GUS

Indeks 371262
e-ISSN 2657-9545
ISSN 0033-2372

Przegląd Statystyczny
Statistical Review

Vol. 70 | No. 4 | 2023

GŁÓWNY URZĄD STATYSTYCZNY
STATISTICS POLAND

INFORMATION FOR AUTHORS

Przegląd Statystyczny. Statistical Review publishes original research papers on theoretical and empirical topics in statistics, econometrics, mathematical economics, operational research, decision science and data analysis. The manuscripts considered for publication should significantly contribute to the theoretical aspects of the aforementioned fields or shed new light on the practical applications of these aspects. Manuscripts reporting important results of research projects are particularly welcome. Review papers, shorter papers reporting on major conferences in the field, and reviews of seminal monographs are eligible for submission on the Editor-in-Chief's request.

Since 1 May 2019, the journal has been publishing articles in English.

Any spelling style is acceptable as long as it is consistent within the manuscript.

All work should be submitted to the journal through the Editorial System (https://www.editorialsystem.com/pst).

For details of the submission process and editorial requirements please visit https://ps.stat.gov.pl/ForAuthors.

Przegląd Statystyczny
Statistical Review

Vol. 70 No. 4 2023

EDITORIAL OFFICE'S ADDRESS

University of Gdańsk, ul. Armii Krajowej 101, 81-824 Sopot, Poland

Language editing: Scientific Journals Division, Statistics Poland
Technical editing and typesetting: Statistical Publishing Establishment – team supervised by Maciej Adamowicz

Zakład Wydawnictw Statystycznych

Printed and bound: Statistical Publishing Establishment
al. Niepodległości 208, 00-925 Warsaw, Poland, zws.stat.gov.pl

Website: ps.stat.gov.pl

ISSN 0033-2372
e-ISSN 2657-9545
Index 371262

Information on the sales of the journal: Statistical Publishing Establishment
phone no. +48 22 608 32 10, +48 22 608 38 10

Order no. 333/2023

CONTENTS

Przegląd Statystyczny. Statistical Review, 2023, vol. 70, 4, 1–36 https://doi.org/10.59139/ps.2023.04.1
Received: 24.05.2024, revised: 6.07.2024, accepted: 9.07.2024

Modified Cramer-von Mises goodness-of-fit test for normality

Piotr Sulewski,[a] Damian Stoltmann[b]

Abstract. The first goal of the article is to apply the modified Cramer-von Mises (*CM*) goodness-of-fit test for normality to a practical problem. The modification of the test involves varying the formula for calculating the empirical distribution function (*EDF*). The critical values are obtained using the Monte Carlo method for sample sizes $n = 10{,}20$ and at a significance level of α=0.05. The second goal is to calculate the power of several tests for appropriately selected alternative distributions. The article shows that the values of constants α, β in the *EDF* formula affect the power of the *CM* test. The effectiveness of the new proposal is illustrated by the analysis of real data sets.

Keywords: empirical distribution function, goodness-of-fit test, Cramer-von Mises test, power of test
JEL: C02, C12, C46, G00

1. Introduction

Numerous goodness-of-fit tests (*GoFTs*) for normality have been considered and applied in many fields of science, including medicine, quality control and hydrology. *GoFTs* for normality are also very popular in economics and finance. They are used to analyse market behaviour (the distribution of rates of return, trading volume or asset prices), assess market efficiency, identify deviations from ideal market conditions and analyse stochastic processes (asset prices or changes in commodity prices). In econometrics, normality tests are used to check whether regression errors are normally distributed. This is important for the proper evaluation of regression models as the violation of the assumption of normality can lead to erroneous statistical conclusions. In demography, on the other hand, the fertility curve is almost normally distributed.

One of the most common normality testing procedures available in statistical software is the Cramer-von Mises (*CM*) test (Cramér, 1928; von Mises, 1931), which belongs to the group of empirical distribution function (*EDF*) tests. Other popular *EDF* tests include the Kolmogorov-Smirnov (*KS*) test (Kolmogorov, 1933; Smirnov, 1948), the Lilliefors (*LF*) test (Lilliefors, 1967), the Kuiper (*K*) test (Kuiper, 1960), the

[a] Pomeranian University, Institute of Exact and Technical Sciences, ul. Arciszewskiego 22a, 76-200 Słupsk, Poland, e-mail: piotr.sulewski@upsl.edu.pl, ORCID: https://orcid.org/0000-0002-0788-6567.
[b] Pomeranian University, Institute of Exact and Technical Sciences, ul. Arciszewskiego 22a, 76-200 Słupsk, Poland, e-mail: damian.stoltmann@upsl.edu.pl, ORCID: https://orcid.org/0000-0001-7053-2684.

Watson (*W*) test (Watson, 1962) and the Anderson-Darling (*AD*) test (Anderson & Darling, 1952).

Recently, many articles have been devoted to goodness-of-fit tests (*GoFTs*) for normality. Table 1 shows the authors of works created in the 21st century.

Table 1. Articles devoted to normal *GoFTs* created in the 21st century

Article	Sample sizes	Article	Sample sizes
Bonett and Seier (2002)	**10, 20, …, 50**, 100	Afeez et al. (2018)	**10, 30, 50**, 100, 300, 500, 1000
Aliaga et al. (2003)	X	Marange and Qin (2019)	**15, 30, 50**, 80, 100, 150, 200
Bontemps and Meddahi (2005)	100, 250, 500, 1000	Sulewski (2019)	**10, 12, …, 30, 40, 50**
Luceño (2006)	100	Tavakoli et al. (2019)	**5, 6, …, 15, 20, 25, 30, 40, 50,** …,100
Yazici and Yolacan (2007)	**20, 30, 40, 50**	Mishra et al. (2019)	**n<30, n>30**
Gel et al. (2007)	**20, 50,** 100	Kellner and Celisse (2019)	**50**, 75, 100, 200, 300, 400
Coin (2008)	**20, 50, 200**	Wijekularathna et al. (2020)	**5, 10, 20, 30, 50,** 75, 100, 200, 500, 1000, 2000
Brys et al. (2008)	100, 1000	Sulewski (2022)	**10, 14, 20**
Gel and Gastwirth (2008)	**30, 50**, 100	Hernandez (2021)	**5, 10, …, 30**
Romão et al. (2010)	**25, 50,** 100	Khatun (2021)	**10, 20, 25, 30, 40, 50,**100, 200, 300
Razali and Wah (2011)	20, 30, 50, 100, 200, …, 500, 1000, 2000	Arnastauskaitė et al. (2021)	**2^5**,2^6,…,2^10
Noughabi and Arghami (2011)	10, 20, 30, 50	Bayoud (2021)	**10, 20 ,…, 50**, 60, 80, 100
Yap and Sim (2011)	**10, 20, 30, 50,** 100, 300, 500, 1000, 2000	Uhm and Yi (2021)	**10, 20, 30,** 100, 200, 300
Chernobai et al. (2012)	X	Sulewski (2021)	**20, 50,** 100
Ahmad and Khan (2015)	**10, 20, …, 50,** 100, 200, 500	Desgagné et al. (2022)	**20, 50,** 100, 200
Mbah and Paothong (2015)	**10, 20, 30, 50,** 100, 200, …, 500, 1000, 2500, 5000	Uyanto (2022)	**10, 30, 50**, 70, 100
Torabi et al. (2016)	**10, 20, 50,** 100, 1000	Ma et al. (2024)	**10, 30, 50**
Feuerverger (2016)	200	Giles (2024)	**10, 25, 50,** 100, 250 ,500, 1000
Nosakhare and Bright (2017)	**5, 10, …, 50,** 100	Borrajo et al. (2024)	**50**, 100, 200, 500
Desgagné and Lafaye de Micheaux (2018)	**10, 12, …, 20, 50**, 100, 200	Terán-García and Pérez-Fernández (2024)	**25,** 900

Note. Sample sizes $n \leq 50$ are in bold.
Source: authors' work.

There are, of course, articles dedicated to the *CM* test. Durbin and Knott (1972) compared the *CM* test with *AD* and *W GoFTs*. Pettitt and Stephens (1976) used the *CM* test for a censored sample. Scott (2000) presented tables of unweighted *CM* statistics for one and two samples, and compared them to the limiting distribution.

Scott and Stewart (2011) presented tables for the Lilliefors distribution and the *CM* distribution, which are used to test for normality when the population mean and variance are unknown.

The small samples that dominate in Table 1 can be used in experimental economics where published papers can be found that describe samples of a dozen or so people in a group. In situations like these, strong testing proves very useful. In terms of the results, a hypothesis accepted in the original paper may be rejected if a more powerful test is applied.

Let $x_{(1)}, x_{(2)}, \dots, x_{(n)}$ be independent and identically distributed observations from unknown continuous cumulative distribution function (*CDF*) $F(x)$. We wish to determine whether $F(x)$ coincides with the *CDF* of normal distribution $\Phi(x)$. Then, we are interested in testing hypothesis $H_0: F(x) = \Phi(x)$ against hypothesis $H_1: F(x) \neq \Phi(x)$. The *EDF* is given by $F_n(x) = \frac{1}{n}\sum_{i=1}^{n} \theta(x - x_i)$, where $\theta(x) = 1$ for $x \geq 0$ and $\theta(x) = 0$ for $x < 0$.

The δ-corrected *KS* test (Harter et al., 1984), investigated further by Khamis (1990, 1992, 1993) redefines the value of the *EDF* at the data points and compares the redefined *EDF* to the *CDF* at the data points. Let the *EDF* at the *i*-th data point be given by

$$F_\delta(x) = \frac{i-\delta}{n-2\delta+1}, 0 \leq \delta \leq 1 \tag{1}$$

Harter et al. (1984) selected $\delta = 0,0.5,1$ for their study.

Bloom (1958) proposed α, β transformation

$$F_{\alpha,\beta}\left(x_{(i)}\right) = \frac{i-\alpha}{n-\alpha-\beta+1}, \alpha, \beta \leq 1 \tag{2}$$

to decrease the MSE of certain statistics. Note that $F_{\delta,\delta}(x) = F_\delta(x)$. This transformation was used to create *GoFTs*.

Sulewski (2022) used the Bloom formula to create the one-component Lilliefors *GoFT* with statistic

$$LF_1 = \underbrace{max}_{i}\left\{\left|F_{\alpha,\beta}\left(x_{(i)}\right) - \Phi\left(x_{(i)}\right)\right|\right\}. \tag{3}$$

We know perfectly well that the greatest discrepancy between the theoretical and empirical distribution functions may occur at different positions in the series. The probability of this discrepancy occurring for a given positional statistic r becomes smaller the more extreme the r is. Hence the idea of a two-component test statistic described in Sulewski (2021). The first component is, as in the original *LF* test, the absolute value of the greatest discrepancy between sample and population distributions. The second component is the position in an ordered sample at which this discrepancy is located. The two-component Lilliefors statistic is given by

$$LF_2(r) = \underbrace{max}_{i} \{|F_{\alpha,\beta}(x_{(i)}) - \Phi(x_{(i)})|\}. \tag{4}$$

Simulation studies for the one- and two-component Lilliefors tests were carried out for the following methods of calculating $F_{\alpha,\beta}(x_{(i)})$ $(\alpha, \beta \leq 1)$:
1. $F_{0,1}(x_{(i)}) = \frac{i}{n}$ – occurs in the *KS* statistic;
2. $F_{1,0}(x_{(i)}) = \frac{i-1}{n}$ – occurs in the *KS* statistic;
3. $F_{0.5,0.5}(x_{(i)}) = \frac{i-0.5}{n}$ – occurs in the *CM* statistic;
4. $F_{0,0}(x_{(i)}) = \frac{i}{n+1}$ – the mean value of *i*-th order statistics of the beta distribution;
5. $F_{0.3,0.3}(x_{(i)}) = \frac{i-0.3}{n+0.4}$ – the median of *i*-th order statistics of the beta distribution;
6. $F_{0.375,0.375}(x_{(i)}) = \frac{i-0.375}{n+0.25}$ – the mean value of *i*-th order statistics of the normal distribution;
7. $F_{0.3175,0.3175}(x_{(i)}) = \frac{i-0.3175}{n+0.365}$ – founded by Filliben (1975);
8. $F_{1,1}(x_{(i)}) = \frac{i-1}{n-1}$ – founded by Harter et al. (1984).

In six of the *EDF* definitions listed above (except $F_{0,1}$ and $F_{1,0}$), $\alpha = \beta$.

The first goal of this paper is to propose a modified *CM* test for normality. The second goal is to calculate the power of several tests for appropriately selected alternative distributions (alternatives).

The rest of this paper is structured as follows. In Section 2, we define a new version of the *CM* test. The similarity measure of the alternative to the normal distribution is described in Section 3. Section 4 presents the alternatives divided into nine groups according to their skewness and excess kurtosis. Simulation studies are presented in Section 5 and real data examples are provided in Section 6. Finally, the concluding remarks are presented in Section 7. Additional material can be found in the Appendix.

2. Modified Cramer-Von Mises test for normality

The *CM* statistic belongs to the class of quadratic *EDF* statistics with measure

$$n \int_{-\infty}^{\infty} [F_n(x) - F(x)]^2 \omega(x) dF(x), \tag{5}$$

where $\omega(x)$ is a weighting function. When the weighting function is $\omega(x) = 1$, the *CM* statistic is obtained:

$$n \int_{-\infty}^{\infty} [F_n(x) - F(x)]^2 dF(x). \tag{6}$$

The *CM* test for normality is defined based on the following statistics:

$$CM = n \int_{-\infty}^{\infty} [F_n(x) - \Phi(x)]^2 \ \phi(x; \mu, \sigma) dx, \tag{7}$$

where $\phi(x), \Phi(x)$ are the *PDF* and *CDF* of the normal distribution, respectively. The simpler form of the *CM* statistic is

$$CM = \frac{1}{12n} + \sum_{i=1}^{n} \left[\Phi\left(x_{(i)}\right) - \frac{i-0.5}{n}\right]^2. \tag{8}$$

The *CM* test is also presented in the second version, namely (Stephens, 1974)

$$CM_s = \left(1 + \frac{1}{2n}\right) CM. \tag{9}$$

We define the modified *CM* (*MCM*) statistic using the Bloom formula. The *MCM* statistic is given by

$$MCVM(\alpha, \beta) = \frac{1}{12n} + \sum_{i=1}^{n} \left[\Phi\left(x_{(i)}\right) - F_{\alpha,\beta}\left(x_{(i)}\right)\right]^2, \tag{10}$$

where $\alpha, \beta \in [0, 1]$. Note that $MCVM(0.5, 0.5) = CVM$. H_0 is rejected for large $MCVM$ statistic values.

3. Similarity measure

Let us assume that

$$\begin{gathered} \bar{x} = \frac{1}{n} \sum_{i=1}^{n} x_i \, , s^2 = \frac{1}{n-1} \sum_{i=1}^{n} (x_i - \bar{x})^2 \, , z_{(i)} = \frac{(x_{(i)} - \bar{x})}{s}, \\ m_k = \frac{1}{n} \sum_{i=1}^{n} \left(x_{(i)} - \bar{x}\right)^k , \gamma_1 = \frac{m_3}{s^3}, \bar{\gamma}_2 = \frac{m_4}{s^4} - 3. \end{gathered} \tag{11}$$

Let us remember that the Malachov inequality is defined as $\bar{\gamma}_2 \geq {\gamma_1}^2 - 2$.

A review of recent statistical literature shows that the small skewness γ_1 and excess kurtosis $\bar{\gamma}_2$ values do not dominate in testing for normality. It is very interesting to see how the *GoFTs* respond to samples coming from alternatives close to the normal distribution.

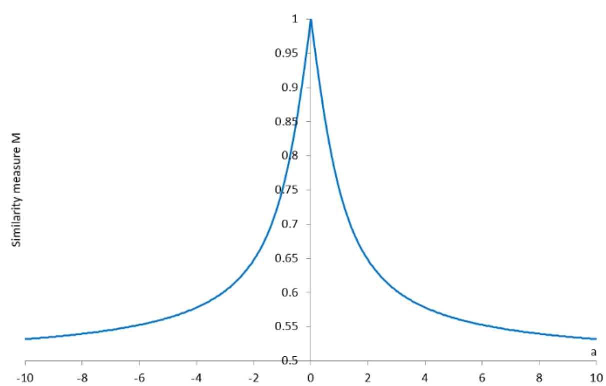
Let $f(x;\mathbf{\theta})$ be a *PDF* of the alternative with vector of parameters $\mathbf{\theta}$. Similarity measure M of alternative (A) to the normal distribution is defined as (Sulewski, 2022)

$$M_A(\boldsymbol{\theta};\mu,\sigma) = \int_{-\infty}^{\infty} min[f(x;\boldsymbol{\theta}), \phi(x;\mu,\sigma)]dx\,, \tag{12}$$

where $\phi(x;\mu,\sigma)$ is the PDF of the normal distribution. $M_A(\mathbf{\theta};\mu,\sigma)$ takes on the values of $[0, 1]$. $M_A(\mathbf{\theta};\mu,\sigma) = 1$ when *PDFs* are identical.

Figure 1 shows the values of similarity measure (12) when an alternative is the skew normal (*SN*) distribution (Azzalini, 1985) with *PDF* $f_{SN}(x;a) = 2\phi(x;0,1)\Phi(ax;0,1)$ $(a \in R)$. Note that if $a \to \mp\infty$, then $M_{SN}(a;0,1) = 0.5$.

Figure 1. Similarity measure $M_{SN}(a;0,1)$ for the skew normal distribution



Source: authors' work.

4. Alternative distributions

As mentioned earlier, there are many articles devoted to testing for normality. In these articles, a lot of alternative distributions (alternatives) were used, including both asymmetric and symmetric ones. Symmetric distributions with undefined γ_1 and $\bar{\gamma}_2$ are Cauchy and slash distributions.

According to the statistical literature, alternatives can be divided into four groups, depending on the support and shape of their densities (see e.g. Esteban et al., 2001; Torabi et al., 2016). These groups include symmetric alternatives with support $(-\infty,\infty)$, asymmetric alternatives with support $(-\infty,\infty)$, alternatives with support

$(0, \infty)$ and alternatives with support $(0, 1)$. Gan and Koehler (1990), Krauczi (2009) and Torabi et al. (2016) divided alternatives into five groups: asymmetric short-tailed, asymmetric long-tailed, symmetric short-tailed, symmetric close to normal and symmetric long-tailed alternatives.

Our idea is to divide alternatives into nine groups according to their γ_1 and $\bar{\gamma}_2$ signs. Groups O–H are defined in Table 2.

Table 2. Groups of alternatives with signs of γ_1 and $\bar{\gamma}_2$

Group	γ_1	$\bar{\gamma}_2$
O	zero	zero
A	positive	positive
B	negative	positive
C	zero	positive
D	zero	negative
E	positive	negative
F	negative	negative
G	positive	zero
H	negative	zero

Source: authors' work.

The main criterion for selecting an alternative for the Monte Carlo simulation is that γ_1 and $\bar{\gamma}_2$ calculated for the alternative parameters belong to the O, A–H groups. This criterion is fulfilled by distributions defined in an infinite domain such as:

- the Edgeworth series (*ES*) with parameters γ_1 and $\bar{\gamma}_2$ as a monolithic distribution;
- the Pearson distribution (*P*) with parameters γ_1 and $\bar{\gamma}_2$ as a monolithic distribution;
- the normal mixture distribution (*NM*) with 5 parameters as a mixture of two normal distributions;
- the normal logistic mixture distribution (*NLM*) with 5 parameters as a mixture of normal and non-normal distributions;
- the normal distribution with a plasticising component (*NDPC*) with six parameters as a mixture of two various distributions;
- the plasticising component mixture (*PCM*) with seven parameters as a mixture of two identical non-normal distributions that characterise multimodality.

We chose values of alternative parameters to obtain desired similarity measure values of the alternative to the normal distribution. These analysed values, if possible, are $0.5, 0.75, 0.9$.

The appendix presents Tables 1A–6A with vectors of alternative parameter $\boldsymbol{\theta}$, mean μ_a, standard deviation σ_a, skewness γ_1, excess kurtosis $\bar{\gamma}_2$ and similarity measure M for the analysed alternatives. The *PDF* formulas and *PDF* curves (see Figures 1A–6A) for the alternative $\boldsymbol{\theta}$ values are also provided in the Appendix. As can be seen in Figure 1A, the *ES* distribution is not suitable for simulation studies as negative *PDF* values are observed even though the normalisation condition is met. Figure 2A indicates very interesting bimodal shapes. Figure 3A and Figure 5A present both unimodal and bimodal shapes. Figure 4A shows unimodal shapes, while very interesting multimodal shapes may be observed in Figure 6A.

5. Power comparisons

The new $MCM(\alpha, \beta)$ test, where $MCM(0.5, 0.5) = CM$ was compared with the CM_S, AD, LF and the Shapiro-Wilk (SW) (Shapiro & Wilk, 1965) tests. To study the power of each of the discussed tests, critical values $cv_{0.05}$ (type I error equals $\alpha = 0.05$) were estimated using $m = 10^6$ order statistics. The power of tests (*PoTs*) was calculated based on $rep = 10^5$ test statistic values.

Table 3 shows critical values and test sizes of the analysed *GoFTs* for sample sizes $n = 10, 20$ and $\alpha = 0.05$. The TS values are close to 0.05, so the simulation procedures are correct.

Table 3. Critical values (*CV*) and test sizes (*TS*) of the analysed *GoFTs* for sample sizes $n = 10, 20$

No	*GoFT*	*CV*		*TS*	
		$n = 10$	$n = 20$	$n = 10$	$n = 20$
1	$MCM(0, 1)$	0.15247	0.14008	0.051	0.051
2	$MCM(1, 0)$	0.15232	0.13992	0.051	0.051
3	$MCM(0, 0)$	0.11423	0.11992	0.052	0.051
4	$MCM(0.3, 0.3)$	0.11487	0.12035	0.052	0.051
5	$MCM(1, 1)$	0.14934	0.13820	0.052	0.052
6	$MCM(0.375, 0.375)$	0.11612	0.12101	0.052	0.051
7	$MCM(0.3175, 0.3175)$	0.11510	0.12048	0.052	0.051
8	$MCM(0.5, 0.5) = CM$	0.11922	0.12285	0.051	0.051
9	CM_S	0.12518	0.12593	0.051	0.051
10	AD	0.68511	0.72118	0.052	0.051
11	LF	0.26186	0.19187	0.051	0.050
12	SW^*	0.84451	0.90441	0.051	0.050

Note. H_0 is rejected for small statistical values.
Source: authors' work.

Tables 7A–14A show how a selection of the *EDF* influences the power of the *MCM* test. The alternatives are indexed. The larger the index (ID), the more the distribution resembles a normal distribution, i.e. the PoTs should decrease as the index value increases. ID = 1 denotes similarity measure $M = 0.5$, while ID = 4 denotes similarity measure $M = 0.95$. The highest *PoTs* of the $MCM(\alpha, \beta)$ values are underlined. The highest *PoTs* for all the analysed tests are in bold.

The simulation results in Tables 7A–14A show that the *MCM* test with the analysed *EDFs* is the most powerful for all the considered similarity measures, alternatives and $n = 10, 20$ ($n = 10$) in 83.75% (94.38%.). The new proposal for group of alternatives A and $n = 10, 20$, ($n = 10$) is the most powerful in 82.5% (100%), for group B and $n = 10, 20$, ($n = 10$) is the most powerful in 85% (100%), for group C and $n =$

$10, 20$, ($n = 10$) is the most powerful in 67.50% (90%), for group D and $n = 10, 20$, ($n = 10$) is the most powerful in 77.5% (85%), for group E and $n = 10, 20$, ($n = 10$) is the most powerful in 92.5% (95%), for group F and $n = 10, 20$, ($n = 10$) is the most powerful in 90% (95%), for group G and $n = 10, 20$, ($n = 10$) is the most powerful in 87.5% (95%) and for group H and $n = 10, 20$, ($n = 10$) is the most powerful in 87.5% (95%). $MCM(0,1)$ with $F_{0,1}(x_{(i)}) = \frac{i}{n}$ dominates for groups A, E and G. $MCM(1,0)$ with $F_{1,0}(x_{(i)}) = \frac{i-1}{n}$ dominates for groups B, F and H. $MCM(0,0)$ with $F_{1,1}(x_{(i)}) = \frac{i-1}{n+1}$ dominates for group C and $MCM(0,0)$ with $F_{0,0}(x_{(i)}) = \frac{i}{n+1}$ dominates for group D. Powers of the *CM* and CM_S tests are the same. We assume that a *GoFT* detects abnormal samples if its power is at least 0.06. Thanks to similarity measures $M = 0.9, 0.95$, this situation occurs in 30%. For alternatives *P*, *NM*, *NLM*, *NDPC* and *PCM*, the test power is less than 0.6 in 14%, 45%, 4%, 44% and 43% of cases, respectively. For alternative groups A – H, the test power is less than 0.6 in 20%, 18%, 22%, 49%, 32%, 35%, 27% and 37% of cases, respectively.

6. Real data examples

In this section, we present an application of the *MCM* test in eight real data sets to illustrate its potentiality. Details related to examples I – VIII are presented in Table 4.

Table 4. Real data examples with sources, sample size, skewness and excess kurtosis values

Ex	Description	Source	n	γ_1	$\bar{\gamma}_2$
I	Strength measured in GPA for single carbon fibres and impregnated 1,000-carbon fibre tows (gauge lengths of 20 mm):1.312, 1.314, 1.479, 1.552, 1.7, 1.803, 1.861, 1.865, 1.944, 1.958, 1.966, 1.997, 2.006, 2.021, 2.027, 2.055, 2.063, 2.098, 2.14, 2.179, 2.224, 2.24, 2.253, 2.27, 2.272, 2.274, 2.301, 2.301, 2.359, 2.382, 2.382, 2.426, 2.434, 2.435, 2.478, 2.49, 2.511, 2.514, 2.535, 2.554, 2.566, 2.57, 2.586, 2.629, 2.633, 2.642, 2.648, 2.684, 2.697, 2.726, 2.77, 2.773, 2.8, 2.809, 2.818, 2.821, 2.848, 2.88, 2.809, 2.818, 2.821, 2.848, 2.88, 2.954, 3.012, 3.067, 3.084, 3.09, 3.096, 3.128, 3.233, 3.433, 3.585, 3.585.	Nofal et al., 2017	74	−0.154	−0.049
II	Macroeconomic data set with information on the number of members of the armed forces	R package longley[4]	16	−0.404	−0.949
III	Socio-economic data (percentage of males involved in agriculture as occupation) for 47 French-speaking provinces of Switzerland	R package swiss[2]	47	−0.331	−0.793

Table 4. Real data examples with sources, sample size, skewness and excess kurtosis values (cont.)

IV	Socio-economic data (percentage of draftees receiving the highest mark on army examination) for 47 French-speaking provinces of Switzerland	R package swiss[3]	47	0.461	−0.011
V	Average heights for American women aged 30–39	R package women[1]	15	0	−1.211
VI	Percent proportion of favourable responses to the question on promotion opportunities from a survey of the clerical employees of a large financial organisation; the data are aggregated from the questionnaires of approximately 35 employees for each of the 30 (randomly selected) departments.	R package attitude[6]	30	−0.911	0.388
VII	Percent proportion of favourable responses to the question on opportunities to learn from a survey of the clerical employees of a large financial organisation; the data are aggregated from the questionnaires of the approximately 35 employees for each of the 30 (randomly selected) departments.	R package attitude[3]	30	0.399	−0.229
VIII	Percent proportion of favourable responses to the question whether the organisation does not allow special privileges from a survey of the clerical employees of a large financial organisation; the data are aggregated from the questionnaires of the approximately 35 employees for each of the 30 (randomly selected) departments.	R package attitude[2]	30	−0.227	−0.514

Source: authors' work.

When fitting the normal distribution to the data, we calculate the p-values for the analysed *GoFTs* based on 10^5 statistic values (see Table 5).

Table 5. The *p*-values for the *GoFTs* related to examples I–VIII

GoFT	I	II	III	IV	V	VI	VII	VIII
$MCM(0,1)$	0.829	0.237	0.304	**0.250**	0.906	0.027	**0.215**	0.638
$MCM(1,0)$	**0.660**	**0.090**	**0.149**	0.493	0.907	**0.005**	0.455	**0.516**
$MCM(0,0)$	0.775	0.119	0.171	0.380	0.797	0.011	0.315	0.533
$MCM(0.3,0.3)$	0.763	0.127	0.192	0.362	0.897	0.010	0.312	0.563
$MCM(1,1)$	0.724	0.162	0.256	0.327	0.996	0.010	0.313	0.630
$MCM(0.375,0.375)$	0.759	0.129	0.198	0.358	0.918	0.010	0.312	0.571
$MCM(0.3175,0.3175)$	0.762	0.127	0.193	0.361	0.902	0.010	0.312	0.565
$MCM(0.5,0.5)=CM$	0.753	0.134	0.208	0.352	0.946	0.010	0.311	0.584
CM_S	0.753	0.134	0.208	0.352	0.946	0.010	0.311	0.584
AD	0.756	0.106	0.195	0.362	0.925	0.015	0.417	0.567
LF	0.826	0.094	0.231	0.287	0.996	0.028	0.529	0.568
SW	0.728	0.111	0.191	0.254	**0.729**	0.034	0.640	0.552

Note. The optimal *MCM(α,β)* test is underlined. The lowest *p*-value for all the analysed tests are in bold.
Source: authors' work.

The optimal $MCM(\alpha,\beta)$ test for examples I-III, VI and VIII is the $MCM(1,0)$ test. The obtained result is consistent with the simulation results showing that for groups B ($\gamma_1 < 0, \bar{\gamma}_2 > 0$) and F ($\gamma_1 < 0, \bar{\gamma}_2 < 0$), the most powerful is $MCM(1,0)$. Non-normality is the most pronounced by the $MCM(1,0)$ test.

The optimal $MCM(\alpha,\beta)$ test for examples IV and VII is the $MCM(0,1)$ test. The obtained result is consistent with the simulation results indicating that for groups E ($\gamma_1 > 0, \bar{\gamma}_2 < 0$), the most powerful is $MCM(0,1)$. The non-normality is the most pronounced by the $MCM(0,1)$ test.

The optimal $MCM(\alpha,\beta)$ test for example V is the $MCM(0,0)$ test. The obtained result is consistent with the simulation results stating that for groups D ($\gamma_1 = 0, \bar{\gamma}_2 < 0$), the most powerful is $MCM(0,0)$. The non-normality is the most pronounced by the SW test.

7. Real data examples

The obtained results show that the methods of calculating the *EDF* depend on the nature of the non-normal (alternative) distribution. There is a method of calculating the *EDF* $F_{\alpha,\beta}(x_i)$ for groups A – H which maximises the power of the $MCM(\alpha,\beta)$ test. $MCM(0,1)$ dominates for alternative groups A, E and G, $MCM(1,0)$ for groups B, F and H, $MCM(1,1)$ for group C and $MCM(0,0)$ for group D.

The new proposal is the most powerful in 95% of the cases for $n = 10$ and in 84% of the cases for $n = 20$.

The new *GoFT* is the best in 100% of cases in groups A and B, in 90% of cases in group C, in 85% of cases in group D and in 95% of cases in groups E – H.

For the normal logistic mixture and Pearson distributions, the analysed tests do not detect abnormal samples in only 4% and 14% of cases, respectively. With regards to the alternative groups, the best results are achieved by groups B, A and C (asymmetric alternatives with a positive excess kurtosis), while the worst by group D (symmetrical alternatives with a negative excess kurtosis).

The good performance of the *MCM* test against other most popular *GoFTs* is illustrated through the analysis of real data sets.

Appendix

Edgeworth series distribution

PDF of the Edgeworth series (*ES*) with parameters γ_1 and $\bar{\gamma}_2$ is given by

$$f_{ES}(x;\gamma_{1,},\bar{\gamma}_2) = \phi(x;0,1)\left(1 + \frac{1}{3!}\gamma_{1,}(x^3 - 3x) + \frac{1}{4!}\bar{\gamma}_2(x^4 - 6x^2 + 3)\right) \ (x \in R),$$

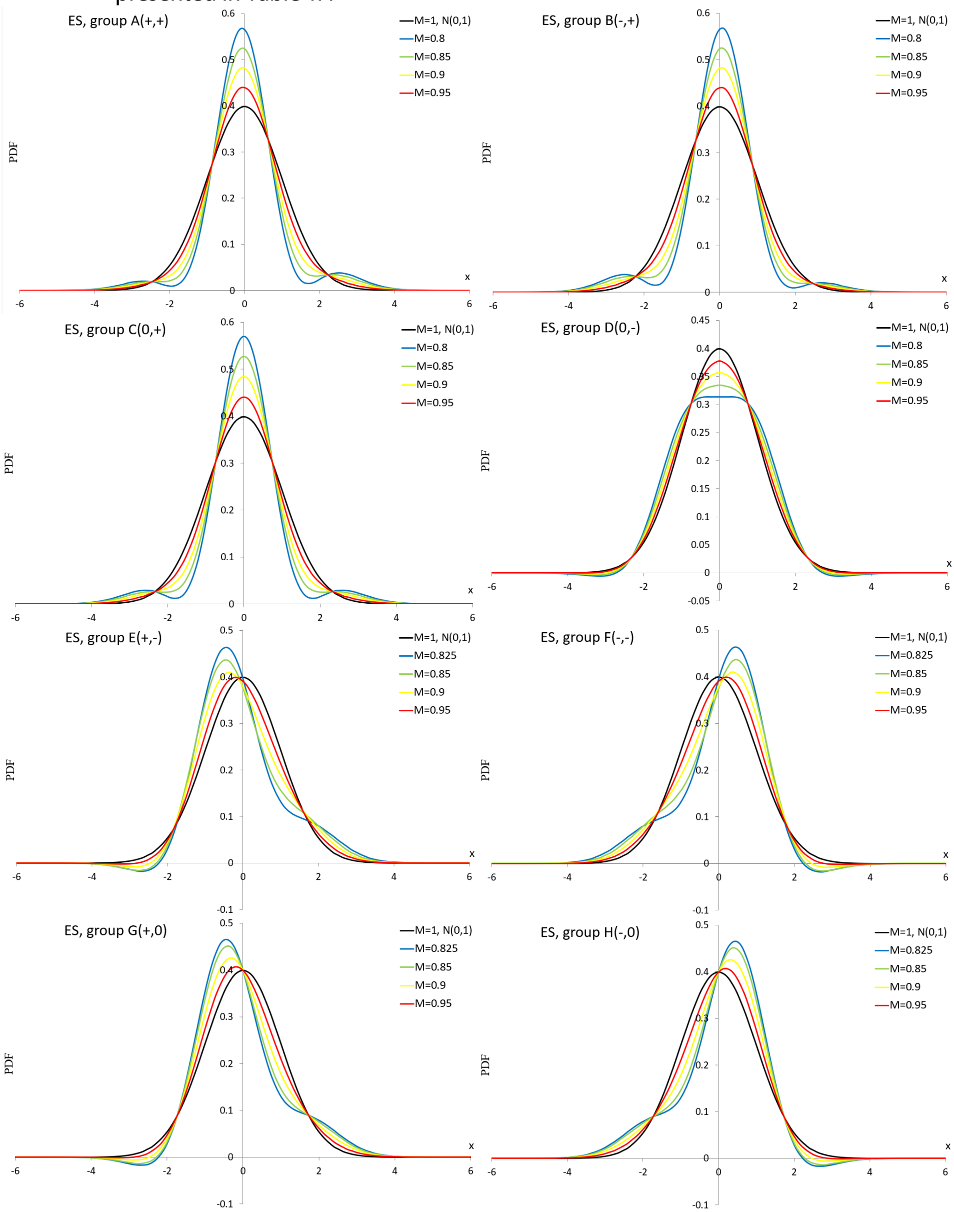
where $\gamma_{1,} \in R, \bar{\gamma}_2 \geq -2$.

Table 1A. Vectors of *ES* parameter $\boldsymbol{\theta}$, mean μ_a, standard deviation σ_a, skewness γ_1, excess kurtosis $\bar{\gamma}_2$ and similarity measure *M*. Groups O, A–H

Group	$\boldsymbol{\theta} = (\gamma_{1,}, \bar{\gamma}_2)$	μ_a	σ_a	$\gamma_{1,}$	$\bar{\gamma}_2$	$M(\boldsymbol{\theta}; \mu, \sigma)$
O	$(0, 0)$	0	1	0	0	$M(\boldsymbol{\theta}; 0,1) = 1$
A	0.4, 3.33	0	1	0.4	3.33	$M(\boldsymbol{\theta}; 0,1) = 0.8$
	0.3, 2.499	0	1	0.3	2.499	$M(\boldsymbol{\theta}; 0,1) = 0.85$
	0.2, 1.666	0	1	0.2	1.666	$M(\boldsymbol{\theta}; 0,1) = 0.9$
	0.1, 0.833	0	1	0.1	0.833	$M(\boldsymbol{\theta}; 0,1) = 0.95$
B	-0.4, 3.33	0	1	-0.4	3.33	$M(\boldsymbol{\theta}; 0,1) = 0.8$
	-0.3, 2.499	0	1	-0.3	2.499	$M(\boldsymbol{\theta}; 0,1) = 0.85$
	-0.2, 1.666	0	1	-0.2	1.666	$M(\boldsymbol{\theta}; 0,1) = 0.9$
	-0.1, 0.833	0	1	-0.1	0.833	$M(\boldsymbol{\theta}; 0,1) = 0.95$
C	0, 3.428	0	1	0	3.428	$M(\boldsymbol{\theta}; 0,1) = 0.8$
	0, 2.571	0	1	0	2.571	$M(\boldsymbol{\theta}; 0,1) = 0.85$
	0, 1.71	0	1	0	1.71	$M(\boldsymbol{\theta}; 0,1) = 0.9$
	0, 0.85	0	1	0	0.85	$M(\boldsymbol{\theta}; 0,1) = 0.95$
D	0, -3.428	0	1	0	-3.428	$M(\boldsymbol{\theta}; 0,1) = 0.8$
	0, -2.571	0	1	0	-2.571	$M(\boldsymbol{\theta}; 0,1) = 0.85$
	0, -1.71	0	1	0	-1.71	$M(\boldsymbol{\theta}; 0,1) = 0.9$
	0, -0.85	0	1	0	-0.85	$M(\boldsymbol{\theta}; 0,1) = 0.95$
E	1.39, -0.067	0	1	1.39	-0.067	$M(\boldsymbol{\theta}; 0,1) = 0.825$
	1.175, -0.46	0	1	1.175	-0.46	$M(\boldsymbol{\theta}; 0,1) = 0.85$
	0.775, -0.408	0	1	0.775	-0.408	$M(\boldsymbol{\theta}; 0,1) = 0.9$
	0.39, -0.15	0	1	0.39	-0.15	$M(\boldsymbol{\theta}; 0,1) = 0.95$
F	-1.39, -0.067	0	1	-1.39	-0.067	$M(\boldsymbol{\theta}; 0,1) = 0.825$
	-1.175, -0.46	0	1	-1.175	-0.46	$M(\boldsymbol{\theta}; 0,1) = 0.85$
	-0.775, -0.408	0	1	-0.775	-0.408	$M(\boldsymbol{\theta}; 0,1) = 0.9$
	-0.39, -0.15	0	1	-0.39	-0.15	$M(\boldsymbol{\theta}; 0,1) = 0.95$
G	1.391, 0	0	1	1.391	0	$M(\boldsymbol{\theta}; 0,1) = 0.825$
	1.19, 0	0	1	1.19	0	$M(\boldsymbol{\theta}; 0,1) = 0.85$
	0.795 ,0	0	1	0.795	0	$M(\boldsymbol{\theta}; 0,1) = 0.9$
	0.4, 0	0	1	0.4	0	$M(\boldsymbol{\theta}; 0,1) = 0.95$
H	-1.391, 0	0	1	-1.391	0	$M(\boldsymbol{\theta}; 0,1) = 0.825$
	-1.19, 0	0	1	-1.19	0	$M(\boldsymbol{\theta}; 0,1) = 0.85$
	-0.795, 0	0	1	-0.795	0	$M(\boldsymbol{\theta}; 0,1) = 0.9$
	-0.4, 0	0	1	-0.4	0	$M(\boldsymbol{\theta}; 0,1) = 0.95$

Source: authors' work.

Figure 1A. *PDF* curves of the Edgeworth series distribution for parameter values presented in Table 1A



Source: authors' work.

Pearson distribution

Let $a = \frac{2\bar{\gamma}_2 - 3\gamma_1^2}{10\bar{\gamma}_2 - 5\gamma_1^2 + 12}$, $b = \frac{|\gamma_1|(\bar{\gamma}_2 + 6)}{10\bar{\gamma}_2 - 5\gamma_1^2 + 12}$, $c = \frac{4\bar{\gamma}_2 - 3\gamma_1^2 + 12}{10\bar{\gamma}_2 - 5\gamma_1^2 + 12}$, $\Delta = b^2 - 4ac$, then the *PDF* of the Pearson (P) distribution is given by

$$f_P(x;\gamma_{1,},\bar{\gamma}_2) = \begin{cases} \dfrac{exp\left[\dfrac{2ab-b}{a(2ax+b)}\right]}{C_1(2ax+b)^{1/a}} & \Delta = 0 \\ \dfrac{exp\left[\dfrac{b-2ab}{a\sqrt{4ac-b^2}}tan^{-1}\left(\dfrac{2ax+b}{\sqrt{4ac-b^2}}\right)\right]}{C_2(ax^2+bx+c)^{1/(2a)}} & \Delta < 0 \\ \dfrac{\left(\dfrac{2ax+b-\sqrt{b^2-4ac}}{2ax+b+\sqrt{b^2-4ac}}\right)^{\frac{b-2ab}{2a\sqrt{b^2-4ac}}}}{C_3(ax^2+bx+c)^{1/(2a)}} & \Delta > 0 \end{cases}$$

$$f_P(x;\gamma_{1,},\bar{\gamma}_2) = \begin{cases} \dfrac{exp\left[\dfrac{2ab-b}{a(2ax+b)}\right]}{C_1(2ax+b)^{1/a}} & \Delta = 0 \\ \dfrac{exp\left[\dfrac{b-2ab}{a\sqrt{4ac-b^2}}tan^{-1}\left(\dfrac{2ax+b}{\sqrt{4ac-b^2}}\right)\right]}{C_2(ax^2+bx+c)^{1/(2a)}} & \Delta < 0, \\ \dfrac{\left(\dfrac{2ax+b-\sqrt{b^2-4ac}}{2ax+b+\sqrt{b^2-4ac}}\right)^{\frac{b-2ab}{2a\sqrt{b^2-4ac}}}}{C_3(ax^2+bx+c)^{1/(2a)}} & \Delta > 0 \end{cases}$$

where $x \in R, \gamma_{1,} \in R, \bar{\gamma}_2 \geq -2$ and C_1, C_2, C_3 are normalising constants defined as

$$C_1 = \int_{-\infty}^{\infty} \frac{exp\left[\dfrac{2ab-b}{a(2ax+b)}\right]}{(2ax+b)^{1/a}} dx\,,$$

$$C_2 = \int_{-\infty}^{\infty} \frac{exp\left[\dfrac{b-2ab}{a\sqrt{4ac-b^2}}tan^{-1}\left(\dfrac{2ax+b}{\sqrt{4ac-b^2}}\right)\right]}{(ax^2+bx+c)^{1/(2a)}} dx,$$

$$C_3 = \int_{-\infty}^{\infty} \frac{\left(\dfrac{2ax+b-\sqrt{\Delta}}{2ax+b+\sqrt{\Delta}}\right)^{\frac{b-2ab}{2a\sqrt{\Delta}}}}{C_8(ax^2+bx+c)^{1/(2a)}} dx.$$

Table 2A. Vectors of the Pearson distribution parameter $\boldsymbol{\theta}$, mean μ_a, standard deviation σ_a, skewness γ_1, excess kurtosis $\bar{\gamma}_2$ and similarity measure *M*. Groups O, A–H

Group	$\boldsymbol{\theta} = (\gamma_1, \bar{\gamma}_2)$	μ_a	σ_a	$\gamma_{1,}$	$\bar{\gamma}_2$	$M(\boldsymbol{\theta}; \mu, \sigma)$
O	(0, 0)	0	1	0	0	$M(\boldsymbol{\theta}; 0,1) = 1$
A	(2.04, 4.1)	0	1	2.04	4.1	$M(\boldsymbol{\theta}; 0,1) = 0.5$
	(1.62, 3.845)	0	1	1.62	3.845	$M(\boldsymbol{\theta}; 0,1) = 0.75$
	(0.9, 2)	0	1	0.9	2	$M(\boldsymbol{\theta}; 0,1) = 0.9$
	(0.4, 0.94)	0	1	0.4	0.94	$M(\boldsymbol{\theta}; 0,1) = 0.95$
B	(-2.04, 4.1)	0	1	-2.04	4.1	$M(\boldsymbol{\theta}; 0,1) = 0.5$
	(-1.62, 3.845)	0	1	-1.62	3.845	$M(\boldsymbol{\theta}; 0,1) = 0.75$
	(-0.9, 2)	0	1	-0.9	2	$M(\boldsymbol{\theta}; 0,1) = 0.9$
	(-0.4, 0.94)	0	1	-0.4	0.94	$M(\boldsymbol{\theta}; 0,1) = 0.95$
C	(0, 11.2)	0	1	0	11.2	$M(\boldsymbol{\theta}; 0,1) = 0.9$
	(0, 3.65)	0	1	0	3.65	$M(\boldsymbol{\theta}; 0,1) = 0.925$
	(0, 1.521)	0	1	0	1.521	$M(\boldsymbol{\theta}; 0,1) = 0.95$
	(0, 0.55)	0	1	0	0.55	$M(\boldsymbol{\theta}; 0,1) = 0.975$
D	(0, -1.695)	0	1	0	-1.695	$M(\boldsymbol{\theta}; 0,1) = 0.5$
	(0, -1.315)	0	1	0	-1.315	$M(\boldsymbol{\theta}; 0,1) = 0.75$
	(0, -0.89)	0	1	0	-0.89	$M(\boldsymbol{\theta}; 0,1) = 0.9$
	(0, -0.588)	0	1	0	-0.588	$M(\boldsymbol{\theta}; 0,1) = 0.95$
E	(0.985, -0.5)	0	1	0.985	-0.5	$M(\boldsymbol{\theta}; 0,1) = 0.5$
	(0.715, -0.475)	0	1	0.715	-0.475	$M(\boldsymbol{\theta}; 0,1) = 0.75$
	(0.515, -0.2)	0	1	0.515	-0.2	$M(\boldsymbol{\theta}; 0,1) = 0.9$
	(0.315, -0.16)	0	1	0.315	-0.16	$M(\boldsymbol{\theta}; 0,1) = 0.95$
F	(-0.985, -0.5)	0	1	-0.985	-0.5	$M(\boldsymbol{\theta}; 0,1) = 0.5$
	(-0.715, -0.475)	0	1	-0.715	-0.475	$M(\boldsymbol{\theta}; 0,1) = 0.75$
	(-0.515, -0.2)	0	1	-0.515	-0.2	$M(\boldsymbol{\theta}; 0,1) = 0.9$
	(-0.315, -0.16)	0	1	-0.315	-0.16	$M(\boldsymbol{\theta}; 0,1) = 0.95$
G	(1.164, 0)	0	1	1.164	0	$M(\boldsymbol{\theta}; 0,1) = 0.5$
	(0.879, 0)	0	1	0.879	0	$M(\boldsymbol{\theta}; 0,1) = 0.75$
	(0.578, 0)	0	1	0.578	0	$M(\boldsymbol{\theta}; 0,1) = 0.9$
	(0.354, 0)	0	1	0.354	0	$M(\boldsymbol{\theta}; 0,1) = 0.95$
H	(-1.164, 0)	0	1	-1.164	0	$M(\boldsymbol{\theta}; 0,1) = 0.5$
	(-0.879, 0)	0	1	-0.879	0	$M(\boldsymbol{\theta}; 0,1) = 0.75$
	(-0.578, 0)	0	1	-0.578	0	$M(\boldsymbol{\theta}; 0,1) = 0.9$
	(-0.354, 0)	0	1	-0.354	0	$M(\boldsymbol{\theta}; 0,1) = 0.95$

Source: authors' work.

Figure 2A. *PDF* curves of the Pearson distribution for parameter values presented in Table 2A

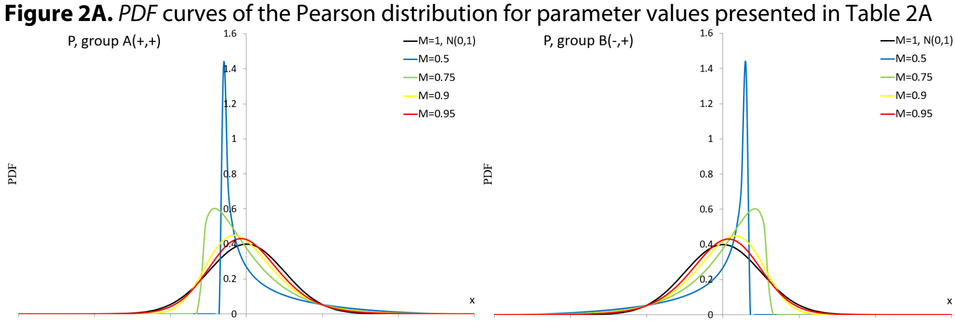
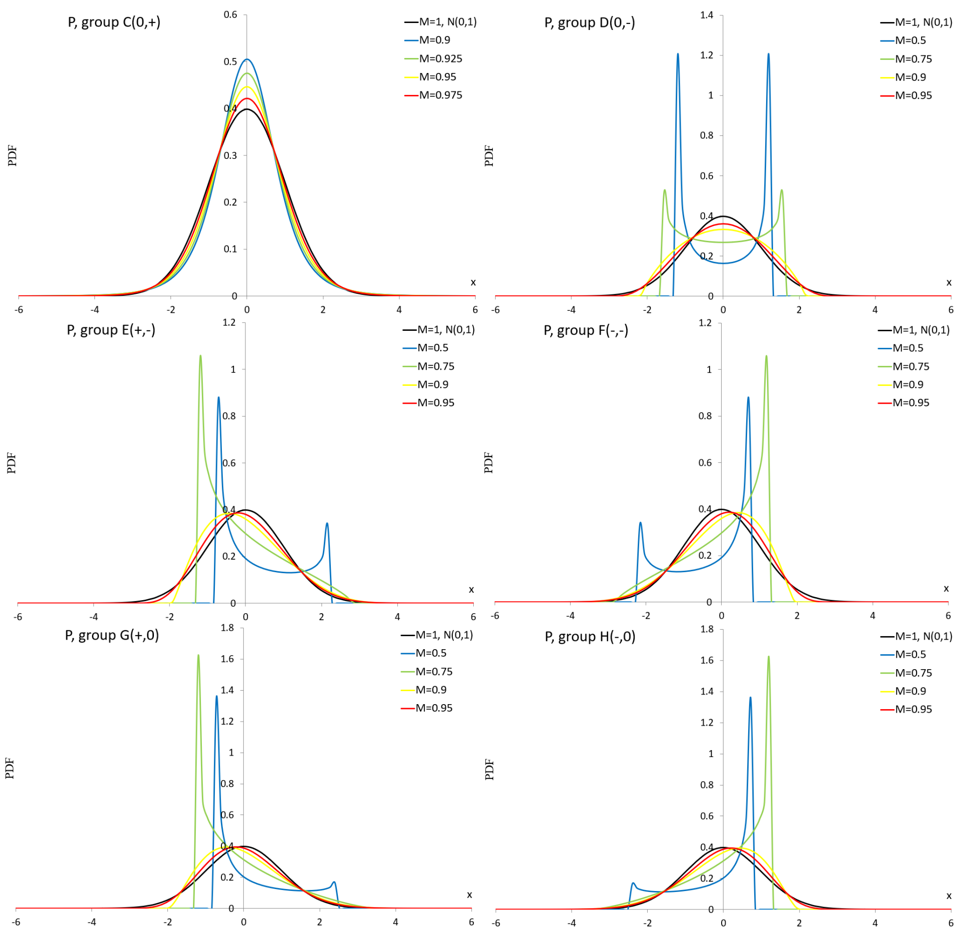


Figure 2A. *PDF* curves of the Pearson distribution for parameter values presented in Table 2A (cont.)



Source: authors' work.

Normal mixture distribution

PDF of the normal mixture (*NM*) distribution is given by

$$f_{NM}(x;\boldsymbol{\theta}) = \omega\phi(x;\mu_1,\sigma_1) + (1-\omega)\phi(x;\mu_2,\sigma_2)\ (x \in R),$$

where $\boldsymbol{\theta} = (\mu_1,\sigma_1,\mu_2,\sigma_2,\omega)$ and $\mu_1,\mu_2 \in R, \sigma_1,\sigma_2 > 0, \omega \in [0,1]$. Special cases of the *NM* distribution are:

- normal $N(\mu_1,\sigma_1)$ for $\omega = 1$, $N(\mu_2,\sigma_2)$ for $\omega = 0$;
- location contaminated normal (*LCN*)

$f_{LCM}(x;\mu_1,\omega) = f_{NM}(x;\mu_1,1{,}0{,}1,\omega)$;

- scale contaminated normal (*SCN*)

$f_{SCM}(x;\ \sigma_1,\omega) = f_{NM}(x;0,\sigma_1,0{,}1,\omega)$.

Table 3A. Vectors of the *NM* parameter $\boldsymbol{\theta}$, mean μ_a, standard deviation σ_a, skewness γ_1, excess kurtosis $\bar{\gamma}_2$ and similarity measure *M*. Groups O, A–H

Group	$\boldsymbol{\theta} = (\mu_1, \sigma_1, \mu_2, \sigma_2, \omega)$	μ_a	σ_a	$\gamma_{1,}$	$\bar{\gamma}_2$	$M(\boldsymbol{\theta}; \mu, \sigma)$
O	$(\mu_1, \sigma_1, \mu_2, \sigma_2, 1)$	0	1	0	0	$M(\boldsymbol{\theta}; \mu_1, \sigma_1) = 1$
	$(\mu_1, \sigma_1, \mu_2, \sigma_2, 0)$	0	1	0	0	$M(\boldsymbol{\theta}; \mu_2, \sigma_2) = 1$
A	0.572, 2.472, 5.614, 3.454, 0.787	1.646	3.408	0.685	0.755	$M(\boldsymbol{\theta}; 0,1) = 0.5$
	-0.215, 1.254, 1.979, 1.99, 0.639	0.577	1.883	0.645	0.502	$M(\boldsymbol{\theta}; 0,1) = 0.75$
	0.497, 1.376, -0.268, 0.884, 0.612	0.2	1.265	0.287	0.249	$M(\boldsymbol{\theta}; 0,1) = 0.9$
	-0.098, 0.857, 0.31, 1.007, 0.767	-0.003	0.911	0.09	0.099	$M(\boldsymbol{\theta}; 0,1) = 0.95$
B	0.502, 2.019, 1.708, 0.953, 0.36	1.274	1.544	-0.748	1.502	$M(\boldsymbol{\theta}; 0,1) = 0.5$
	0.06, 1.437, 1.004, 0.609, 0.634	0.406	1.285	-0.5	0.499	$M(\boldsymbol{\theta}; 0,1) = 0.75$
	0.709, 0.368, -0.072, 1.115, 0.193	0.079	1.06	-0.301	0.15	$M(\boldsymbol{\theta}; 0,1) = 0.9$
	0.159, 0.955, -0.123, 1.158, 0.271	-0.047	1.114	-0.05	0.059	$M(\boldsymbol{\theta}; 0,1) = 0.95$
C	0.519, 6.599, 0.519, 1.058, 0.665	0.519	5.416	0	1.398	$M(\boldsymbol{\theta}; 0,1) = 0.5$
	0.137, 0.581, 0.137, 2.391, 0.294	0.137	2.034	0	1.054	$M(\boldsymbol{\theta}; 0,1) = 0.75$
	0.225, 1.106, 0.225, 0.335, 0.89	0.225	1.049	0	0.299	$M(\boldsymbol{\theta}; 0,1) = 0.9$
	-0.09, 1.029, -0.09, 1.37, 0.825	-0.09	1.096	0	0.201	$M(\boldsymbol{\theta}; 0,1) = 0.95$
D	2.303, 0.51, 0.624, 0.481, 0.515	1.489	0.975	0	-1.099	$M(\boldsymbol{\theta}; 0,1) = 0.5$
	2.707, 0.013, 0.017, 1.125, 0.238	0.657	1.509	0	-1.001	$M(\boldsymbol{\theta}; 0,1) = 0.75$
	1.243, 0.621, -0.39, 0.811, 0.347	0.111	1.09	0	-0.63	$M(\boldsymbol{\theta}; 0,1) = 0.9$
	-1.112, 0.794, 0.023, 0.974, 0.13	0	0.897	0	-0.329	$M(\boldsymbol{\theta}; 0,1) = 0.95$
E	-0.475, 2.22, 5.318, 2.427, 0.721	1.141	3.457	0.5	-0.204	$M(\boldsymbol{\theta}; 0,1) = 0.5$
	-0.019, 1.369, 2.979, 1.15, 0.829	0.494	1.748	0.339	-0.1	$M(\boldsymbol{\theta}; 0,1) = 0.75$
	0.077, 0.844, 1.108, 0.779, 0.845	0.237	0.914	0.074	-0.035	$M(\boldsymbol{\theta}; 0,1) = 0.9$
	1.091, 0.969, -0.111, 1.056, 0.1	0.009	1.108	0.05	-0.01	$M(\boldsymbol{\theta}; 0,1) = 0.95$
F	-0.692, 0.705, 2.1, 0.679, 0.324	1.195	1.476	-0.542	-0.852	$M(\boldsymbol{\theta}; 0,1) = 0.5$
	-0.055, 1.277, 1.781, 0.443, 0.775	0.358	1.377	-0.3	-0.5	$M(\boldsymbol{\theta}; 0,1) = 0.75$
	-0.09, 1.08, -1.581, 0.92, 0.9	-0.239	1.155	-0.071	-0.042	$M(\boldsymbol{\theta}; 0,1) = 0.9$
	0.386, 0.845, -0.145, 0.918, 0.1	-0.092	0.925	-0.01	-0.011	$M(\boldsymbol{\theta}; 0,1) = 0.95$
G	2.686, 3.099, -0.964, 2.217, 0.471	0.755	3.232	0.4	0	$M(\boldsymbol{\theta}; 0,1) = 0.5$
	-0.56, 1.465, 1.411, 1.45, 0.8	-0.166	1.661	0.151	0	$M(\boldsymbol{\theta}; 0,1) = 0.75$
	-0.286, 1.114, 0.984, 1.105, 0.801	-0.033	1.222	0.101	0	$M(\boldsymbol{\theta}; 0,1) = 0.9$
	-0.232, 0.938, 0.727, 0.897, 0.878	-0.115	0.984	0.051	0	$M(\boldsymbol{\theta}; 0,1) = 0.95$
H	2.425, 1.101, 0.272, 1.693, 0.526	1.404	1.775	-0.499	0	$M(\boldsymbol{\theta}; 0,1) = 0.5$
	0.864, 1.125, -1.339, 1.241, 0.735	0.28	1.511	-0.386	0	$M(\boldsymbol{\theta}; 0,1) = 0.75$
	0.429, 1.078, -0.364, 1.228, 0.434	-0.02	1.23	-0.1	0	$M(\boldsymbol{\theta}; 0,1) = 0.9$
	0.108, 1.088, -0.524, 1.073, 0.879	0.032	1.106	-0.01	0	$M(\boldsymbol{\theta}; 0,1) = 0.95$

Source: authors' work.

Figure 3A. *PDF* curves of the normal mixture distribution for parameter values presented in Table 3A

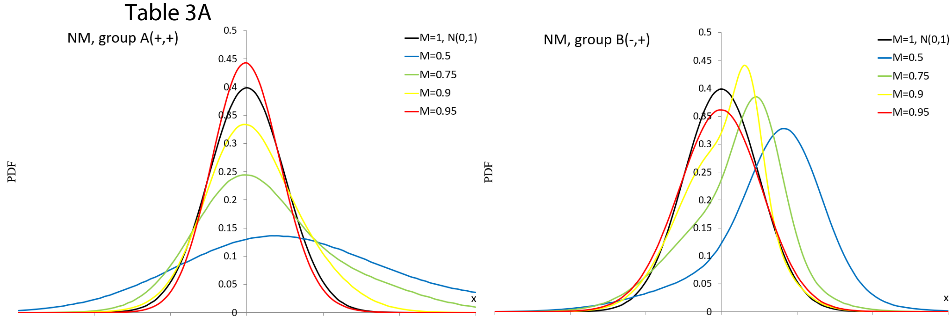
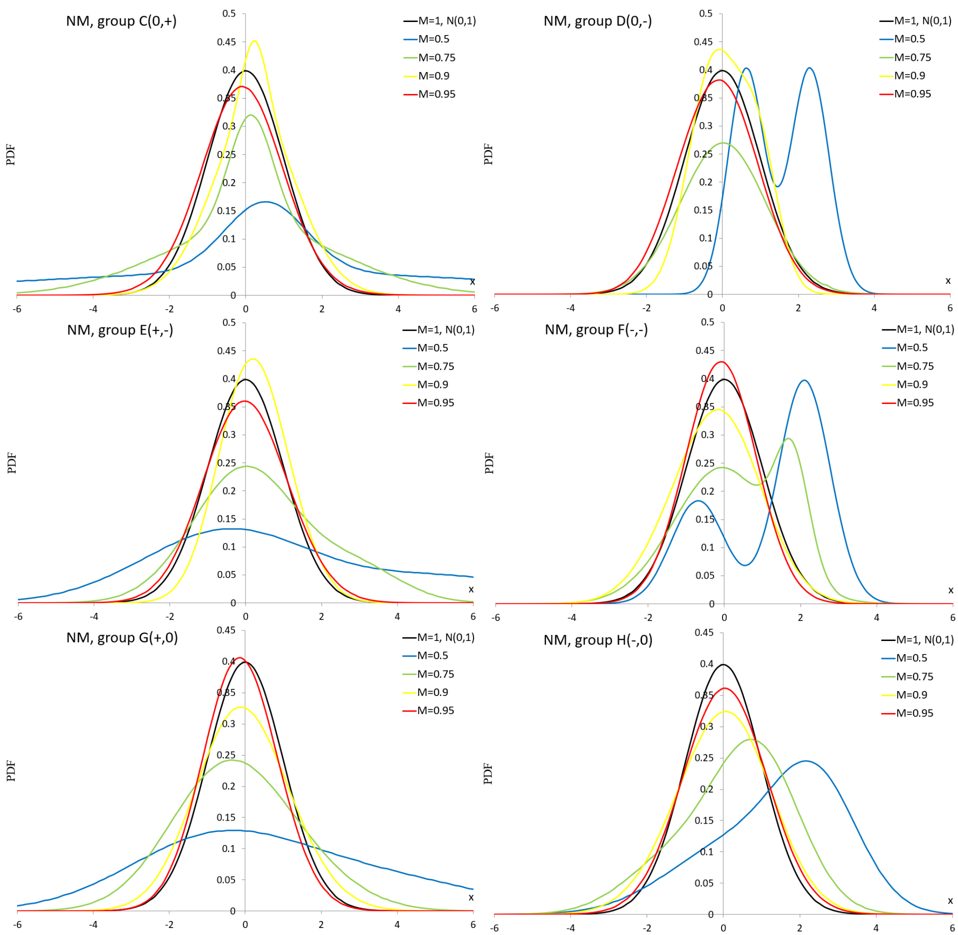


Figure 3A. *PDF* curves of the normal mixture distribution for parameter values presented in Table 3A (cont.)



Source: authors' work.

Normal logistic mixture distribution

PDF of the normal logistic mixture (*NLM*) distribution is given by

$$f_{NLM}(x;\boldsymbol{\theta}) = \omega\phi(x;\mu_1,\sigma_1) + (1-\omega)\frac{exp[-(x-\mu_2)/\sigma_2]}{\sigma_2\{1+xp[-(x-\mu_2)/\sigma_2]\}^2}\ (x \in R),$$

where $\boldsymbol{\theta} = (\mu_1, \sigma_1, \mu_2, \sigma_2, \omega)$ and $\mu_1, \mu_2 \in R, \sigma_1, \sigma_2 > 0, \omega \in [0,1]$. Special cases of the *NLM* distribution are:

- normal $N(\mu_1, \sigma_1)$ for $\omega = 1$;
- logistic (L) $f_L(x;\mu_2,\sigma_2)$ for $\omega = 0$.

Table 4A. Vectors of *NLM* parameter $\boldsymbol{\theta}$, mean μ_a, standard deviation σ_a, skewness γ_1, excess kurtosis $\bar{\gamma}_2$ and similarity measure *M*. Groups O, A–H

Group	$\boldsymbol{\theta} = (\mu_1, \sigma_1, \mu_2, \sigma_2, \omega)$	μ_a	σ_a	$\gamma_{1,}$	$\bar{\gamma}_2$	$M(\boldsymbol{\theta}; \mu, \sigma)$
O	$(\mu_1, \sigma_1, \mu_2, \sigma_2, 1)$	0	1	0	0	$M(\boldsymbol{\theta}; \mu_1, \sigma_1) = 1$
A	0.003, 0.223, 1.17, 0.701, 0.823	0.209	0.571	2.012	4.307	$M(\boldsymbol{\theta}; 0,1) = 0.5$
	2.179, 2.001, -0.212, 0.917, 0.15	0.146	1.429	1.408	3.475	$M(\boldsymbol{\theta}; 0,1) = 0.75$
	0.089, 1.119, 0.66, 1.78, 0.9	0.146	1.214	0.173	0.585	$M(\boldsymbol{\theta}; 0,1) = 0.9$
	-0.063, 1.019, 1.533, 0.962, 0.9	0.096	1.121	0.173	0.093	$M(\boldsymbol{\theta}; 0,1) = 0.95$
B	1.305, 0.572, 0.281, 1.747, 0.545	0.839	1.351	-0.851	1.731	$M(\boldsymbol{\theta}; 0,1) = 0.5$
	0.449, 0.572, 0.281, 1.669, 0.479	0.361	1.271	-0.151	1.731	$M(\boldsymbol{\theta}; 0,1) = 0.75$
	0.146, 0.843, 0.128, 1.669, 0.73	0.141	1.127	-0.015	1.578	$M(\boldsymbol{\theta}; 0,1) = 0.9$
	0.034, 0.974, 0.038, 0.767, 0.535	0.035	0.884	-0.002	0.158	$M(\boldsymbol{\theta}; 0,1) = 0.95$
C	1.382, 0.88, 1.382, 1.769, 0.242	1.382	1.6	0	0.465	$M(\boldsymbol{\theta}; 0,1) = 0.5$
	0.536, 1.412, 0.536, 1.855, 0.9	0.536	1.462	0	0.124	$M(\boldsymbol{\theta}; 0,1) = 0.75$
	0.107, 0.595, 0.107, 0.528, 0.142	0.107	0.538	0	0.025	$M(\boldsymbol{\theta}; 0,1) = 0.9$
	0.021, 0.875, 0.021, 0.756, 0.9	0.021	0.864	0	0.018	$M(\boldsymbol{\theta}; 0,1) = 0.95$
D	2.664, 1.103, 0.28, 1.333, 0.352	1.119	1.696	0	-0.458	$M(\boldsymbol{\theta}; 0,1) = 0.5$
	0.772, 1.253, -1.236, 0.706, 0.899	0.569	1.353	0	-0.268	$M(\boldsymbol{\theta}; 0,1) = 0.75$
	0.952, 0.489, 0.008, 0.682, 0.12	0.121	0.729	0	-0.176	$M(\boldsymbol{\theta}; 0,1) = 0.9$
	0.321, 0.631, -0.209, 0.665, 0.268	-0.066	0.697	0	-0.035	$M(\boldsymbol{\theta}; 0,1) = 0.95$
E	2.917, 1.445, 0.207, 1.5, 0.292	0.999	1.929	0.201	-0.237	$M(\boldsymbol{\theta}; 0,1) = 0.5$
	0.277, 0.569, 1.43, 0.334, 0.9	0.392	0.650	0.162	-0.205	$M(\boldsymbol{\theta}; 0,1) = 0.75$
	0.425, 0.956, -0.714, 0.775, 0.673	0.053	1.047	0.106	-0.199	$M(\boldsymbol{\theta}; 0,1) = 0.9$
	-0.19, 0.909, 0.928, 0.889, 0.778	0.058	1.017	0.107	-0.034	$M(\boldsymbol{\theta}; 0,1) = 0.95$
F	0.757, 0.486, 1.444, 0.459, 0.45	1.134	0.582	-0.105	-0.24	$M(\boldsymbol{\theta}; 0,1) = 0.5$
	0.263, 0.905, 1.505, 0.832, 0.571	0.795	1.069	-0.040	-0.232	$M(\boldsymbol{\theta}; 0,1) = 0.75$
	-0.435, 0.71, 0.495, 0.765, 0.286	0.229	0.859	-0.038	-0.123	$M(\boldsymbol{\theta}; 0,1) = 0.9$
	0.004, 1.011, -1.433, 0.77, 0.9	-0.139	1.079	-0.038	-0.115	$M(\boldsymbol{\theta}; 0,1) = 0.95$
G	0.131, 0.191, 2.38, 1.824, 0.393	1.497	1.8	0.807	0	$M(\boldsymbol{\theta}; 0,1) = 0.5$
	1.419, 1.558, -0.332, 0.321, 0.649	0.804	1.520	0.688	0	$M(\boldsymbol{\theta}; 0,1) = 0.75$
	-0.049, 0.982, 1.642, 0.994, 0.781	0.322	1.208	0.275	0	$M(\boldsymbol{\theta}; 0,1) = 0.9$
	-0.104, 0.984, 0.585, 0.981, 0.802	0.032	1.021	0.028	0	$M(\boldsymbol{\theta}; 0,1) = 0.95$
H	1.558, 1.02, -2.125, 0.982, 0.794	0.8	1.8	-0.8	0	$M(\boldsymbol{\theta}; 0,1) = 0.5$
	0.769, 1.037, -1.115, 1.083, 0.762	0.320	1.320	-0.320	0	$M(\boldsymbol{\theta}; 0,1) = 0.75$
	-0.071, 0.854, 0.669, 0.714, 0.496	0.302	0.869	-0.186	0	$M(\boldsymbol{\theta}; 0,1) = 0.9$
	0.176, 0.841, -0.229, 0.858, 0.64	0.030	0.869	-0.019	0	$M(\boldsymbol{\theta}; 0,1) = 0.95$

Source: authors' work.

Figure 4A. *PDF* curves of the normal logistic mixture distribution for parameter values presented in Table 4A

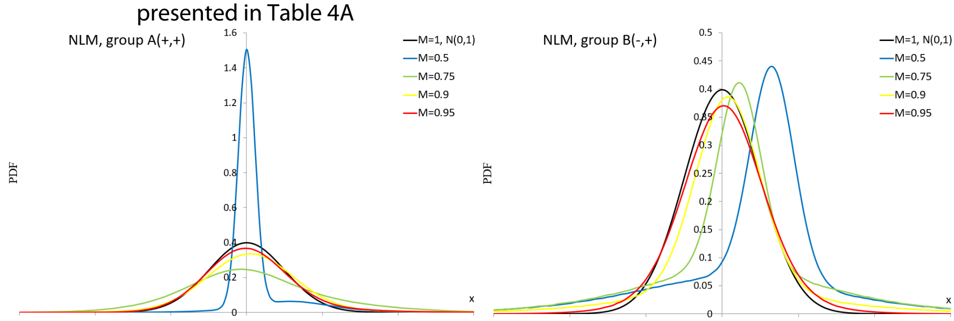
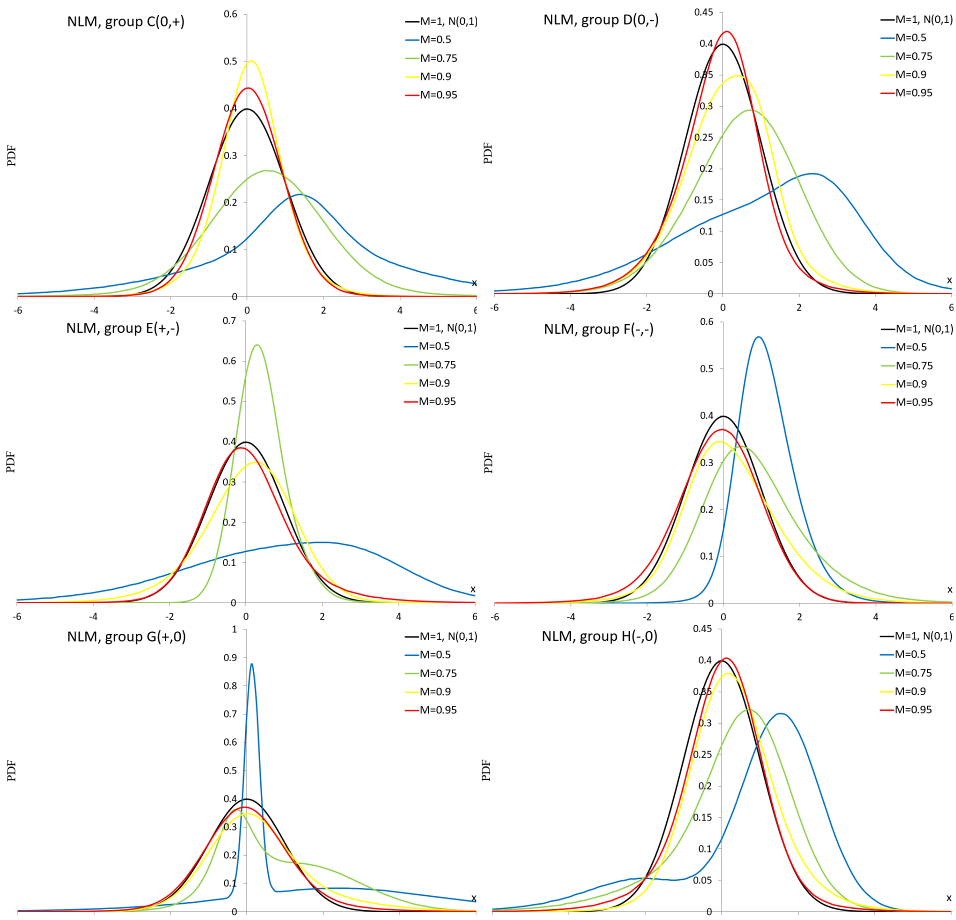


Figure 4A. *PDF* curves of the normal logistic mixture distribution for parameter values presented in Table 4A (cont.)



Source: authors' work.

Normal distribution with plasticising component

PDF of the normal distribution with plasticising component (*NDPC*) is given by

$$f_{NDPC}(x;\boldsymbol{\theta}) = \omega\phi(x;\mu_1,\sigma_1) + (1-\omega)\frac{c_2}{\sigma_2}\left|\frac{x-\mu_2}{\sigma_2}\right|^{c_2-1}\phi\left(\left|\frac{x-\mu_2}{\sigma_2}\right|^{c_2};0,1\right)\ (x \in R),$$

where $\boldsymbol{\theta} = (\mu_1,\sigma_1,\mu_2,\sigma_2,c_2,\omega)$ and $\mu_1,\mu_2 \in R, \sigma_1,\sigma_2 > 0, c_2 \geq 1, \omega \in [0,1]$. Special cases of the *NDPC* distribution are:

- $N(\mu_1,\sigma_1)$ for $\omega = 1$ and $N(\mu_2,\sigma_2)$ for $c_2 = 1, \omega = 0$;
- plasticising component (*PC*) $f_{PC}(x;\mu_2,\sigma_2,c_2)$ for $\omega = 0$.

Table 5A. Vectors of *NDPC* parameter $\boldsymbol{\theta}$, mean μ_a, standard deviation σ_a, skewness γ_1, excess kurtosis $\bar{\gamma}_2$ and similarity measure *M*. Groups O, A–H

Group	$\boldsymbol{\theta} = (\mu_1, \sigma_1, \mu_2, \sigma_2, c_2, \omega)$	μ_a	σ_a	$\gamma_{1,}$	$\bar{\gamma}_2$	$M(\boldsymbol{\theta}; \mu, \sigma)$
O	$\mu_1, \sigma_1, \mu_2, \sigma_2, c_2, 1$	0	1	0	0	$M(\boldsymbol{\theta}; \mu_1, \sigma_1) = 1$
	$\mu_1, \sigma_1, \mu_2, \sigma_2, 1, 0$	0	1	0	0	$M(\boldsymbol{\theta}; \mu_2, \sigma_2) = 1$
A	1.948, 1.27, 1.143, 0.793, 1.482, 0.324	1.404	1.009	0.599	0.853	$M(\boldsymbol{\theta}; 0{,}1) = 0.5$
	0.265, 0.415, 0.996, 1.541, 1.16, 0.313	0.767	1.288	0.426	0.152	$M(\boldsymbol{\theta}; 0{,}1) = 0.75$
	0.173, 0.358, 0.289, 1.268, 1.132, 0.198	0.266	1.104	0.056	0.071	$M(\boldsymbol{\theta}; 0{,}1) = 0.9$
	0.047, 1.02, -0.014, 0.872, 1, 0.214	-0.001	0.906	0.012	0.06	$M(\boldsymbol{\theta}; 0{,}1) = 0.95$
B	0.895, 0.421, -0.141, 1.9, 3.141, 0.872	0.762	0.804	-2	5.085	$M(\boldsymbol{\theta}; 0{,}1) = 0.5$
	0.539, 0.632, -1.078, 2.061, 1.174, 0.741	0.12	1.34	-1.499	2.986	$M(\boldsymbol{\theta}; 0{,}1) = 0.75$
	-0.966, 1.824, 0.259, 0.889, 1.1, 0.26	-0.059	1.305	-0.899	1.999	$M(\boldsymbol{\theta}; 0{,}1) = 0.9$
	-0.099, 0.938, 0.399, 0.646, 1.204, 0.831	-0.015	0.911	-0.125	0.036	$M(\boldsymbol{\theta}; 0{,}1) = 0.95$
C	1.592, 1.867, 1.596, 1.215, 1.2, 0.249	1.595	1.365	-0.002	0.528	$M(\boldsymbol{\theta}; 0{,}1) = 0.5$
	0.571, 1.023, 0.571, 1.962, 1.15, 0.505	0.571	1.508	0	0.325	$M(\boldsymbol{\theta}; 0{,}1) = 0.75$
	-0.097, 1.332, -0.097, 1.058, 1.1, 0.614	-0.097	1.223	0	0.101	$M(\boldsymbol{\theta}; 0{,}1) = 0.9$
	0.003, 1.135, 0.003, 0.95, 1.05, 0.874	0.003	1.112	0	0.026	$M(\boldsymbol{\theta}; 0{,}1) = 0.95$
D	-0.692, 2.203, -0.692, 2.544, 1.759, 0.25	-0.692	2.265	0	-1	$M(\boldsymbol{\theta}; 0{,}1) = 0.5$
	0.323, 1.312, 0.605, 1.335, 1.2, 0.01	0.602	1.266	0	-0.587	$M(\boldsymbol{\theta}; 0{,}1) = 0.75$
	0.179, 0.494, 0.179, 1.163, 1.426, 0.443	0.179	0.862	0	-0.202	$M(\boldsymbol{\theta}; 0{,}1) = 0.9$
	0.195, 0.96, -0.719, 0.858, 1.109, 0.918	0.12	0.983	0	-0.05	$M(\boldsymbol{\theta}; 0{,}1) = 0.95$
E	0.675, 0.284, 2.122, 1.968, 2.104, 0.374	1.581	1.565	0.749	-0.849	$M(\boldsymbol{\theta}; 0{,}1) = 0.5$
	0.423, 1.032, 1.058, 2.077, 1.815, 0.494	0.744	1.544	0.311	-0.667	$M(\boldsymbol{\theta}; 0{,}1) = 0.75$
	0.159, 0.389, 0.296, 1.257, 1.649, 0.326	0.251	0.96	0.116	-0.597	$M(\boldsymbol{\theta}; 0{,}1) = 0.9$
	1.081, 0.621, -0.216, 0.755, 1, 0.24	0.095	0.912	0.1	-0.298	$M(\boldsymbol{\theta}; 0{,}1) = 0.95$
F	1.609, 0.59, 0.322, 2.194, 1.609, 0.309	0.72	1.784	-0.491	-0.728	$M(\boldsymbol{\theta}; 0{,}1) = 0.5$
	0.617, 0.737, 0.129, 1.752, 1.465, 0.332	0.291	1.395	-0.239	-0.526	$M(\boldsymbol{\theta}; 0{,}1) = 0.75$
	0.189, 0.39, 0.007, 1.336, 1.739, 0.414	0.082	0.957	-0.195	-0.422	$M(\boldsymbol{\theta}; 0{,}1) = 0.9$
	0.155, 0.882, 0.019, 1.184, 1.175, 0.581	0.098	0.995	-0.05	-0.188	$M(\boldsymbol{\theta}; 0{,}1) = 0.95$
G	1.877, 0.829, 0.573, 0.583, 2.562, 0.383	1.072	0.912	0.669	-0.001	$M(\boldsymbol{\theta}; 0{,}1) = 0.5$
	0.362, 1.583, -1.112, 1.026, 1.283, 0.608	-0.216	1.55	0.405	0	$M(\boldsymbol{\theta}; 0{,}1) = 0.75$
	0.055, 0.702, 0.474, 1.586, 1.328, 0.473	0.276	1.191	0.31	0	$M(\boldsymbol{\theta}; 0{,}1) = 0.9$
	0.212, 1.443, -0.012, 1.057, 1.088, 0.1	0.01	1.079	0.05	0	$M(\boldsymbol{\theta}; 0{,}1) = 0.95$
H	1.064, 0.408, 0.687, 0.908, 2, 0.699	0.951	0.587	-0.601	0	$M(\boldsymbol{\theta}; 0{,}1) = 0.5$
	0.186, 0.259, -0.072, 0.973, 1.697, 0.499	0.057	0.659	-0.473	0	$M(\boldsymbol{\theta}; 0{,}1) = 0.75$
	0.116, 1.127, -0.443, 1.588, 1.225, 0.816	0.013	1.223	-0.143	0	$M(\boldsymbol{\theta}; 0{,}1) = 0.9$
	0.029, 1.105, 0.161, 0.958, 1.036, 0.329	0.118	1.003	-0.028	0	$M(\boldsymbol{\theta}; 0{,}1) = 0.95$

Source: authors' work.

Figure 5A. *PDF* curves of the *NDPC* for parameter values presented in Table 5A

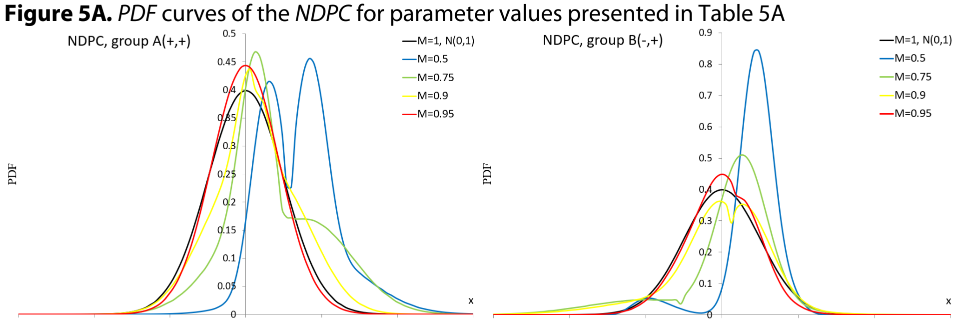
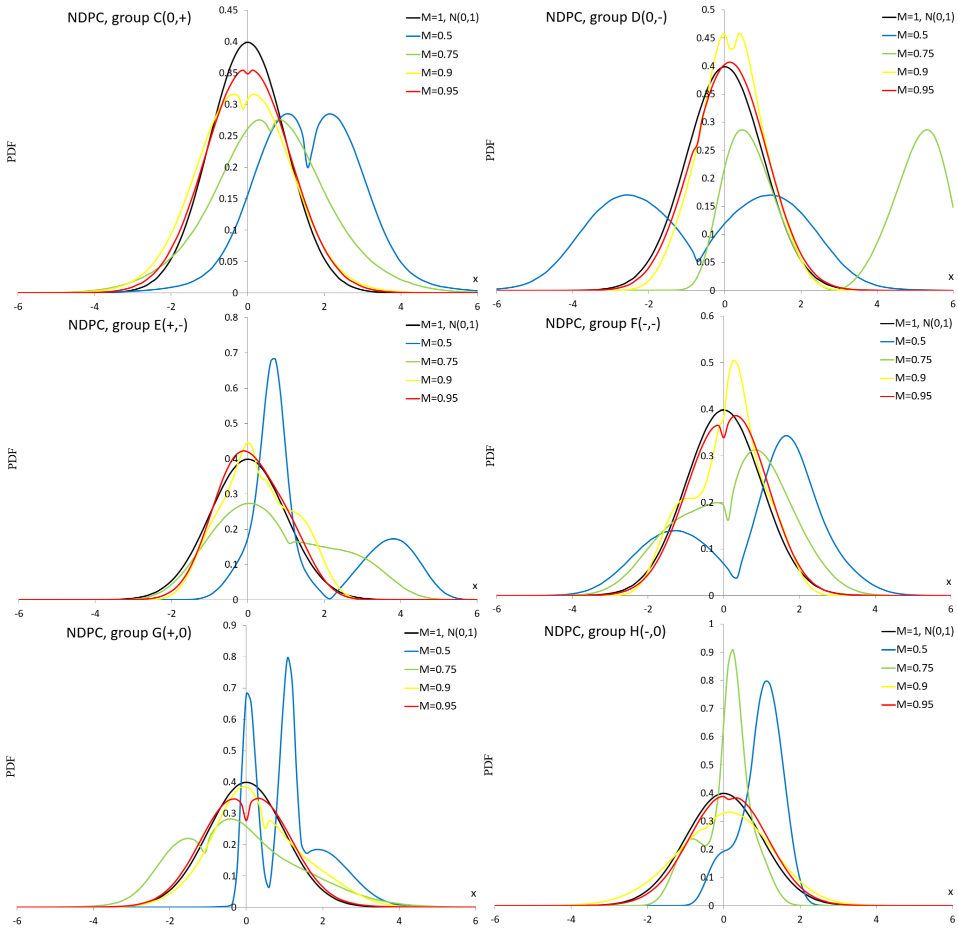


Figure 5A. *PDF* curves of the *NDPC* for parameter values presented in Table 5A (cont.)



Source: authors' work.

Plasticising component mixture distribution

PDF of the plasticising component mixture (*PCM*) distribution is given by

$$f_{PCM}(x;\boldsymbol{\theta}) = \omega f_{PC}(x;\mu_1,\sigma_1,c_1) + (1-\omega) f_{PC}(x;\mu_2,\sigma_2,c_2)\ (x \in R),$$

where $f_{PC}(x;\mu,\sigma,c) = \frac{c}{\sigma}\left|\frac{x-\mu}{\sigma}\right|^{c-1}\phi\left(\left|\frac{x-\mu}{\sigma}\right|^{c};0,1\right)$ $(x \in R)$ and
$\boldsymbol{\theta} = (\mu_1,\sigma_1,c_1,\mu_2,\sigma_2,c_2,\omega)$, $\mu_1,\mu_2 \in R$,
$\sigma_1,\sigma_2 > 0$, $c_1,c_2 \geq 1$, $\omega \in [0,1]$. Special cases of the *PCM* distribution are:

- $N(\mu_1,\sigma_1)$ for $c_1 = 1, \omega = 1$; $N(\mu_2,\sigma_2)$ for $c_2 = 1, \omega = 0$,
- plasticising component $PC(\mu_1,\sigma_1,c_1)$, $PC(\mu_2,\sigma_2,c_2)$ for $\omega = 1, \omega = 0$, respectively.

Table 6A. Vectors of *PCM* parameter $\boldsymbol{\theta}$, mean μ_a, standard deviation σ_a, skewness γ_1, excess kurtosis $\bar{\gamma}_2$ and similarity measure *M*. Groups O, A–H

Group	$\boldsymbol{\theta} = (\mu_1, \sigma_1, c_1, \mu_2, \sigma_2, c_2, \omega)$	μ_a	σ_a	$\gamma_{1,}$	$\bar{\gamma}_2$	$M(\boldsymbol{\theta}; \mu, \sigma)$
O	$\mu_1, \sigma_1, 1, \mu_2, \sigma_2, c_2, 1$	0	1	0	0	$M(\boldsymbol{\theta}; \mu_1, \sigma_1) = 1$
	$\mu_1, \sigma_1, c_1, \mu_2, \sigma_2, 1{,}0$	0	1	0	0	$M(\boldsymbol{\theta}; \mu_2, \sigma_2) = 1$
A	1.415, 1.684, 2.194, 11.252, 5.474, 2.331, 0.9	2.399	3.622	2.647	7.663	$M(\boldsymbol{\theta}; 0{,}1) = 0.5$
	0.444, 0.899, 1.602, 1.653, 2.506, 1.876, 0.64	0.879	1.604	0.913	0.412	$M(\boldsymbol{\theta}; 0{,}1) = 0.75$
	-0.076, 1.056, 1.1, 0.701, 1.646, 1.095, 0.71	0.149	1.268	0.374	0.374	$M(\boldsymbol{\theta}; 0{,}1) = 0.9$
	0.026, 1.078, 1.001, 0.701, 1.646, 1.174, 0.95	0.06	1.117	0.099	0.148	$M(\boldsymbol{\theta}; 0{,}1) = 0.95$
B	1.366, 0.572, 1.11, 0.502, 1.669, 1.253, 0.658	1.071	1.099	-0.978	1.565	$M(\boldsymbol{\theta}; 0{,}1) = 0.5$
	0.67, 0.425, 1.576, -0.323, 1.696, 1.05, 0.349	0.024	1.444	-0.569	0.606	$M(\boldsymbol{\theta}; 0{,}1) = 0.75$
	-0.204, 2.209, 1.205, 0.133, 1.139, 1.05, 0.076	0.107	1.224	-0.122	0.457	$M(\boldsymbol{\theta}; 0{,}1) = 0.9$
	0.121, 0.936, 1.05, -0.17, 1.917, 1.411, 0.95	0.106	0.982	-0.1	0.204	$M(\boldsymbol{\theta}; 0{,}1) = 0.95$
C	1.597, 2.518, 1.263, 1.596, 0.856, 1.285, 0.526	1.597	1.797	0	0.601	$M(\boldsymbol{\theta}; 0{,}1) = 0.5$
	0.012, 0.274, 1.256, 0.012, 2.046, 1.01, 0.183	0.012	1.846	0	0.598	$M(\boldsymbol{\theta}; 0{,}1) = 0.75$
	0.127, 1.089, 1.01, 0.127, 0.183, 1.01, 0.863	0.127	1.01	0	0.401	$M(\boldsymbol{\theta}; 0{,}1) = 0.9$
	0.075, 0.973, 1.01, 0.075, 1.964, 1.362, 0.867	0.075	1.119	0	0.387	$M(\boldsymbol{\theta}; 0{,}1) = 0.95$
D	1.631, 0.893, 1.05, 1.632, 2.104, 1.554, 0.498	1.632	1.488	0	-0.268	$M(\boldsymbol{\theta}; 0{,}1) = 0.5$
	0.639, 1.576, 1.167, 0.64, 1.085, 1.199, 0.163	0.64	1.12	0	-0.251	$M(\boldsymbol{\theta}; 0{,}1) = 0.75$
	0.666, 1.123, 4.041, 0.233, 1.069, 1.05, 0.01	0.237	1.052	0	-0.198	$M(\boldsymbol{\theta}; 0{,}1) = 0.9$
	0.225, 1.087, 1.05, -0.067, 1.094, 1.05, 0.233	0.001	1.081	0	-0.18	$M(\boldsymbol{\theta}; 0{,}1) = 0.95$
E	1.472, 0.782, 1.11, 0.236, 0.291, 3.203, 0.692	1.091	0.861	0.38	-0.8	$M(\boldsymbol{\theta}; 0{,}1) = 0.5$
	-0.196, 0.341, 1.064, 0.613, 0.758, 1.204, 0.153	0.489	0.734	0.201	-0.7	$M(\boldsymbol{\theta}; 0{,}1) = 0.75$
	0.722, 0.703, 1.304, -0.57, 0.598, 1.05, 0.455	0.018	0.893	0.179	-0.617	$M(\boldsymbol{\theta}; 0{,}1) = 0.9$
	0.584, 1.171, 9.804, -0.016, 1.024, 1.076, 0.05	0.014	1.013	0.028	-0.351	$M(\boldsymbol{\theta}; 0{,}1) = 0.95$
F	0.261, 1.419, 1.909, 3.099, 0.744, 1.567, 0.57	1.481	1.757	-0.3	-1.107	$M(\boldsymbol{\theta}; 0{,}1) = 0.5$
	0.037, 1.295, 1.076, 1.316, 1.171, 1.654, 0.485	0.696	1.326	-0.204	-0.4	$M(\boldsymbol{\theta}; 0{,}1) = 0.75$
	0.201, 0.121, 1.573, 0.184, 1.177, 1.161, 0.066	0.185	1.087	-0.003	-0.331	$M(\boldsymbol{\theta}; 0{,}1) = 0.9$
	0.049, 1.063, 1.088, 1.392, 0.511, 1.05, 0.99	0.062	1.038	-0.008	-0.328	$M(\boldsymbol{\theta}; 0{,}1) = 0.95$
G	1.088, 0.894, 3.782, 1.969, 2.71, 1.792, 0.55	1.484	1.793	0.6	0	$M(\boldsymbol{\theta}; 0{,}1) = 0.5$
	1.515, 2.553, 3.55, 0.07, 1.328, 1.619, 0.07	0.171	1.359	0.501	0	$M(\boldsymbol{\theta}; 0{,}1) = 0.75$
	-0.034, 1.072, 1.159, 1.146, 1.51, 1.301, 0.756	0.254	1.238	0.401	0	$M(\boldsymbol{\theta}; 0{,}1) = 0.9$
	0.825, 1.615, 1.868, 0.067, 0.934, 1.05, 0.141	0.174	1.044	0.336	0	$M(\boldsymbol{\theta}; 0{,}1) = 0.95$
H	0.816, 1.867, 1.24, 1.787, 1.272, 1.05, 0.278	1.517	1.475	-0.302	0	$M(\boldsymbol{\theta}; 0{,}1) = 0.5$
	-0.364, 1.889, 1.057, 0.29, 1.413, 1.05, 0.527	-0.055	1.682	-0.154	0	$M(\boldsymbol{\theta}; 0{,}1) = 0.75$
	0.286, 0.405, 1.27, -0.263, 1.261, 1.05, 0.112	-0.202	1.188	-0.128	0	$M(\boldsymbol{\theta}; 0{,}1) = 0.9$
	-0.153, 1.344, 1.349, -0.024, 0.539, 1.05, 0.565	-0.097	1	-0.12	0	$M(\boldsymbol{\theta}; 0{,}1) = 0.95$

Source: authors' work.

Figure 6A. *PDF* curves of the *PCM* distribution for parameter values presented in Table 6A

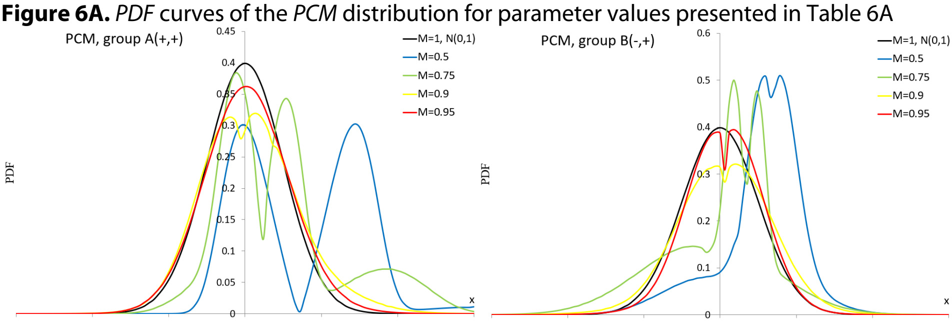
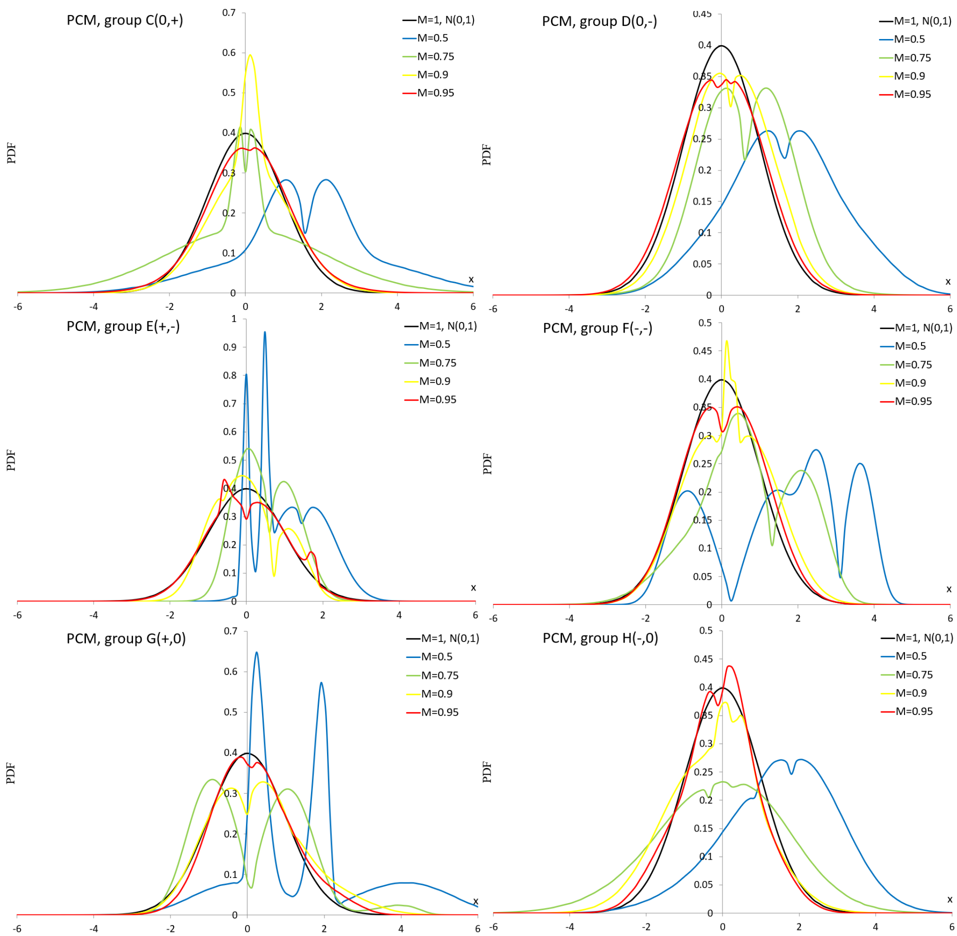


Figure 6A. *PDF* curves of the *PCM* distribution for parameter values presented in Table 6A (cont.)



Source: authors' work.

Table 7A. The power of *GoFTs* for group of alternatives A (ALTs)

ALT	N	Numbered *GoFT* (see Table 3)											
		1	2	3	4	5	6	7	8	9	10	11	12
P_1	10	**0.853**	0.614	0.789	0.792	0.764	0.791	0.792	0.789	0.789	0.817	0.658	0.847
	20	0.993	0.965	0.987	0.987	0.984	0.987	0.987	0.987	0.987	0.993	0.957	**0.997**
P_2	10	**0.374**	0.126	0.260	0.270	0.269	0.272	0.270	0.273	0.273	0.293	0.219	0.316
	20	0.650	0.380	0.533	0.544	0.541	0.545	0.544	0.546	0.546	0.599	0.421	**0.670**
P_2	10	**0.149**	0.049	0.093	0.099	0.107	0.101	0.100	0.103	0.103	0.108	0.092	0.111
	20	**0.227**	0.092	0.149	0.158	0.174	0.161	0.159	0.164	0.164	0.181	0.138	0.208
P_4	10	**0.092**	0.048	0.064	0.069	0.076	0.070	0.069	0.072	0.072	0.074	0.067	0.076
	20	**0.117**	0.060	0.080	0.086	0.100	0.088	0.087	0.091	0.091	0.098	0.081	0.108
NM_1	10	**0.146**	0.045	0.087	0.094	0.105	0.096	0.095	0.099	0.099	0.103	0.090	0.105
	20	**0.221**	0.087	0.142	0.153	0.170	0.156	0.153	0.159	0.159	0.171	0.136	0.185
NM_2	10	**0.143**	0.042	0.087	0.092	0.098	0.093	0.092	0.095	0.095	0.099	0.089	0.100
	20	**0.224**	0.084	0.147	0.155	0.165	0.157	0.156	0.159	0.159	0.169	0.138	0.179
NM_2	10	**0.084**	0.041	0.059	0.061	0.065	0.062	0.062	0.063	0.063	0.063	0.061	0.063
	20	**0.102**	0.048	0.071	0.073	0.079	0.075	0.074	0.075	0.075	0.078	0.071	0.081
NM_4	10	**0.057**	0.048	0.052	0.053	0.054	0.053	0.053	0.054	0.054	0.054	0.051	0.053
	20	**0.061**	0.047	0.053	0.053	0.055	0.054	0.054	0.054	0.054	0.055	0.053	0.056
NLM_1	10	**0.628**	0.469	0.552	0.577	0.604	0.581	0.578	0.588	0.588	0.596	0.541	0.591
	20	**0.881**	0.832	0.846	0.860	0.879	0.863	0.861	0.867	0.867	0.878	0.832	**0.885**
NLM_2	10	**0.112**	0.059	0.075	0.083	0.098	0.086	0.084	0.089	0.089	0.092	0.083	0.092
	20	**0.153**	0.084	0.102	0.115	0.139	0.118	0.116	0.123	0.123	0.131	0.106	0.137
NLM_2	10	0.138	0.116	0.114	0.127	**0.149**	0.130	0.128	0.135	0.135	0.142	0.121	0.145
	20	0.220	0.193	0.181	0.201	0.239	0.206	0.202	0.214	0.214	0.237	0.181	**0.267**
NLM_4	10	**0.113**	0.058	0.076	0.083	0.096	0.085	0.084	0.087	0.087	0.093	0.081	0.095
	20	**0.157**	0.087	0.106	0.118	0.141	0.121	0.119	0.126	0.126	0.138	0.107	**0.159**
$NDPC_1$	10	**0.117**	0.058	0.093	0.092	0.088	0.091	0.092	0.091	0.091	0.096	0.084	0.100
	20	0.170	0.082	0.127	0.128	0.128	0.129	0.128	0.129	0.129	0.150	0.115	**0.179**
$NDPC_2$	10	**0.183**	0.065	0.131	0.135	0.131	0.135	0.135	0.135	0.135	0.131	0.126	0.121
	20	**0.306**	0.157	0.243	0.246	0.242	0.247	0.246	0.246	0.246	0.231	0.224	0.188
$NDPC_2$	10	**0.063**	0.049	0.053	0.056	0.059	0.056	0.056	0.057	0.057	0.057	0.056	0.054
	20	**0.071**	0.052	0.056	0.060	0.067	0.061	0.060	0.062	0.062	0.060	0.062	0.053
$NDPC_4$	10	**0.052**	0.051	0.051	0.051	**0.052**	0.051	0.051	0.051	0.051	**0.052**	0.051	0.051
	20	**0.053**	0.051	0.051	0.051	**0.053**	0.051	0.051	0.052	0.052	0.052	0.051	**0.053**
PCM_1	10	0.594	0.476	**0.608**	0.595	0.540	0.590	0.594	0.581	0.581	0.593	0.527	0.594
	20	0.888	0.821	0.874	0.874	0.854	0.874	0.874	0.871	0.871	0.895	0.814	**0.896**
PCM_2	10	**0.296**	0.081	0.192	0.201	0.203	0.202	0.201	0.204	0.204	0.219	0.172	0.228
	20	**0.549**	0.267	0.416	0.431	0.439	0.433	0.431	0.437	0.437	0.481	0.342	0.494
PCM_2	10	**0.085**	0.041	0.061	0.063	0.065	0.063	0.063	0.064	0.064	0.066	0.062	0.067
	20	**0.104**	0.048	0.073	0.076	0.079	0.076	0.076	0.077	0.077	0.083	0.070	0.093
PCM_4	10	**0.058**	0.049	0.053	0.053	0.055	0.054	0.053	0.054	0.054	0.055	0.053	0.054
	20	**0.062**	0.049	0.053	0.054	0.057	0.055	0.054	0.055	0.055	0.056	0.054	0.058

Note. The highest *PoTs* of the *MCM(α,β)* values are underlined. The highest *PoTs* for all the analysed tests are in bold.
Source: authors' work.

Table 8A. The power of *GoFTs* for group of alternatives B (ALTs)

ALT	N	Numbered *GoFT* (see Table 3)											
		1	2	3	4	5	6	7	8	9	10	11	12
P_1	10	0.612	**0.853**	0.788	0.790	0.762	0.789	0.790	0.786	0.786	0.815	0.657	0.846
	20	0.964	**0.993**	0.986	0.987	0.984	0.987	0.987	0.986	0.986	0.993	0.957	**0.997**
P_2	10	0.125	**0.372**	0.260	0.268	0.267	0.269	0.269	0.271	0.271	0.292	0.217	0.315
	20	0.380	0.655	0.537	0.545	0.542	0.547	0.546	0.548	0.548	0.603	0.423	**0.673**
P_2	10	0.047	**0.149**	0.092	0.099	0.107	0.100	0.099	0.102	0.102	0.107	0.091	0.112
	20	0.093	**0.228**	0.150	0.160	0.175	0.163	0.161	0.166	0.166	0.181	0.140	0.207
P_4	10	0.048	**0.092**	0.065	0.069	0.076	0.070	0.069	0.072	0.072	0.074	0.067	0.074
	20	0.059	**0.119**	0.079	0.086	0.100	0.088	0.087	0.091	0.091	0.098	0.081	0.107
NM_1	10	0.068	**0.182**	0.112	0.125	0.144	0.128	0.126	0.132	0.132	0.138	0.116	0.139
	20	0.141	**0.293**	0.194	0.214	0.249	0.219	0.215	0.226	0.226	0.244	0.189	0.256
NM_2	10	0.049	**0.151**	0.097	0.102	0.107	0.103	0.102	0.105	0.105	0.107	0.096	0.106
	20	0.101	**0.237**	0.165	0.173	0.182	0.175	0.174	0.177	0.177	0.180	0.152	0.175
NM_2	10	0.042	**0.108**	0.077	0.078	0.076	0.078	0.078	0.078	0.078	0.078	0.073	0.077
	20	0.066	**0.153**	0.114	0.114	0.110	0.114	0.114	0.113	0.113	0.113	0.102	0.108
NM_4	10	0.048	**0.055**	0.051	0.051	0.052	0.051	0.051	0.051	0.051	0.052	0.052	0.051
	20	0.046	**0.056**	0.050	0.051	0.052	0.051	0.051	0.051	0.051	0.052	0.051	0.052
NLM_1	10	0.318	0.446	0.369	0.404	**0.451**	0.412	0.406	0.423	0.423	0.416	0.377	0.382
	20	0.683	0.758	0.695	0.729	**0.776**	0.736	0.730	0.746	0.746	0.735	0.666	0.659
NLM_2	10	0.301	0.322	0.289	0.326	**0.378**	0.333	0.328	0.346	0.346	0.338	0.310	0.305
	20	0.593	0.609	0.562	0.604	**0.671**	0.613	0.606	0.628	0.628	0.613	0.549	0.534
NLM_2	10	0.248	0.250	0.225	0.255	**0.302**	0.262	0.257	0.272	0.272	0.281	0.237	0.275
	20	0.458	0.460	0.414	0.455	**0.524**	0.464	0.458	0.478	0.478	0.504	0.402	0.508
NLM_4	10	0.075	0.075	0.067	0.073	**0.084**	0.075	0.074	0.077	0.077	0.080	0.072	0.081
	20	0.094	0.094	0.080	0.090	0.111	0.092	0.091	0.096	0.096	0.106	0.083	**0.120**
$NDPC_1$	10	0.265	**0.439**	0.348	0.378	0.414	0.384	0.380	0.393	0.393	0.410	0.343	0.415
	20	0.581	0.684	0.609	0.638	0.680	0.644	0.639	0.653	0.653	0.686	0.590	**0.714**
$NDPC_2$	10	0.197	**0.436**	0.310	0.338	0.370	0.344	0.340	0.351	0.351	0.364	0.304	0.365
	20	0.532	**0.722**	0.612	0.640	0.683	0.647	0.642	0.656	0.656	0.681	0.573	0.688
$NDPC_2$	10	0.067	**0.181**	0.111	0.124	0.143	0.127	0.125	0.131	0.131	0.140	0.116	0.146
	20	0.132	**0.284**	0.185	0.205	0.240	0.210	0.206	0.217	0.217	0.245	0.177	0.282
$NDPC_4$	10	0.044	**0.060**	0.052	0.052	0.053	0.052	0.052	0.052	0.052	0.053	0.053	0.052
	20	0.043	**0.062**	0.053	0.053	0.053	0.053	0.053	0.053	0.053	0.054	0.052	0.055
PCM_1	10	0.103	**0.290**	0.184	0.204	0.230	0.209	0.205	0.215	0.215	0.224	0.186	0.224
	20	0.289	**0.507**	0.380	0.408	0.450	0.414	0.410	0.423	0.423	0.444	0.351	0.441
PCM_2	10	0.095	**0.251**	0.179	0.187	0.189	0.189	0.188	0.190	0.190	0.188	0.174	0.177
	20	0.258	**0.442**	0.358	0.369	0.375	0.371	0.370	0.374	0.374	0.359	0.335	0.307
PCM_2	10	0.052	**0.062**	0.056	0.057	0.058	0.057	0.057	0.058	0.058	0.059	0.056	0.059
	20	0.055	0.070	0.061	0.063	0.066	0.063	0.063	0.064	0.064	0.069	0.059	**0.077**
PCM_4	10	0.048	**0.058**	0.054	0.053	0.053	0.053	0.053	0.053	0.053	0.054	0.052	0.055
	20	0.050	**0.062**	0.057	0.056	0.056	0.056	0.056	0.056	0.056	0.058	0.054	**0.063**

Note. As in Table 7A.
Source: authors' work.

Table 9A. The power of *GoFTs* for group of alternatives C (ALTs)

ALT	N	Numbered *GoFT* (see Table 3)											
		1	2	3	4	5	6	7	8	9	10	11	12
P_1	10	0.110	0.110	0.095	0.108	**0.130**	0.111	0.109	0.116	0.116	0.121	0.105	0.123
	20	0.170	0.170	0.143	0.163	0.202	0.169	0.165	0.176	0.176	0.191	0.148	**0.208**
P_2	10	0.092	0.092	0.079	0.090	**0.108**	0.093	0.091	0.097	0.097	0.101	0.088	0.101
	20	0.130	0.130	0.108	0.124	0.155	0.128	0.125	0.134	0.134	0.147	0.115	**0.162**
P_2	10	0.076	0.075	0.067	0.074	**0.085**	0.075	0.074	0.078	0.078	0.080	0.072	0.081
	20	0.092	0.094	0.079	0.089	0.110	0.092	0.089	0.096	0.096	0.104	0.084	**0.116**
P_4	10	0.061	0.061	0.057	0.061	**0.067**	0.062	0.061	0.063	0.063	0.064	0.060	0.063
	20	0.067	0.067	0.060	0.065	0.075	0.066	0.065	0.067	0.067	0.071	0.063	**0.078**
NM_1	10	0.194	0.196	0.179	0.204	**0.238**	0.209	0.205	0.217	0.217	0.205	0.203	0.174
	20	0.366	0.368	0.335	0.369	**0.427**	0.377	0.371	0.389	0.389	0.356	0.351	0.265
NM_2	10	0.122	0.121	0.105	0.122	**0.147**	0.125	0.122	0.131	0.131	0.128	0.121	0.115
	20	0.200	0.200	0.171	0.197	**0.244**	0.203	0.198	0.213	0.213	0.202	0.186	0.162
NM_2	10	0.061	0.061	0.056	0.060	**0.067**	0.061	0.060	0.063	0.063	0.062	0.061	0.060
	20	0.068	0.069	0.059	0.066	**0.079**	0.067	0.066	0.069	0.069	0.069	0.066	0.066
NM_4	10	0.053	0.054	0.051	0.053	**0.055**	0.054	0.053	0.054	0.054	0.055	0.052	0.055
	20	0.056	0.055	0.053	0.055	**0.059**	0.055	0.055	0.056	0.056	0.057	0.055	0.059
NLM_1	10	0.132	0.133	0.113	0.132	**0.163**	0.136	0.133	0.142	0.142	0.144	0.129	0.137
	20	0.223	0.224	0.188	0.218	**0.272**	0.225	0.220	0.236	0.236	0.239	0.198	0.224
NLM_2	10	0.094	0.095	0.086	0.095	**0.108**	0.096	0.095	0.099	0.099	0.105	0.090	0.107
	20	0.139	0.142	0.120	0.136	0.164	0.139	0.137	0.146	0.146	0.162	0.123	**0.186**
NLM_2	10	0.082	0.081	0.071	0.079	**0.093**	0.081	0.080	0.085	0.085	0.088	0.077	0.088
	20	0.105	0.105	0.086	0.099	0.126	0.103	0.100	0.108	0.108	0.117	0.093	**0.130**
NLM_4	10	0.060	0.062	0.058	0.061	**0.066**	0.062	0.061	0.063	0.063	0.065	0.060	0.064
	20	0.071	0.071	0.064	0.069	0.078	0.070	0.069	0.072	0.072	0.077	0.067	**0.086**
$NDPC_1$	10	0.061	0.061	0.065	0.063	0.061	0.063	0.063	0.062	0.062	0.065	0.061	**0.066**
	20	0.072	0.072	0.076	0.074	0.072	0.074	0.074	0.073	0.073	0.080	0.070	**0.092**
$NDPC_2$	10	0.059	0.060	0.054	0.057	**0.062**	0.057	0.057	0.059	0.059	0.060	0.057	0.060
	20	0.064	0.064	0.058	0.062	**0.071**	0.063	0.062	0.065	0.065	0.067	0.061	0.068
$NDPC_2$	10	0.053	0.052	0.053	0.053	0.052	0.053	0.053	0.053	0.053	**0.054**	0.052	0.053
	20	0.054	0.054	0.055	0.055	0.054	0.055	0.055	0.054	0.054	0.055	0.053	**0.057**
$NDPC_4$	10	0.051	0.051	0.050	0.050	**0.051**	0.050	0.050	0.050	0.050	0.051	0.051	0.050
	20	0.051	0.050	0.050	0.050	**0.051**	0.050	0.050	0.050	0.050	0.050	0.051	0.051
PCM_1	10	0.080	0.081	0.069	0.078	**0.092**	0.080	0.079	0.083	0.083	0.084	0.078	0.082
	20	0.103	0.103	0.086	0.099	**0.125**	0.103	0.100	0.108	0.108	0.110	0.093	0.101
PCM_2	10	0.101	0.101	0.091	0.102	**0.119**	0.104	0.102	0.108	0.108	0.102	0.107	0.089
	20	0.151	0.152	0.132	0.150	**0.181**	0.154	0.151	0.161	0.161	0.144	0.156	0.110
PCM_2	10	0.075	0.074	0.066	0.073	**0.085**	0.075	0.074	0.077	0.077	0.075	0.076	0.069
	20	0.096	0.096	0.082	0.093	**0.114**	0.096	0.094	0.100	0.100	0.094	0.096	0.079
PCM_4	10	0.060	0.059	0.056	0.058	**0.062**	0.059	0.058	0.060	0.060	0.061	0.058	0.061
	20	0.062	0.061	0.056	0.059	0.067	0.060	0.059	0.061	0.061	0.065	0.058	**0.071**

Note. As in Table 7A.
Source: authors' work.

Table 10A. The power of *GoFTs* for group of alternatives D (ALTs)

ALT	N	Numbered *GoFT* (see Table 3)											
		1	2	3	4	5	6	7	8	9	10	11	12
P_1	10	0.468	0.467	0.629	0.572	0.390	0.556	0.568	0.526	0.526	0.600	0.367	**0.667**
	20	0.871	0.871	0.932	0.910	0.815	0.903	0.909	0.890	0.890	0.950	0.715	**0.980**
P_2	10	0.090	0.093	**0.154**	0.125	0.064	0.118	0.123	0.105	0.105	0.118	0.082	0.127
	20	0.212	0.211	0.311	0.262	0.149	0.251	0.260	0.230	0.230	0.287	0.148	**0.350**
P_2	10	0.044	0.042	**0.061**	0.052	0.033	0.049	0.051	0.046	0.046	0.045	0.045	0.043
	20	0.057	0.056	**0.085**	0.070	0.039	0.066	0.069	0.059	0.059	0.063	0.053	0.059
P_4	10	0.040	0.041	**0.050**	0.045	0.035	0.044	0.045	0.041	0.041	0.041	0.043	0.039
	20	0.043	0.043	**0.057**	0.050	0.034	0.048	0.049	0.044	0.044	0.044	0.045	0.038
NM_1	10	0.098	0.099	**0.160**	0.132	0.073	0.125	0.131	0.113	0.113	0.112	0.101	0.104
	20	0.229	0.233	**0.330**	0.284	0.172	0.272	0.281	0.251	0.251	0.247	0.195	0.212
NM_2	10	0.157	0.128	0.208	0.184	0.116	0.177	0.183	0.167	0.167	0.201	0.142	**0.216**
	20	0.311	0.262	0.365	0.333	0.241	0.325	0.331	0.309	0.309	0.432	0.325	**0.477**
NM_2	10	0.051	0.043	**0.062**	0.054	0.039	0.052	0.054	0.049	0.049	0.048	0.049	0.046
	20	0.064	0.055	**0.082**	0.069	0.045	0.067	0.069	0.062	0.062	0.060	0.058	0.054
NM_4	10	0.047	0.048	**0.049**	0.048	0.047	0.048	0.048	0.048	0.048	0.048	0.048	0.047
	20	0.047	0.048	**0.050**	0.049	0.046	0.049	0.049	0.048	0.048	0.047	0.048	0.045
NLM_1	10	0.046	**0.130**	0.090	0.092	0.092	0.092	0.092	0.092	0.092	0.096	0.084	0.097
	20	0.082	**0.197**	0.139	0.143	0.146	0.144	0.144	0.145	0.145	0.156	0.123	0.170
NLM_2	10	0.044	**0.069**	0.056	0.057	0.058	0.057	0.057	0.057	0.057	0.058	0.056	0.058
	20	0.045	**0.080**	0.060	0.062	0.066	0.063	0.062	0.063	0.063	0.066	0.061	0.072
NLM_2	10	0.058	**0.089**	0.067	0.072	0.081	0.073	0.072	0.075	0.075	0.079	0.070	0.080
	20	0.076	0.116	0.085	0.093	0.110	0.096	0.094	0.098	0.098	0.108	0.086	**0.122**
NLM_4	10	0.069	**0.106**	0.078	0.087	0.101	0.088	0.087	0.091	0.091	0.095	0.083	0.096
	20	0.097	0.148	0.103	0.117	0.145	0.121	0.118	0.126	0.126	0.138	0.108	**0.150**
$NDPC_1$	10	0.097	0.096	**0.155**	0.128	0.072	0.121	0.126	0.109	0.109	0.109	0.098	0.101
	20	0.223	0.224	**0.321**	0.276	0.165	0.264	0.273	0.243	0.243	0.237	0.189	0.204
$NDPC_2$	10	0.052	0.051	**0.069**	0.060	0.042	0.058	0.060	0.055	0.055	0.054	0.055	0.051
	20	0.068	0.068	**0.095**	0.081	0.051	0.078	0.081	0.072	0.072	0.069	0.069	0.059
$NDPC_2$	10	0.047	0.046	**0.048**	0.047	0.045	0.046	0.046	0.046	0.046	0.046	0.047	0.045
	20	0.046	0.045	**0.048**	0.046	0.043	0.046	0.046	0.045	0.045	0.044	0.048	0.040
$NDPC_4$	10	0.050	0.050	**0.051**	0.050	0.049	0.051	0.051	0.051	0.051	0.050	0.051	0.049
	20	0.047	**0.050**	**0.049**	0.049	0.048	0.048	0.048	0.048	0.048	0.048	0.049	0.047
PCM_1	10	0.046	0.046	**0.050**	0.048	0.043	0.047	0.048	0.046	0.046	0.045	0.046	0.043
	20	0.045	0.044	**0.049**	0.046	0.041	0.046	0.046	0.045	0.045	0.043	0.047	0.039
PCM_2	10	0.054	0.054	**0.067**	0.061	0.047	0.059	0.060	0.056	0.056	0.055	0.056	0.054
	20	0.065	0.066	**0.084**	0.074	0.054	0.072	0.074	0.068	0.068	0.066	0.068	0.062
PCM_2	10	0.048	0.047	**0.053**	0.049	0.044	0.049	0.049	0.048	0.048	0.048	0.049	0.046
	20	0.049	0.047	**0.055**	0.051	0.043	0.050	0.051	0.048	0.048	0.047	0.050	0.045
PCM_4	10	0.046	0.047	**0.051**	0.048	0.044	0.047	0.048	0.046	0.046	0.046	0.047	0.046
	20	0.046	0.048	**0.053**	0.050	0.043	0.049	0.050	0.048	0.048	0.046	0.048	0.045

Note. As in Table 7A.
Source: authors' work.

Table 11A. The power of *GoFTs* for group of alternatives E (ALTs)

ALT	N	Numbered *GoFT* (see Table 3)											
		1	2	3	4	5	6	7	8	9	10	11	12
P_1	10	0.753	0.537	0.739	0.719	0.635	0.712	0.717	0.700	0.700	0.737	0.579	**0.775**
	20	0.972	0.927	0.972	0.967	0.942	0.965	0.966	0.961	0.961	0.979	0.901	**0.990**
P_2	10	**0.264**	0.081	0.212	0.200	0.162	0.196	0.199	0.189	0.189	0.206	0.150	0.224
	20	0.496	0.259	0.440	0.419	0.352	0.414	0.418	0.402	0.402	0.459	0.296	**0.530**
P_2	10	**0.106**	0.031	0.074	0.071	0.064	0.070	0.071	0.069	0.069	0.072	0.065	0.074
	20	**0.158**	0.052	0.116	0.110	0.096	0.109	0.110	0.106	0.106	0.115	0.096	0.127
P_4	10	**0.073**	0.033	0.057	0.054	0.050	0.054	0.054	0.054	0.054	0.054	0.053	0.054
	20	**0.086**	0.036	0.067	0.064	0.058	0.064	0.064	0.062	0.062	0.064	0.060	0.066
NM_1	10	**0.139**	0.036	0.094	0.094	0.086	0.093	0.094	0.092	0.092	0.093	0.086	0.092
	20	**0.227**	0.085	0.169	0.166	0.152	0.166	0.166	0.163	0.163	0.165	0.141	0.157
NM_2	10	**0.094**	0.035	0.065	0.065	0.065	0.065	0.065	0.066	0.066	0.066	0.063	0.065
	20	**0.124**	0.048	0.088	0.088	0.086	0.088	0.088	0.088	0.088	0.089	0.082	0.085
NM_2	10	**0.054**	0.045	0.050	0.050	0.049	0.050	0.050	0.049	0.049	0.049	0.049	0.048
	20	**0.056**	0.044	0.050	0.051	0.049	0.051	0.051	0.050	0.050	0.051	0.051	0.051
NM_4	10	**0.054**	0.047	0.052	0.051	0.051	0.051	0.051	0.051	0.051	0.052	0.050	0.051
	20	**0.054**	0.046	0.050	0.050	0.050	0.050	0.050	0.050	0.050	0.050	0.050	0.050
NLM_1	10	0.044	**0.094**	0.068	0.070	0.072	0.070	0.070	0.071	0.071	0.074	0.066	0.075
	20	0.057	**0.122**	0.086	0.089	0.095	0.090	0.089	0.091	0.091	0.099	0.082	0.114
NLM_2	10	**0.092**	0.043	0.063	0.068	0.073	0.068	0.068	0.070	0.070	0.071	0.066	0.072
	20	**0.119**	0.053	0.079	0.085	0.096	0.086	0.085	0.087	0.087	0.093	0.080	0.101
NLM_2	10	0.050	**0.104**	0.070	0.076	0.084	0.077	0.076	0.079	0.079	0.081	0.074	0.083
	20	0.068	**0.137**	0.090	0.099	0.116	0.102	0.100	0.105	0.105	0.115	0.091	0.129
NLM_4	10	**0.131**	0.062	0.086	0.096	0.111	0.099	0.097	0.102	0.102	0.106	0.092	0.109
	20	**0.194**	0.103	0.129	0.144	0.172	0.147	0.145	0.152	0.152	0.169	0.130	0.192
$NDPC_1$	10	0.703	0.463	**0.716**	0.689	0.567	0.679	0.687	0.660	0.660	0.648	0.597	0.603
	20	0.979	0.949	**0.983**	0.979	0.954	0.977	0.978	0.974	0.974	0.969	0.940	0.940
$NDPC_2$	10	**0.105**	0.033	0.088	0.080	0.061	0.077	0.079	0.074	0.074	0.074	0.071	0.072
	20	**0.172**	0.072	0.154	0.139	0.104	0.136	0.138	0.129	0.129	0.128	0.111	0.115
$NDPC_2$	10	**0.060**	0.033	0.057	0.051	0.041	0.050	0.051	0.048	0.048	0.047	0.048	0.044
	20	**0.073**	0.038	0.071	0.063	0.046	0.061	0.063	0.058	0.058	0.056	0.056	0.048
$NDPC_4$	10	**0.057**	0.038	0.053	0.049	0.044	0.048	0.049	0.047	0.047	0.048	0.048	0.046
	20	**0.062**	0.038	0.057	0.053	0.045	0.052	0.053	0.051	0.051	0.050	0.052	0.047
PCM_1	10	**0.146**	0.058	0.143	0.125	0.084	0.120	0.124	0.113	0.113	0.116	0.105	0.117
	20	0.272	0.151	**0.280**	0.248	0.172	0.241	0.247	0.226	0.226	0.243	0.208	0.255
PCM_2	10	0.085	0.043	**0.089**	0.077	0.053	0.074	0.077	0.069	0.069	0.068	0.067	0.065
	20	0.137	0.077	**0.148**	0.128	0.083	0.123	0.127	0.114	0.114	0.111	0.102	0.097
PCM_2	10	**0.071**	0.031	0.062	0.057	0.045	0.055	0.056	0.053	0.053	0.052	0.052	0.050
	20	**0.091**	0.040	0.082	0.073	0.054	0.072	0.073	0.068	0.068	0.070	0.061	0.064
PCM_4	10	0.054	0.041	**0.055**	0.052	0.043	0.050	0.051	0.049	0.049	0.049	0.049	0.046
	20	0.059	0.042	**0.061**	0.055	0.043	0.054	0.055	0.052	0.052	0.049	0.053	0.044

Note. As in Table 7A.
Source: authors' work.

Table 12A. The power of *GoFTs* for group of alternatives F (ALTs)

ALT	N	Numbered *GoFT* (see Table 3)											
		1	2	3	4	5	6	7	8	9	10	11	12
P_1	10	0.538	0.755	0.741	0.721	0.636	0.714	0.720	0.701	0.701	0.740	0.582	**0.777**
	20	0.926	0.971	0.972	0.966	0.942	0.964	0.965	0.961	0.961	0.979	0.900	**0.990**
P_2	10	0.081	0.268	0.214	0.203	0.164	0.199	0.202	0.192	0.192	0.209	0.152	0.227
	20	0.261	0.498	0.441	0.420	0.356	0.415	0.419	0.404	0.404	0.461	0.298	**0.532**
P_2	10	0.030	**0.106**	0.075	0.072	0.064	0.071	0.071	0.070	0.070	0.072	0.066	0.073
	20	0.052	**0.156**	0.115	0.110	0.095	0.108	0.110	0.106	0.106	0.114	0.095	0.127
P_4	10	0.032	**0.074**	0.056	0.054	0.051	0.054	0.054	0.053	0.053	0.054	0.052	0.053
	20	0.036	**0.086**	0.065	0.062	0.056	0.062	0.062	0.061	0.061	0.063	0.059	0.064
NM_1	10	0.120	**0.326**	0.297	0.274	0.204	0.267	0.272	0.255	0.255	0.254	0.216	0.240
	20	0.447	**0.656**	0.642	0.614	0.517	0.606	0.612	0.591	0.591	0.584	0.476	0.530
NM_2	10	0.041	**0.087**	0.086	0.075	0.053	0.072	0.074	0.068	0.068	0.069	0.064	0.069
	20	0.067	**0.129**	0.129	0.114	0.077	0.110	0.113	0.103	0.103	0.109	0.091	0.110
NM_2	10	0.045	**0.055**	0.051	0.050	0.050	0.050	0.050	0.050	0.050	0.050	0.050	0.050
	20	0.044	**0.057**	0.051	0.051	0.050	0.050	0.051	0.050	0.050	0.050	0.050	0.048
NM_4	10	0.049	0.049	**0.050**	0.049	0.049	0.049	0.049	0.049	0.049	0.049	0.049	0.049
	20	0.049	**0.051**	**0.051**	0.051	0.051	0.051	0.051	0.051	0.051	0.051	0.050	0.050
NLM_1	10	**0.112**	0.047	0.074	0.079	0.087	0.080	0.080	0.082	0.082	0.086	0.077	0.087
	20	**0.162**	0.074	0.108	0.116	0.130	0.118	0.117	0.121	0.121	0.131	0.105	0.146
NLM_2	10	**0.125**	0.050	0.080	0.087	0.097	0.089	0.087	0.091	0.091	0.094	0.083	0.096
	20	**0.181**	0.082	0.118	0.129	0.148	0.131	0.130	0.135	0.135	0.148	0.116	0.163
NLM_2	10	**0.108**	0.057	0.075	0.081	0.091	0.082	0.081	0.085	0.085	0.088	0.078	0.090
	20	**0.149**	0.080	0.101	0.111	0.130	0.114	0.112	0.117	0.117	0.128	0.101	0.142
NLM_4	10	0.050	**0.086**	0.062	0.067	0.073	0.068	0.067	0.069	0.069	0.071	0.065	0.072
	20	0.059	**0.103**	0.073	0.078	0.089	0.080	0.079	0.082	0.082	0.088	0.074	0.100
$NDPC_1$	10	0.101	**0.281**	0.251	0.234	0.181	0.228	0.233	0.219	0.219	0.211	0.202	0.192
	20	0.365	**0.567**	0.549	0.523	0.436	0.516	0.522	0.502	0.502	0.473	0.439	0.394
$NDPC_2$	10	0.034	**0.091**	0.076	0.070	0.056	0.068	0.070	0.066	0.066	0.065	0.064	0.062
	20	0.057	**0.133**	0.119	0.108	0.083	0.106	0.108	0.101	0.101	0.097	0.095	0.082
$NDPC_2$	10	0.034	**0.079**	0.061	0.060	0.056	0.060	0.060	0.059	0.059	0.057	0.058	0.052
	20	0.044	**0.104**	0.081	0.079	0.072	0.078	0.079	0.077	0.077	0.073	0.075	0.058
$NDPC_4$	10	0.045	0.052	**0.053**	0.050	0.046	0.049	0.050	0.048	0.048	0.048	0.049	0.047
	20	0.043	**0.055**	0.054	0.051	0.045	0.050	0.051	0.049	0.049	0.048	0.049	0.046
PCM_1	10	0.065	0.150	**0.158**	0.135	0.084	0.129	0.134	0.119	0.119	0.128	0.101	0.130
	20	0.183	0.321	0.336	0.298	0.201	0.288	0.296	0.271	0.271	0.322	0.195	**0.340**
PCM_2	10	0.051	0.055	**0.065**	0.060	0.047	0.058	0.059	0.056	0.056	0.055	0.055	0.052
	20	0.062	0.068	**0.082**	0.075	0.056	0.073	0.074	0.069	0.069	0.070	0.066	0.066
PCM_2	10	0.045	0.046	**0.051**	0.048	0.042	0.047	0.047	0.046	0.046	0.045	0.047	0.043
	20	0.043	0.045	**0.050**	0.047	0.039	0.046	0.046	0.044	0.044	0.043	0.046	0.040
PCM_4	10	0.047	0.047	**0.054**	0.050	0.042	0.049	0.050	0.048	0.048	0.047	0.048	0.046
	20	0.049	0.051	**0.060**	0.055	0.042	0.053	0.054	0.051	0.051	0.049	0.052	0.046

Note. As in Table 7A.
Source: authors' work.

Table 13A. The power of *GoFTs* for group of alternatives G (ALTs)

ALT	N	Numbered *GoFT* (see Table 3)											
		1	2	3	4	5	6	7	8	9	10	11	12
P_1	10	0.790	0.565	0.760	0.747	0.678	0.742	0.746	0.733	0.733	0.767	0.612	**0.800**
	20	0.981	0.940	0.977	0.974	0.959	0.973	0.973	0.971	0.971	0.984	0.922	**0.992**
P_2	10	**0.298**	0.089	0.224	0.217	0.189	0.215	0.217	0.210	0.210	0.229	0.167	0.250
	20	0.551	0.288	0.467	0.456	0.410	0.453	0.455	0.445	0.445	0.505	0.330	**0.584**
P_2	10	**0.114**	0.031	0.077	0.075	0.070	0.075	0.075	0.074	0.074	0.078	0.069	0.080
	20	**0.172**	0.057	0.121	0.119	0.109	0.118	0.119	0.117	0.117	0.127	0.102	0.143
P_4	10	**0.079**	0.034	0.058	0.057	0.056	0.057	0.057	0.056	0.056	0.058	0.055	0.057
	20	**0.095**	0.038	0.068	0.068	0.065	0.067	0.068	0.067	0.067	0.070	0.064	0.074
NM_1	10	**0.096**	0.033	0.065	0.065	0.064	0.065	0.065	0.065	0.065	0.066	0.064	0.066
	20	**0.128**	0.047	0.088	0.089	0.086	0.089	0.089	0.089	0.089	0.091	0.081	0.092
NM_2	10	**0.061**	0.043	0.053	0.052	0.052	0.052	0.052	0.052	0.052	0.052	0.052	0.052
	20	**0.068**	0.042	0.055	0.054	0.054	0.055	0.054	0.055	0.055	0.055	0.054	0.055
NM_2	10	**0.057**	0.045	0.051	0.050	0.050	0.050	0.050	0.050	0.050	0.051	0.050	0.050
	20	**0.059**	0.043	0.050	0.051	0.051	0.051	0.051	0.051	0.051	0.052	0.050	0.052
NM_4	10	**0.054**	0.048	0.051	0.051	0.050	0.051	0.051	0.051	0.051	0.051	0.050	0.049
	20	**0.054**	0.047	0.052	0.051	0.052	0.052	0.052	0.051	0.051	0.052	0.051	0.052
NLM_1	10	**0.526**	0.351	0.472	0.482	0.468	0.482	0.482	0.481	0.481	0.473	0.454	0.431
	20	**0.803**	0.709	0.767	0.777	0.779	0.778	0.777	0.780	0.780	0.766	0.753	0.672
NLM_2	10	**0.189**	0.056	0.143	0.137	0.118	0.136	0.137	0.133	0.133	0.134	0.119	0.133
	20	**0.332**	0.156	0.277	0.268	0.233	0.265	0.268	0.260	0.260	0.263	0.222	0.254
NLM_2	10	**0.170**	0.066	0.105	0.117	0.135	0.119	0.118	0.123	0.123	0.130	0.109	0.134
	20	**0.265**	0.129	0.176	0.195	0.226	0.199	0.196	0.205	0.205	0.224	0.173	0.251
NLM_4	10	**0.106**	0.068	0.076	0.084	0.099	0.086	0.084	0.089	0.089	0.094	0.081	0.097
	20	0.148	0.099	0.105	0.119	0.145	0.122	0.120	0.127	0.127	0.141	0.109	**0.161**
$NDPC_1$	10	**0.180**	0.071	0.144	0.141	0.124	0.139	0.140	0.137	0.137	0.144	0.133	0.145
	20	0.333	0.140	0.257	0.256	0.234	0.255	0.256	0.251	0.251	0.301	0.236	**0.336**
$NDPC_2$	10	**0.087**	0.032	0.060	0.060	0.059	0.060	0.060	0.060	0.060	0.061	0.059	0.061
	20	**0.113**	0.042	0.078	0.078	0.075	0.078	0.078	0.078	0.078	0.083	0.071	0.092
$NDPC_2$	10	**0.097**	0.037	0.067	0.068	0.068	0.068	0.068	0.068	0.068	0.068	0.067	0.067
	20	**0.130**	0.053	0.092	0.093	0.093	0.094	0.094	0.094	0.094	0.093	0.088	0.087
$NDPC_4$	10	0.050	0.047	**0.052**	0.050	0.045	0.049	0.050	0.048	0.048	0.049	0.049	0.049
	20	0.056	0.051	**0.060**	0.057	0.051	0.056	0.057	0.055	0.055	0.056	0.054	0.058
PCM_1	10	**0.260**	0.106	0.203	0.204	0.189	0.203	0.204	0.202	0.202	0.206	0.198	0.194
	20	**0.484**	0.265	0.400	0.405	0.391	0.406	0.405	0.405	0.405	0.409	0.402	0.339
PCM_2	10	0.143	0.088	**0.165**	0.144	0.097	0.139	0.143	0.129	0.129	0.131	0.118	0.127
	20	0.278	0.194	**0.302**	0.275	0.197	0.268	0.273	0.253	0.253	0.261	0.210	0.260
PCM_2	10	**0.084**	0.034	0.061	0.060	0.058	0.060	0.060	0.060	0.060	0.061	0.058	0.061
	20	**0.106**	0.041	0.074	0.074	0.071	0.074	0.074	0.074	0.074	0.078	0.070	0.086
PCM_4	10	**0.083**	0.036	0.060	0.060	0.060	0.060	0.060	0.060	0.060	0.061	0.059	0.061
	20	**0.102**	0.041	0.071	0.072	0.071	0.072	0.072	0.072	0.072	0.075	0.069	0.077

Note. As in Table 7A.
Source: authors' work.

Table 14A. The power of *GoFTs* for group of alternatives H (ALTs)

ALT	n	Numbered *GoFT* (see Table 3)											
		1	2	3	4	5	6	7	8	9	10	11	12
P_1	10	0.565	0.792	0.761	0.748	0.680	0.743	0.747	0.733	0.733	0.767	0.612	**0.801**
	20	0.942	0.982	0.978	0.975	0.960	0.974	0.975	0.972	0.972	0.985	0.924	**0.993**
P_2	10	0.086	**0.299**	0.222	0.217	0.190	0.215	0.217	0.210	0.210	0.229	0.164	0.250
	20	0.284	0.552	0.467	0.456	0.408	0.452	0.455	0.445	0.445	0.504	0.329	**0.583**
P_2	10	0.032	**0.116**	0.080	0.078	0.072	0.077	0.077	0.076	0.076	0.079	0.072	0.081
	20	0.057	**0.169**	0.120	0.118	0.108	0.117	0.118	0.116	0.116	0.125	0.101	0.142
P_4	10	0.034	**0.078**	0.058	0.057	0.055	0.056	0.057	0.056	0.056	0.058	0.055	0.058
	20	0.039	**0.096**	0.069	0.069	0.066	0.069	0.069	0.068	0.068	0.070	0.065	0.075
NM_1	10	0.033	**0.119**	0.080	0.079	0.076	0.079	0.079	0.079	0.079	0.080	0.074	0.080
	20	0.065	**0.182**	0.130	0.129	0.121	0.128	0.129	0.128	0.128	0.132	0.113	0.134
NM_2	10	0.034	**0.097**	0.066	0.066	0.066	0.066	0.066	0.067	0.067	0.067	0.064	0.067
	20	0.047	**0.129**	0.088	0.089	0.088	0.089	0.089	0.089	0.089	0.091	0.082	0.092
NM_2	10	0.045	**0.058**	0.052	0.052	0.051	0.052	0.052	0.051	0.051	0.052	0.050	0.050
	20	0.044	**0.060**	0.051	0.051	0.051	0.051	0.051	0.052	0.052	0.052	0.051	0.051
NM_4	10	0.049	**0.051**	0.050	0.049	0.050	0.049	0.049	0.049	0.049	0.050	0.049	0.049
	20	0.049	**0.051**	0.050	0.050	0.050	0.050	0.050	0.050	0.050	0.050	0.050	0.050
NLM_1	10	0.119	**0.360**	0.241	0.257	0.269	0.260	0.258	0.264	0.264	0.273	0.229	0.274
	20	0.378	**0.638**	0.510	0.529	0.548	0.533	0.530	0.538	0.538	0.559	0.456	0.567
NLM_2	10	0.067	**0.184**	0.115	0.127	0.145	0.130	0.128	0.135	0.135	0.141	0.118	0.144
	20	0.143	**0.297**	0.199	0.218	0.251	0.223	0.219	0.230	0.230	0.250	0.192	0.276
NLM_2	10	**0.094**	0.058	0.068	0.074	0.086	0.075	0.074	0.078	0.078	0.082	0.072	0.084
	20	0.127	0.074	0.087	0.096	0.116	0.099	0.097	0.102	0.102	0.113	0.090	**0.128**
NLM_4	10	0.082	**0.115**	0.085	0.097	0.115	0.099	0.098	0.103	0.103	0.109	0.093	0.111
	20	0.122	0.167	0.121	0.138	0.173	0.142	0.139	0.149	0.149	0.165	0.127	**0.186**
$NDPC_1$	10	0.041	**0.157**	0.098	0.101	0.103	0.102	0.101	0.103	0.103	0.106	0.094	0.104
	20	0.100	**0.261**	0.181	0.186	0.187	0.187	0.186	0.188	0.188	0.196	0.160	0.193
$NDPC_2$	10	0.066	**0.188**	0.130	0.136	0.140	0.137	0.136	0.139	0.139	0.135	0.129	0.121
	20	0.166	**0.328**	0.250	0.260	0.265	0.261	0.260	0.263	0.263	0.247	0.237	0.194
$NDPC_2$	10	0.042	**0.064**	0.053	0.052	0.053	0.052	0.052	0.053	0.053	0.053	0.052	0.052
	20	0.042	**0.068**	0.056	0.055	0.055	0.055	0.055	0.055	0.055	0.055	0.055	0.055
$NDPC_4$	10	0.049	**0.052**	0.051	0.051	0.050	0.050	0.051	0.050	0.050	0.051	0.050	0.050
	20	0.048	**0.052**	0.051	0.051	0.050	0.051	0.051	0.051	0.051	0.051	0.050	0.051
PCM_1	10	0.036	**0.077**	0.057	0.057	0.056	0.057	0.057	0.057	0.057	0.058	0.055	0.057
	20	0.039	**0.094**	0.067	0.067	0.065	0.067	0.067	0.067	0.067	0.068	0.064	0.070
PCM_2	10	0.042	**0.063**	0.053	0.053	0.052	0.053	0.053	0.053	0.053	0.054	0.051	0.053
	20	0.041	**0.067**	0.055	0.055	0.053	0.055	0.055	0.054	0.054	0.055	0.054	0.056
PCM_2	10	0.042	**0.073**	0.058	0.058	0.058	0.058	0.058	0.058	0.058	0.058	0.057	0.056
	20	0.046	**0.087**	0.069	0.068	0.066	0.068	0.068	0.068	0.068	0.066	0.066	0.063
PCM_4	10	0.041	**0.067**	0.052	0.054	0.056	0.054	0.054	0.055	0.055	0.055	0.054	0.054
	20	0.044	**0.074**	0.057	0.059	0.062	0.059	0.059	0.060	0.060	0.059	0.059	0.054

Note. As in Table 7A.
Source: authors' work.

References

Afeez, B. M., Maxwell, O., Otekunrin, O. A., & Happiness, O. I. (2018). Selection and Validation of Comparative Study of Normality Test. *American Journal of Mathematics and Statistics*, *8*(6), 190–201. https://doi.org/10.5923/j.ajms.20180806.05.

Ahmad, F., & Khan, R. A. (2015). Power Comparison of Various Normality Tests. Pakistan *Journal of Statistics and Operation Research*, *11*(3), 331–345. https://doi.org/10.18187/pjsor.v11i3.845.

Aliaga, A. M., & Martínez-González, E., & Cayón, L., & Argüeso, F., & Sanz, J. L., & Barreiro, R .B. (2003). Goodness-of-fit tests of Gaussianity: constraints on the cumulants of the MAXIMA data. *New Astronomy Reviews*, *47*(8–10), 821–826. https://doi.org/10.1016/j.newar.2003.07.010.

Anderson, T. W., & Darling, D. A. (1952). Asymptotic theory of certain "goodness of fit" criteria based on stochastic processes. *Annals of Mathematical Statistics*, *23*(2), 193–212. https://doi.org /10.1214/aoms/1177729437.

Arnastauskaitė, J., Ruzgas, T., & Bražėnas, M. (2021). An Exhaustive Power Comparison of Normality Tests. *Mathematics*, *9*(7), 1–20. https://doi.org/10.3390/math9070788.

Azzalini, A. (1985). A Class of Distributions which Includes the Normal Ones. *Scandinavian Journal of Statistics*, *12*(2), 171–178.

Bayoud, H. A. (2021). Tests of normality: new test and comparative study. *Communications in Statistics – Simulation and Computation*, *50*(12), 4442–4463. https://doi.org/10.1080/03610918.2019.1643883.

Bloom, G. (1958). *Statistical estimates and transformed beta-variables*. John Wiley & Sons.

Bonett, D. G., & Seier, E. (2002). A test of normality with high uniform power. *Computational Statistics and Data Analysis*, *40*(3), 435–445. https://doi.org/10.1016/S0167-9473(02)00074-9.

Bontemps, C., & Meddahi, N. (2005). Testing normality: a GMM approach. *Journal of Econometrics*, *124*(1), 149–186. https://doi.org/10.1016/j.jeconom.2004.02.014.

Borrajo, M. I., González-Manteiga, W., & Martínez-Miranda, M. D. (2024). Goodness-of-fit test for point processes first-order intensity. *Computational Statistics* & *Data Analysis*, *194*, 1–16. https://doi.org/10.1016/j.csda.2024.107929.

Brys, G., Hubert, M., & Struyf, A. (2008). Goodness-of-fit tests based on a robust measure of skewness. *Computational Statistics*, *23*(3), 429–442. https://doi.org/10.1007/s00180-007-0083-7.

Chernobai, A., Rachev, S. T., & Fabozzi, F. J. (2012). Composite Goodness-of-Fit Tests for Left-Truncated Loss Samples. In C. F. Lee (Ed.), *Handbook of Financial Econometrics and Statistics* (pp. 575–596). Springer.

Coin, D. (2008). A goodness-of-fit test for normality based on polynomial regression. *Computational Statistics and Data Analysis*, *52*(4), 2185–2198. https://doi.org/10.1016/j.csda.2007.07.012.

Cramér, H. (1928). On the Composition of Elementary Errors. First paper: Mathematical deductions. *Scandinavian Actuarial Journal*, (1), 13–74. https://doi.org/10.1080/03461238.1928.10416862.

Desgagné, A., & Lafaye de Micheaux, P. (2018). A powerful and interpretable alternative to the Jarque-Bera test of normality based on 2nd-power Skewness and Kurtosis, using the Rao's Score Test on the APD family. *Journal of Applied Statistics*, *45*(13), 2307–2327. https://doi.org/10.1080 /02664763.2017.1415311.

Desgagné, A., Lafaye de Micheaux, P., & Ouimet, F. (2022). Goodness-of-fit tests for Laplace, Gaussian and exponential power distributions based on λ-th power skewness and kurtosis. Statistics. *A Journal of Theoretical and Applied Statistics*, *57*(1), 94–122. https://doi.org/10.1080/02331888.2022.2144859.

Durbin, J., & Knott, M. (1972). Components of Cramér-von Mises statistics. I. *Journal of the Royal Statistical Society: Series B (Methodological)*, *34*(2), 290–307.

Esteban, M. D., Castellanos, M. E., Morales, D., & Vajda, I. (2001). Monte Carlo comparison of four normality tests using different entropy estimates. *Communications in Statistics – Simulation and Computation*, *30*(4), 761–785.

Feuerverger, A. (2016). On Goodness of Fit for Operational Risk. *International Statistical Review*, *84*(3), 434–455.

Filliben, J. J. (1975). The Probability Plot Correlation Coefficient Test for Normality. *Technometrics*, *17*(1), 111–117. https://doi.org/10.1080/00401706.1975.10489279.

Gan, F. F., & Koehler, K. J. (1990). Goodness-of-Fit Tests Based on P-P Probability Plots. *Technometrics*, *32*(3), 289–303. https://doi.org/10.2307/1269106.

Gel, Y. R., & Gastwirth, J. L. (2008). A robust modification of the Jarque-Bera test of normality. *Economics Letters*, *99*(1), 30–32. https://doi.org/10.1016/j.econlet.2007.05.022.

Gel, Y. R., Miao, W., & Gastwirth, J. L. (2007). Robust directed tests of normality against heavy-tailed alternatives. *Computational Statistics & Data Analysis*, *51*(5), 2734–2746. https://doi.org/10.1016/j.csda.2006.08.022.

Giles, D. E. (2024). New Goodness-of-Fit Tests for the Kumaraswamy Distribution. *Stats*, *7*(2), 373–389. https://doi.org/10.3390/stats7020023.

Harter, H. L., Khamis, H. J., & Lamb, R. E. (1984). Modified Kolmogorov-Smirnov tests of goodness of fit. *Communications in Statistics – Simulation and Computation*, *13*(3), 293–323. https://doi.org/10.1080/03610918408812378.

Hernandez, H. (2021). *Testing for Normality: What is the Best Method?* (ForsChem Research Reports, vol. 6, 2021-05). https://doi.org/10.13140/RG.2.2.13926.14406.

Kellner, J., & Celisse, A. (2019). A one-sample test for normality with kernel methods. *Bernoulli*, *25*(3), 1816–1837. https://doi.org/10.3150/18-BEJ1037.

Khamis, H. J. (1990). The δ-corrected Kolmogorov-Smirnov test for goodness-of-fit. *Journal of Statistical Planning and Inference*, *24*(3), 317–335. https://doi.org/10.1016/0378-3758(90)90051-U.

Khamis, H. J. (1992). The Δ-corrected Kolmogorov-Smirnov test with estimated parameters. *Journal of Nonparametric Statistics*, *2*(1), 17–27. https://doi.org/10.1080/10485259208832539.

Khamis, H. J. (1993). A comparative study of the &-corrected Kolmogorov-Smirnov test. *Journal of Applied Statistics*, *20*(3), 401–421. https://doi.org/10.1080/02664769300000040.

Khatun, N. (2021). Applications of Normality Test in Statistical Analysis. *Open Journal of Statistics*, *11*(1), 113–122. https://doi.org/10.4236/ojs.2021.111006.

Kolmogorov, A. (1933). Sulla determinazione empirica di unalegge di distributione. *Giornale dell' Istituto Italiano degli Attuari*, (4), 83–91.

Krauczi, É. (2009). A study of the quantile correlation test for normality. TEST. *An Official Journal of the Spanish Society of Statistics and Operations Research*, *18*(1), 156–165. https://doi.org/10.1007/s11749-007-0074-6.

Kuiper, N. H. (1960). Tests concerning random points on a circle. *Indagationes Mathematicae (Proceedings)*, *63*, 38–47. https://doi.org/10.1016/S1385-7258(60)50006-0.

Lilliefors, H. W. (1967). On the Kolmogorov-Smirnov Test for Normality with Mean and Variance Unknown. *Journal of the American Statistical Association*, *62*(318), 399–402. https://doi.org/10.1080/01621459.1967.10482916.

Luceño, A. (2006). Fitting the generalized Pareto distribution to data using maximum goodness-of-fit estimators. *Computational Statistics* & *Data Analysis*, *51*(2), 904–917. https://doi.org/10.1016/j.csda.2005.09.011.

Ma, Y., Kitani, M., & Murakami, H. (2024). On modified Anderson-Darling test statistics with asymptotic properties. *Communications in Statistics – Theory and Methods*, *53*(4), 1420–1439. https://doi.org/10.1080/03610926.2022.2101121.

Marange, C. S., & Qin, Y. (2019). An Empirical Likelihood Ratio Based Comparative Study on Tests for Normality of Residuals in Linear Models. *Metodoloski zvezki*, *16*(1), 1–16. https://doi.org/10.51936/ramh7128.

Mbah, A. K., & Paothong, A. (2015). Shapiro-Francia test compared to other normality test using expected *p*-value. *Journal of Statistical Computation and Simulation*, *85*(15), 3002–3016. https://doi.org/10.1080/00949655.2014.947986.

von Mises, R. (1931). Über einige Abschätzungen von Erwartungswerten. *Journal für die reine und angewandte Mathematik*, *1931*(165), 184–193. https://doi.org/10.1515/crll.1931.165.184.

Mishra, P., Pandey, C. M., Singh, U., Gupta, A., Sahu, C., & Keshri, A. (2019). Descriptive Statistics and Normality Tests for Statistical Data. *Annals of Cardiac Anaesthesia*, *22*(1), 67–72. https://doi.org/10.4103/aca.ACA_157_18.

Nofal, Z. M., Afify, A. Z., Yousof, H. M., & Cordeiro, G. M. (2017). The generalized transmuted-G family of distributions. *Communications in Statistics – Theory and Methods*, *46*(8), 4119–4136. https://doi.org/10.1080/03610926.2015.1078478.

Nosakhare, U. H., & Bright, A. F. (2017). Evaluation of Techniques for Univariate Normality Test Using Monte Carlo Simulation. *American Journal of Theoretical and Applied Statistics*, *6*(5–1), 51–61. https://doi.org/10.11648/j.ajtas.s.2017060501.18.

Noughabi, H. A., & Arghami, N. R. (2011). Monte Carlo comparison of seven normality tests. *Journal of Statistical Computation and Simulation*, *81*(8), 965–972. https://doi.org /10.1080/00949650903580047.

Pettitt, A. N., & Stephens, M. A. (1976). Modified Cramér-von Mises statistics for censored data. *Biometrika*, *63*(2), 291–298. https://doi.org/10.1093/biomet/63.2.291.

Razali, N. M., & Wah, Y. B. (2011). Power comparisons of Shapiro-Wilk, Kolmogorov-Smirnov, Lilliefors and Anderson-Darling tests. *Journal of Statistical Modeling and Analytics*, *2*(1), 21–33.

Romão, X., Delgado, R., & Costa, A. (2010). An empirical power comparison of univariate goodness-of-fit tests for normality. *Journal of Statistical Computation and Simulation*, *80*(5), 545–591. https://doi.org/10.1080/00949650902740824.

Scott, W. F. (2000). Tables of the Cramér-von Mises distributions. *Communications in Statistics – Theory and Methods*, *29*(1), 227–235. https://doi.org/10.1080/03610920008832479.

Scott, W. F., & Stewart, B. (2011). Tables for the Lilliefors and Modified Cramer-von Mises Tests of Normality. *Communications in Statistics – Theory and Methods*, *40*(4), 726–730. https://doi.org/10.1080/03610920903453467.

Shapiro, S. S., & Wilk, M. B. (1965). An analysis of variance test for normality (complete samples). *Biometrika*, *52*(3–4), 591–611. https://doi.org/10.1093/biomet/52.3-4.591.

Smirnov, N. (1948). Table for estimating the goodness of fit of empirical distributions. *Annals of Mathematical Statistics*, *19*(2), 279–281. https://doi.org/10.1214/aoms/1177730256.

Stephens, M. A. (1974). EDF Statistics for Goodness of Fit and Some Comparisons. *Journal of the American Statistical Association, 69*(347), 730–737. https://doi.org/10.1080/01621459.1974.10480196.

Sulewski, P. (2019). Modification of Anderson-Darling goodness-of-fit test for normality. *Afinidad. Journal of Chemical Engineering Theoretical and Applied Chemistry*, *76*(588), 270–277. https://www.raco.cat/index.php/afinidad/article/view/361876.

Sulewski, P. (2021). Two component modified Lilliefors test for normality. *Equilibrium. Quarterly Journal of Economics and Economic Policy*, *16*(2), 429–455. https://doi.org/10.24136/eq.2021.016.

Sulewski, P. (2022). Modified Lilliefors goodness-of-fit test for normality. *Communications in Statistics – Simulation and Computation*, *51*(3), 1199–1219. https://doi.org/10.1080/03610918.2019.1664580.

Tavakoli, M., Arghami, N., & Abbasnejad, M. (2019). A Goodness of Fit Test For Normality Based on Balakrishnan-Sanghvi Information. *Journal of The Iranian Statistical Society*, *18*(1), 177–190. https://doi.org/10.29252/jirss.18.1.177.

Terán-García, E., & Pérez-Fernández, R. (2024). A robust alternative to the Lilliefors test of normality. *Journal of Statistical Computation and Simulation*, *94*(7), 1494–1512. https://doi.org/10.1080/00949655.2023.2291138.

Torabi, H., Montazeri, N. H., & Grané, A. (2016). A test for normality based on the empirical distribution function. *SORT*, *40*(1), 55–88. https://doi.org/10.2436/20.8080.02.35.

Uhm, T., & Yi, S. (2021). A comparison of normality testing methods by empirical power and distribution of *P*-values. *Communications in Statistics – Simulation and Computation*, *52*(9), 4445–4458. https://doi.org/10.1080/03610918.2021.1963450.

Uyanto, S. S. (2022). An Extensive Comparisons of 50 Univariate Goodness-of-fit Tests for Normality. *Austrian Journal of Statistics*, *51*(3), 45–97. https://doi.org/10.17713/ajs.v51i3.1279.

Watson, G. S. (1962). Goodness-of-fit tests on a circle. II. *Biometrika*, *49*(1–2), 57–63. https://doi.org/10.1093/biomet/49.1-2.57.

Wijekularathna, D. K., Manage, A. B. W., & Scariano, S. M. (2020). Power analysis of several normality tests: A Monte Carlo simulation study. *Communications in Statistics – Simulation and Computation*, *51*(3), 757–773. https://doi.org/10.1080/03610918.2019.1658780.

Yap, B. W., & Sim, C. H. (2011). Comparisons of various types of normality tests. *Journal of Statistical Computation and Simulation*, *81*(12), 2141–2155. https://doi.org/10.1080/00949655.2010.520163.

Yazici, B., & Yolacan, S. A. (2007). A comparison of various tests of normality. *Journal of Statistical Computation and Simulation*, *77*(2), 175–183. https://doi.org/10.1080/10629360600678310.

Przegląd Statystyczny. Statistical Review, 2023, vol. 70, 3, 37–72 https://doi.org/10.59139/ps.2023.04.2
Received: 29.06.2024, revised: 2.09.2024, accepted: 7.10.2024

Reduce extreme losses and retain extreme profits through hedging with gold and cryptocurrencies: A global stock market perspective

Krzysztof Echaust,[a] Małgorzata Just[b]

Abstract. The study focuses on the safe-haven and hedging properties of gold and selected cryptocurrencies against stock markets' extreme risk observed during the COVID-19 pandemic and the Russian invasion of Ukraine. The loss reduction is compared with the profit sacrifice obtained through hedging in terms of the tail thickness of the return distribution. The findings show that gold is able to reduce extreme losses more intensively than extreme profits. Tether reduces volatility and tail risk the most effectively but it is characterised by the worst profit/risk ratio. Bitcoin and Ether increase investment risk; thus, they fail to act as an effective hedge or a safe haven. On the other hand, these cryptocurrencies added to the stock portfolio increase the probability of extreme profits more than extreme losses. The paper provides new insights into the benefits of safe-haven or hedging strategies.

Keywords: gold, cryptocurrencies, conditional value at risk, distribution tail, hedging, safe haven

JEL: C13, C58, G11, G15

1. Introduction

In December 2019, a new virus called SARS-CoV-2 had started to spread rapidly all over the globe, causing the COVID-19 disease and about seven million deaths. A series of unprecedented government interventions to control the infection brought the economy to a standstill. Financial markets also replied with crashes to an extent that had not been observed since the global financial crisis of 2008. In February 2022, the Russian invasion of Ukraine threw global financial and commodity markets into further turmoil. At times like these, investors avoid risky stocks searching for safe-haven assets which, when added to their investments, protect the portfolio from enormous losses. The term ‘safe haven’, introduced by Baur and Lucey (2010), delineates an instrument that exhibits either uncorrelated or negatively correlated behaviour with the held assets during a market crash. If it is uncorrelated or negatively correlated with the held assets on average, it is called

[a] Poznań University of Economics and Business, Department of Operations Research and Mathematical Economics, al. Niepodległości 10, 61-875 Poznań, Poland, e-mail: krzysztof.echaust@ue.poznan.pl, ORCID: https://orcid.org/0000-0002-3855-256X.
[b] Poznań University of Life Sciences, Department of Finance and Accounting, ul. Wojska Polskiego 28, 60-637 Poznań, Poland, e-mail: malgorzata.just@up.poznan.pl, ORCID: https://orcid.org/0000-0001-7655-6046.

a hedge. In this paper, the term ‘hedging’ is employed in a broad sense, encompassing events occurring in extreme market conditions. Finally, if an instrument is positively (but not perfectly) correlated with the held assets on average, it is called a diversifier.

Gold is most commonly considered as a safe-haven asset for stock markets (Baur & McDermott, 2010; Beckmann et al., 2015; Boubaker et al., 2020; Gürgün & Ünalmiş, 2014) due to its relatively low volatility and low correlation with stock markets. Additionally, the fundamental price of gold depends on the economic state of other markets (Baur & Glover, 2014). However, its ability to become a shelter against stock market risk weakened during the COVID-19 period (Akhtaruzzaman et al., 2021; Al-Nassar et al., 2023; Chemkha et al., 2021; Echaust & Just, 2022; Hasan et al., 2021; Salisu et al., 2021). At that time, cryptocurrencies have gained enormous popularity as candidates for ‘a port in the storm’ (Corbet et al., 2020). The potential of cryptocurrencies stems from the fact that they are independent from central authorities. Mariana et al. (2021) find that the two largest cryptocurrencies (Bitcoin and Ether) are suitable as short-term safe-haven assets as their daily returns tended to be negatively correlated with S&P 500 returns during the pandemic. Bouri et al. (2020) indicate cryptocurrencies as safe-haven assets for the aggerate US equity index and selected sectors, whereas Będowska-Sójka and Kliber (2021) as weak safe havens for selected equity indices. On the other hand, the opposite results were presented by Conlon et al. (2020), Conlon and McGee (2020) and Long et al. (2021). They found that Bitcoin and Ether failed in the role of a safe-haven asset. Baur et al. (2022) show that for extreme levels of volatility, Bitcoin does not reduce the risk when added to a benchmark stock portfolio. They consider portfolios consisting of an optimal allocation of Bitcoin (that aims for a minimum variance or a maximum Sharpe ratio) relative to holding the underlying S&P 500. Moreover, Conlon et al. (2020) show evidence of increased downside risk for portfolios consisting of any weight of Bitcoin and Ether relative to the stock index (MSCI World, S&P 500, FTSE 100, FTSE MIB, IBEX). Recently, Just and Echaust (2024) examined the role of five cryptocurrencies as safe havens against the G7 and BRICS stock market risk. The authors find that the conditional probability that Bitcoin can reduce at least 10% of volatility given that index returns fell below the 1st percentile ranges only from 2% to 28% for various stock indices. Moreover, the probabilities calculated for other cryptocurrencies are lower. Xu and Kinkyo (2023) suggest that investing in G7 stocks and Bitcoin in the short term as well as investing in stocks and gold in the long term are reasonable investments for investors. Gold provides higher hedging effectiveness and downside risk reduction than Bitcoin in the long term. The third cryptocurrency considered in our empirical study whose safe-haven properties are widely discussed is Tether. It belongs to a category of cryptocurrencies called stablecoins which aim to keep its valuation stable. The low volatility of Tether is perceived as a desired property of safe havens. Cheema et al. (2022) show that during a pandemic, investors should look for liquid and stable assets rather than gold. Meanwhile, Tether could act as a safe-haven investment for global stock markets (Conlon et al., 2020; Kliber, 2022).

Although cryptocurrencies may potentially yield high returns, which is encouraging for investors, they also entail high volatility and downside risk compared to gold and other conventional asset classes (Iqbal et al., 2023). Indeed, investors with a limited risk-aversion do not restrict their perception to risk reduction and they also consider profits for investment in their hedging decisions. Given the limited possibility of Bitcoin and Ether to reduce risk, their ability to generate profits in extreme market conditions may be the main argument for their application in hedging strategies. The implementation of any such strategy, which is essentially intended to protect the investment from losses, is bound to have an unfavourable impact on the profit potential. The prospect theory of Kahneman and Tversky (1979) indicates that people are more sensitive to losses than gains. Loss aversion means that investors perceive more disutility from losses than utility from gains of equal size. The preferable hedge strategy should offer asymmetry between risk and profit reductions. Based on stocks traded in China, Japan, Korea and Taiwan, Eom et al. (2021) investigate a trade-off relationship between loss avoidance and a profit sacrifice through a portfolio diversification strategy. According to their results, investors reduce the likelihood of high losses through portfolio diversification; however, their potential for higher profits is thus sacrificed. We adopt this concept to verify the relative benefits from a hedging strategy against stock market tail risk using gold and cryptocurrencies. We study the unexplored relationship between high loss avoidance and high profit sacrifice relative to the hedging strategy.

The first aim of this research is to compare the ability of gold and cryptocurrencies (Bitcoin, Ether and Tether) to act as safe-haven assets against global stock market risk in the context of the COVID-19 pandemic and the Russia–Ukraine war. The second and most important aim is to verify the asymmetry between risk reduction and profit sacrifice. The third aim is to compare two methods of the asymmetry analysis. The approach based on tail thickness is compared to the conditional value at risk (CVaR). To achieve these objectives, we first follow Kroner and Ng's (1998) optimisation procedure based on Engle's (2002) dynamic conditional correlation (DCC) model. We thus estimate the optimal weights of a hedging portfolio. We then calculate the reduction in volatility offered by pair-wise portfolios and estimate the asymmetry between downside and upside risk. To obtain a comprehensive view of the extrema, we focus on the tails of return distribution which represent extreme losses and extreme profits on investment. We propose to use the peaks over threshold (POT) method to compare the tail behaviour of hedged and unhedged trading positions as well as the upper (profits) and lower (losses) tails of the return distribution. Subsequently, we compare the tail behaviour approach with the estimates of the CVaR, which is recognised as a credible tool for the assessment of the diversification benefits (Conlon et al., 2020; Conlon & McGee, 2020).

Our main findings can be summarised as follows. Firstly, in line with the literature, gold and Tether are found to have been effective hedges in the research period, whereas Bitcoin and Ether increased the investment risk, thus failing to act as an effective shelter against stock market risk. Secondly, and most importantly, gold is the only asset which is able to reduce extreme losses more intensively than extreme profits for chosen indices. Bitcoin and Ether, added to the stock portfolio, increase the probability of extreme profits more than extreme losses. The return distributions of the hedged portfolios consisting of Bitcoin or Ether indicate much fatter upper tails relative to the lower tails. Therefore, cryptocurrencies provide investors with a valuable profit–loss relationship in terms of extreme values. Thirdly, we found that an inference based on CVaR may lead investors to misleading conclusions and decisions. The investigation of the tail behaviour outperforms the approach based on CVaR offering a broader and more reliable view of extreme losses in relation to extreme profits in a trading strategy.

The remainder of the paper is organised as follows. The next section provides data and an in-depth description of the methodological approach adopted in the empirical part of the paper. Section 3 shows detailed results relating to the research objectives. The final section summarises the main findings and presents the conclusions.

2. Data and methodology

2.1. Data

We analyse the log-returns of global stock indices, gold and three cryptocurrencies. Our sample includes stock indices from the world's largest exchanges by market capitalisation (World Federation of Exchanges, 2023). We selected two indices from each of the three regions: the Americas (S&P 500, SPX – United States; S&P/TSX Composite Index, TSX – Canada), the Asia-Pacific region (Shanghai Composite Index, SHC – China; Nikkei 225, NKX – Japan), and Europe, the Middle East and Africa (CAC 40, CAC – France; FTSE 250, FTM – Great Britain). We focus on the leading stock markets to represent viable investors' interests. Our sample includes three cryptocurrencies with the largest market capitalisation: two non-stable (Bitcoin, BTC-USD; Ether, ETH-USD) and one stable (Tether, USDT-USD). These cryptocurrencies are most often considered as hedging assets or safe havens. The analysis is based on daily data for the period from 2nd January 2019 to 14th August 2023. Gold is quoted on the London Metal Exchange and its prices are sourced from kitco.com. The cryptocurrency prices and the stock indices are obtained from finance.yahoo.com and stooq.pl, respectively.

The descriptive statistics in Table 1 show that Ether has the highest mean return. On the other hand, Tether is the only asset which has a negative mean/median return. Tether and gold have the lowest volatility among the considered assets, whereas the volatility of Bitcoin and Ether exceeds the volatility level of the indices several times over. Gold, Bitcoin and Ether are positively correlated with indices, while Tether negatively (Table 2). Tether appears to have the desired characteristics of a hedging instrument with low volatility and a negative correlation. However, at the same time, it shows a negative mean return, much lower than the other candidates for hedging instruments. It can be a good hedge for risk-averse investors but its usefulness seems to be undermined when return on investment gains importance.

Table 1. Descriptive statistics of asset returns

Asset	Min	Median	Mean	Max	SD
SPX	-12.77	0.0954	0.0511	8.97	1.40
TSX	-13.18	0.1001	0.0305	11.29	1.19
SHC	-8.04	0.0302	0.0233	5.55	1.09
NKX	-6.27	0.0829	0.0453	7.73	1.24
CAC	-13.10	0.1062	0.0387	8.06	1.35
FTM	-9.82	0.0440	0.0056	8.04	1.24
Gold	-5.26	0.0531	0.0339	5.13	0.96
USDT	-5.26	-0.0010	-0.0021	5.34	0.35
BTC	-46.47	0.1048	0.1726	20.30	4.42
ETH	-55.07	0.1205	0.2127	34.35	5.65

Source: authors' work.

Table 2. Correlation between asset returns

Asset/ Index	SPX	TSX	SHC	NKX	CAC	FTM
Gold	0.1420***	0.2179***	0.1097***	0.0309	0.0826***	0.1073***
USDT	-0.2034***	-0.2428***	-0.0051	-0.0067	-0.1498***	-0.0839***
BTC	0.3242***	0.3492***	0.0459	0.0709**	0.2516***	0.2205***
ETH	0.3459***	0.3630***	0.0959***	0.0980***	0.2680***	0.2401***

Note. Correlation means the Pearson correlation between indices and assets; *** and ** indicate significance at the level of 1% and the level of 5%, respectively.

Source: authors' work.

2.2. Methodology

Hedging stocks implies a combined position consisting of stocks and a hedging instrument. In our setting, investors hold stocks and wish to hedge the stock market risk by adding a hedge instrument in a long position. We can model the hedge portfolio as:

$$r_{p,t} = (1 - w_{i,t})\, r_{Index,t} + w_{i,t} r_{i,t}, \tag{1}$$

where $r_{Index,t}$ is the return of the stock market at time t, and $r_{i,t}$ is the return of the i-th hedge price at time t, $w_{i,t}$ represents the time-varying weights of the i-th hedging instrument.

If investors add gold or cryptocurrencies to their stock portfolios with the aim to reduce risk, the optimal weights of the individual assets ($w_{i,t}, i$ = GOLD, USDT, BTC, ETH) to obtain minimum risk portfolios are calculated from the following formula (Kroner & Ng, 1998):

$$w_{i,t} = \frac{h_{Index,t} - h_{Index/i,t}}{h_{Index,t} - 2h_{Index/i,t} + h_{i,t}}, \tag{2}$$

where $h_{Index,t}$ and $h_{i,t}$ are the conditional variances of the index returns and safe-haven candidate returns, respectively. $h_{Index/i,t}$ is the covariance between index and gold, Bitcoin, Ether and Tether returns on the t-th day. We use the DCC model[1] (Engle, 2002) to compute the conditional variance and covariance. Moreover, we assume no short selling of the assets.

We employ the GARCH(1,1) model[2] (Bollerslev, 1986) to obtain conditional volatility ($h_{i,t}$) of asset returns ($r_{i,t}$):

$$r_{i,t} = \mu_i + e_{i,t}, e_{i,t} = \sqrt{h_{i,t}}\varepsilon_{i,t}, \varepsilon_{i,t} \sim \mathrm{N}(0,1), \tag{3}$$

$$h_{i,t} = \omega_i + \alpha_i e_{i,t-1}^2 + \beta_i h_{i,t-1}, \tag{4}$$

where $\omega_i, \alpha_i, \beta_i > 0, \alpha_i + \beta_i < 1$.

[1] Akhtaruzzaman et al. (2021) used four models: DCC, asymmetric DCC, corrected DCC and DCC–DECO to describe the relationship between gold and major stock indices during the COVID-19 pandemic, and obtained consistent results. Therefore, we use the simple DCC model.

[2] Although the GARCH(1,1) model is relatively simple, it provides relatively good estimates and predictions of volatility compared to much more complex models (see e.g. Hansen & Lunde, 2005).

Then, we estimate the bivariate DCC model parameters. Let us denote the two-dimensional vector by $\boldsymbol{e}_t = \left(e_{1,t}, e_{2,t}\right)'$. The DCC model assumes that (Engle, 2002):

$$\boldsymbol{e}_t | I_{t-1} \sim \mathrm{N}(\boldsymbol{0}, \boldsymbol{H}_t), \boldsymbol{H}_t = \boldsymbol{D}_t \boldsymbol{R}_t \boldsymbol{D}_t, \tag{5}$$

where I_{t-1} is the information set available at time t-1, $\boldsymbol{D}_t = \mathrm{diag}\left(\sqrt{h_{1,t}}, \sqrt{h_{2,t}}\right)$ and conditional variance $h_{i,t}$ is modelled using the GARCH model. In turn, conditional correlation matrix $\boldsymbol{R}_t$ is given by

$$\boldsymbol{R}_t = \left(\mathrm{diag}(\boldsymbol{Q}_t)\right)^{-1/2} \boldsymbol{Q}_t \left(\mathrm{diag}(\boldsymbol{Q}_t)\right)^{-1/2} \tag{6}$$

with

$$\boldsymbol{Q}_t = (1 - a - b)\bar{\boldsymbol{Q}} + a\boldsymbol{\varepsilon}_{t-1}\boldsymbol{\varepsilon}'_{t-1} + b\boldsymbol{Q}_{t-1}, \tag{7}$$

where $\bar{\boldsymbol{Q}}$ is the unconditional covariance matrix of $\boldsymbol{\varepsilon}_t$ $\left(\varepsilon_{i,t} = e_{i,t}/\sqrt{h_{i,t}}\right)$; a, b are parameters such that $a, b \geq 0$ and $a + b < 1$.

In the second stage of the analysis, we investigate how adding gold or cryptocurrencies to a portfolio can reduce the risk of a stock portfolio. We examine the conditional variance of portfolio returns by applying the following formula:

$$h_{Portfolio_i,t} = (1 - w_{i,t})^2 h_{Index,t} + w_{i,t}{}^2 h_{i,t} + 2(1 - w_{i,t}) w_{i,t} h_{Index/i,t}. \tag{8}$$

In the third stage, we focus on extreme returns, i.e. using an extreme-value-theory-based method, we compare the distribution tails and the values of the tail risk measure of stock indices and two-component portfolios consisting of these indices and hedging assets. We need to adequately fit the tails of the return distributions to compare the relationship between high loss avoidance and high profit sacrifice. They can be easily modelled using the peaks over threshold method. This method is an approach of the extreme value theory (EVT) that allows modelling all observations in a sample that exceed a high threshold using the generalised Pareto (GP) distribution.

Let R be a random variable of returns with unknown cumulative distribution function (cdf) F, and excess distribution function F_u over high threshold u is defined by

$$F_u(y) = P(R - u \leq y | R > u) = \frac{F(y+u) - F(u)}{1 - F(u)} \text{ for } 0 \leq y \leq r_0 - u, \tag{9}$$

where $r_0 \leq \infty$ is the right endpoint of F.

According to the Pickands-Balkema-de Haan Theorem (Balkema & de Haan, 1974) for a large class of underlying distribution functions F and high enough threshold u, a function $\beta(u)$ exists so that:

$$\lim_{u \to r_0} \sup_{0 \leq y \leq r_0 - u} \left| F_u(y) - G_{\xi,\beta(u)}(y) \right| = 0. \tag{10}$$

As an after-effect of the theorem, F_u can be approximated by a GP distribution, which is defined as:

$$G_{\xi,\beta}(y) = \begin{cases} 1 - \left(1 + \xi \dfrac{y}{\beta}\right)^{-\frac{1}{\xi}} & \text{for } \xi \neq 0 \\ 1 - \exp\left(-\dfrac{y}{\beta}\right) & \text{for } \xi = 0, \end{cases} \tag{11}$$

where $\beta > 0$, $y \geq 0$ for $\xi \geq 0$ and $0 \leq y \leq -\beta/\xi$ for $\xi < 0$. Parameters of the GP distribution are scale parameter β and shape parameter ξ.

An approximation of cdf for returns exceeding a sufficiently high threshold can be obtained by transforming (9) and (11):

$$F(r) = G_{\xi,\beta}(r - u)(1 - F(u)) + F(u) \text{ for } r > u. \tag{12}$$

In order to obtain a useful closed form of distribution (12), $F(u)$ can be simply replaced with the empirical estimator of exceedance over threshold u. The estimator is given by $(n - k_u)/n$, where n represents the total number of observations (returns), and k_u is the number of observations exceeding threshold u.

The log-likelihood method is used to estimate the parameters of the GP distribution. The estimator of cumulative distribution F is then given as:

$$\hat{F}(r) = \begin{cases} 1 - \dfrac{k_u}{n}\left(1 + \hat{\xi}\dfrac{r - u}{\hat{\beta}}\right)^{-\frac{1}{\hat{\xi}}} & \text{for } \xi \neq 0 \\ 1 - \dfrac{k_u}{n}\exp\left(-\dfrac{r - u}{\hat{\beta}}\right) & \text{for } \xi = 0. \end{cases} \tag{13}$$

Value at risk (VaR) and conditional value at risk (CVaR) are the most commonly used measures of tail risk. These measures differ in terms of their mathematical properties, stability of statistical estimation, simplicity of optimisation procedures and acceptance by regulators (Sarykalin et al., 2008), which determine the choice of their application. We use CVaR to measure the tail risk for long and short positions at a 95% confidence level as it provides an adequate picture of the risks reflected in the extreme tails (Sarykalin et al., 2008). Since CVaR is defined in terms of VaR, we begin by presenting VaR, which can be seen as a quantile of F. Therefore, the q-quantile of the GP distribution for a sample size of length n is calculated as:

$$\widehat{VaR}_q = \begin{cases} u + \frac{\hat{\beta}}{\hat{\xi}}\left(\left(\frac{n}{k_u}(1-q)\right)^{-\hat{\xi}} - 1\right) \text{ for } \xi \neq 0 \\ u + \hat{\beta}\ln\left(\frac{n}{k_u}(1-q)\right) \text{ for } \xi = 0. \end{cases} \tag{14}$$

CVaR provides the expected size of return that exceeds VaR:

$$CVaR_q = \mathrm{E}\left(R \middle| R \geq VaR_q\right). \tag{15}$$

Hence, CVaR is given as (Dowd, 2005):

$$\widehat{CVaR}_q = \frac{\widehat{VaR}_q}{1-\hat{\xi}} - \frac{\hat{\beta}-\hat{\xi}u}{1-\hat{\xi}} \text{ for } \xi \neq 0 \tag{16}$$

and

$$\widehat{CVaR}_q = \widehat{VaR}_q + \hat{\beta} \text{ for } \xi = 0. \tag{17}$$

3. Empirical results

3.1. Optimal hedging

Figure 1 displays the time-varying weights for all the portfolios considered in this study. The optimal weight for gold behaves in a different way than for cryptocurrencies. It rose significantly in the first quarter of 2020, when global financial markets suffered high losses and then displayed a significant downward trend. The same phenomenon occurred during the Russian invasion of Ukraine in February 2022. The result indicates that in the first phase of the COVID-19 pandemic and later, at the beginning of the war in Europe, investors should have held more gold to reduce risk, thereby the cost of an optimal hedging strategy was relatively high during those periods. Optimal weights for Ether and Bitcoin are small and often equal to zero. Both cryptocurrencies are not able to reduce risk effectively when added to an equity portfolio. Such a result confirms the findings of Baur et al. (2022). The authors prove that the benefits of Bitcoin in the portfolio come from the expected returns and can enhance the risk-return relationship but do not substantially lower the risk. Tether has different characteristics compared to the foregoing assets. It negatively correlates to most indices and demonstrates low volatility. The effect is that the optimal weight in the portfolio is close to one for most of the time from the second half of 2020. Such a result shows that investors should sell off equity portfolios and replace them with Tether. However, the benefit of this strategy is questionable since it generates a high cost of hedging and entirely changes the investment profile. Additionally, we must not overlook the importance of Tether's lowest median returns compared to other assets.

Figure 1. Optimal weights for hedging instruments

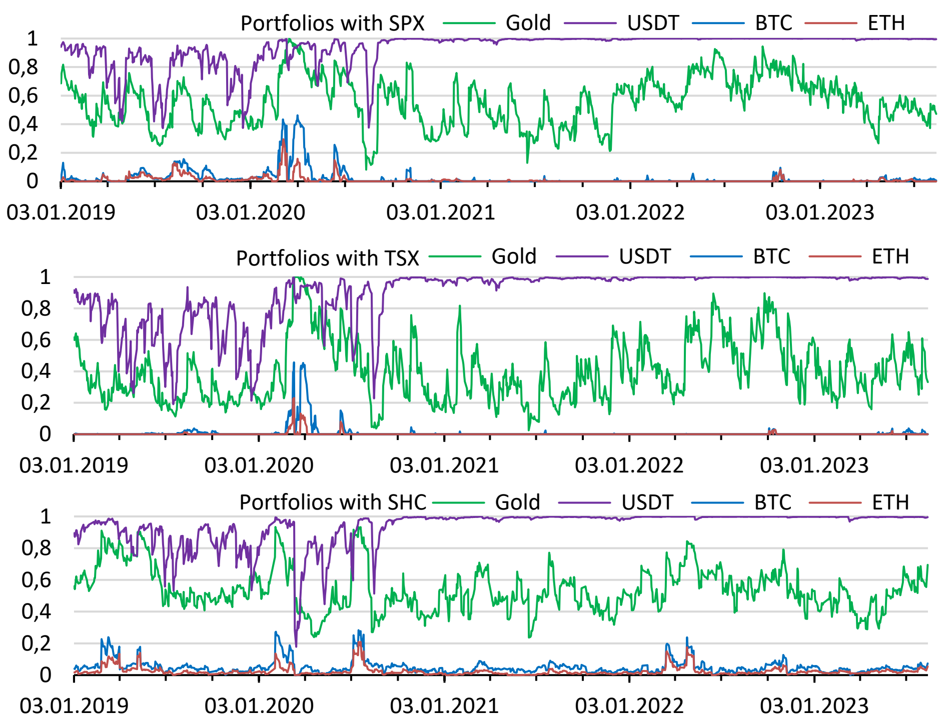
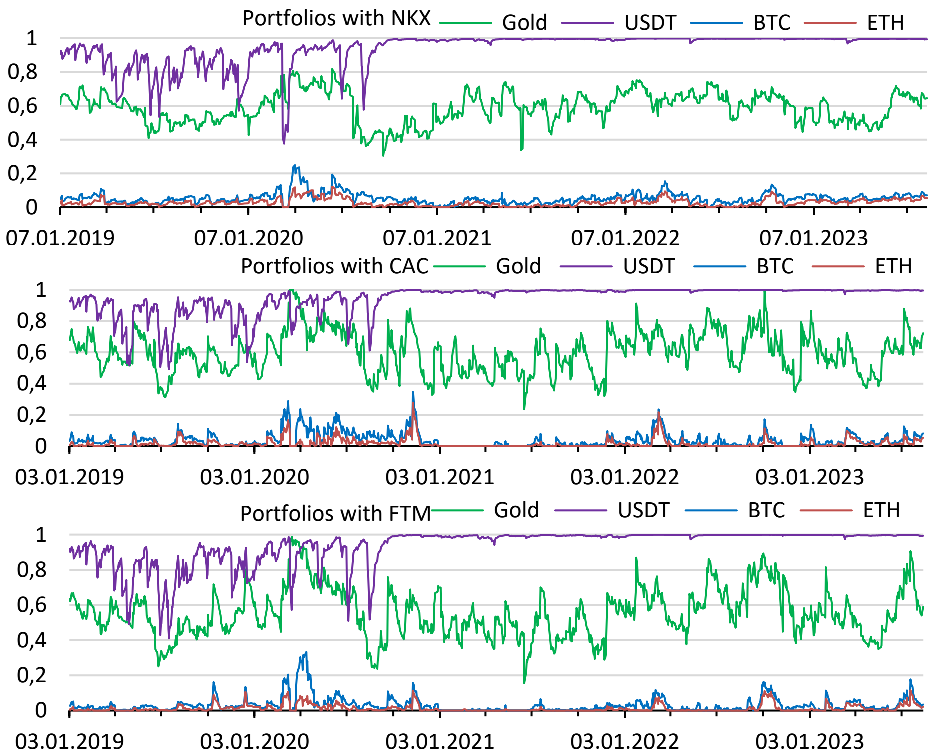


Figure 1. Optimal weights for hedging instruments (cont.)



Source: authors' work.

Table 3 shows the descriptive statistics for optimal portfolios. Adding Bitcoin or Ether to the base portfolio significantly increases downside risk expressed in the minimum return. Gold in the portfolio substantially lowers volatility and range. Note that the portfolio that consists of gold is the only one which can improve the ratio of index performance to volatility (mean/SD) for all the considered assets. On the other hand, Tether is the only hedge instrument which destroys the ratio for all assets. The low risk of the portfolio corresponds to a negative median return and finally makes the mean/risk ratio unattractive. Statistics for the portfolio with Tether are close to those for Tether itself since it almost replaces equities in the optimal hedge strategy.

Table 3. Descriptive statistics of optimal portfolio returns

Portfolio	Min	Median	Mean	Max	SD	Mean/SD
SPX	-12.77	0.0954	0.0511	8.97	1.40	0.0366
SPX + Gold	-5.51	0.0846	0.0469	5.34	0.77	0.0612
SPX + USDT	-3.04	-0.0006	0.0050	5.28	0.32	0.0157
SPX + BTC	-24.24	0.1054	0.0421	8.88	1.52	0.0277
SPX + ETH	-18.82	0.1015	0.0455	8.96	1.48	0.0307

Table 3. Descriptive statistics of optimal portfolio returns (cont.)

Portfolio	Min	Median	Mean	Max	SD	Mean/SD
TSX	-13.18	0.1001	0.0305	11.29	1.19	0.0257
TSX + Gold	-5.41	0.0567	0.0352	5.30	0.70	0.0501
TSX + USDT	-3.52	-0.0008	0.0044	5.26	0.31	0.0140
TSX + BTC	-28.40	0.0961	0.0190	10.36	1.38	0.0138
TSX + ETH	-22.86	0.1001	0.0214	11.29	1.31	0.0164
SHC	-8.04	0.0302	0.0233	5.55	1.09	0.0213
SHC + Gold	-4.17	0.0420	0.0218	3.67	0.74	0.0293
SHC + USDT	-2.89	-0.0002	0.0035	5.09	0.33	0.0106
SHC + BTC	-6.49	0.0224	0.0328	5.30	1.08	0.0303
SHC + ETH	-7.05	0.0233	0.0342	5.64	1.09	0.0313
NKX	-6.27	0.0829	0.0453	7.73	1.24	0.0364
NKX + Gold	-5.01	0.0213	0.0407	5.87	0.78	0.0522
NKX + USDT	-5.85	-0.0013	0.0030	5.02	0.36	0.0084
NKX + BTC	-7.77	0.0740	0.0499	7.48	1.24	0.0404
NKX + ETH	-6.64	0.0644	0.0514	7.58	1.24	0.0413
CAC	-13.10	0.1062	0.0387	8.06	1.35	0.0286
CAC + Gold	-5.98	0.0402	0.0321	5.31	0.77	0.0418
CAC + USDT	-4.12	-0.0005	0.0013	5.11	0.33	0.0038
CAC + BTC	-19.80	0.1093	0.0452	8.06	1.42	0.0319
CAC + ETH	-17.71	0.1148	0.0404	8.06	1.40	0.0288
FTM	-9.82	0.0440	0.0056	8.04	1.24	0.0045
FTM + Gold	-5.94	0.0277	0.0221	5.30	0.75	0.0296
FTM + USDT	-4.12	-0.0009	-0.0014	5.06	0.35	-0.0040
FTM + BTC	-14.41	0.0454	0.0073	7.87	1.27	0.0057
FTM + ETH	-12.42	0.0415	0.0064	8.04	1.26	0.0050

Source: authors' work.

Figure 2. Reduction in conditional volatility (in pp.) by optimal portfolios relative to indices

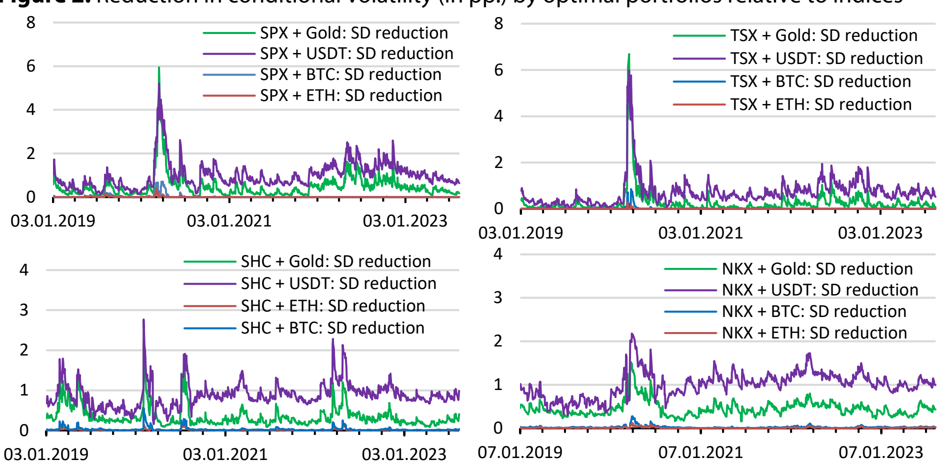
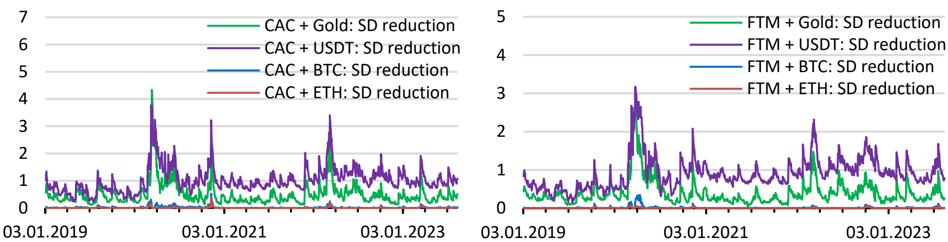


Figure 2. Reduction in conditional volatility (in pp.) by optimal portfolios relative to indices (cont.)



Source: authors' work.

Figure 2 demonstrates how much the conditional volatility of an equity portfolio is reduced by applying an optimal hedging strategy. Gold and Tether can significantly decrease volatility (the Wilcoxon signed rank test for median equality with significance level of 1%), especially in times when the financial market collapsed. Bitcoin and Ether can decrease the risk only to a small extent. Even in periods of a volatility explosion, in March 2020, when optimal weights were relatively high, both cryptocurrencies were not able to significantly reduce the risk. This result proves that Bitcoin and Ether cannot be considered as safe-haven assets.

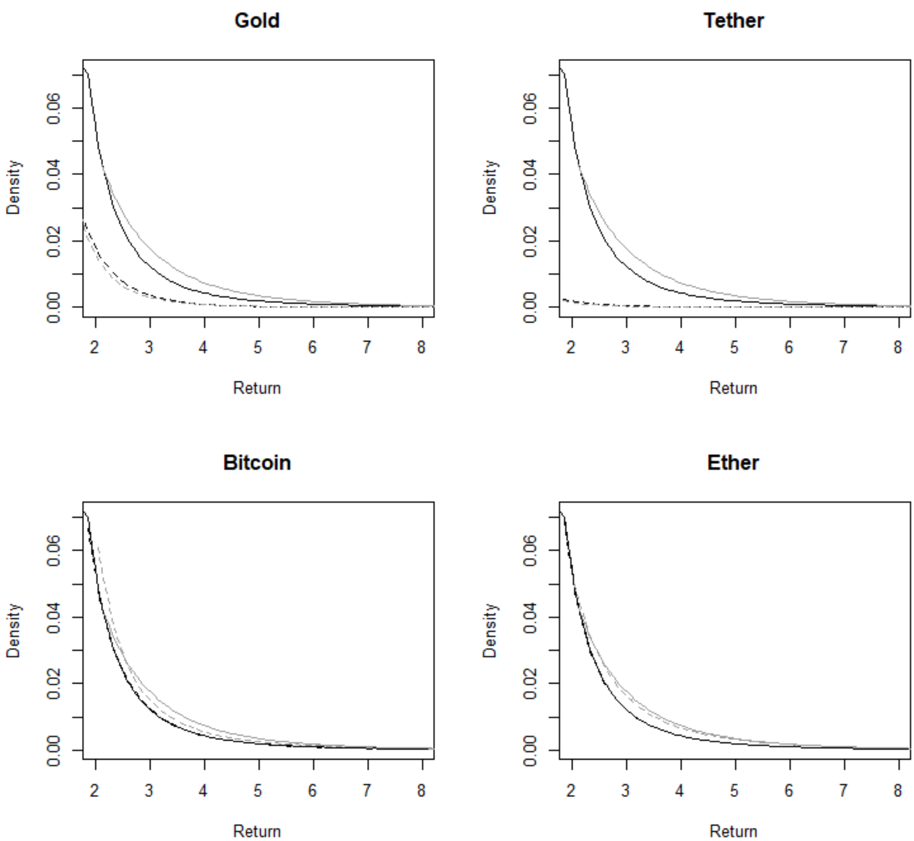
3.2. Extreme losses versus extreme profits in optimal hedging

Risk-averse investors have no motivation to apply Bitcoin or Ether in the hedging role. This result confirms the findings of Echaust et al. (2024) who compared the hedging effectiveness of cryptocurrencies in a short hedge strategy with favoured index futures contracts. However, the ability of cryptocurrencies to generate abnormal profits may be the key argument to consider in hedging decisions. Their independence from financial market fundamentals might on the one hand provide diversification benefits and generate high returns during financial crashes on the other. The aim is to indicate to what extent hedging strategies reduce extreme risk compared to extreme profits. We propose the comparison of the tail behaviour of the considered assets and portfolios, and compare the findings with the results based on the CVaR. The tails of return distributions are estimated using the peaks over threshold method. The first task involves choosing the appropriate tail threshold which separates the extrema from the middle part of the return distribution. The appropriate selection of a threshold level is considered to be a complex and challenging task. While there are many concepts and approaches presented in the literature, none of them have been indicated to be suitable, and there is no single answer as to where the distribution tail begins. Searching for the tail of the distribution is always a trade-off between bias and variance. If the chosen threshold is too low, the tail estimates indicate a high bias. The more the threshold is away from the tail, the more the empirical distribution of extrema deviates from the theoretical model. On the other hand, a too high threshold results in high variance of the model estimates since not much data exceeds the threshold. Numerous authors applied a fixed percentile of the total sample size as the threshold, usually 10%, 5% or 1% of the upper statistics (Bee et al., 2016; Echaust, 2021; Fernandez, 2005; Gençay et al., 2003; Longin, 2000; McNeil & Frey, 2000; Totić & Božović, 2016). More sophisticated approaches use a threshold selection based on graphical techniques based on a mean excess plot (Aboura, 2014; Cifter, 2011; Gilli & Këllezi, 2006; Łuczak & Just, 2020) or the graphical representation of the Hill estimator (Hill, 1975). However, the choice procedures of the graphical-based threshold require the identification of the stable regions in the graphs; therefore, they are highly subjective and difficult to apply in empirical studies. Finally, in some studies, the choice of the threshold is based on optimisation procedures. An extensive overview of such methods is provided in Caeiro and Gomes (2016) and Danielsson et al. (2016). A simulation study of Just and Echaust (2021) showed that most of the optimisation algorithms return a too high threshold to calculate tail risk measures according to the requirements of the Basel Committee. We decided to apply the widely accepted in the literature thresholds equal to the 5th percentile for the lower tail and the 95th for the upper tail. Parameters of the general Pareto distribution (Formula 11) for threshold exceedances are computed with the evir R package. For the sake of brevity, we do not present the estimations of the parameters and they are available from the authors upon request. Finally, according to Formula (13), we obtained the estimate of the upper tail of the unconditional distribution for returns. A lower tail can be considered in the same way as the upper tail after the multiplication of returns by minus one.

Figure 3 shows the return distribution tails for the S&P 500 index as well as the tails of an optimal pair-wise portfolio combined with an index and a hedging instrument. The solid black line represents the upper tail of the index returns, the solid grey line represents the lower tail of the index, while the dashed lines show the tails of the portfolio distribution (in black for the upper tail and in grey for the lower tail). Since the tails for different assets represent different data ranges, we decided to show the entire upper tail area of 5% for each index and the other tails against its background. Figures A1–A5 in the Appendix show the same for the other considered stock indices. As expected, most index distributions exhibit a fatter lower tail than the upper one. The highest differences are evident for the S&P 500 and CAC indices. From an intuitive point of view, it seems to be clear that the lower tails for financial markets must be heavier than the upper tails. This is due to the fact that the growth trends are built over long-time horizons but the crashes are more volatile and the price movements in absolute value are much larger. However, from a statistical point of view, the problem is much more complex. In empirical studies, the differences are not significant or distributions exhibit symmetry as the Canadian TSX returns distribution has shown in Figure A1. For instance, Longin (1996) found the equality of tail thickness for the S&P 500 returns. Similarly, Jondeau and Rockinger (2003) did not find statistically significant asymmetry in mature, Asian, Eastern European and Latin American markets. The similarity between the lower and upper tails of returns has been reported in e.g. Chen and Ibragimov (2019), Daníelsson and de Vries (1997) and Koedijk and Kool (1992). Only the minority of studies report heavier lower tails than upper tails, e.g. Gregory-Allen et al. (2012) or Hartmann et al. (2004). Contrary to those studies, we do not compare the estimates of the tail index which is a measure of the tail behaviour. The tail index cannot measure the extreme risk level independently. To properly quantify the probability of extreme events, both threshold and tail behaviour must be taken into account. Graphical analysis enables us to consider the tails in a more complex way than any estimate of tail fatness. Tails for hedge portfolios are presented in Figure 3 and Figures A1–A5 as the dashed lines demonstrate symmetry between lower and upper tails, especially for the highest extrema. The differences between solid and dashed lines represent the effectiveness of the hedge. We can notice significant differences between the tails of the hedged and unhedged trading positions for gold and Tether, which confirms the ability of both assets in the extreme risk reduction. The result supports our findings for volatility presented in Section 3.1. Moreover, Tether reduces extreme risk the most effectively. The highest reduction in the probability of extreme losses compared to the reduction in the probability of extreme profits is noticeable for the S&P 500 and CAC indices. The result is the effect of heavier lower tails of return distribution for these indices. The tails for the portfolios with Bitcoin and Ether do not indicate any differences in relation to the index. Weight close to zero for these hedge instruments makes them have only a minor effect on the extreme returns of the pair-wise portfolios. A more interesting issue is to check their usefulness in generating profits in the suboptimal portfolios described in Section 3.4.

For comparison, we carried out a similar analysis based on the CVaR measure, which takes an average of returns in the tail of distribution. We calculate the CVaR under the generalised Pareto distribution (Formulas 16 and 17) with a threshold equal to the 5th percentile. The results of the computations are presented in Table 4; the second and third column shows the CVaRs for the lower and upper tails, respectively, while the next two columns exhibit the change in CVaR as an effect of the hedging strategy. The inference based on the CVaR is mostly the same as that based on the distribution tails. However, findings regarding the asymmetry between the reduction in losses and profits indicate differences. Based on CVaR, we find that both gold and Tether reduce losses more effectively than profits for all the considered indices, excluding NKX. Both hedging instruments against the TSX index demonstrate the highest asymmetry of reduction in the extreme returns, which was not reflected in tail plots. TSX returns have the highest extrema among the considered returns and have not too many outliers; therefore, the CVaR defined as an expected value of VaR exceedances is highly affected by extreme returns. Bitcoin and Ether in an optimal portfolio only slightly change CVaR and rather increase the risk in the lower tail.

Figure 3. Distribution tails (pdf) of the SPX returns and an optimal portfolio consisting of an index and hedge asset. Black line – 5% upper tail, grey line – lower tail (symmetrical image), solid line – index, dashed line – optimal portfolio



Source: authors' work.

Table 4. Reduction in 5% CVaR (in pp.) by optimal portfolios relative to indices

Portfolio	CVaR Lower tail	CVaR Upper tail	ΔCVaR Lower tail	ΔCVaR Upper tail
SPX	3.47	3.04		
SPX + Gold	1.74	1.76	1.73	1.28
SPX + Tether	0.69	0.76	2.78	2.28
SPX + Bitcoin	3.57	2.91	-0.10	0.13
SPX + Ether	3.60	3.01	-0.14	0.03
TSX	3.09	2.33		
TSX + Gold	1.68	1.57	1.41	0.76
TSX + Tether	0.65	0.73	2.44	1.60
TSX + Bitcoin	3.14	2.27	-0.05	0.06
TSX + Ether	3.29	2.30	-0.20	0.03
SHC	2.62	2.35		
SHC + Gold	1.80	1.61	0.82	0.74
SHC + Tether	0.94	0.79	1.68	1.56
SHC + Bitcoin	2.58	2.40	0.05	-0.05
SHC + Ether	2.61	2.38	0.01	-0.03
NKX	2.77	2.82		
NKX + Gold	1.76	1.81	1.01	1.01
NKX + Tether	0.74	0.83	2.03	2.00
NKX + Bitcoin	2.82	2.74	-0.05	0.08
NKX + Ether	2.81	2.80	-0.04	0.03
CAC	3.41	2.94		
CAC + Gold	1.80	1.71	1.61	1.23
CAC + Tether	0.75	0.78	2.66	2.16
CAC + Bitcoin	3.50	2.92	-0.10	0.02
CAC + Ether	3.50	3.00	-0.09	-0.07
FTM	3.03	2.82		
FTM + Gold	1.71	1.73	1.32	1.09
FTM + Tether	0.78	0.80	2.24	2.03
FTM + Bitcoin	3.07	2.77	-0.05	0.05
FTM + Ether	3.07	2.85	-0.04	-0.03

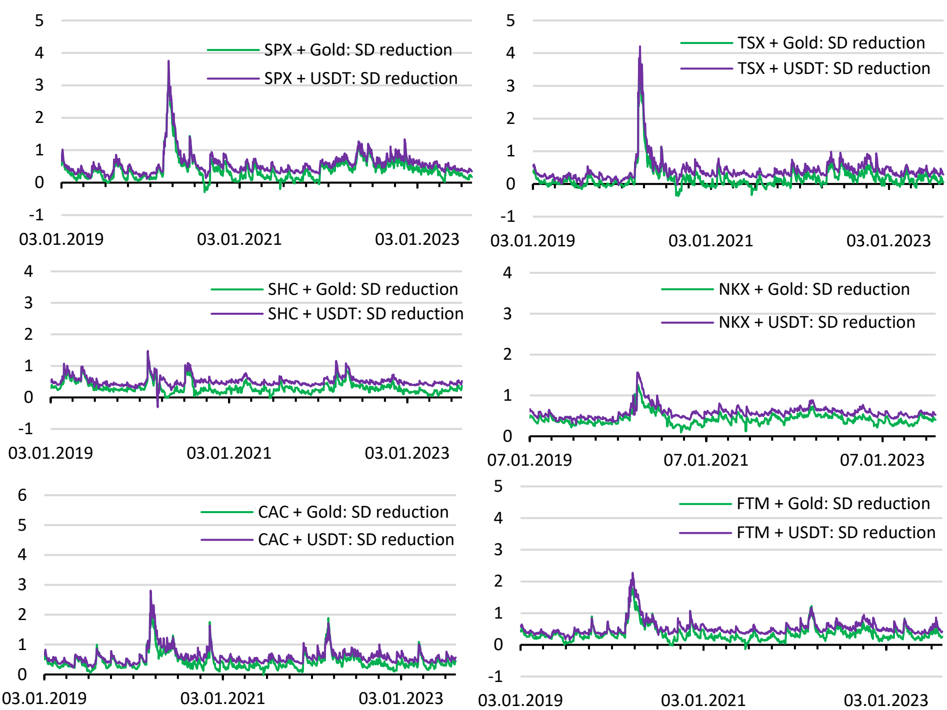
Source: authors' work.

3.3. Hedging for an equal-weighted portfolio

As indicated in Section 3.1, optimal hedging seems to be a reasonable investment strategy only for gold. This section provides a similar analysis for equal-weighted portfolios. There is no consensus in the literature on the proportion of hedge assets that should be included in a strategic portfolio allocation (Akhtaruzzaman et al., 2021; Lucey et al., 2021). An equal-weighted portfolio is only one example portfolio chosen for analysis; however, such a choice enables capturing the important features of the considered instruments. Table 5 shows the descriptive statistics for these portfolios. Statistics for gold do not differ substantially from the statistics presented in Table 3 for the optimal portfolio since the optimal weight oscillates around 50%. Tether is still the most effective hedge and reduces volatility and extrema the most efficiently. Nevertheless, Tether is the only hedge that performs worse than the index in terms of the mean/SD ratio. Bitcoin and Ether added to an index significantly increase the volatility and extrema. The overall and extreme risks of the portfolio far exceed the risks of the base investment. However, the profit-to-risk ratio has improved. Therefore, it seems reasonable to check the relative benefits from the trading strategy.

Figure 4 shows the reduction in conditional volatility of indices hedged with gold and Tether. There is no reason to present the results for Bitcoin and Ether since their application in the portfolio substantially increases the volatility relative to index. The results coincide with the descriptive statistics. Tether significantly outperforms gold in the hedging role since it reduces volatility to a greater extent. We confirmed the finding with the Wilcoxon signed rank test for the medians at the significance level of 1%. The analysis leads us to the conclusion that Tether is a better hedge and safe-haven asset than gold for all stock markets. Risk reduction may be the primary criterion for investment strategy in turbulent times; however, Tether is the worst option in normal market conditions when profits become the goal of an investment.

Figure 4. Reduction in conditional volatility (in pp.) by equal-weighted portfolios relative to indices



Source: authors' work.

Table 5. Descriptive statistics of equal-weighted portfolio returns

Portfolio	Min	Median	Mean	Max	SD	Mean/SD
SPX	-12.77	0.0954	0.0511	8.97	1.40	0.0366
SPX + Gold	-8.84	0.0738	0.0429	7.05	0.90	0.0475
SPX + USDT	-6.51	0.0409	0.0245	4.67	0.68	0.0358
SPX + BTC	-28.23	0.1288	0.1139	10.07	2.54	0.0448
SPX + ETH	-32.53	0.1479	0.1345	16.43	3.17	0.0425
TSX	-13.18	0.1001	0.0305	11.29	1.19	0.0257
TSX + Gold	-9.16	0.0617	0.0326	8.21	0.84	0.0387
TSX + USDT	-5.39	0.0435	0.0142	5.83	0.58	0.0246
SPX + BTC	-29.82	0.1038	0.1037	10.24	2.49	0.0417
SPX + ETH	-34.12	0.1305	0.1242	17.45	3.12	0.0399
SHC	-8.04	0.0302	0.0233	5.55	1.09	0.0213
SHC + Gold	-4.19	0.0550	0.0298	3.72	0.77	0.0385
SHC + USDT	-3.99	0.0188	0.0105	3.05	0.57	0.0183
SPX + BTC	-24.00	0.0490	0.1038	10.78	2.37	0.0437
SPX + ETH	-28.30	0.0574	0.1252	17.60	3.03	0.0413
NKX	-6.27	0.0829	0.0453	7.73	1.24	0.0364
NKX + Gold	-4.83	0.0229	0.0408	6.01	0.81	0.0506
NKX + USDT	-5.77	0.0262	0.0218	3.77	0.65	0.0338
SPX + BTC	-25.49	0.0831	0.1157	10.62	2.43	0.0476
SPX + ETH	-29.79	0.1261	0.1363	15.91	3.09	0.0441
CAC	-13.10	0.1062	0.0387	8.06	1.35	0.0286
CAC + Gold	-9.13	0.0558	0.0363	6.59	0.86	0.0422
CAC + USDT	-4.35	0.0587	0.0183	4.21	0.67	0.0272
SPX + BTC	-29.79	0.1338	0.1058	9.81	2.47	0.0428
SPX + ETH	-34.09	0.1738	0.1259	17.51	3.08	0.0409
FTM	-9.82	0.0440	0.0056	8.04	1.24	0.0045
FTM + Gold	-7.49	0.0509	0.0197	6.59	0.82	0.0240
FTM + USDT	-4.19	0.0188	0.0017	4.20	0.63	0.0027
SPX + BTC	-28.15	0.0978	0.0891	9.62	2.42	0.0368
SPX + ETH	-32.45	0.1340	0.1091	17.30	3.03	0.0360

Source: authors' work.

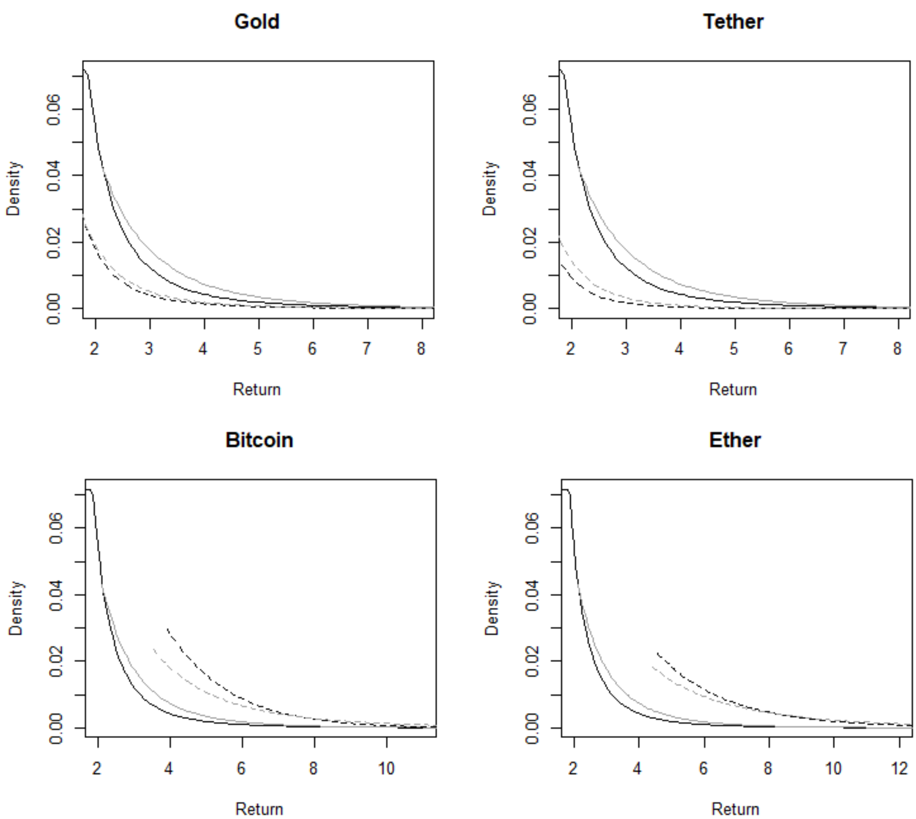
3.4. Extreme losses versus extreme profits in an equal-weighted portfolio

In Section 3.2, we take an optimal weight which is close to zero for Bitcoin and Ether and close to one for Tether. In the case of the first two assets, the optimal portfolios do not differ significantly from the index, whereas in the case of the third one, Tether dominates in the portfolio. Such an assumption gives us a nearly one-component portfolio and prevents us from comparing the relative benefits from hedging. By using an equal-weighted portfolio, we are able to verify the potential of the considered assets for extreme risk reduction in comparison to extreme profit sacrifice.

Figure 5 and Figures A6–A10 show the tails of the return distributions for the indices and equal-weighted portfolios. In the same way as in the previous figures, we present the index right tail area of 5% and the other tails against its background. The distribution tails for portfolios consisting of particular indices and Tether are the thinnest; thus, Tether outperforms other assets in terms of extrema reduction. The asymmetry between risk reduction and profit sacrifice (for gold and Tether) is interpreted as the inequality of differences between the grey lines and the black lines, respectively. More precisely, the difference shows how much the hedging instrument added to a portfolio makes the return distribution tail thinner than the tail for the index. It is more profitable for investors to reduce the lower tail (difference in grey lines) more intensively than the upper tail (difference in black lines), which is interpreted as a higher risk reduction over profit sacrifice. The beneficial asymmetry is the most visible for the S&P 500–gold and FTM–gold pairs, whereas adverse asymmetry is for the NKX–Tether pair. Bitcoin and Ether increase substantially the tail risk of a base investment. Distribution tails of hedge portfolios begin at much higher return levels and decay at a slower rate than the tails for indices. For instance, for the NKX–Ether portfolio, the upper tail begins in a place where the upper tail for the index disappears. In terms of the relative benefits of investment in pair-wise portfolios with Bitcoin or Ether, portfolios with Bitcoin indicate much greater differences between the black lines than between the grey lines for almost all indices. The results show that Bitcoin added to the index increases potential profits much more than the downside risk. It is evident that along with the inclusion of Bitcoin in the portfolio, investors benefit from higher probability of extreme profits in relation to the probability of extreme losses. We observed a similar effect to the Bitcoin case in the Ether portfolio with S&P 500 or TSX. For SHC, CAC and FTM, the same relation holds, however, the differences are not as distinct as for the former indices. The result is ambiguous only for NKX.

As in the previous section, we have done computations of CVaR which are presented in Table 6. The results based on the tail measure do not coincide with those based on the distribution tails. Even when gold and Tether decrease the CVaR more in the lower tail than in an upper tail, the differences are not as clear as shown in the tail plots (e.g. FTM–gold pair). The highest discrepancy between the results from the used methods is noticeable for Bitcoin and Ether. CVaR yields the opposite results relative to the tail analysis for SPX, NKX, CAC and FTM indices in pair with both Bitcoin or Ether, which suggests a higher risk increase of extreme losses compared to the potential extreme profits. These exceptions indicate the need for caution when interpreting results based on a single risk measure. CVaR provides an adequate picture of the risks reflected in the most extreme values in the tail, but it fails to properly capture most of the data from the tail. An analysis based on the entire tail of the return distribution is more general and reliable.

Figure 5. Distribution tails (pdf) of the SPX returns and an equal-weighted portfolio consisting of an index and hedge asset. Black line – 5% upper tail, grey line – lower tail (symmetrical image), solid line – index, dashed line – equal-weighted portfolio



Source: authors' work.

Table 6. Reduction in 5% CVaR (in pp.) by equal-weighted portfolios relative to indices

Portfolio	CVaR Lower tail	CVaR Upper tail	ΔCVaR Lower tail	ΔCVaR Upper tail
SPX	3.47	3.04		
SPX + Gold	2.08	1.99	1.39	1.05
SPX + Tether	1.69	1.44	1.78	1.60
SPX + Bitcoin	6.17	5.55	-2.70	-2.51
SPX + Ether	7.77	6.86	-4.30	-3.82
TSX	3.09	2.33		
TSX + Gold	2.01	1.73	1.08	0.60
TSX + Tether	1.46	1.14	1.63	1.19
TSX + Bitcoin	5.99	5.55	-2.90	-3.22
TSX + Ether	7.56	6.81	-4.47	-4.48
SHC	2.62	2.35		
SHC + Gold	1.83	1.69	0.79	0.66
SHC + Tether	1.41	1.22	1.22	1.13
SHC + Bitcoin	5.56	5.47	-2.94	-3.12
SHC + Ether	7.17	6.98	-4.55	-4.63

Table 6. Reduction in 5% CVaR (in pp.) by equal-weighted portfolios relative to indices (cont.)

Portfolio	CVaR Lower tail	CVaR Upper tail	ΔCVaR Lower tail	ΔCVaR Upper tail
NKX	2.77	2.82		
NKX + Gold	1.85	1.85	0.92	0.97
NKX + Tether	1.43	1.40	1.34	1.42
NKX + Bitcoin	5.97	5.43	-3.20	-2.61
NKX + Ether	7.86	6.86	-5.09	-4.03
CAC	3.41	2.94		
CAC + Gold	2.09	1.80	1.31	1.14
CAC + Tether	1.71	1.50	1.70	1.43
CAC + Bitcoin	5.98	5.43	-2.57	-2.49
CAC + Ether	7.48	6.74	-4.07	-3.80
FTM	3.03	2.82		
FTM + Gold	1.93	1.81	1.09	1.01
FTM + Tether	1.55	1.42	1.48	1.40
FTM + Bitcoin	5.91	5.34	-2.88	-2.52
FTM + Ether	7.45	6.62	-4.43	-3.80

Source: authors' work.

4. Conclusions

Hedging strategies against the risk of six global stock markets are considered in this paper. We compare the effectiveness of two hedging strategies, i.e. optimal hedging and equal-weighted portfolio hedging using gold and cryptocurrencies in the research period covering the COVID-19 pandemic and the Russia-Ukraine war. The empirical study provides an examination of the relative risk reduction in the lower and upper tails of the return distribution through the analysis of the portfolio tail thickness. We are able to verify how much the extreme risk is reduced with the hedging instrument relative to the profit sacrifice. We find several results that shed new light on the benefits of hedging with cryptocurrencies and gold, and thus provide findings relevant for individual and institutional investors.

The optimal hedge strategy is appropriate only when gold is applied as a hedge or a safe-haven instrument. The optimal weight for Tether is close to one; thus, it almost replaces equities from the portfolio built on variance minimisation. On the other hand, optimal weights for Bitcoin and Ether are close to zero. The high volatility of both assets does not allow for an effective risk reduction. Meanwhile, gold provides a good shelter for stock markets, since it reduces volatility, downside risk and provides the highest profit/risk ratio. Moreover, gold is able to reduce the probability of extreme losses more intensively than the probability of extreme profits.

In the equal-weighted portfolio strategy, Tether can still reduce the risk more effectively than gold. However, Tether is an asset which demonstrates the lowest profit/risk ratio among the considered hedges. Moreover, along with the reduction of volatility both extreme losses and extreme profits are reduced at the same rate. Identifying which one is a better shelter for the global stock markets is ambiguous and highly depends on investor preferences. Bitcoin and Ether fail to act as effective hedges or safe-haven assets since they substantially increase volatility and the downside risk. However, we provide convincing arguments that the latter instruments added to stock market indices increase the extreme potential profits on investments more intensively than extreme losses for most of the returns from the tail area of 5%.

Our empirical findings have significant implications for the financial market participants. We address the key question: Is it possible to reduce extreme losses and save extreme profits in a hedging strategy? The answer is negative, safe-haven assets added to the base portfolio always reduce the potential extreme profits along with the unwanted huge losses. Profit sacrifice is an alternative cost of downside risk reduction. However, the relationship between losses and profits depends on the type of the hedging strategy and the hedging assets. The empirical results presented in this study reveal which popular safe-haven candidates offer a beneficial profit/loss relationship. Using the findings, investors can improve their asset allocation and hedging effectiveness by taking into account the asymmetry between profits and losses according to individual expectations and risk tolerance.

References

Aboura, S. (2014). When the U.S. Stock Market Becomes Extreme?. *Risks*, *2*(2), 211–225. https://doi.org/10.3390/risks2020211.

Akhtaruzzaman, M., Boubaker, S., Lucey, B. M., & Sensoy, A. (2021). Is gold a hedge or a safe-haven asset in the COVID-19 crisis?. *Economic Modelling*, *102*, 1–26. https://doi.org/10.1016/j.econmod.2021.105588.

Al-Nassar, N. S., Boubaker, S., Chaibi, A., & Makram, B. (2023). In search of hedges and safe havens during the COVID-19 pandemic: Gold versus Bitcoin, oil, and oil uncertainty. *The Quarterly Review of Economics and Finance*, *90*, 318–332. https://doi.org/10.1016/j.qref.2022.10.010.

Balkema, A. A., & de Haan, L. (1974). Residual Life Time at Great Age. *The Annals of Probability*, *2*(5), 792–804. https://doi.org/10.1214/aop/1176996548.

Baur, D. G., & Glover, K. J. (2014). Heterogeneous expectations in the gold market: Specification and estimation. *Journal of Economic Dynamics and Control*, *40*, 116–133. https://doi.org/10.1016/j.jedc.2014.01.001.

Baur, D. G., Hoang, L. T., & Hossain, M. Z. (2022). Is Bitcoin a Hedge? How Extreme Volatility Can Destroy the Hedge Property. *Finance Research Letters*, *47B*, 1–27. https://doi.org/10.1016/j.frl.2021.102655.

Baur, D. G., & Lucey, B. M. (2010). Is Gold a Hedge or a Safe Haven? An Analysis of Stocks, Bonds and Gold. *The Financial Review*, *45*(2), 217–229. https://doi.org/10.1111/j.1540-6288.2010.00244.x.

Baur, D. G., & McDermott, T. K. (2010). Is gold a safe haven? International evidence. *Journal of Banking & Finance*, *34*(8), 1886–1898. https://doi.org/10.1016/j.jbankfin.2009.12.008.

Beckmann, J., Berger, T., & Czudaj, R. (2015). Does gold act as a hedge or a safe haven for stocks? A smooth transition approach. *Economic Modelling*, *48*, 16–24. https://doi.org/10.1016/j.econmod.2014.10.044.

Bee, M., Dupuis, D. J., & Trapin, L. (2016). Realizing the extremes: Estimation of tail-risk measures from a high-frequency perspective. *Journal of Empirical Finance*, *36*, 86–99. https://doi.org/10.1016/j.jempfin.2016.01.006.

Będowska-Sójka, B., & Kliber, A. (2021). Is there one safe-haven for various turbulences? The evidence from gold, Bitcoin and Ether. *The North American Journal of Economics and Finance*, *56*, 1–12. https://doi.org/10.1016/j.najef.2021.101390.

Bollerslev, T. (1986). Generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics*, *31*(3), 307–327. https://doi.org/10.1016/0304-4076(86)90063-1.

Boubaker, H., Cunado, J., Gil-Alana, L. A., & Gupta, R. (2020). Global Crises and Gold as a Safe Haven: Evidence from Over Seven and a Half Centuries of Data. *Physica A: Statistical Mechanics and its Applications*, *540*, 1–13. https://doi.org/10.1016/j.physa.2019.123093.

Bouri, E., Shahzad, S. J. H., & Roubaud, D. (2020). Cryptocurrencies as hedges and safe-havens for US equity sectors. *The Quarterly Review of Economics and Finance*, *75*, 294–307. https://doi.org/10.1016/j.qref.2019.05.001.

Caeiro, F., & Gomes, M. I. (2016). Threshold Selection in Extreme Value Analysis. In D. K. Dey & J. Yan (Eds.), *Extreme Value Modeling and Risk Analysis. Methods and Applications* (pp. 71–89). Chapman and Hall/CRC. https://doi.org/10.1201/b19721.

Cheema, M. A., Faff, R. W., & Szulczyk, K. R. (2022). The 2008 Global Financial Crisis and COVID-19 Pandemic: How Safe are the Safe Haven Assets?. *International Review of Financial Analysis*, *83*, 1–13. https://doi.org/10.1016/j.irfa.2022.102316.

Chemkha, R., BenSaïda, A., Ghorbel, A., & Tayachi, T. (2021). Hedge and safe haven properties during COVID-19: Evidence from Bitcoin and gold. *The Quarterly Review of Economics and Finance*, *82*, 71–85. https://doi.org/10.1016/j.qref.2021.07.006.

Chen, Z., & Ibragimov, R. (2019). One country, two systems? The heavy-tailedness of Chinese A- and H-share markets. *Emerging Markets Review*, *38*, 115–141. https://doi.org/10.1016/j.ememar.2018.11.007.

Cifter, A. (2011). Value-at-risk estimation with wavelet-based extreme value theory: Evidence from emerging markets. *Physica A: Statistical Mechanics and its Applications*, *390*(12), 2356–2367. https://doi.org/10.1016/j.physa.2011.02.033.

Conlon, T., Corbet, S., & McGee, R. J. (2020). Are cryptocurrencies a safe haven for equity markets? An international perspective from the COVID-19 pandemic. *Research in International Business and Finance*, *54*, 1–10. https://doi.org/10.1016/j.ribaf.2020.101248.

Conlon, T., & McGee, R. (2020). Safe haven or risky hazard? Bitcoin during the Covid-19 bear market. *Finance Research Letters*, *35*, 1–5. https://doi.org/10.1016/j.frl.2020.101607.

Corbet, S., Hou, Y. G., Hu, Y., Larkin, C., & Oxley, L. (2020). Any port in a storm: Cryptocurrency safe-havens during the COVID-19 pandemic. *Economics Letters*, *194*, 1–7. https://doi.org/10.1016/j.econlet.2020.109377.

Daníelsson, J., & de Vries, C. G. (1997). Tail index and quantile estimation with very high frequency data. *Journal of Empirical Finance*, *4*(2–3), 241–257. https://doi.org/10.1016/S0927-5398(97)00008-X.

Danielsson, J., Ergun, L. M., de Haan, L., & de Vries, C. G. (2016). *Tail Index Estimation: Quantile Driven Threshold Selection* (SRC Discussion Paper No 58). https://doi.org/10.2139/ssrn.2717478.

Dowd, K. (2005). *Measuring Market Risk* (2nd edition). John Wiley & Sons. https://doi.org/10.1002/9781118673485.

Echaust, K. (2021). Asymmetric tail dependence between stock market returns and implied volatility. *The Journal of Economic Asymmetries*, *23*, 1–13. https://doi.org/10.1016/j.jeca.2020.e00190.

Echaust, K., & Just, M. (2022). Is gold still a safe haven for stock markets? New insights through the tail thickness of portfolio return distributions. *Research in International Business and Finance*, *63*, 1–19. https://doi.org/10.1016/j.ribaf.2022.101788.

Echaust, K., Just, M., & Kliber, A. (2024). To hedge or not to hedge? Cryptocurrencies, gold and oil against stock market risk. *International Review of Financial Analysis*, *94*, 1–22. https://doi.org/10.1016/j.irfa.2024.103292.

Engle, R. (2002). Dynamic Conditional Correlation: A Simple Class of Multivariate Generalized Autoregressive Conditional Heteroskedasticity Models. *Journal of Business* & *Economic Statistics*, *20*(3), 339–350. https://doi.org/10.1198/073500102288618487.

Eom, C., Kaizoji, T., Livan, G., & Scalas, E. (2021). Limitations of portfolio diversification through fat tails of the return Distributions: Some empirical evidence. *The North American Journal of Economics and Finance*, *56*, 1–22. https://doi.org/10.1016/j.najef.2020.101358.

Fernandez, V. (2005). Risk management under extreme events. *International Review of Financial Analysis*, *14*(2), 113–148. https://doi.org/10.1016/j.irfa.2004.06.012.

Gençay, R., Selçuk, F., & Ulugülyağci, A. (2003). High volatility, thick tails and extreme value theory in value-at-risk estimation. *Insurance: Mathematics and Economics*, *33*(2), 337–356. https://doi.org/10.1016/j.insmatheco.2003.07.004.

Gilli, M., & Këllezi, E. (2006). An application of extreme value theory for measuring financial risk. *Computational Economics*, *27*(2–3), 207–228. https://doi.org/10.1007/s10614-006-9025-7.

Gregory-Allen, R., Lu, H., & Stork, P. (2012). Asymmetric extreme tails and prospective utility of momentum returns. *Economics Letters*, *117*(1), 295–297. https://doi.org/10.1016 /j.econlet.2012.05.040.

Gürgün, G., & Ünalmiş, İ. (2014). Is gold a safe haven against equity market investment in emerging and developing countries?. *Finance Research Letters*, *11*(4), 341–348. https://doi.org/10.1016/j.frl.2014.07.003.

Hansen, P. R., & Lunde, A. (2005). A forecast comparison of volatility models: does anything beat a GARCH (1,1)?. *Journal of Applied Econometrics*, *20*(7), 873–889. https://doi.org/10.1002/jae.800.

Hartmann, P., Straetmans, S., & de Vries, C. G. (2004). Asset market linkages in crisis periods. *The Review of Economics and Statistics*, *86*(1), 313–326. https://doi.org/10.1162/003465304323023831.

Hasan, Md. B., Hassan, M. K., Rashid, Md. M., & Alhenawi, Y. (2021). Are safe haven assets really safe during the 2008 global financial crisis and COVID-19 pandemic?. *Global Finance Journal*, *50*, 1–11. https://doi.org/10.1016/j.gfj.2021.100668.

Hill, B. M. (1975). A Simple General Approach to Inference About the Tail of a Distribution. *The Annals of Statistics*, *3*(5), 1163–1174. https://doi.org/10.1214/aos/1176343247.

Iqbal, F., Zahid, M., & Koutmos, D. (2023). Cryptocurrency Trading and Downside Risk. *Risks*, *11*(7), 1–18. https://doi.org/10.3390/risks11070122.

Jondeau, E., & Rockinger, M. (2003). Testing for differences in the tails of stock-market returns. *Journal of Empirical Finance*, *10*(5), 559–581. https://doi.org/10.1016/S0927-5398(03)00005-7.

Just, M., & Echaust, K. (2021). An Optimal Tail Selection in Risk Measurement. *Risks*, *9*(4), 1–16. https://doi.org/10.3390/risks9040070.

Just, M., & Echaust, K. (2024). Cryptocurrencies against stock market risk: New insights into hedging effectiveness. *Research in International Business and Finance*, *67*(A), 1–26. https://doi.org/10.1016/j.ribaf.2023.102134.

Kahneman, D., & Tversky, A. (1979). Prospect theory: An analysis of decision under risk. *Econometrica*, *47*(2), 263–292. https://doi.org/10.2307/1914185.

Kliber, A. (2022). Looking for a safe haven against American stocks during COVID-19 pandemic. *The North American Journal of Economics and Finance*, *63*, 1–21. https://doi.org/10.1016/j.najef.2022.101825.

Koedijk, K. G., & Kool, C. J. M. (1992). Tail Estimates of East European Exchange Rates. *Journal of Business & Economic Statistics*, *10*(1), 83–96. https://doi.org/10.1080/07350015.1992.10509889.

Kroner, K. F., & Ng, V. K. (1998). Modeling Asymmetric Comovements of Asset Returns. *The Review of Financial Studies*, *11*(4), 817–844. https://doi.org/10.1093/rfs/11.4.817.

Long, S., Pei, H., Tian, H., & Lang, K. (2021). Can both Bitcoin and gold serve as safe-haven assets? – A comparative analysis based on the NARDL model. *International Review of Financial Analysis*, *78*, 1–12. https://doi.org/10.1016/j.irfa.2021.101914.

Longin, F. M. (1996). The Asymptotic Distribution of Extreme Stock Market Returns. *The Journal of Business*, *69*(3), 383–408. https://doi.org/10.1086/209695.

Longin, F. M. (2000). From value at risk to stress testing: The extreme value approach. *Journal of Banking & Finance*, *24*(7), 1097–1130. https://doi.org/10.1016/S0378-4266(99)00077-1.

Lucey, B. M., Peat, M., Šević, A., & Vigne, S. A. (2021). What is the optimal weight for gold in a portfolio? *Annals of Operations Research*, *297*(1–2), 277–291. https://doi.org/10.1007/s10479-019-03496-5.

Łuczak, A., & Just, M. (2020). The positional MEF-TOPSIS method for the assessment of complex economic phenomena in territorial units. *Statistics in Transition new series*, *21*(2), 157–172. https://doi.org/10.21307/STATTRANS-2020-018.

Mariana, C. D., Ekaputra, I. A., & Husodo, Z. A. (2021). Are Bitcoin and Ethereum safe-havens for stocks during the COVID-19 pandemic? *Finance Research Letters*, *38*, 1–7. https://doi.org/10.1016/j.frl.2020.101798.

McNeil, A. J., & Frey, R. (2000). Estimation of tail-related risk measures for heteroscedastic financial time series: An extreme value approach. *Journal of Empirical Finance*, *7*(3–4), 271–300. https://doi.org/10.1016/S0927-5398(00)00012-8.

Salisu, A. A., Raheem, I. D., & Vo, X. V. (2021). Assessing the safe haven property of the gold market during COVID-19 pandemic. *International Review of Financial Analysis*, *74*, 1–7. https://doi.org/10.1016/j.irfa.2021.101666.

Sarykalin, S., Serraino, G., & Uryasev, S. (2008). Value-at-Risk vs. Conditional Value-at-Risk in Risk Management and Optimization. In Z.-L. Chen, S. Raghavan & P. Gray (Eds.), *State-of-the-Art Decision-Making Tools in the Information-Intensive Age* (pp. 270–294). Informs. https://doi.org/10.1287/educ.1080.0052.

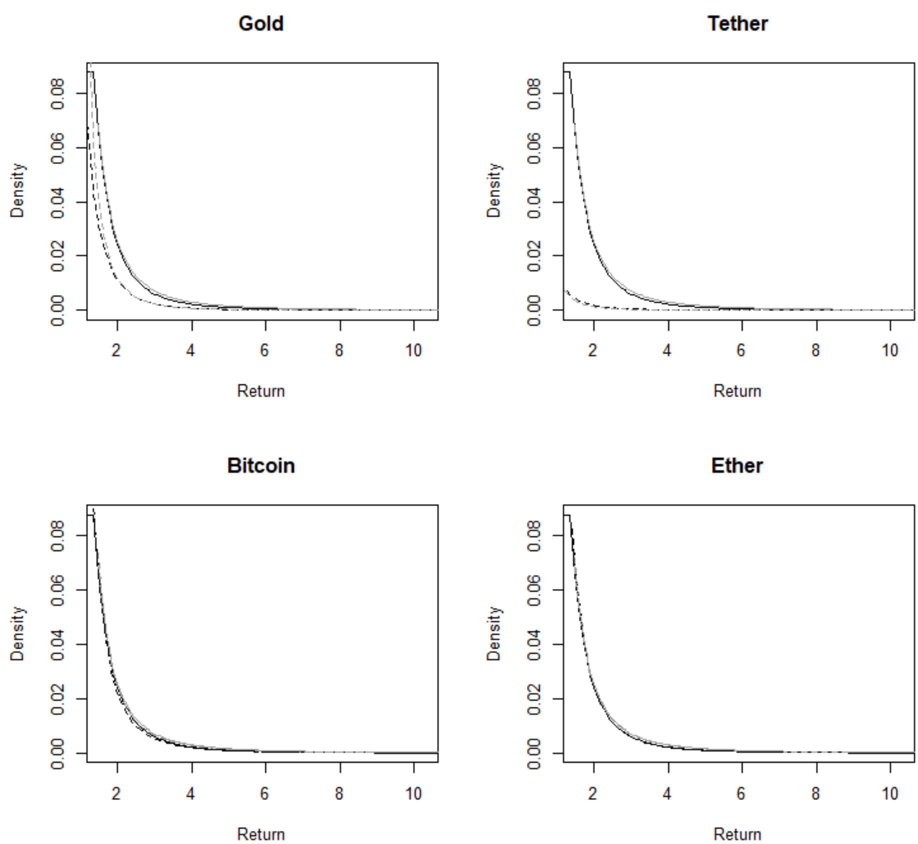
Totić, S., & Božović, M. (2016). Tail risk in emerging markets of Southeastern Europe. *Applied Economics*, *48*(19), 1785–1798. https://doi.org/10.1080/00036846.2015.1109037.

World Federation of Exchanges. (2023). *Market Statistics*. https://focus.world-exchanges.org/issue/march-2023/market-statistics.

Xu, L., & Kinkyo, T. (2023). Hedging effectiveness of bitcoin and gold: Evidence from G7 stock markets. *Journal of International Financial Markets, Institutions and Money*, *85*, 1–18. https://doi.org/10.1016/j.intfin.2023.101764.

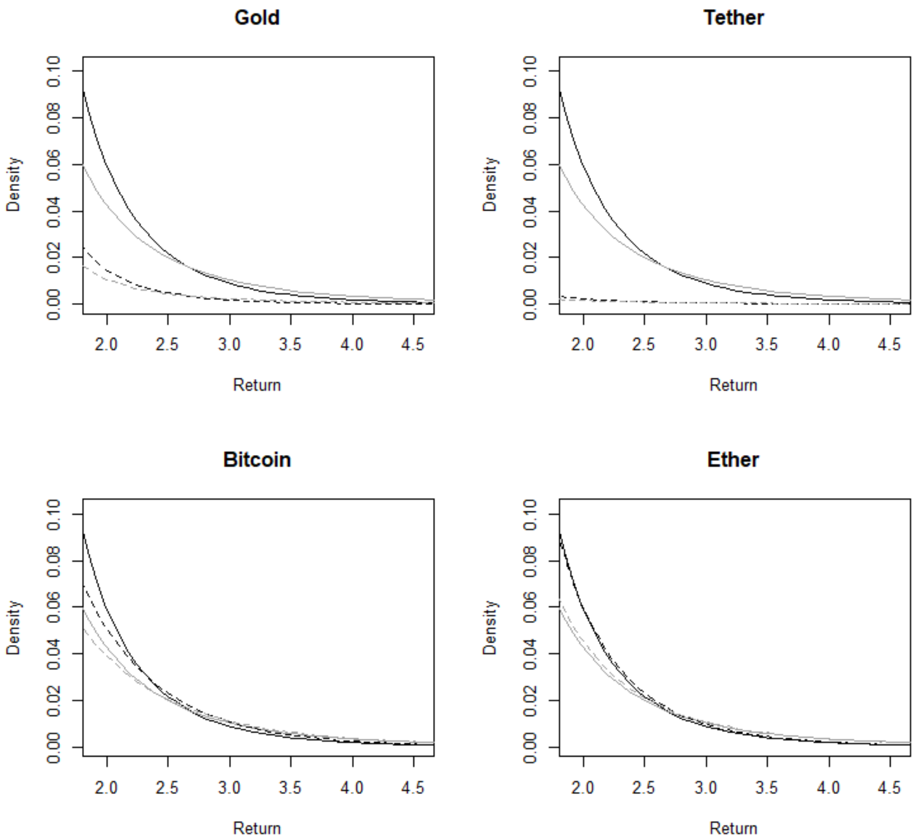
Appendix

Figure A1. Distribution tails (pdf) of the TSX returns and an optimal portfolio consisting of an index and hedge asset. Black line – 5% upper tail, grey line – lower tail (symmetrical image), solid line – index, dashed line – optimal portfolio



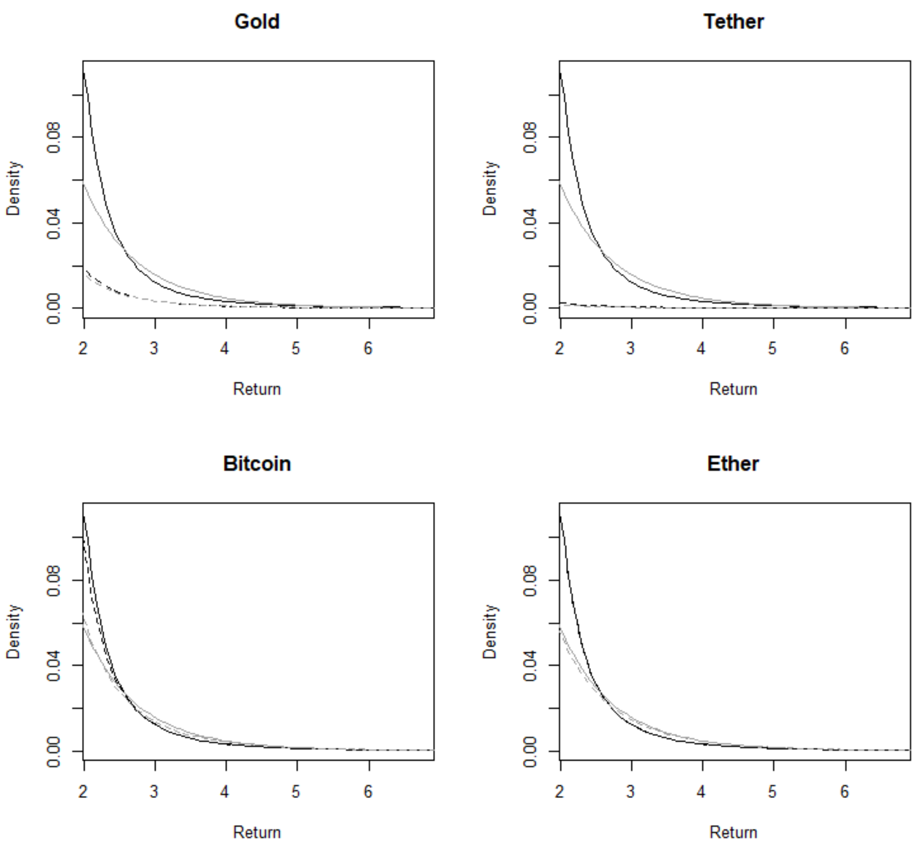
Source: authors' work.

Figure A2. Distribution tails (pdf) of the SHC returns and an optimal portfolio consisting of an index and hedge asset. Black line – 5% upper tail, grey line – lower tail (symmetrical image), solid line – index, dashed line – optimal portfolio



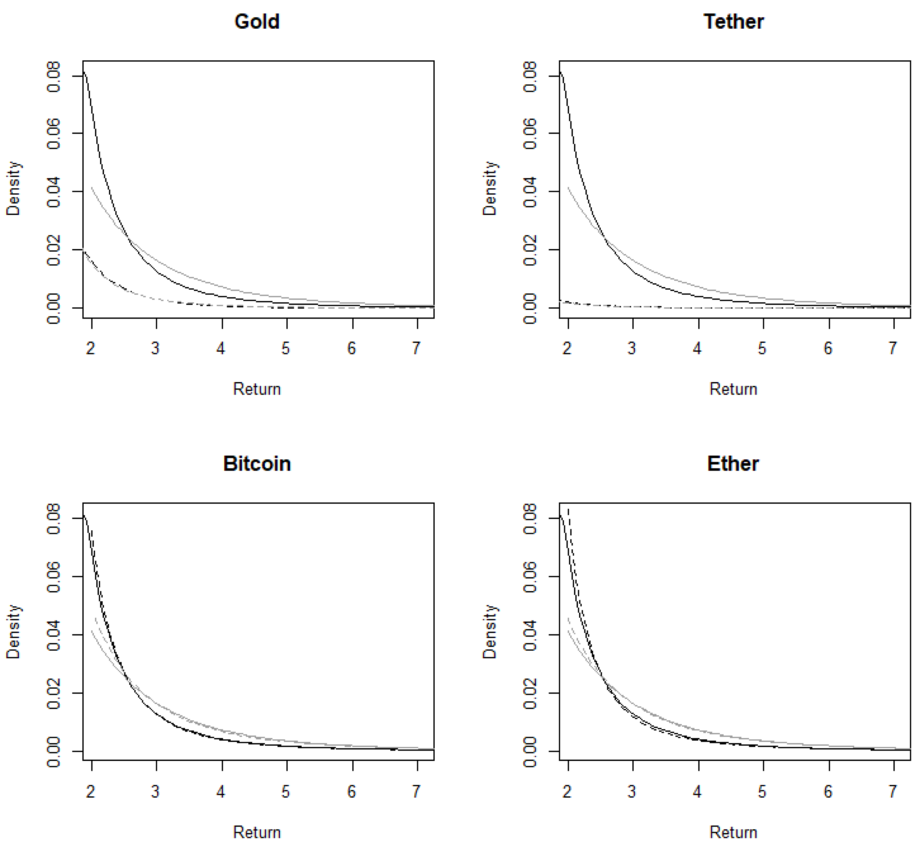
Source: authors' work.

Figure A3. Distribution tails (pdf) of the NKX returns and an optimal portfolio consisting of an index and hedge asset. Black line – 5% upper tail, grey line – lower tail (symmetrical image), solid line – index, dashed line – optimal portfolio



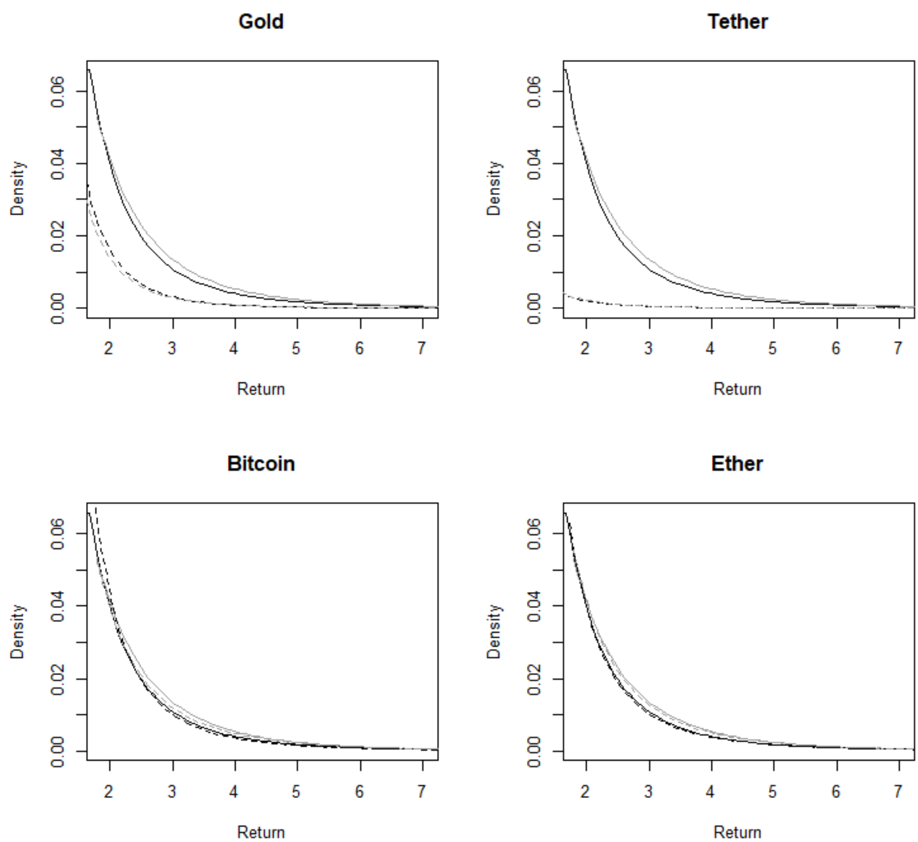
Source: authors' work.

Figure A4. Distribution tails (pdf) of the CAC returns and an optimal portfolio consisting of an index and hedge asset. Black line – 5% upper tail, grey line – lower tail (symmetrical image), solid line – index, dashed line – optimal portfolio



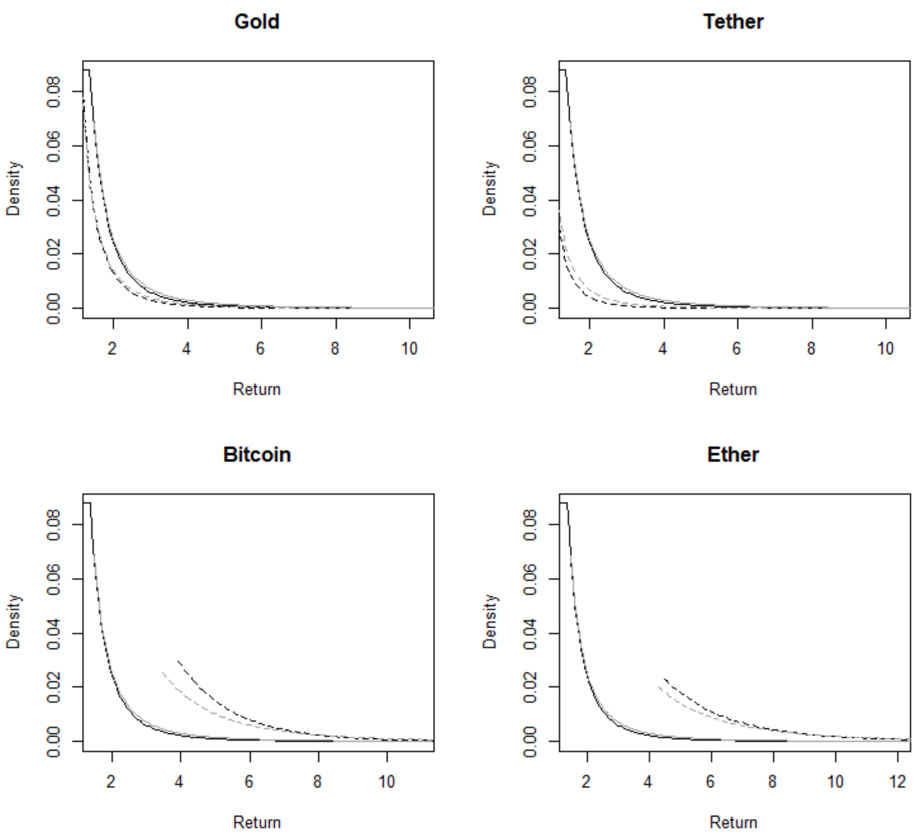
Source: authors' work.

Figure A5. Distribution tails (pdf) of the FTM returns and an optimal portfolio consisting of an index and hedge asset. Black line – 5% upper tail, grey line – lower tail (symmetrical image), solid line – index, dashed line – optimal portfolio



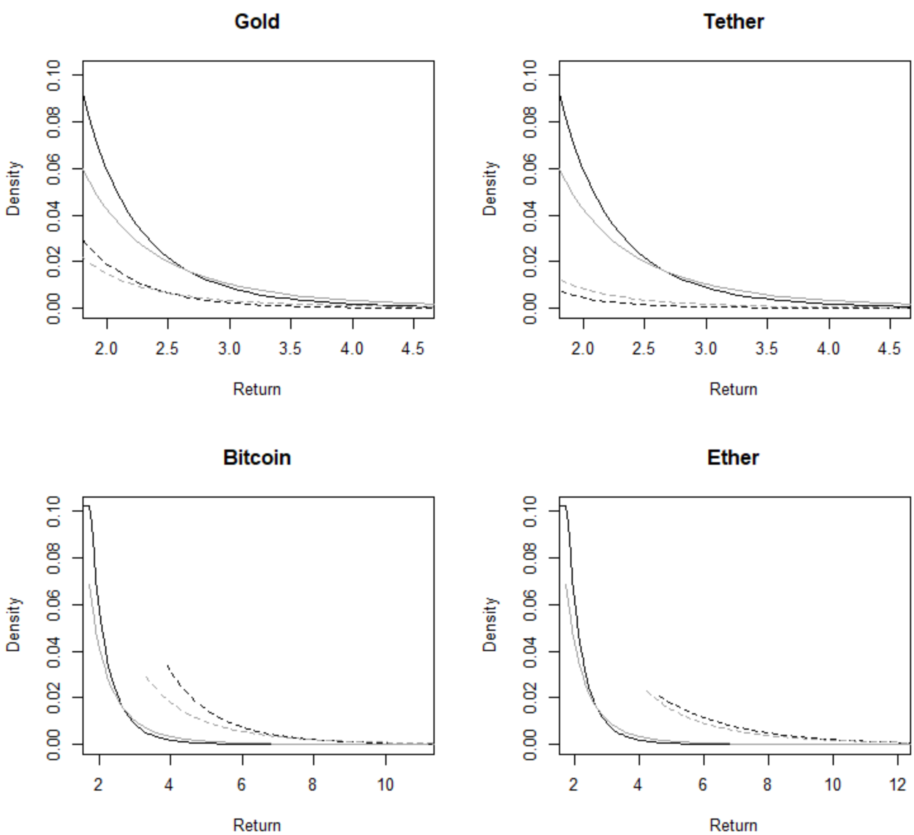
Source: authors' work.

Figure A6. Distribution tails (pdf) of the TSX returns and an equal-weighted portfolio consisting of an index and hedge asset. Black line – 5% upper tail, grey line – lower tail (symmetrical image), solid line – index, dashed line – equal-weighted portfolio



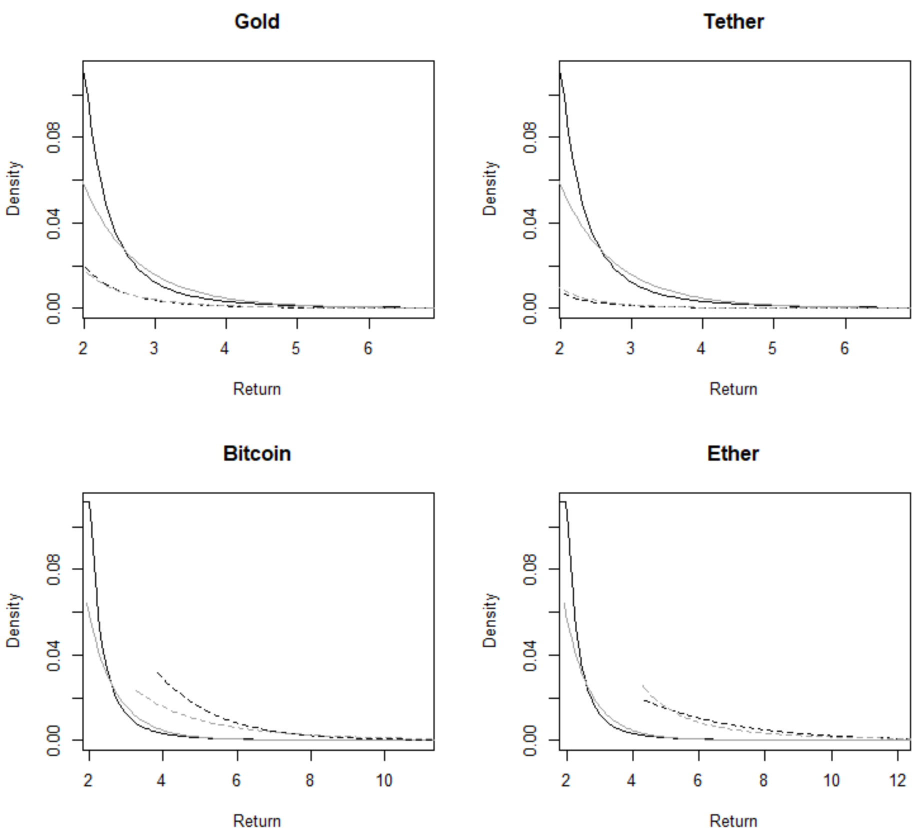
Source: authors' work.

Figure A7. Distribution tails (pdf) of the SHC returns and an equal-weighted portfolio consisting of an index and hedge asset. Black line – 5% upper tail, grey line – lower tail (symmetrical image), solid line – index, dashed line – equal-weighted portfolio



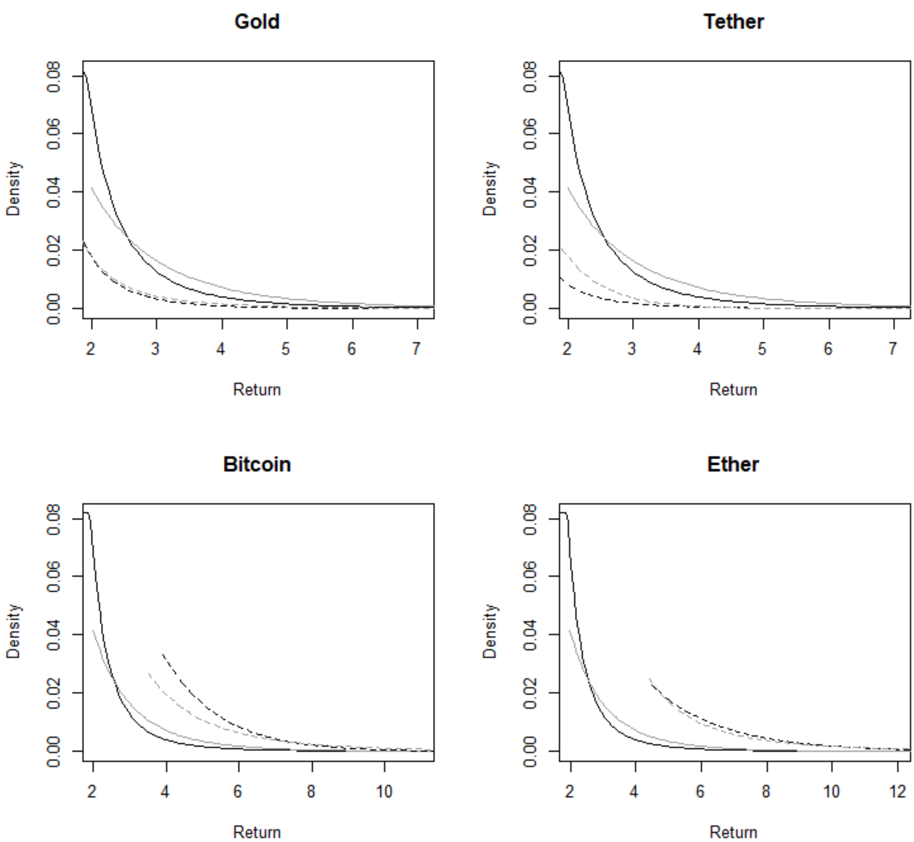
Source: authors' work.

Figure A8. Distribution tails (pdf) of the NKX returns and an equal-weighted portfolio consisting of an index and hedge asset. Black line – 5% upper tail, grey line – lower tail (symmetrical image), solid line – index, dashed line – equal-weighted portfolio



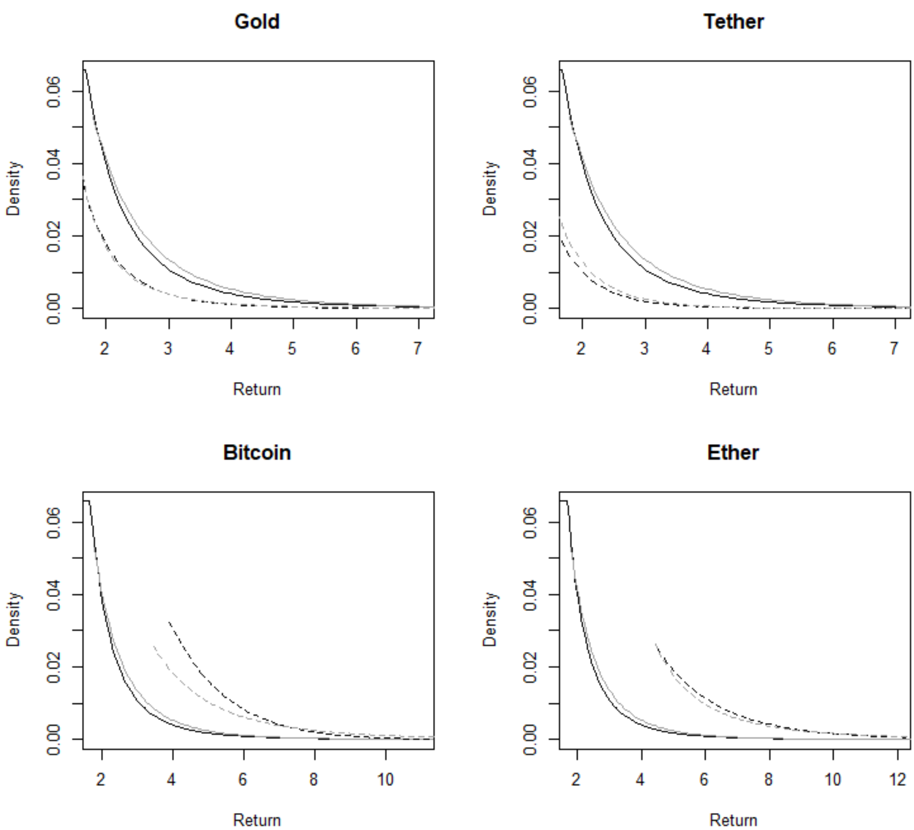
Source: authors' work.

Figure A9. Distribution tails (pdf) of the CAC returns and an equal-weighted portfolio consisting of an index and hedge asset. Black line – 5% upper tail, grey line – lower tail (symmetrical image), solid line – index, dashed line – equal-weighted portfolio



Source: authors' work.

Fig. A10. Distribution tails (pdf) of the FTM returns and an equal-weighted portfolio consisting of an index and hedge asset. Black line – 5% upper tail, grey line – lower tail (symmetrical image), solid line – index, dashed line – equal-weighted portfolio



Source: authors' work.

Przegląd Statystyczny. Statistical Review, 2023, vol. 70, 4, 73–85 https://doi.org/10.59139/ps.2023.04.3
Received: 25.06.2024, revised: 14.10.2024, accepted: 15.10.2024

Computational problems in variable selection for multilevel models using stepwise regression

Hanna Wdowicka[a]

Abstract. Multilevel modelling is a methodology that allows the consideration of variability in the level of the studied variables and the nature of the relationships between them, depending on the affiliation of study units to higher-level units (groups). Additionally, by dividing the studied population into groups, it is possible to explain part of the variability of the estimated characteristic using higher-level characteristics. The usefulness of multilevel modelling in estimating socio-economic characteristics was investigated in the author's previous works. However, with large populations characterised by a multilevel structure, a significant drawback of this approach is its high computational complexity, often resulting in unacceptably long computation times. The main objective of the article is to propose a simplification in the algorithm of forward stepwise multilevel regression, allowing a significant reduction in the time required for variable selection in the model. The considerations will be illustrated by constructing a multilevel model to examine the determinants of daily flows related to employment based on the matrix of employment-related population flows developed from the 2021 National Census of Population and Housing (NSP 2021).

Keywords: multilevel modelling, multilevel structure, random effects, cross-model, commuting to work

JEL: C51, C52, C55

1. Introduction

The idea of multilevel modelling emerged in the early 1970s, when attention was drawn to the fact students within the same school and students from different schools displayed different levels of academic achievement. D. Lindley and A. Smith developed general frameworks for studying nested data with complex structures of random errors (Lindley & Smith, 1972). Incorporating dependencies among units at the first level belonging to the same units at higher levels significantly improves estimation precision compared to classical linear regression, provided the estimated variable has a multilevel structure (Hox, 2010). The applicability of multilevel modelling to estimating socio-economic characteristics was analysed in works such as Gruchociak (2012b), Suchecka and Łaszkiewicz (2017) or Węziak (2007).

[a] Poznań University of Economics and Business, Department of Statistics, al. Niepodległości 10, 61-875 Poznań, Poland, e-mail: hanna.wdowicka@ue.poznan.pl, ORCID: https://orcid.org/0000-0003-2045-9831.

The aim of this paper is to present the computational challenges associated with constructing highly complex multilevel models, along with the proposed solutions. To achieve this, the first part of the article will introduce the empirical problem that underlies the construction of this kind of models. A brief overview of the concept of multilevel modelling will follow. Subsequently, we will explain why this method was chosen to address the specific empirical problem. We will then discuss the construction of appropriate models from a formal point of view. In the final part, we will show any computational difficulties we encountered and will propose relevant solutions.

The empirical aim of the study was to analsye the determinants of daily commuting flows and the relationships between them.

In Poland, the analysis of daily commutes to work is possible thanks to the results of the study carried out periodically by the City Statistics Center of the Statistical Office in Poznań. So far, the results of four editions of the study have been published, for the years: 2006, 2011, 2016 and 2021 (Filas-Przybył & Stachowiak, 2019; Kowalewski, 2014, 2024; Kruszka, 2010). The primary data sources for the first three editions were the Ministry of Finance's tax registers. Data from the National Census of Population and Housing in 2011 and from the Social Insurance Institution were also used (Filas-Przybył & Stachowiak, 2019; Kowalewski, 2024). The National Census of Population and Housing 2021 applied a mixed-method approach, combining data collected from respondents with administrative data (Łysoń, 2024), which was also the case regarding the latest edition of the commuting survey. All four editions of the survey present data at the gmina (the smallest administrative unit in Poland, alternatively referred to as a commune) level, and more specifically for pairs of gminas, with further breakdowns to urban and rural data in the case of urban-rural gminas.

In the current study, the latest edition of the above-mentioned survey was analysed, specifically the 'Commuting to work' survey, based on the results of the National Census of Population and Housing 2021 (Kowalewski, 2024).

2. Multilevel Modelling Methodology

The methodology of multilevel modelling allows the consideration of similarities among units at the first level of the analysis that belong to the same groups formed by a grouping variable at the second and higher levels (Bates, 2010; Biecek, 2011; Raudenbush & Bryk, 2002). Unlike classical linear regression, multilevel modelling does not assume that all observations are independent; instead, it acknowledges the dependence among units at the first level that belong to the same units at higher levels (Twisk, 2010). Failing to account for such dependencies leads to underestimated standard errors (Hox, 2010; Klimanek, 2003). Thus, multilevel models capture two types of variability: differences among units at the first level belonging to the same units at higher levels, and differences among higher-level units themselves (Frątczak & Mianowska, 2012). Incorporating dependencies among units at the first level that belong to the same higher-level units significantly improves the precision of estimation compared to classical linear regression, provided the estimated variable has a multilevel structure. This improvement occurs regardless of whether the classical regression model disregards the division of first-level units entirely or treats observations belonging to the same groups as a whole (with estimates conducted for entire groups) (Twisk, 2010).

Additionally, by dividing the studied population into groups, it becomes possible to explain some of the variability of the estimated characteristic using the characteristics from the higher levels. The need for aggregating information available at different levels is also mentioned in the work of Bołt et al. (1985).

It should be emphasised that the use of multilevel modelling methodology is justified only for specific populations and variables. Both the population and the variable should have a multilevel structure. Regarding the population and the classical multilevel model, this means that the population can be divided into a finite number of distinct and collectively exhaustive groups that cover all units at the first level (alternatively referred to as units at the second level) (Goldstein, 2003; Hox, 2010; Łaszkiewicz, 2016).

In the case of a model with more than two levels, units at the second level can also be divided into distinct and collectively exhaustive groups that cover the entire population (further referred to as units at the third level), and so on.

If a model has two grouping criteria (known as a cross-level model), units at the second level are defined twice, with each first-level unit belonging to exactly one second-level unit defined by each of the two grouping criteria. For units at the second level defined by both criteria, higher-level units can also be defined (Bates, 2010; Biecek, 2011).

Using a multilevel model is justified when the estimated variable is of a multilevel structure. This means that its value should significantly differ between groups, i.e. units, at each of the higher levels. This variability can stem from a direct relationship between the variable of interest and the fact that the first-level unit belongs to higher-level units. A classic example from the literature illustrating such a scenario is the variability in academic achievement, caused both by individual students' abilities and predispositions (factors at the first, individual level) and teacher qualifications and teaching methods (factors at the second, group level) (Goldstein, 2003; Hox, 2010).

Another reason for the variability of the studied variable between groups can be the relationship between this variable and the division into groups based on a certain hidden, often unmeasurable, variable. An example of this is the relationship between the economic activity and the level of regional development (Gruchociak, 2012a, 2012b). Geographical distribution of factors determining labor demand, such as the presence of natural resources, industrial facilities, development of technical, communication, and educational infrastructure, are significant determinants influencing the economic activity of the population, alongside factors affecting the degree of entrepreneurship among individuals.

If the dependent variable has a multilevel structure, employing an appropriate multilevel model can significantly improve the quality of the analysis (Goldstein, 2003; Hox, 2010; Raudenbush & Bryk, 2002; Twisk, 2010).

3. Empirical problem

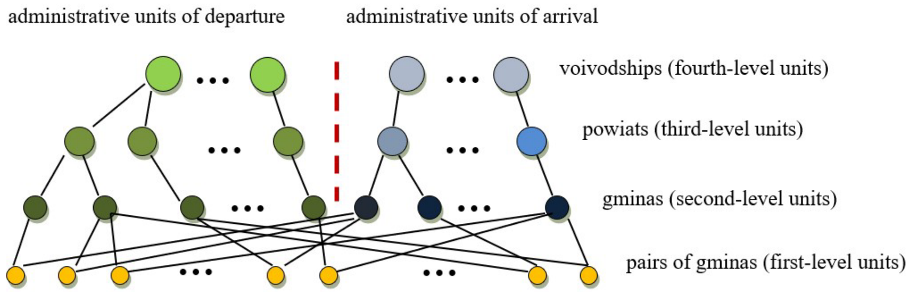
3.1. Dependent variable and multilevel structure

The study focuses on daily commuting patterns between pairs of gminas. As mentioned earlier, the analysis of daily employee mobility is based on the results of the 'Commuting to work' survey using data from the 2021 National Census of Population and Housing (Kowalewski, 2024). Information regarding the number of individuals commuting to work is available at the gmina level, with a breakdown to an urban and a rural part in the case of urban-rural gminas. There were 95,890 pairs of gminas[1] with non-zero[2] employment-related flows, and these pairs are treated in our study as first-level units. The administrative divisions of the country naturally create a multilevel structure (voivodships consist of powiats, which further down consist of gminas). Given that the first-level units of the analysis are defined as pairs of gminas in this study, we are dealing with a cross-level multilevel structure with two grouping criteria defined at the second level. The second-level units can be identified as the gminas from which the commuting flows originate and the gminas of destination, thereby establishing two grouping criteria at the second level. Powiats and voivodships are considered as third- and fourth-level units respectively, due to both grouping criteria (see Figure 1).

[1] Throughout this article, a gmina is understood as a commune divided into urban and rural parts.

[2] Due to statistical confidentiality, the flows between communes with fewer than three people were disregarded.

Figure 1. Multilevel structure of administrative units in a cross-model with two grouping criteria defined the at the second level



Source: author's work.

The above-described structure of the studied community makes it possible to try to explain the number of individuals commuting between the gminas using a cross-level multilevel model with two grouping criteria (related to residential and workplace territorial units), and subsequent levels defined as successive territorial units (gmina, powiat, voivodship) determined by both grouping criteria. Obviously, the inclusion of the multilevel structure in the model will be preceded by verifying whether the studied characteristic of the commuting patterns has a multilevel structure, based on the described grouping criteria.

3.2. Explanatory variables

The distance between the place of residence and the place of work is regarded as one of the most crucial factors influencing the commuting intensity (Gumuła et al., 2007). Therefore, the distance between the gmina of residence and the gmina of workplace was adopted as the explanatory variable at the first level.

The number of persons commuting to work between gminas is naturally affected by the sizes of both gminas. Therefore, the characteristics related to these sizes constitute the second group of potential explanatory variables for the commuting intensity (see Table). The volume of industry within a gmina might also significantly impact commuting patterns, and thus the third group of potential explanatory variables could emerge. It is also plausible that the characteristics related to the attractiveness of a particular place and the relative ease of commuting from there would significantly influence the commuting intensity, so we put these factors together as the fourth group of potential explanatory variables (see Table).

Variations in the intensity of commuting to the workplace might also depend on the attractiveness of the employment conditions both in the 'departure' and the 'destination' gminas. According to the literature, a comprehensive set of characteristics illustrating the labour market situation should comprise variables representing both the labour demand and the supply (Gołata, 2004). Therefore, the set of potential explanatory variables has been expanded to include three additional groups describing the labour market, focusing in particular on the supply, demand and pricing characteristics (see Table).

Table. A set of potential explanatory variables for commuting to work between gminas

distance	
X	distance between gminas[3]
gmina size	
Cga1 ,Cgb1	area of the gmina in square kilometers
Cga2 ,Cgb2	number of people of working age
Cga3 ,Cgb3	number of working people
industry volume	
Ega1 ,Egb1	number of national economy entities from the public sector
Ega2 ,Egb2	number of national economy entities from the private sector
Ega3 ,Egb3	number of commercial companies with foreign capital
Ega4 ,Egb4	number of natural persons conducting economic activity
Ega5 ,Egb5	number of national economy entities
Epa1 ,Epb1	enterprises' investment outlays per person of working age
location	
Lga1 ,Lgb1	distance from the nearest voivodship capital
Lga2 ,Lgb2	distance from the nearest metropolis[4]
Lga3 ,Lgb3	number of parking lots in the Park & Ride system
Lpa1 ,Lpb1	investment outlays of enterprises in thousands of PLN
Lva1 ,Lvb1	number of railway lines

[3] Measured in a straight line between the centroids of gminas (Kopczewska, 2006).
[4] The metropolis were selected using the following procedure: all gminas were ranked in descending order according to the number of people employed. In the first step, the set of central centers was defined as the city with the largest number of employees, then the set was expanded by subsequent gminas ranked according to the number of employees. For the sets of large cities defined in this way in subsequent stages, the correlation between the distance and the intensity of trips to work was examined. Then we checked for which of the defined sets of large cities specified in individual steps the relationship between the distance and the intensity of trips to work was the strongest. In line with Thünen's theory that central centers stimulate the development of areas surrounding them, it was assumed that the distance from each gmina to the nearest central center should influence the intensity of trips to work. Ultimately, a set of large cities was selected for which the correlation relationship was the strongest. According to the above procedure, the following seven metropolis were identified: Warsaw, Krakow, Wroclaw, Poznan, Lodz, Gdansk and Katowice.

Table. A set of potential explanatory variables for commuting to work between gminas (cont.)

demand	
Dga1 ,Dgb1	ratio of working people to the working-age population
Dga2 ,Dgb2	number of working people per square kilometer
Dga3 ,Dgb3	number of people of working age per square kilometer
Dva1 ,Dvb1	share of graduates of public universities among the working-age population
Dva2 ,Dvb2	share of university graduates among the working-age population
supply	
Sga1 ,Sgb1	intensity[5] of national economy entities from the public sector
Sga2 ,Sgb2	intensity of national economy entities from the private sector
Sga3 ,Sgb3	intensity of commercial companies with foreign capital
Sga4 ,Sgb4	share of natural persons conducting economic activity in the working-age population
Sga5 ,Sgb5	intensity of national economy entities
Sga6 ,Sgb6	share of the unemployed among the working-age population
price	
Ppa1 ,Ppb1	average monthly gross salary

Note. The first letter of the subscript indicates the level at which this characteristic is available; g- gmina, d-powiat, v-voivodship. The second letter of the subscript determines the grouping criterion of a given territorial unit; a-territorial unit of residence, b-territorial unit of the workplace.

Source: author's work based on data sets published by GUS.

Among the potential determinants of commuting distances to work, various characteristics available at different levels of aggregation have been considered. Therefore, it seems particularly justified to attempt an analysis of the determinants of labour mobility using the modelling that takes into account the multilevel structure of the labour market.

4. Research procedure

To explain how the number of commutes between pairs of gminas is determined, a four-level cross model with two grouping criteria at the second level was constructed. Units at the first level were defined as pairs of gminas, while units at the second level were gminas from which commutes originated and the destination ones, thereby defining two grouping criteria at the second level. Powiats and voivodships were adopted as third- and fourth-level units taking into account both grouping criteria (see Figure 1). Thus, we considered the incorporation of up to six different grouping criteria defined by the gmina, powiat and voivodship of residence, as well as the gmina, powiat and voivodship the workplace. Additionally, a set of 53 potential explanatory variables defined at various levels (see Table),

[5] Understood here as the ratio of the number of national economy entities from the public sector to the number of working-age inhabitants; it is understood analogically elsewhere in Table.

interactions among these variables, and variations in the influence of individual explanatory variables on the number of commuters depending on their affiliation with higher-level grouping units were considered.

As we can see, designing an optimal multilevel model involves considering numerous potential model extensions for this research problem. Applying stepwise regression to all these extensions simultaneously is impossible due to the extensive computational time required.

This is a common problem in multilevel modelling, therefore the construction of a multilevel model is usually carried out in stages, which makes it possible to shorten the time of selecting extensions for the model (Bliese, 2022; Hox, 2010; Twisk, 2010). The analysis of determinants of commuting to work was conducted in the stages described below.

At first, expanding the fixed part of the model (i.e., without accounting for random components arising from the multilevel data structure) was carried out in five steps. In the following three stages, a random part of the model associated with grouping effects was included. This aimed to initially account for effects that can be measured (i.e. fixed effects). It appears that if the level of commuting varies across certain territorial units, but these differences can be explained by measurable explanatory variables, it should not be considered as a random factor, but it rather should be explained using these variables. Another advantage of this sequential model expansion is that after incorporating random components, the estimation time for the subsequent models with considered extensions significantly increases. Therefore, adding them in the final stages allows a substantial reduction in the total computation time.

The construction of the multilevel model was preceded by verifying the hypothesis of the multilevel structure of the estimated variable, conducted using an analysis-of-variance test (Krzyśko, 1996).

STAGE 1

In the first step, explanatory variables were introduced independently of the aggregation level for which they were defined. A total of 53 explanatory variables were analysed.

STAGE 2

In the second step, we introduced the interactions between the distance between the gminas of residence and of the workplace and the remaining potential explanatory variables defined for territorial units specified by both the place of residence and the workplace. This was done regardless of whether they were included or not in step 1. If the influence of a particular explanatory variable on commuting propensity varies depending on the commuting distance, including their interactions should improve the model fit. In such cases, we can conclude that the distance moderates the impact of an explanatory variable on the number of people commuting to work. A total of 52 interactions were analysed in this stage.

STAGE 3

In stage 3, the interaction between the characteristics of the gmina of residence and the gmina of work was taken into consideration. This enabled us to indicate in what ways the characteristics of the gmina of residence moderate the impact of the characteristics of the gmina of work on attracting commuters to it. Interactions between each pair of explanatory variables were considered, i.e. 20 x 20 = 400 potential extensions.

STAGE 4

Then, we were considering the inclusion of interactions between the characteristics of the same gmina (residence or work). This enabled us to indicate how certain characteristics of the gmina of residence affect the impact of other characteristics of the same gmina on the intensity of departures from this gmina, and how some characteristics of the gmina of the workplace affect the impact of other characteristics of the same gmina on the intensity of arrivals to this gmina. Interactions between each pair of explanatory variables were analysed, both for the gminas of residence and of the workplace, i.e. 2 x 20 x 19 / 2 = 380 potential extensions.

STAGE 5

Then we were considering taking into account the interactions between the characteristics of territorial units from different levels but according to the same grouping criterion (i.e. for territorial units of residence or of the workplace). This enabled us to determine in what ways the characteristics of higher-level residence units affect the impact of the characteristics of lower-level residence units on the intensity of trips to work, and how the characteristics of higher-level work units affects the impact of the characteristics of lower-level work units on the intensity of trips to work. Therefore, interactions between each pair of explanatory variables were analysed, both for territorial units of residence and of the workplace, i.e. 2 x (3 x 20 + 3 x 20 + 3 x 3) = 258 potential extensions.

STAGE 6

Then the construction of the random part of the model began. In stage 6, the impact of each grouping factor on the level of commuting to work, i.e. six potential random effects, were examined.

STAGE 7

In this stage, the differences in the impact of the commuting distance on the number of people commuting to work across the territorial units of residence and of the workplace of commuters were taken into account. Therefore, we were considering the introduction of six further random effects.

STAGE 8

In the last stage, we were contemplating taking into account the differences in the impact of other explanatory variables on the number of people commuting to work across higher-level territorial units created due to the same grouping criteria as the examined explanatory variable. Therefore, we were considering the introduction of another 2 x (2 x 20 + 3) = 86 random effects.

Within the individual stages, extensions were introduced according to forward stepwise regression procedure, with the result of the likelihood ratio test as the improvement criterion. In most of the stages (except stages 2, 3, 4 and 5), the improvement was considered statistically significant if the p-value did not exceed 0.05. The above-mentioned exceptions occurred because during including interactions between variables in the model, special caution was recommended, so for these stages the significance level was 0.01.

After having completed each stage, we checked whether any of the previously added (in this stage or any previous one) extensions still improves the quality of the model fit (measured using the likelihood ratio test, p-value = 0.05) in a statistically significant way. If any of the previously-added model extensions did not meet this condition, they were removed from the model in steps. Among other developmants, this approach allows, at least to some extent, the replacement of a model extension introduced in an earlier stage by the extension from a later stage that better explains a given part of the variability in commuting.

Before starting the calculations, all explanatory variables were centered, which is a standard procedure in multilevel modelling. In addition, variables measured on very different scales were rescaled to avoid convergence problems.

The calculations were performed in the R program using the lme4 library (Bates, 2013; Bates et al., 2015; Bates, 2024), in which the parameters of multilevel models are estimated using sparse matrices. The code of the described algorithm is available in the appendix of this article.

5. Calculational problems and proposed solutions

Despite the construction of the model divided into eight stages, starting with the construction of the less time-consuming constant part of the model and the use of the effective lme4 library of the R program, the number of extensions considered, especially the interactions between the potential explanatory variables, resulted in unacceptably long calculation times[6]. Therefore, we decided to use a simplified

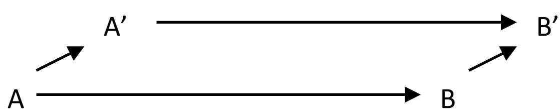
[6] The computation time has been estimated at many billions of years.

forward stepwise regression procedure in selected stages (those concerning the interactions between variables).

In order to shorten the calculation time, we tried to modify the forward stepwise regression procedure used. Namely, we assumed that if a certain extension of the model does not significantly improve the quality of the model fitting, then after enriching the model with other elements, adding the same extension as before, will not improve the quality of its fitting either.

The adopted simplification can also be written in a formal way. Let model B be an extension of model A, and models A' and B' be extensions of models A and B, respectively, with the same component. If the likelihood ratio test does not indicate a significant improvement of model A' in relation to model A, then model B' is not significantly better than model B according to this criterion (see Figure 2).

Figure 2. A diagram for the gradual expansion of multilevel models and the relationships between them



Source: author's work.

It should be noted that this simplification was adopted only in the stages in which interactions between variables were considered (i.e. stages 2, 3, 4, 5, 7 and 8), while in the remaining stages, the classic forward stepwise regression algorithm was used. It seems that the risk of disregarding an interaction between variables that would become significant only after taking into account another interaction at the same stage is relatively small. Additionally, it has to be remembered that no stepwise regression procedure can guarantee the selection of the optimal model; to be sure, models with all possible subsets of the considered extensions should be considered. Also, extending multilevel models within stages, although widely used (Bliese, 2022; Hox, 2010; Twisk, 2010), involves the risk of disregarding extensions that could become important at some point in the construction of the model.

It is worth remembering that thanks to the use of such a simplification, the number of models estimated successively during the forward stepwise regression procedure, e.g. in the third stage, decreased from 400! to less than a thousand, and in stage 8 instead of 86! multilevel models[7], it was enough to estimate just over 100

[7] Estimation of one multilevel model with a complexity level corresponding to stage 8 takes about 10 minutes, so it can be calculated that the estimation 86! such models would take many trillions of years.

such models, which allowed the calculations to be carried out in a finite time. In this context, the adoption of the above-mentioned simplification seems to be totally justified.

6. Conclusions

In conclusion, it can be said that the procedure of model construction that has been carefully thought over and the use of the proposed simplification in the forward stepwise regression procedure allowed the calculations to be carried out in an acceptable time. A discussion of the results obtained regarding the determinants of commuting between gminas will be discussed in detail in a separate article.

References

Bates, D. (2010). *lme4: Mixed-effects modeling with R*. Springer.

Bates, D. (2024). *Computational methods for mixed models*. http://cran.r-project.org/web/packages/lme4/vignettes/Theory.pdf.

Bates, D. (2013). *Linear mixed model implementation in lme4*.

Bates, D., Mächler, M., Bolker, B. M., & Walker, S. C. (2015). Fitting Linear Mixed-Effects Models Using lme4. *Journal of Statistical Software*, *67*(1), 1–48. https://doi.org/10.18637/jss.v067.i01.

Biecek, P. (2011). *Analiza danych z programem R. Modele liniowe z efektami stałymi, losowymi i mieszanymi*. Wydawnictwo Naukowe PWN.

Bliese, P. (2022). *Multilevel Modeling in R (2. 7). A Brief Introduction to R, the multilevel package and the nlme package*. https://cran.r-project.org/doc/contrib/Bliese_Multilevel.pdf.

Bołt, T. W., Krauze, K., & Kulawczuk, T. (1985). *Agregacja modeli ekonometrycznych*. Państwowe Wydawnictwo Ekonomiczne.

Filas-Przybył, S., & Stachowiak, D. (2019). *Przepływy ludności związane z zatrudnieniem w 2016 r.* Urząd Statystyczny w Poznaniu. https://stat.gov.pl/obszary-tematyczne/rynek-pracy/opracowania/przeplywy-ludnosci-zwiazane-z-zatrudnieniem-w-2016-r-,20,1.html.

Frątczak, E., & Mianowska, M. (2012). Modele mieszane. In E. Frątczak (Ed.), *Zaawansowane metody analiz statystycznych* (pp. 223–380). Oficyna Wydawnicza SGH.

Goldstein, H. (2003). *Multilevel Statistical Models* (3rd edition). Edward Arnold.

Gołata, E. (2004). *Estymacja pośrednia bezrobocia na lokalnym rynku pracy*. Wydawnictwo Akademii Ekonomicznej w Poznaniu.

Gruchociak, H. (2012a). Konstrukcja estymatora regresyjnego dla danych o strukturze dwupoziomowej. *Prace Naukowe Uniwersytetu Ekonomicznego we Wrocławiu*, (242), 590–600.

Gruchociak, H. (2012b). Możliwości zastosowania modelowania dwupoziomowego w badaniach ekonomicznych. *Przegląd Statystyczny. Statistical Review*, *59*(4), 409–433. https://doi.org/10.59139/ps.2012.04.4.

Gumuła, W., Socha, J., & Wojciechowski, W. (2007). *Presja płacowa oraz niedopasowanie strukturalne na rynku pracy w świetle badań NBP*. Narodowy Bank Polski. https://static.nbp.pl/publikacje/materialy-i-studia/ms219.pdf.

Hox, J. J. (2010). *Multilevel Analysis. Techniques and Applications* (2nd edition). Routledge. https://doi.org/10.4324/9780203852279.

Klimanek, T. (2003). *Wielopoziomowa analiza struktury agrarnej gminy w systemie Geo-Info* [rozprawa doktorska, Akademia Ekonomiczna w Poznaniu]. https://bazawiedzy.ue.poznan.pl /info/phd/UEP844a7d3f538c48f4ba20dbebc9ee7803/.

Kopczewska, K. (2006). *Ekonometria i statystyka przestrzenna z wykorzystaniem programu R-CRAN*. CeDeWu.

Kowalewski, J. (Ed.). (2014). *Dojazdy do pracy. Narodowy Spis Powszechny Ludności i Mieszkań 2011*. Główny Urząd Statystyczny. https://stat.gov.pl/spisy-powszechne/nsp-2021/nsp-2021-wyniki-ostateczne/dojazdy-do-pracy-nsp-2011,8,1.html.

Kowalewski, J. (Ed.). (2024). *Dojazdy do pracy w świetle wyników Narodowego Spisu Powszechnego Ludności i Mieszkań 2021*. Główny Urząd Statystyczny, Urząd Statystyczny w Poznaniu. https://stat.gov.pl/spisy-powszechne/nsp-2021/nsp-2021-wyniki-ostateczne/dojazdy-do-pracy-w-swietle-wynikow-nsp-2021,8,2.html.

Kruszka, K. (Ed.). (2010). *Dojazdy do pracy w Polsce. Terytorialna identyfikacja przepływów ludności związanych z zatrudnieniem*. Główny Urząd Statystyczny, Urząd Statystyczny w Poznaniu. https://stat.gov.pl/obszary-tematyczne/rynek-pracy/opracowania/dojazdy-do-pracy-w-polsce-2010-na-podstawie-danych-administracyjnych,8,1.html.

Krzyśko, M. (1996). *Statystyka matematyczna*. Wydawnictwo Naukowe Uniwersytetu im. Adama Mickiewicza w Poznaniu.

Lindley, D. V., & Smith, A. F. M. (1972). Bayes Estimates for the Linear Model. *Journal of the Royal Statistical Society, Series B (Methodological)*, *34*(1), 1–41.

Łaszkiewicz, E. (2016). *Ekonometria przestrzenna III. Modele wielopoziomowe – teoria i zastosowania*. Wydawnictwo C.H. Beck.

Łysoń, P. (Ed.). (2024). *Ludność na rynku pracy w świetle wyników Narodowego Spisu Powszechnego Ludności i Mieszkań 2021 – wersja uaktualniona*. Główny Urząd Statystyczny. https://stat.gov.pl/spisy-powszechne/nsp-2021/nsp-2021-wyniki-ostateczne/ludnosc-na-rynku-pracy-w-swietle-wynikow-narodowego-spisu-powszechnego-ludnosci-i-mieszkan-2021-wersja-uaktualniona,11,3.html.

Raudenbush, S. W., & Bryk, A. S. (2002). *Hierarchical Linear Models. Applications and Data Analysis Methods* (2nd edition). Sage Publications.

Suchecka, J., & Łaszkiewicz, E. (2017). Wykorzystanie modelowania wielopoziomowego w analizie jakości życia. *Studia Ekonomiczne. Zeszyty Naukowe Uniwersytetu Ekonomicznego w Katowicach*, (344), 133–144. https://www.ue.katowice.pl/fileadmin/user_upload/wydawnictwo /SE_Artyku%C5%82y_341_360/SE_344/08.pdf.

Twisk, J. W. R. (2010). *Analiza wielopoziomowa – przykłady zastosowań. Praktyczny podręcznik biostatyki i epidemiologii*. Oficyna Wydawnicza SGH.

Węziak, D. (2007). Wielopoziomowe modelowanie regresyjne w analizie danych. *Wiadomości Statystyczne*, *52*(9), 1–12.