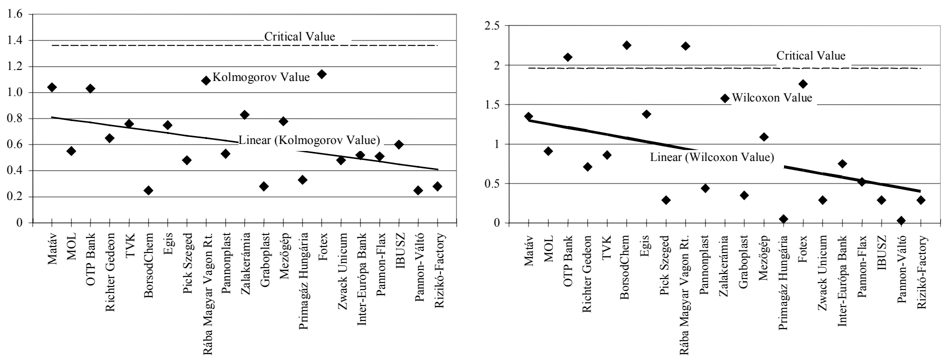
Q^+ and Q^- are the test statistics calculated for the adequate tail ranges. Q^+ and Q^- are characterised by a χ^2 distribution with 21 degrees of freedom. At the 5 percent significance level the critical value is 32.67 in both cases. The test value is 26.81 in the positive range and 17.78 in the negative one. Consequently, the hypothesis that the Hill-index values do not differ from each other significantly as capitalisation changes cannot be rejected.

ROBUST TESTS ON SYMMETRY

In the first part of this paper the hypothesis of the normality of stock returns was rejected. Consequently, other robust methods seem to be adequate to test symmetry. These robust methods do not depend on the distribution. The two tests described in the following had been created originally to test the identity of two distributions. In our case the given distribution is split into two parts; positive and negative returns are treated as separate distributions. By multiplying the negative values by –1, the identity of the two split distributions can be tested. The acceptance of the hypothesis of identity means the acceptance of the hypothesis of symmetry; and, in reverse, the rejection of the hypothesis of identity means the acceptance of the hypothesis of asymmetry. The results of the Kolmogorov-Smirnov test and the results of the Wilcoxon tests are demonstrated in Figure 7 respectively.

In both cases the hypothesis of symmetry can be accepted at the 5 percent significance level. In the case of two stocks (OTP Bank and Rába Magyar Vagon Rt.) the hypothesis of symmetry must be rejected according to the Wilcoxon test. Furthermore, Figures 7 demonstrate that the asymmetry of stock return distributions tends to decrease as capitalisation decreases. This result is very similar to the case of normality.

Figure 7. The result of the Kolmogorov-Smirnov test and the Wilcoxon test



As far as capitalisation decreases, the variance and the standard deviation of returns increase, so the risk of a risk-avoiding investor grows. This effect is reduced by the favourable change in the symmetry of stock returns. This result is in accordance with the extended, three-moment based CAPM (*Gamba–Rossi*; 1998), which involves skewness into the model.

RESULTS ANALYSED BY RANK-CORRELATION

The correlation between the position in the capitalisation list and the different distribution indicators are analysed in this section. The Spearman rank-correlation is calculated according to equation 8:

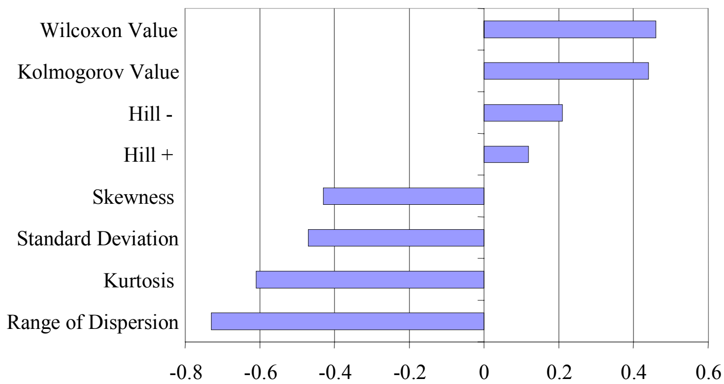
$$\rho = 1 - \frac{6\sum_{i=1}^{n}\left(x_i - y_i\right)}{n\left(n^2 - 1\right)}, \qquad /8/$$

where

x_i – is the rank value of stock '*i*' as capitalisation decreases,
y_i – is the rank value of stock '*i*' as for the given distribution character, and
n – is the number of stocks under discussion, in our case it is 21.

The results are summarised in Figure 8. The results demonstrated in the following are in accordance with the results shown by the graphic methods described earlier. Besides the Hill-indices, a strong or medium correlation can be found between the rank in decreasing capitalisation and the different characters of distribution.

Figure 8. Rank correlation among the different distribution characters and the decreasing capitalisation



*

Summarising the results of the investigations the following statements can be regarded as proven.

1. Presuming the normal theoretical distribution of stock returns, as capitalisation decreases, the empirical distributions tend to have higher and higher kurtosis values, thus exhibiting greater departures from normality. The case is the reverse when analysing skewness; as capitalisation decreases the distributions tend to become more similar to the normal distribution.

2. *Discarding the assumption at a normal distribution* of stock returns, the picture is different from the previous case. Analysing the tails of the distributions, the fatness of the tails was proven the same in all cases, so capitalisation has no effect on that. Robust symmetry tests showed that the hypothesis of asymmetry of distributions must be rejected in almost all cases; however, the symmetry tends to grow as capitalisation decreases.

The risk of a *risk avoiding investor* grows as he/she restructures his/her portfolio towards less capitalised stocks due to the increasing variance and possibly growing kurtosis. This effect is reduced by the favourable change in asymmetry while going less capitalised.

REFERENCES

BACHELIER, L. (1964): Theory of speculation. In: COOTNER, P. (ed.), *The random character of stock market prices.* Massachusetts Institute of Technology Press, Cambridge, MA.

BODIE – KANE – MARCUS (1996): *Befektetések*. Tanszék Kft. Budapest.

GAMBA, A – ROSSI, F. A. (1998): A three-moment based portfolio selection model. *Rivista di Matematica per le Scienze Sociali*, No. 1. p. 25–48.

HILL, B. M. (1975): A simple general approach to interference about the tail of a distribution. *Annals of Statistics*, No. 3, p. 1163–1173.

LAU, H. SH. – WINGENDER, J. R. – LAU, A. H. L. (1989): On estimating skewness in stock returns. *Management Scienze*, September No. 9.

KAHLER, J. (1993): *On the modelling of speculative prices by stable paretian distributions and regularly varying tails*. ZEW University of Mannheim, mimeo.

KOEDIJK, K. G. – SCHAFGANS, M. M. A. – DE VRIES, C. G. (1990): The tail index of exchange rate returns. *Journal of International Economics*, No. 29. p. 93–108.

KOEDIJK, K. G. – STORK, P. A. – DE VRIES, C. G. (1992): Differences between foreign exchange regimes: The view from the tails. *Journal of International Money & Finance*, No. 11. p. 462–473.

LÉVY, P. (1924): Théorie des erreurs, la loi de Gauss et les loi exceptionelles. *Bulletin of Social Mathematics*, No. 52. p. 49–85.

LUX, T. – VARGA, J. (1996): A Pareto hipotézis vizsgálata – Értékpapírpiaci hozamok és az extremális hozamok eloszlása. *Szigma*, Vol. 27. No. 4. p. 1–23.

MARTON, R. (2001): A magyar tőkepiac hatékonyságának vizsgálata. *Bankszemle*, Vol. 45. Április–május p. 72–88.

PALÁGYI, Z (1999): Árfolyamingadozások és kockázatbecslés a Budapesti Értéktőzsdén. *Szigma*, Vol. 30. No. 1–2. p. 27–32.

RAPPAI, G. – VARGA, J. (1997): Applicability of the CAPM on the Hungarian Stock Market, In: *New operational approaches for financial modelling (Contribution to management science*). Physica Verlag, Heidelberg–New York.

VARGA, J. (1998): On distributions for stock returns: A survey of empirical investigations. In: PARDALOS, P. – ZOPOUNIDIS, C. (eds.) *Managing in uncertainty: Theory and practice.* Kluwer Academic Publishers, Dordrecht, p. 139–151.

VARGA, J. (1999): Stock return distribution: A survey of empirical investigations. *Hungarian Statistical Review*, Vol. 77, Special Number. p. 23–34.

VARGA, J. (2001): Pénz-és tőkepiaci idősorok sztochasztikus volatilitás modelljei. *Szigma*, Vol. 32. No. 1–2. p. 71–86.