USING GED (GENERALIZED ERROR DISTRIBUTION) FOR MODELING DISTRIBUTION OF THE RATES OF RETURN

Rafał Czyżycki

Abstract

Distribution of the rates of return is one of the most commonly analyzed and used concepts at the capital market. While using density function of the rate of return, a classical assumption is frequently adopted that a given rate of return demonstrates normal distribution. This article presents possible use of GED, which is nonclassical estimation method, for modeling distribution of the rates of return.

Keywords: rate of return, density function, nonclassical methods, GED.

1. Introduction

The rate of return, besides risk, is one of the most important expressions in the theory and practice of finance. It is a basic factor i.a. in making investment decisions. The most often, the investors' decision related to buying or selling the securities depends on the level of probability to reach a specified profit level or incurring possible loss. While estimating the level of such probability in practice, an assumption is very often adopted that the rates of return demonstrate normal distribution. Such assumption is also very often adopted in various types of models, which describe mechanisms of capital market, or which are involved in supporting investment decisions. It occurs in the Black Scholes stock option valuation method (1973), the capital asset pricing model *[CAMP]* (refer to Sharpe (1964)), or in case of some methods for establishing a value at risk (VaR). Assumption on normal distribution of the rates of return is mainly adopted in order to accelerate, simplify and facilitate conducting specified calculations. However, a significant deviation of real rates of return from the ones adopted in assumptions may result in many negative consequences. It may be the base for questioning reliability, hence applicability of many techniques, methods and models used for analyses, diagnoses and forecasts of the capital market.

2. Investigating normality of the rates of return

The rate of return is the most frequently defined as arithmetic or logarithmic rate of return. In case of analyzing the arithmetic rate of return R_t, its value is established by using the formula

$$R_t = \frac{P_t - P_{t-1} + D_t}{P_{t-1}} \tag{1}$$

but the logarithmic rate of return R^*_t may be established as

$$R^*_t = ln\frac{P_t + D_t}{P_{t-1}} \tag{2}$$

where:
  P_t- security price in period *t;*
  P_t-i- security price in period *t-1;*
  D_t- value of a dividend paid in period *t.*

Depending on assumed time frame, the rates of return may be analyzed on daily, weekly, monthly or annual basis. For the purpose of the article, the analysis was limited only

to daily arithmetic rates of return and it was performed for the WIG20 index. Time frame includes the period from 18.04.1994 (the first quotation of the index) to 18.11.2013, i.e. 4871 quotations. The basic numerical characteristics describing formation of described rates of returns in aforementioned period is presented in the table 1.

Table 1. Basic statistics for formation of daily rates of return of the WIG20 index.

	WIG20
N	4871
average	0,00037
standard deviation	0,01917
kurtosis	4,69242
skewed	0,03615
min	-0,13204
max	0,15993

Source: own elaboration.

Investigating normality of the rates of return involves setting up two hypotheses: the null hypothesis, which refers to conformity of a distribution function for a given rate of return with normal distribution function, and the alternative hypothesis to the null hypothesis assuming that distribution function for the rate of return is not a normal distribution function. Aforementioned hypotheses are verified with suitable conformance tests. In this article, the Kolmogorov-Smirnov test (K-S test) for normality with the Lilliefors correction, and chi-square test, were used for this purpose. Value of K-S test was 0.062195, and chi-square was 647.56043 and in both cases, the hypothesis on normal distribution of the rates of return should be rejected. While performing analyses of distribution of the rates of return, in majority of cases, rejection of the hypothesis on normal distribution of the rates of return results from occurrence of so-called fat-tail distribution and then, suitable distributions are used for their description, i.a. such as α-stable distribution, Student's t-distribution, or GED.

3. GED - Generalized Error Distribution

GED density function, which is also referred to as GGD (Generalized Gaussian Distribution), is described with the formula (Purczynski 2003)

$$f(x) = \frac{\lambda \cdot s}{2 \cdot \Gamma(\frac{1}{s})} \cdot \exp\left(-\lambda^s \cdot |x - \mu|^s\right) \qquad \textbf{(3).}$$

where:
   $\Gamma(z)$ - Euler function,
   s - shape parameter,
   λ - scale parameter,
   μ - location parameter.

For s=l, GED transforms into the Laplace distribution (the double exponential distribution)

$$f(x) = \frac{\lambda}{2} \cdot \exp\left(-\lambda \cdot |x - \mu|\right), \qquad (4),$$

but for s=2, normal distribution is obtained

$$f(x) = \frac{\lambda}{\sqrt{\pi}} \cdot \exp\left(-\lambda^2 \cdot (x - \mu)^2\right) \qquad (5).$$

Suitable method for estimation of the GED parameters is a maximum likelihood estimation method (MLE) (Bednarz 2012). In order to simplify calculation process, centering sequence of X_j values may be done by subtracting $\hat{\mu},$

$$\hat{\mu} = \frac{1}{N}\Sigma_{i=1}^{N} x_i \tag{6},$$

so the density function transforms into formula

$$f(x) = \frac{\lambda \cdot s}{2 \cdot \Gamma(\frac{1}{s})} \cdot \exp\left(-|\lambda \cdot x|^s\right) \tag{7}$$

and then, using MLE, the logarithm of the likelihood function is determined as a formula

$$\ln\bigl(L(\lambda, s)\bigr) = N \cdot \ln(\lambda) + N \cdot \ln\left(\frac{s}{2\Gamma(\frac{1}{s})}\right) - \Sigma_{i=1}^{N}|\lambda \cdot x_i|^s, \tag{8},$$

based on which, with conditions

$$\frac{\partial \ln\,(L(\lambda,s))}{\partial \lambda} = 0, \quad \frac{\partial \ln\,(L(\lambda,s))}{\partial s} = 0 \tag{9},$$

results

$$\lambda = \sqrt[s]{\frac{N}{s \cdot \Sigma_{i=1}^{N}|x_i|^s}} \tag{10}$$

and

$$s + \psi\left(\tfrac{1}{s}\right) + \ln\left(\tfrac{s}{N}\Sigma_{i=1}^{N}|x_i|^s\right) - \frac{s \cdot \Sigma_{i=1}^{N}|x_i|^s \ln|x_i|}{\Sigma_{i=1}^{N}|x_i|^s} = 0 \tag{11}$$

are obtained,
where $\psi(z) = \frac{d}{dz}[ln\Gamma(z)]$.

I. Study results

Conducted attempt to model distribution of the rates of return using GED within the whole analyzed period did not bring satisfactory results. Based on the formula (11), value of s parameter was estimated at the level of 1.06414, and after replacing obtained value into the formula (10), the result was λ=67.78024. Then, chi-square test was conducted in order to verify hypothesis on conformity of real distribution of the rates of returns on the WIG20 index with assumed GED. Obtained chi-squared=50.92579, which means that at the statistical significance of p=0.015585, aforementioned hypothesis should be rejected.

In the second part of the studies, the analyzed period was divided into more uniform periods of bull and bear markets. Within the period from 18.04.1994 to 18.11.2013, 4 bull markets and 4 bear markets may be classified at the Warsaw Stock Exchange (compare with Tomasik 2011):

- bull market 1 (B1) lasting from 18.04.1994 to 29.03.1995;
- bear market 1 (H1) lasting from 29.03.1995 to 03.03.1998;
- bull market 2 (B2) lasting from 03.03.1998 to 12.10.1998;
- bear market 2 (H2) lasting from 12.10.1998 to 27.03.2000;
- bull market 3 (B3) lasting from 27.03.2000 to 04.10.2001;
- bear market 3 (H3) lasting from 04.10.2001 to 06.07.2007;
- bull market 4 (B4) lasting from 06.07.2007 to 18.02.2009;
- bear market 4 (H4) lasting from 18.02.2009 to 18.11.2013

Conducting aforementioned classification undoubtedly improved efficiency of using GED for describing the rates of return on the WIG20 index. There was no basis (at p=0.1) for rejection of the hypothesis on conformity of the real distribution of the rates of return with the distribution described by GED. However, an attention should be paid to large changes in estimated values of parameters describing this distribution. Within the periods of bear market,

comparing to the periods of the bull market, it may be noticed that increase in value of shape parameter *s* and decrease in value of scale parameter λ take place. Numerical characteristics describing development of aforementioned parameters together with information on value of the chi-square test for uniformity is presented in the table 2.

Table 2. Basic statistics describing GED in analyzed subperiods.

period		s	λ	chi-squared (chi2)
18.04.1994r. - 29.03.1995r.	B1	1,41489	24,19981	chi2=1,94644; p=0,856505
29.03.1995r. - 03.03.1998r.	H1	1,14925	56,44365	chi2=11,47523; p=0,176198
03.03.1998r. - 12.10.1998r.	B2	1,51227	27,66117	chi2=2,51433; p=0,642072
12.10.1998r. - 27.03.2000r.	H2	1,27543	47,82500	chi2=10,91912; p=0,206324
27.03.2000r. - 04.10.2001r.	B3	1,68096	41,97033	chi2=5,81487; p=0,830569
04.10.2001r. - 06.07.2007r.	H3	1,27532	72,23224	chi2=18,41355; p=0,782499
06.07.2007r. - 18.02.2009r.	B4	1,34577	43,68876	chi2=14,00531; p=0,122136
18.02.2009r. - 18.11.2013r.	H4	1,13660	80,37132	chi2=24,36023; p=0,143577

Source: own elaboration.

In addition, it should be emphasized that within the periods of the bear market, the level of p-value was definitely higher than it was within the periods of the bull market (except for the period B4).

In order to check whether the change of the respective parameters describing GED is significant, chi-square test for uniformity was conducted. For this purpose, theoretical values of the rates of return, which are described by GED using parameters of a given subperiod, were established for each subperiod. Then, hypothesis on a substantial difference between obtained distribution functions was verified with aforementioned test. Obtained results are presented in the Table 3, where values of chi-square statistics were placed above the main diagonal, and p-values were provided under the main diagonal.

Table 3. Values of chi-square test for uniformity together with the levels of p-value corresponding to the hypothesis on identity of GED functions in analyzed subperiods.

period	B1	H1	B2	H2	B3	H3	B4	H4
B1	x	2,79678	0,52019	1,34772	1,19295	4,56362	1,43245	4,46269
H1	0,99857	x	0,15458	0,41688	0,38556	1,58314	0,45090	1,44118
B2	1,00000	1	x	0,05278	1,50274	5,68984	0,05400	5,59120
H2	0,99998	1	1,00000	x	0,71658	2,86547	1,34363	2,67795
B3	0,99999	1	0,99962	1,00000	x	2,64144	0,77646	2,51537
H3	0,98359	1,00000	0,89324	0,99998	0,99999	x	0,21626	0,72799
B4	0,99997	1	1,00000	0	1,00000	1	x	2,42850
H5	0,98520	1,00000	0,89920	0,99999	0,99999	1	1,00000	x

Source: own elaboration.

Results presented in the table above unanimously indicate that each analyzed subperiod reveals no basis for rejecting assumption that the rates of returns within this period demonstrate the same distribution as they are in other periods. However, study results obtained earlier, based on which the hypothesis on conformity of distribution of the rates of return on the WIG20 index should be rejected within the whole analyzed period (18.04.1994 - 18.11.2013) with GED, suggest that it might have been caused by i.a. significant changes in *s* and λ parameters within the respective studied subperiods. Therefore, the last element of the study with possible use of GED in modeling the rates of return will include analyzing changes of the respective parameters in time. For this purpose, 100 first quotations were used within analyzed period (i.e. from 18.04.1994 to 25.10.1994) and based on the formulas (10) and (11),

s and *λ* parameters, which describe GED in this period, were established. Then, the period of analysis was moved by one quotation (19.04.1994-26.10.1994) and the aforementioned parameters were established again. This procedure was repeated until the moment, when subject of the analysis included the period of the last 100 quotations (i.e. the period 24.08.2013-18.11.2013). Values of *s* and *λ* parameters obtained this way are presented in the chart 1.

Chart 1. Values of *s* and *k* parameters of GED within the period of 18.04.1994 - 18.11.2013.
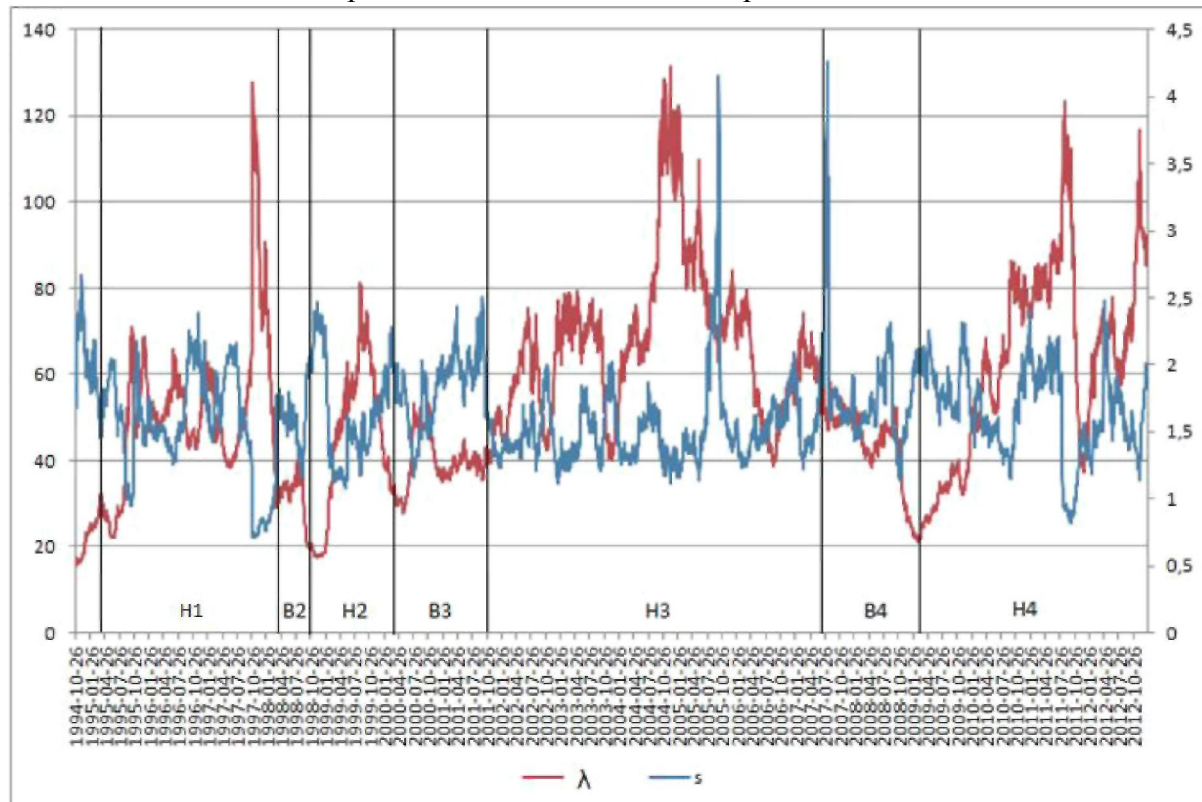
Source: own elaborations.

In the chart above, significant changes may be noticed in forming values of *s* and *λ* parameters in the analyzed period. The most often, reversal of the bull markets into the bear markets took place after a significant fall of the value of *λ* scale parameter of GED, and at the same time, each change (reversal of the bull markets into the bear markets as well as reversal of the bear markets into the bull markets) was accompanied by forming value of *s* shape parameter at the local maximum.

Conclusions
Based on deductions made in the article, the following conclusions may be drawn:

- preliminary assumption, which is frequently adopted in conducting studies on the capital market that the rates of return are subject to normal distribution, is not justified. The results presented in this article only relate to the WIG20 index, but other studies conducted by the author confirmed that the rates of return for many other companies quoted at the Warsaw Stock Exchange failed to create normal distribution,

- in order to perform a proper analysis of distribution of the rates of return on stocks, non-classical methods, which include i.a. GED methods presented in this article, should be used in majority of cases. To a larger extent, it is supported by a constant increase in computational

power of computers and, which is more important, in availability of the specialist statistical software, which is dedicated to non-classical methods (including simulation software),

- if using GED, significant changes in values of s and λ parameters should be considered, depending on the period, which is subject of the analysis, and its length,

- analysis of changes in developing s and λ parameters of GED may be used as an element of technical analysis, i.a. within the scope of predicting changes in stock market cycles.

Sources
1. Black, F., Scholes, M. *The Pricing of Options and Corporate Liabilities.* The Journal of Political Economy. Vol. 81 Issue 3,1973. 637-654 p.
2. Czyżycki, R. Społki Teleinformatyczne jako społki defensywne na Giełdzie Papierów Wartościowych. Wydawnictwo Naukowe Uniwersytetu Szczecińskiego, 2012
3. Sharpe, WF. *A Theory of Market Equilibrium under Conditions of Risk.* The Journal of Finance. 19(3), 1964, 425-442p.
4. Klóska, R., Czyżycki, R. *Wybrane zagadnienia ze statystyki. ECONOMICUS, Szczecin* (2008).
5. Tomiasik, E. *Analiza rozkladów stóp zwrotu z instrumentów finansowych na polskim rynku kapitalowym.* Praca doktorska, Poznan ,2011.
6. Purczyński, J. *Wykorzystanie symulacji komputerowych w estymacji wybranych modeli ekonometrycznych i statystycznych.* Wydawnictwo Naukowe Uniwersytetu Szczecinskiego, 2003.
7. Bednarz, K., Purczynski, J. *Metody estymacji parametrow uogolnionego rozkladu Gaussa.* Technika Transportu Szynowego, (2012), 1367-1375p.

Contact
Dr Rafał Czyżycki, Ph.D.
Department of Quantitative Methods
Faculty of Management and Economics of Services
University of Szczecin
ul. Cukrowa 8, 71-004 Szczecin, Poland
tel. +48 91 444 31 56
email: Rafal.Czyzycki@wzieu.pl