

# INFLUENCE OF THE KURTOSIS VALUE ON THE RESULT OF GOODNESS OF FIT TEST FOR THE DAILY RATE OF RETURN ON SELECTED WIG20 AND WIG30 COMPANIES

KAMILA BEDNARZ-OKRZYŃSKA

University of Szczecin, Faculty of Management and Economics of Services, POLAND  
e-mail: kamila.bednarz@wzieu.pl

RECEIVED	18 January 2018
ACCEPTED	2 September 2018
JEL CLASSIFICATION	C12, C13, D53, G11
KEYWORDS	kurtosis, chi-squared test, Gaussian distribution, Laplace distribution
ABSTRACT	In the paper, the modeling of empirical distributions of return rates on WIG20 and WIG30 companies was conducted. The validity of the hypothesis was tested, which stated that the fitting of theoretical distributions to empirical distributions of return rates – where the fitting was tested by the chi-squared test – depends on the value of kurtosis. In order to prove the validity of the hypothesis, studies on four sets of diversified data were conducted.

---

## Introduction

This work has been inspired by the observations made while working on paper (Bednarz-Okrzyńska, 2016), where modeling of empirical distributions of return rates on WIG20 companies was conducted by means of theoretical distributions.

On the basis of the studies conducted in the aforementioned work, the dependence of the value of excess kurtosis on the time interval applied in measuring return rates was observed. Namely, for the daily data, the excess kurtosis equaled 4.112, and for the monthly data – 1.232. At the same time, the influence of the time interval used in return rate measurements on the results of modeling empirical distributions of return rates using the Gaussian and Laplace distributions was also observed. Namely, the normal distribution provided a better fit for the monthly data, and the Laplace distribution proved more useful for the daily data.

The results obtained in paper (Bednarz-Okrzyńska, 2016) point toward the relationship between the value of excess kurtosis and the result of goodness-of-fit test applied to the above mentioned theoretical distributions. Therefore, this paper concentrates on the issue of influence of the value of excess kurtosis on the result of a goodness-of-fit test.

Below, the basic relations, used later throughout the paper, are presented.

First order moment  $m$  (mean value) is derived from formula

$$m = \frac{1}{N} \times \sum_i x_i \quad (1)$$

where:  $x_i$ - observations;  $i = 1, 2, 3, \dots, N$ .

The variance is given by:

$$V(X) = \frac{1}{N} \times \sum_i (x_i - m)^2 \quad (2)$$

Fourth central moment is given by:

$$M_4 = \frac{1}{N} \times \sum_i (x_i - m)^4 \quad (3)$$

Standardized fourth central moment  $KU$  (kurtosis) is derived from formula (Krishnamoorthy, 2006; Sobczyk, 2004):

$$KU = \frac{M_4}{[V(X)]^2} \quad (4)$$

Another measure derived from the kurtosis is excess  $ex$  (Sobczyk, 2004):

$$ex = KU - 3 \quad (5)$$

where:  $KU$  is given by (4).

Yet another relation describing concentration  $K$  can be found in (Tarczyński, 2002):

$$K = \frac{N^2(N+1)}{(N-1)(N-2)(N-3)} \times KU - \frac{3(N-1)^2}{(N-2)(N-3)} \quad (6)$$

where:  $KU$  is given by (4).

From (5) and (6) follows:  $K \xrightarrow{N \rightarrow \infty} ex$ , which means that for a large sample the values of concentration  $K$  will be close to the value of excess  $ex$ .

In this paper formula (6) will be applied. Papers (Bonato, 2011, pp. 77–87; Kim, White, 2004, pp. 56–73) create certain confusion, since they define kurtosis by means of equation (5). The clarification can be found in (Kotz, Kozubowski, Podgórski; 2001), where equation (5) is referred to as excess kurtosis. It may be assumed that the authors of papers (Bonato, 2011, pp. 77–87; Kim, White, 2004, pp. 56–37) use the phrase ‘kurtosis’ in place of ‘excess kurtosis’ for the sake of simplification. When following papers (Bonato, 2011, pp. 77–87; Kim, White, 2004, pp. 56–77) in this paper, the term ‘kurtosis’ will be used with reference to formula (6).

For the normal distribution, kurtosis takes the value  $K=0$ , however for the Laplace distribution  $K=3$ . This means that the value of kurtosis should have an impact on the results of the goodness-of-fit test when modeling empirical distributions of return rates using the Gaussian and Laplace distributions.

The primary goal of this paper is to determine the validity of the hypothesis which states that the fitting of theoretical distributions to empirical distributions of return rates – measured with the chi-squared goodness-of-fit test – depends on the value of kurtosis.

In order to validate the hypothesis, appropriate studies will be conducted on the sets of data of diversified nature.

Therefore, four sets of data were investigated including return rates on selected WIG companies. Set A included daily, weekly and monthly data on: WIG, WIG20, MWIG40, SWIG80 indexes and WIG20 companies from the period of one year, 2013. It comprised in total 72 observations.

Set B included return rates on the same indexes and companies as Set A, however it comprised only the daily data from the period 2010–2013 including 89 observations.

Set C included daily return rates on WIG and WIG20 indexes as well as WIG20 companies covering the period 4.10.2001–30.08.2017. Within this time interval, four boom subintervals (H4, H5, H6, H7) and three slump subintervals (B4, B5, B6) were identified. The analysis comprised WIG20 companies which were listed on the WIG20 index during each slump and boom period under study. The following boom and slump periods were identified: Boom 4 – from 4.10.2001, comprising 1,444 trading sessions; Slump 4 – from 6.07.2007, comprising 403 sessions; Boom 5 – from 17.02.2009, comprising 541 sessions; Slump 5 – from 7.04.2011, comprising 93 sessions; Boom 6 – from 5.06.2012, comprising 725 sessions; Slump 6 – from 8.05.2015, comprising 176 sessions and Boom 7 – from 27.06.2016, comprising 296 sessions. Set C included 136 observations.

Set D comprised daily return rates on WIG30 bank sector stocks covering the period of three years from 30.05.2014 until 31.05.2017. Within this time interval, three subintervals were identified: Period I from 31.05.2016 until 31.05.2017, Period II from 29.05.2015 until 31.05.2017 and Period III from 30.05.2014 until 31.05.2017- the set comprised in total 30 observations. The following bank sector companies were considered: ALIOR, BZWBK, HANDLOWY, INGBSK, MBANK, MILLENNIUM, PEKAO, PKOBP.

## Results of calculations

The data set A comprised the values of daily, weekly and monthly return rates on WIG20 companies in 2013.

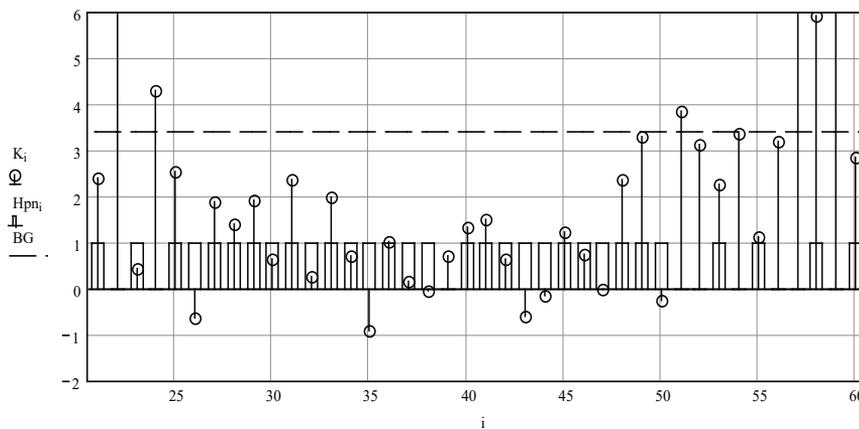
For the sake of clarity, Figure 1 presents the results of calculations for 40 observations out of 72. The lines topped with circles  $K_i$  show the values of concentration. The dashed line BG marks the value of the concentration threshold. Symbol  $H_{pn}$  marks the result of the chi-squared test. Presence of a bar ( $H_{pn} = 1$ ) means a positive result of the test. And the lack of a bar ( $H_{pn} = 0$ ) corresponds to a negative result of the test. It was expected that for the concentration larger than the value of threshold BG, a negative result of the test would be obtained. In fact, there

were six cases when the concentration exceeded the value of the threshold  $BG = 3.4$ , and in five cases a negative result of the goodness-of-fit test was obtained. Hence the value of ratio  $NEGG = \frac{5}{6} = 0.833$ . Generally ratio NEGG is given by:

$$NEGG = \frac{LN}{LG} \tag{7}$$

where:

- $LG$  – number of cases when  $K > BG$  in total,
- $LN$  – number of cases when  $K > BG$  and  $H_{pn} = 0$ .



**Figure 1.** Sample results of calculations for Set A of the data. Concentration  $K$  marked by means of lines topped with circle  $K_i$ . Dashed line  $BG$  marks the threshold value of concentration. Symbol  $H_{pn}$  marks the result of the chi-squared test (bars)

Source: author’s own study.

Figure 2 presents the values of ratio NEGG as a function of threshold value  $B1$ . The largest value of ratio NEGG = 0.833 corresponds to the threshold value  $B1 = 3.4$ .

Figure 3 made for Set A corresponds to the Laplace distribution – labeling the same as in Figure 1. This time it was assumed that for the concentration smaller than the value of threshold  $BL$ , a negative result of the chi-squared test would be obtained. Ratio NEGL is given by:

$$NEGL = \frac{LN}{LG} \tag{8}$$

where:

- $LG$  – number of cases when  $K < BL$  in total,
- $LN$  – number of cases when  $K < BL$  and  $H_{pL} = 0$

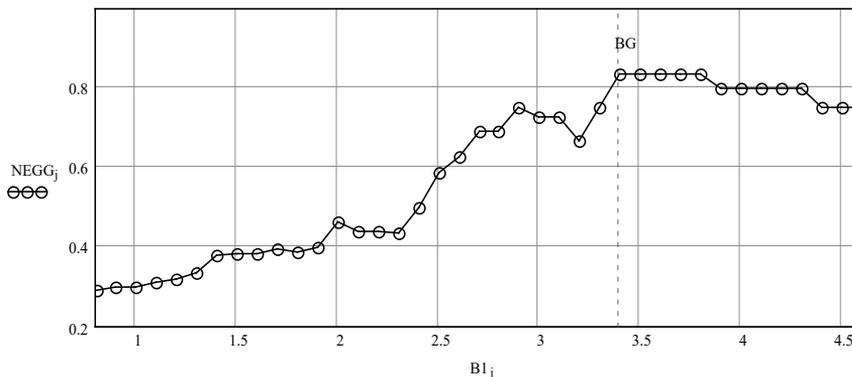


Figure 2. Values of ratio NEGG as a function of the threshold value B1 (Set A)

Source: author's own study.

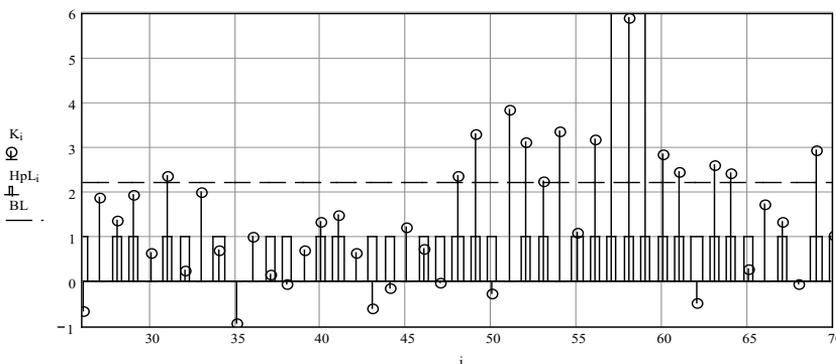


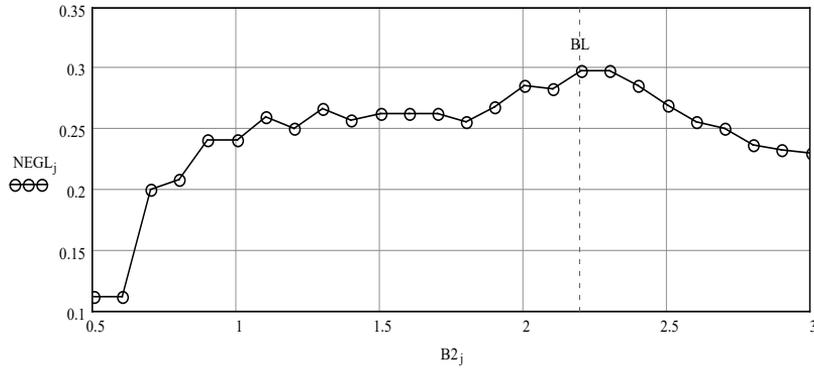
Figure 3. Results of calculations for the Laplace distribution (Set A of the data). Symbol HpL marks the result of the chi-squared test (bars). The rest of the labeling is the same as in Figure 1. The figure was made for BL = 2.2

Source: author's own study.

Figure 4 presents the values of ratio NEGL for the changing threshold value B2. For threshold value BL = 2.2, the largest value of ratio is obtained  $NEGL = \frac{LN}{LG} = \frac{14}{47} = 0.298$ .

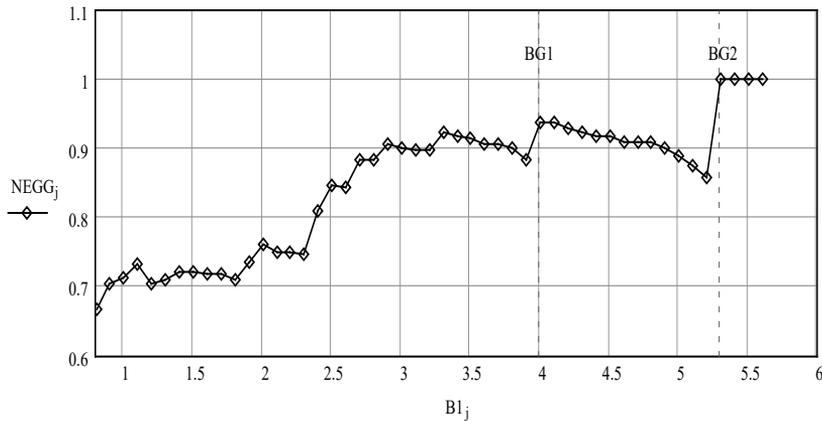
Subsequent calculations were made for Set B – values of daily return rates on WIG20 companies in the period 2010–2013.

In Figure 5 two threshold values BG1 and BG2 were singled out. For BG1 = 4 ratio  $NEGG = \frac{LN}{LG} = \frac{15}{16} = 0.938$  and for BG2 = 5.3 ratio  $NEGG = 1$  (LN = LG = 6). For threshold value BG2, the largest value of NEGG can be observed, and for threshold value B1, the largest number of observations – 15 out of 16 – can be observed.



**Figure 4.** Values of ratio NEGL as a function of threshold value B2 (Set A)

Source: author's own study.



**Figure 5.** Values of ratio NEGG as a function of threshold value B1 (Set B)

Source: author's own study.

In Figure 6 two threshold values can be found BL1 = 0.7 and BL2 = 1.1 for which NEGL = 0.5. However taking into account the number of observations, value BL2 should be rated more highly, for which  $\frac{9}{18} = 0.5$  can be observed, while for BL1, only  $\frac{4}{8} = 0.5$  can be observed.

Figure 7 was made for Set C of data.

In Figure 7 threshold value BG = 3.9 can be found for which ratio NEGG assumes the highest value, NEGG = 0.933.

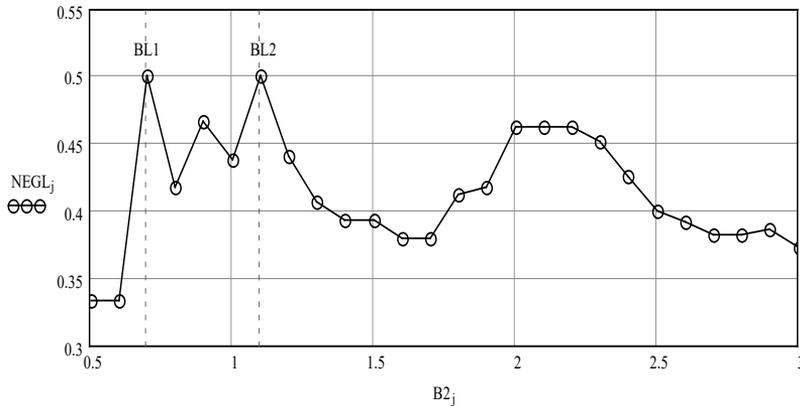


Figure 6. Values of ratio NEGL as a function of threshold value B2 (Set B)

Source: author's own study.

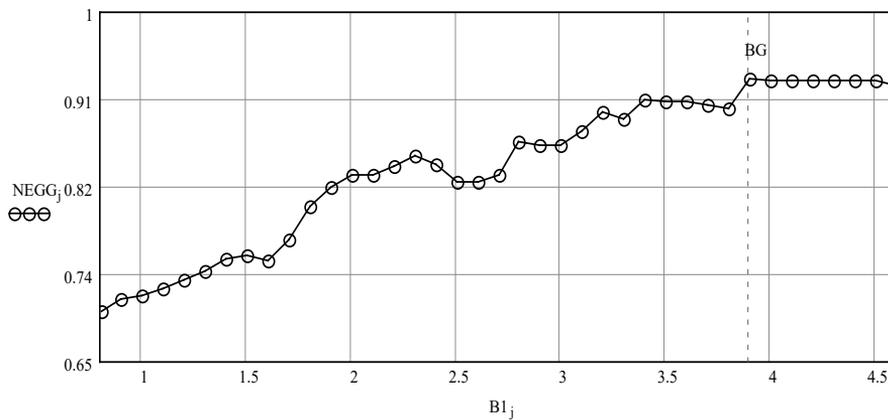


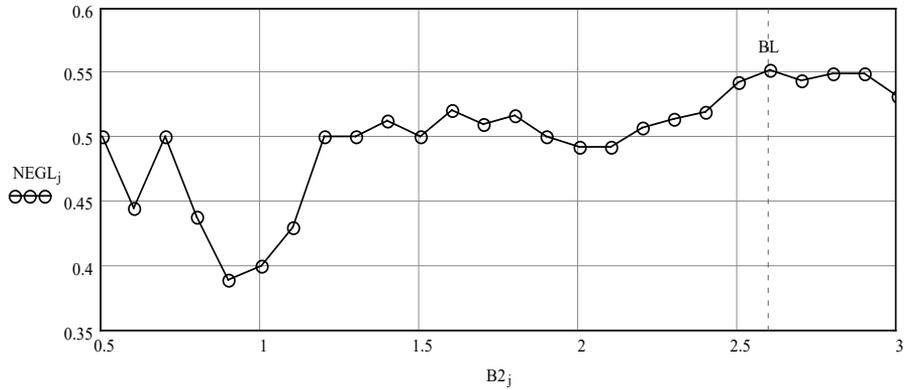
Figure 7. Values of ratio NEGG as a function of threshold value B1 (Set C)

Source: author's own study.

In Figure 8 threshold value BL = 2.6 can be found for which ratio NEGL assumes the highest value, NEGL = 0.552.

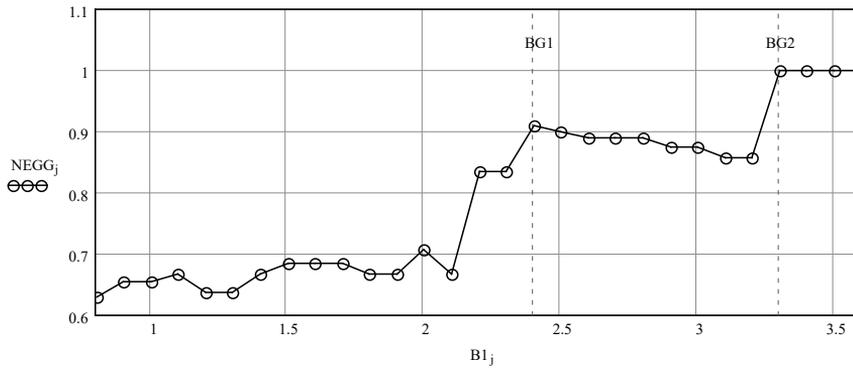
The last set of data (D) comprised the values of daily return rates on WIG30 bank sector stocks. Figure 9 presents the values of ratio NEGG as a function of threshold value B1. From the results presented in Figure 9, two threshold values were determined: BG1 = 2.4 for which  $NEGG1 = \frac{10}{11} = 0,909$ , and BG2=3.3, where  $NEGG2 = \frac{5}{5} = 1.0$ .

In Figure 10 threshold value BL = 1.2 can be found for which ratio NEGL assumes the highest value, NEGL = 0.875.



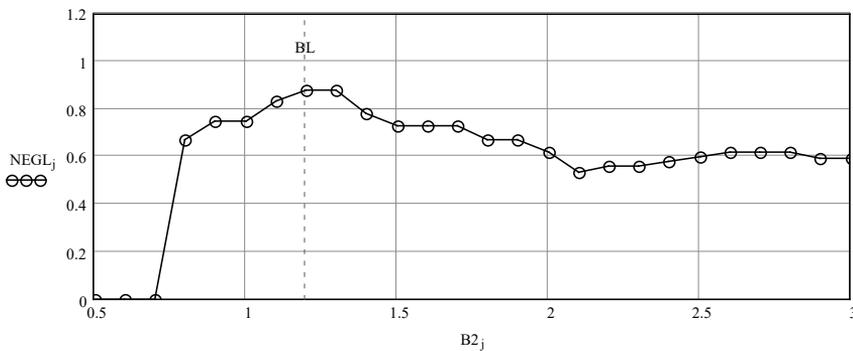
**Figure 8.** Values of ratio NEGL as a function of threshold value B2 (Set C)

Source: author's own study.



**Figure 9.** Values of ratio NEGG as a function of threshold value B1 (Set D)

Source: author's own study.



**Figure 10.** Values of ratio NEGL as a function of threshold value B2 (Set D)

Source: author's own study.

## Conclusions

Table 1 comprises the results of the calculations for particular sets of data.

**Table 1.** Compilation of calculation results for the data in sets A, B, C and D

Set of data	Gaussian distribution		Laplace distribution	
	threshold value	maximum	threshold value	maximum
	bg	negg	bl	negl
A	3.4	0.833	2.2	0.298
B	4.0	0.938	0.7	0.500
	5.3	1.000	1.1	0.500
C	3.9	0.933	2.6	0.552
D	2.4	0.909	1.2	0.875
	3.3	1.000		

Source: author's own study.

The results in Table 1 prove that the order of the data sets: A, B, C, D was not coincidental – the order corresponds to increasing maximum values of ratios NEGG and NEGL. The worst results for Set A stem from the heterogeneity of the data: daily, weekly and monthly values of return rates. As it was already mentioned in Introduction, the value of kurtosis for the monthly data was 3.34 times larger than for the daily data. Set A was included in the paper on purpose, however, so as to warn against combining return rates with different measure time intervals.

Sets B and C yield similar maximum values of ratios NEGG and NEGL, which results from the fact that they comprise homogenous (daily) data.

The best results were obtained for Set D, which comprised the data of doubled homogeneity, since they were daily data and the companies represented the same sector (banking).

As far as the hypothesis put forward in Introduction is concerned, it should be noticed that the hypothesis was validated for the normal distribution, since the maximum values of ratio NEGG > 0.9 – obtained for Sets B, C, D – are satisfying. In the case of the Laplace distribution, no unequivocal answer was found. Sets B and C did not provide decisive results. Only in the case of Set D, the value of ratio NEGL = 0.875 can be accepted as supportive evidence for the hypothesis.

## References

- Bednarz-Okrzyńska, K. (2016). *Wykorzystanie funkcji Laplace'a do modelowania rozkładu stóp zwrotu wybranych indeksów giełdowych i spółek*. Rozprawa doktorska. Szczecin.
- Bonato, M. (2011). Robust estimation of skewness and kurtosis in distributions with infinite higher moments. *Finance Research Letters*, 8, 77–87.
- Kim, T.H., White, H. (2004). On more robust estimation of skewness and kurtosis. *Finance Research Letters*, 1, 56–73.
- Kotz, S., Kozubowski, T.J., Podgórski, K. (2001). *The Laplace Distribution and Generalizations. A Revisit with Applications to Communications, Economics, Engineering and Finance*. Boston: Birkhauser.
- Krishnamoorthy, K. (2006). *Handbook of Statistical Distributions with Applications*. Taylor & Francis Group. LLC.

Sobczyk, M. (2004). *Statystyka*. Warszawa: Wydawnictwo Naukowe PWN.

Tarczyński, W. (2002). *Fundamentalny portfel papierów wartościowych*. Warszawa: PWE.

**Cite this article as:** Bednarz-Okrzyńska, K. (2018). Influence of the kurtosis value on the result of goodness of fit test for the daily rate of return on selected WIG20 and WIG30 companies. *European Journal of Service Management*, 3 (27/2), 27–36. DOI: 10.18276/ejism.2018.27/2-03.