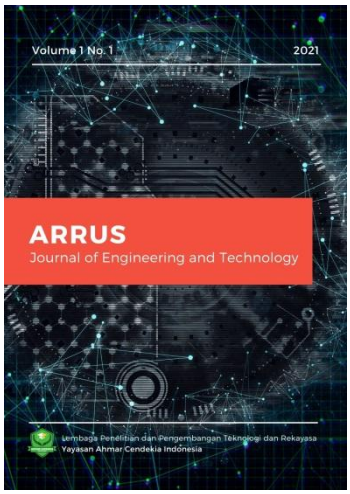
https://doi.org/10.35877/jetech5606

OPEN ACCESS

ISSN : 2776-7914 (Print) / 2807-3045 (Online)



*Corresponding author: Annisa Syalsabila, Department of Statistics, Universitas Negeri Makassar, Makassar, Indonesia

E-mail: annisa.syalsabila@unm.ac.id

RESEARCH ARTICLE

Tail Specification in Value-at-Risk: Evidence from the 2026 MSCI Reclassification Shock

Annisa Syalsabila[1]* & Muh Qodri Alfairus[2]

[1]Department of Statistics, Faculty of Mathematics and Natural Sciences, Universitas Negeri Makassar, Makassar, South Sulawesi
[2]Department of Accounting, Facultyof Economics and Business, Universitas Negeri Makassar, Makassar, Indonesia

Abstract: This study examines how much Value-at-Risk (VaR) performance around a market reclassification shock depends on the tail specification, separately from the volatility specification. It uses daily log returns of the Jakarta Composite Index and the four largest banks on the Indonesia Stock Exchange from January 2022 to October 2026. Models are estimated on a rolling 250-day window updated daily, which gives each asset 902 out-of-sample test days. Four models are compared: normal GARCH(1,1), Student-t GARCH(1,1), GARCH with a peaks over threshold tail, and GARCH with a block maxima tail. The normal, peaks over threshold, and block maxima models share one GARCH estimate, so their conditional variance series are identical and only the residual quantile differs. At the 99% level, the normal tail produces 86 violations against 45.1 expected and is rejected for two of five assets. The three heavy-tailed specifications produce 58, 62, and 64 violations with at most one rejection, and the Diebold-Mariano test does not separate them. At the 99% level, the tail specification appears to matter more than the choice among heavy tails, which these data cannot resolve. Single-start GARCH estimation also stops at a lower optimum than a ten-start search on about one-fifth of daily asset observations.

Keywords: Backtesting, Extreme value theory, GARCH, MSCI reclassification, Value-at-Risk

1. Introduction

The Indonesia Stock Exchange halted trading for 30 minutes on the morning of October 16, 2026, after the Jakarta Composite Index opened 8.01% lower at 5,393 (Investor Trust, 2026a). Media reports attributed the fall to a consultation that MSCI opened on October 15, 2026 (New York time) on a proposal to downgrade Indonesia from Emerging Market to Standalone Market status (Kontan, 2026a; MSCI, 2026). The final decision is scheduled for no later than June 30, 2027 (MSCI, 2026). About Rp760 trillion of market capitalization evaporated at the open (Katadata, 2026a), and the index closed the day 11.57% lower at 5,185 (Investor Trust, 2026b). Over the week the index lost 14.91%, and cumulative foreign net selling in 2026 reached Rp88.19 trillion (Kabar Bursa, 2026).

The October episode was not the first. On January 27, 2026, MSCI announced interim measures for the Indonesian market: a freeze on increases in the Foreign Inclusion Factor and Number of Shares, a freeze on adding new constituents to the Investable Market Indexes, and a freeze on migration between size segments. MSCI also warned that market accessibility

This open access article is distributed under a Creative Commons Attribution (CC-BY-NC) 4.0 license.

https://doi.org/10.35877/jetech5606

ISSN : 2776-7914 (Print) / 2807-3045 (Online)

would be reviewed again if conditions did not improve (Detikcom, 2026). The index closed down 7.35% on January 28, 2026 (Katadata Databoks, 2026), and trading was halted on that day and again on January 29. With the January and October episodes, three of the five trading halts under the current government followed MSCI announcements. The other two followed fiscal concerns in March 2025 and the U.S. tariff decision in April 2025 (Katadata, 2026b).

An index provider's announcement is information that is not always contained in the price or volatility history, yet those two channels are all that a conditional volatility model observes. Its timing is also hard to predict from data. MSCI stated in June 2026 that reform progress would be assessed at the November 2026 Index Review (Kontan, 2026b), but the consultation opened earlier, on October 15 (MSCI, 2026). Studies of stock-level changes in MSCI index composition link price movements to the repositioning of index-tracking funds (Chakrabarti et al., 2005). This study extrapolates that finding to a market-level announcement as a hypothesis and does not test it.

This raises a practical question for managers of Value-at-Risk (VaR) systems for Indonesian equities: when such a shock occurs, how much of model performance is determined by the tail specification, and how much by the volatility specification? The question is hard to answer because the models being compared usually differ in several components at once. A GARCH model with Student-t innovations, for example, differs from a GARCH model with an extreme value tail both in the estimated variance series and in the quantile applied to it. Differences in backtesting results therefore cannot be attributed to either component.

The design of this study separates the two components. Three of the four models are built on a single GARCH(1,1) estimate, so their conditional variance series are numerically identical. The only difference is how the standardized residual quantile is obtained: from the normal distribution, from a generalized Pareto distribution fitted to exceedances over a threshold, or from a generalized extreme value distribution fitted to block maxima. The fourth model uses Student-t innovations and serves as a parametric benchmark.

The study makes three contributions. First, the effect of the tail specification on VaR performance is isolated for a market reclassification shock, using rolling estimation that yields 902 test days per asset rather than a short event window. Second, the two standard approaches of extreme value theory, peaks over threshold and block maxima, are compared in the same design. Third, the estimated tail index is calibrated by Monte Carlo simulation at the same sample size as the application, so that the shape parameter can be interpreted with its bias and variability in view.

2. Literature Review

2.1. *Value-at-Risk and conditional volatility*

Value-at-Risk is the loss that will not be exceeded with probability 1−p over a given horizon, and it remains the reference measure in banking supervision and internal limit systems (Jorion, 2007). For a daily return r_t and information set $\mathcal{F}_{t-1}$, VaR at tail probability p is defined by

$$P(r_t \leq \mathrm{VaR}_t(p) \mid \mathcal{F}_{t-1}) = p \qquad (1)$$

with p equal to 0.01 for 99% VaR and 0.05 for 95% VaR.

VaR accuracy depends on two separate modeling decisions. The first concerns conditional volatility dynamics, which the ARCH family of Engle (1982) and the GARCH generalization of Bollerslev (1986) handle in a standard way. For GARCH(1,1), the return and its conditional variance are modeled as

$$r_t = \mu + \varepsilon_t, \ \varepsilon_t = \sigma_t z_t, \ (\sigma_t)^2 = \omega + \alpha_1(\varepsilon_{t-1})^2 + \beta_1(\sigma_{t-1})^2 \qquad (2)$$

with $\omega>0$, $a_1\geq 0$, $\beta_1\geq 0$, $a_1+\beta_1<1$, and z_t a standardized innovation with mean zero and variance one. VaR then follows from

This open access article is distributed under a Creative Commons Attribution (CC-BY-NC) 4.0 license.

https://doi.org/10.35877/jetech5606

ISSN : 2776-7914 (Print) / 2807-3045 (Online)

$$\mathrm{VaR}_t(p) = \mu + \sigma_t z_p \qquad (3)$$

where z_p is the p-th quantile of the distribution of z_t. The second decision concerns the distribution of z_t itself: the normal assumption understates the frequency of large losses, and Bollerslev (1987) introduced the Student-t distribution precisely to address this.

2.2. *Extreme value theory*

Extreme value theory addresses the second decision in a more targeted way because it models only the tail rather than the whole distribution, and it has two standard formulations. The peaks over threshold approach fits a generalized Pareto distribution (GPD) to exceedances over a high threshold, on the basis of the limit results of Balkema and de Haan (1974) and Pickands (1975). The GPD distribution function with shape parameter ξ and scale $\beta>0$ is

$$G(y) = 1 - (1 + \xi y/\beta)^{-1/\xi}, \; y \geq 0 \qquad (4)$$

and becomes $1-\exp(-y/\beta)$ when $\xi=0$.

The block maxima approach divides the sample into blocks, takes the maximum of each block, and fits a generalized extreme value (GEV) distribution to those maxima. With location m, scale $s>0$, and shape ξ, its distribution function is

$$H(x) = \exp\{-[1 + \xi(x - m)/s]^{-1/\xi}\} \qquad (5)$$

In both formulations, $\xi>0$ indicates a heavy, power-law tail, $\xi=0$ an exponential tail, and $\xi<0$ a tail with a finite upper bound. Embrechts et al. (1997) and Coles (2001) discuss both in detail.

The two formulations use the data differently. Peaks over threshold keeps every observation above the threshold, so a period containing several large losses contributes all of them. Block maxima keeps one value per block and discards the remaining information in the block. On the same sample, the threshold method therefore generally uses more tail observations than the number of block maxima. Gilli and Këllezi (2006) applied both methods to stock index data and concluded that peaks over threshold performs better because it makes better use of the sample information.

Applying extreme value theory directly to returns ignores volatility clustering. McNeil and Frey (2000) therefore proposed a two-stage procedure: a GARCH model is estimated first, the tail model is applied to the standardized residuals, and the two are combined into a conditional quantile. This separation is what makes the comparison in this study possible, because the tail stage can be varied while the volatility stage is held fixed. In simulations with samples of 1,000 Student-t observations, McNeil and Frey report that the GPD estimator remains robust as long as the number of exceedances exceeds roughly 50. In their backtest, the rolling window contains 1,000 observations and the number of exceedances is fixed at 100.

2.3. *Backtesting*

Model adequacy is assessed by backtesting the sequence of VaR violations, and the unconditional coverage test of Kupiec (1995) examines whether the violation rate, meaning the share of days on which the loss exceeds VaR, equals the tail probability p. When N violations are observed in T days, its likelihood ratio statistic is

$$\mathrm{LR}_{uc} = -2\ln[(1-p)^{T-N}p^N] + 2\ln[(1 - N/T)^{T-N}(N/T)^N] \qquad (6)$$

which follows a chi-square distribution with one degree of freedom under the null hypothesis.

The Christoffersen (1998) test adds an independence condition, that is, whether violations arrive in clusters, and combines it with the Kupiec test into a conditional coverage statistic with two degrees of freedom. Backtests based on violation counts have low power in short samples, so forecasts are also compared through the pinball loss function at quantile level p,

This open access article is distributed under a Creative Commons Attribution (CC-BY-NC) 4.0 license.

https://doi.org/10.35877/jetech5606

ISSN : 2776-7914 (Print) / 2807-3045 (Online)

$$L_p(r_t, q_t) = (p - I(r_t < q_t))(r_t - q_t) \qquad (7)$$

with q_t the VaR forecast and I(·) the indicator function. The difference in mean loss between two models is tested with the Diebold and Mariano (1995) test of predictive accuracy. The small-sample correction of Harvey et al. (1997) is added. The original test assumes that forecasts do not depend on parameter estimation, whereas Giacomini and White (2006) extended it to forecasts produced by rolling windows, which fits this design because their framework covers estimation windows of fixed or limited size.

2.4. *Estimation reliability and window length*

One practical issue in VaR applications is the reliability of GARCH likelihood optimization. Zivot (2009) states that standard optimization does not guarantee that the global maximum is reached, and that poor starting values can make the likelihood behave badly and cause convergence problems. He also notes that β_1 is not identified when $a_1=0$, and recommends exploring the likelihood surface by perturbing starting values and re-estimating. Doornik and Ooms (2008) show that the likelihood of a GARCH regression model can be multimodal when the conditional mean contains dummy variables, and the working paper version of their study (Doornik & Ooms, 2003) notes that multimodality is more likely when the dummy is placed before or inside a turbulent period. For the exchange rate data they use, the working paper version found no multimodality in GARCH(1,1) without dummies. This study does not claim a new discovery of the phenomenon. It reports how often it occurs, and how large its effect is on rolling VaR, for Indonesian data.

Estimation window lengths differ across studies. McNeil and Frey (2000) use 1,000 observations, whereas Hendricks (1996) tests windows of 50, 125, 250, 500, and 1,250 days and finds that longer windows produce more stable forecasts under equally weighted moving average approaches. The Basel market risk framework requires an observation period of at least one year for computing VaR (Basel Committee on Banking Supervision [BCBS], 2009). Shorter windows adjust more quickly to changes in volatility but leave fewer observations for tail estimation.

3. Research Method and Materials

3.1. *Data and observation period*

Adjusted daily closing prices were obtained from Yahoo Finance for the Jakarta Composite Index (^JKSE) and the four banks with the largest market capitalization: Bank Central Asia (BBCA), Bank Rakyat Indonesia (BBRI), Bank Negara Indonesia (BBNI), and Bank Mandiri (BMRI). Prices span January 3, 2022 to October 27, 2026 and yield 1,152 daily log returns, $r_t=\ln(P_t/P_{t-1})$, from January 4, 2022 to October 27, 2026 after alignment with the exchange calendar. There are no duplicate dates or missing values. All five series match the index calendar exactly.

Two returns were checked against published reports before any model was estimated. The log return of the composite index on October 16, 2026 is −12.3004%, equivalent to a simple return of −11.57%, and on January 28, 2026 it is −7.6297%, equivalent to −7.35%. Both match the closing declines reported by Investor Trust (2026b) and Katadata Databoks (2026). Trading halt days are treated as ordinary observations.

Five periods were defined by date before any results were examined, so the grouping could not be adjusted to the findings. A period here is a date window set by the researcher, not a statistically estimated regime. The Stress2025 period covers March 18 to April 30, 2025 and contains the two trading halts of that year. The Jan2026 period covers January 28 to February 10, 2026 and contains the first MSCI episode. The Interim period covers February 11 to October 14, 2026. The Oct2026 period covers October 15 to 27, 2026 and contains the reclassification shock. The remaining days form the Baseline period. Data after October 27, 2026 are available but excluded, because the period scheme does not extend beyond that date and lengthening any period would be arbitrary.

This open access article is distributed under a Creative Commons Attribution (CC-BY-NC) 4.0 license.

https://doi.org/10.35877/jetech5606

ISSN : 2776-7914 (Print) / 2807-3045 (Online)

3.2. *Rolling estimation design*

On each test day t, the observation is treated as out-of-sample and the 250 preceding returns form the in-sample estimation window. The model in equation (2) is estimated on that window, the VaR for day t is computed from the estimated parameters, and it is compared with the realized return of day t, which was not used in estimation. The window then moves forward one day and all estimation is repeated. The first window spans January 4, 2022 to January 6, 2023 and produces the forecast for January 9, 2023.

Of the 1,152 returns, the first 250 serve only as the initial estimation sample, which leaves 902 out-of-sample test days for each asset, from January 9, 2023 to October 27, 2026. The distribution of test days across periods is 700 for Baseline, 24 for Stress2025, 10 for Jan2026, 159 for Interim, and 9 for Oct2026.

A 250-day window was chosen because it is close to one year of observations, the minimum observation period in the Basel framework (BCBS, 2009) and one of the window lengths tested by Hendricks (1996). A calendar year on the Indonesia Stock Exchange contains 236 to 245 trading days in 2022 to 2025, so 250 days slightly exceeds one year. The window is much shorter than the 1,000 observations of McNeil and Frey (2000), and this study does not compare other window lengths. As a result, a threshold at the 95th percentile leaves only about 12 exceedances, and the threshold was lowered to the 80th percentile so that exactly 50 exceedances are available, which is the lower robustness limit reported by McNeil and Frey. The choice of percentile is discussed in Section 4.2.

GARCH parameters are estimated by Gaussian quasi-maximum likelihood (QML) for two reasons. First, Gaussian QML gives consistent estimates of the conditional variance parameters even when the true innovations are not normal (Bollerslev & Wooldridge, 1992), and McNeil and Frey (2000) used it, under the name pseudo-maximum likelihood, in the first stage of their two-stage procedure. Second, the estimation does not depend on the tail model, so the normal, peaks over threshold, and block maxima models can share one conditional variance series. QML is less efficient than the correct likelihood when the true innovations are non-normal (Engle & González-Rivera, 1991), and this study does not measure the size of that loss at 250 observations. Model M2, estimated with a Student-t likelihood, serves as the benchmark for its consequences.

Every daily estimation is run by maximum likelihood from ten different starting values, and the estimate with the highest log-likelihood is kept. This follows the advice of Zivot (2009). The GARCH likelihood on short windows can have more than one local maximum. The full rationale and the timing of its introduction are given in Section 4.4.

3.3. *Four tail specifications*

The four models differ only in how the quantile z_p in equation (3) is obtained.

Model M1 uses a normal tail: z_p is the p-th quantile of the standard normal distribution, which is constant over time. The parameters are estimated by Gaussian QML, and this specification is the reference for the heavy-tailed alternatives.

Model M2 uses a Student-t tail. GARCH is re-estimated with Student-t innovations, the degrees of freedom are estimated jointly with the variance parameters each day, and z_p is taken from the estimated standardized t distribution, so only M2 has its own variance series.

Model M3 uses peaks over threshold, with L=−z the standardized residual loss. The threshold u is set at the 80th percentile of L within the estimation window, which leaves exactly 50 exceedances. The GPD in equation (4) is fitted to those exceedances by maximum likelihood, and the loss quantile at tail probability p is obtained from

$$L_p = u + (\beta/\xi)[(p/\lambda)^{-\xi} - 1], \quad \lambda = N_u/N \qquad (8)$$

This open access article is distributed under a Creative Commons Attribution (CC-BY-NC) 4.0 license.

https://doi.org/10.35877/jetech5606

ISSN : 2776-7914 (Print) / 2807-3045 (Online)

where N_u is the number of exceedances, N the sample size, and $z_p = -L_p$. The exponential form is used when ξ is numerically zero.

Model M4 uses block maxima. The same loss L is divided into non-overlapping blocks of n days. Blocks are counted backward from the most recent day in the estimation window so that the latest block is always complete, and the oldest days that do not fill a block are discarded. The GEV in equation (5) is fitted to the block maxima by maximum likelihood. Because the maximum of n independent observations has distribution function F^n, the daily loss quantile at tail probability p is read from the estimated distribution at probability $(1-p)^n$, that is,

$$L_p = m + (s/\xi)[(-n\ln(1-p))^{-\xi} - 1] \qquad (9)$$

The specification used is weekly blocks with n=5, which produces 50 maxima in a 250-day window. This number was deliberately matched to the number of exceedances in M3 so that the two tail estimators work with the same number of tail observations.

Models M1, M3, and M4 are built on the same GARCH-normal estimate for every asset and day, so μ and σ_t in equation (3) are identical across the three and only the residual quantile changes. This was verified rather than assumed. The values of μ and σ_t were recovered from the two VaR levels of M1, whose normal quantiles are known constants, and then recombined with the stored tail parameters. That procedure reproduced the VaR series of M3 and M4 with a maximum difference of zero at a tolerance of 10^{-10} for all 902 days on all five assets.

3.4. *Backtesting and forecast comparison*

For each asset, model, confidence level, and period, the violation sequence is tested in four ways: the Kupiec unconditional coverage statistic in equation (6), an exact binomial test, the Christoffersen independence test, and the Christoffersen conditional coverage test. The exact binomial test is run alongside the Kupiec statistic because the chi-square approximation is less reliable when the number of violations is small. Tests on periods with fewer than 30 days are flagged as low-power and are not read as evidence in either direction. The significance level is 5%.

Forecasts are compared through the pinball loss in equation (7). The test is the Diebold and Mariano (1995) test with the Harvey et al. (1997) correction. Comparisons in which one model records fewer than five violations are excluded, because with so few violations the loss difference is dominated by the difference in the width of the two VaRs and the test statistic is uninformative. This criterion is a choice made by the researcher.

3.5. *Calibration of the tail index*

A shape parameter estimated from a finite sample is not the tail index of the underlying distribution, and the gap can be large. To make the estimated values interpretable, a Monte Carlo simulation was run at the same sample size as the application. A total of 1,500 samples, each containing 250 independent standardized Student-t observations with five degrees of freedom, were generated, so the true tail index equals 0.200 by construction. Each tail estimator was then applied exactly as in the main analysis. The simulation records the bias and sampling standard deviation of each estimator, the probability of obtaining a negative estimate, and the expected difference between the block maxima and peaks over threshold estimators under the same true tail index. A second simulation repeated the exercise for three to thirty degrees of freedom, producing a mapping from the true tail index to the expected estimate.

4. Results and Discussion

4.1. *Results*

4.1.1. *Anatomy of the shock*

This open access article is distributed under a Creative Commons Attribution (CC-BY-NC) 4.0 license.

https://doi.org/10.35877/jetech5606

ISSN : 2776-7914 (Print) / 2807-3045 (Online)

Table 1 presents daily simple returns over the event window. The loss is concentrated on a single day. On October 16 the four banks fell between 12.90% (BBNI) and 14.84% (BBRI), and three of them (BBCA, BBRI, and BMRI) stopped at around −14.8%. That figure coincides with the 15% lower auto rejection limit that has applied to all price ranges since April 8, 2025 (Hukumonline, 2025). The uniform decline therefore may reflect the price limit and cannot be read as a judgment on each bank, and the loss that sellers actually wanted to realize may have been larger. Prices reversed on October 20: the index rose 5.65% and the four banks rose between 7.88% and 9.84%.

Table 1: Daily simple returns over the event window (%). Trading was halted on October 16, 2026.

Date	JKSE	BBCA	BBRI	BBNI	BMRI
Oct 15, 2026	−1.33	−0.82	0.32	0.00	0.51
Oct 16, 2026	−11.57	−14.81	−14.84	−12.90	−14.82
Oct 19, 2026	−3.64	−6.86	−3.79	−0.34	−0.88
Oct 20, 2026	5.65	7.88	9.84	8.78	9.82
Oct 21, 2026	0.94	1.44	0.72	0.93	0.27
Oct 22, 2026	1.12	−1.90	3.20	4.92	4.05
Oct 23, 2026	0.95	−2.90	0.34	0.29	1.30
Oct 26, 2026	−1.22	−1.69	0.00	1.46	0.26
Oct 27, 2026	−1.99	−0.81	−0.34	−4.61	−1.02
Cumulative, Oct 15 to 27	−11.39	−20.00	−6.15	−2.93	−2.27

Note: Author's calculations from adjusted closing prices from Yahoo Finance. Model estimation uses log returns, which on October 16 equal −12.30% (JKSE), −16.03% (BBCA), −16.06% (BBRI), −13.82% (BBNI), and −16.05% (BMRI).

Recovery was far less uniform than the decline. In the seven trading days after the day of the decline, the index recovered 1.56% in simple terms, BMRI rose 14.16%, BBNI 11.45%, and BBRI 9.85%, while BBCA fell another 5.31%. The nine-day cumulative return therefore ranges from −2.27% for BMRI to −20.00% for BBCA, even though the losses on the day of the decline were almost the same. The BBCA difference is reported as an observation, because explaining it would require order flow or ownership data that this study does not use.

The daily standard deviation of log returns over October 15 to 27 is 2.8 to 3.3 times the standard deviation of the Interim period, but much of that gap comes from October 16, and without October 15 and 16 the ratio falls to 1.7 to 2.0. The models therefore face one extreme loss followed by several days of volatility roughly twice as high.

4.1.2. Out-of-sample violations

Table 2 reports 99% VaR violations over the entire test period, with 9.02 violations per asset expected, and the normal tail produces an excess of violations. Model M1 produces 86 violations in aggregate against 45.1 expected, is rejected by the Kupiec test for two of five assets, and does worst on the composite index, with 27 violations against 9.0 expected.

Table 2: 99% VaR violations over 902 out-of-sample test days. Kupiec test p-values in parentheses; expected violations are 9.02 per asset.

Asset	M1 Normal	M2 Student-t	M3 POT/GPD	M4 BM/GEV
JKSE	27 (<0.001)	19 (0.004)	13 (0.212)	15 (0.068)
BBCA	15 (0.068)	9 (0.995)	14 (0.123)	14 (0.123)
BBRI	18 (0.008)	12 (0.343)	13 (0.212)	14 (0.123)
BBNI	13 (0.212)	9 (0.995)	11 (0.522)	10 (0.747)
BMRI	13 (0.212)	9 (0.995)	11 (0.522)	11 (0.522)
Total (expected 45.1)	86	58	62	64
Kupiec rejections	2/5	1/5	0/5	0/5

Note: Author's calculations. Rejections are computed at the 5% level. The exact binomial test gives the same rejections.

The three heavy-tailed specifications produce fewer violations and are hard to separate. The totals are 58 for Student-t, 62 for peaks over threshold, and 64 for block maxima, with one Kupiec rejection for the first and none for the other two. The composite index again gives the sharpest contrast: on the same conditional variance series, the normal quantile is rejected

This open access article is distributed under a Creative Commons Attribution (CC-BY-NC) 4.0 license.

https://doi.org/10.35877/jetech5606

ISSN : 2776-7914 (Print) / 2807-3045 (Online)

with a p-value below 0.001, whereas the GPD and GEV quantiles are not rejected, although the difference between rejection and non-rejection is itself untested. The 24-violation gap between M1 and M3 comes largely from the composite index (14 violations), while across the four banks the gap is 10 violations. The Student-t model is rejected on the composite index with a p-value of 0.004. Because M2 uses its own variance series, its difference from the other three models cannot be attributed to the tail alone.

The three heavy-tailed models still produce more violations than the 45.1 expected. The Christoffersen independence test rejects only one of 40 tests (five assets, four models, two confidence levels), so there is no evidence that violations cluster once conditional volatility is accounted for, although the test has low power with 9 to 27 violations per series. The conditional coverage test rejects five tests, four of which involve M1 or M2 on JKSE and BBRI. Because the conditional coverage statistic is the sum of the unconditional coverage and independence statistics, these rejections come mainly from the excess number of violations and not from clustering. One exception is M1 on BBRI, where both the unconditional coverage and the independence statistics reject.

Table 3 summarizes the 95% results and the robustness checks. At the 95% confidence level the total violations are 240, 255, 240, and 258 against 225.5 expected. The Kupiec test rejects one of five assets for Student-t (JKSE) and for block maxima (BBNI), and none for the normal and peaks over threshold models, so the normal tail is not a problem at this confidence level. Loss-based comparisons support no ranking: of 300 pairwise comparisons (five assets, two confidence levels, five model pairs, and six time scopes, namely the full test period plus five periods), 123 have enough power to be tested, and only 7 of them (5.7%) are significant at the 5% level, roughly the proportion expected by chance, although these tests are not independent. The direction of the differences is mixed, because the normal model is better in three of those seven significant comparisons. On this basis the four forecasts cannot be distinguished by accuracy, and only the number of violations at the 99% confidence level separates the models, and that only descriptively.

Table 3: Pooled VaR 95% results and robustness checks across the five assets. Expected VaR 95% violations are 225.5. The Christoffersen, DM, and shape rows pool both confidence levels where applicable.

Measure	M1 Normal	M2 Student-t	M3 POT/GPD	M4 BM/GEV
VaR 95% violations	240	255	240	258
Kupiec rejections, VaR 95%	0/5	1/5	0/5	1/5
Christoffersen CC rejections	2/10	2/10	0/10	1/10
Significant DM comparisons (of 123)	4	0	4	6
Mean estimated ξ	–	–	−0.050	0.002
Failed tail estimations (days)	0	0	0	0

Note: Author's calculations. The DM row counts significant comparisons at the 5% level that involve the model concerned, and each comparison involves two models. The ξ row gives the mean ξ for M3 and M4, whereas M1 and M2 have no extreme value shape parameter.

The 80th percentile threshold for M3 was chosen so that the number of exceedances meets the McNeil and Frey (2000) robustness limit, namely more than about 50. That limit comes from a sample of 1,000 observations, so 50 exceedances equal 5% of the sample, whereas in a 250-day window the same number equals 20%. In theory, a relatively low threshold adds bias to tail estimation, while a high threshold adds variance (Coles, 2001). In a 250-day window the 90th percentile yields only 25 exceedances, and the consequence is visible directly in the data. The 90th percentile specification suffers GPD estimation failures on eight asset-days, seven of them on BBCA in mid-2023, while the 80th percentile never fails. On JKSE, the 90th percentile yields 16 violations with a Kupiec p-value of 0.035 and is rejected, whereas the 80th percentile yields 13 violations with a p-value of 0.212 and is not rejected. The Monte Carlo simulation in Section 4.5 adds another reason: the sampling standard deviation of the shape estimate falls from 0.301 with 25 exceedances to 0.185 with 50. The simulation does not show a larger bias at the 80th percentile (mean −0.025 versus −0.041 at the 90th percentile). The lowering of this threshold was decided after the results for the 250-day window had been seen, although its literature basis existed beforehand.

This open access article is distributed under a Creative Commons Attribution (CC-BY-NC) 4.0 license.

https://doi.org/10.35877/jetech5606

ISSN : 2776-7914 (Print) / 2807-3045 (Online)

Figure 1 presents the 99% VaR of the four models for the composite index. In calm periods the four series nearly coincide. In 2026, the M1 VaR is narrowest, M2 lies in the middle, and M3 and M4 nearly coincide and are widest, with mean 99% VaR over the Interim period of −0.036, −0.042, −0.052, and −0.050 for M1 to M4. All series widen when volatility rises and widen very sharply after October 16, 2026.

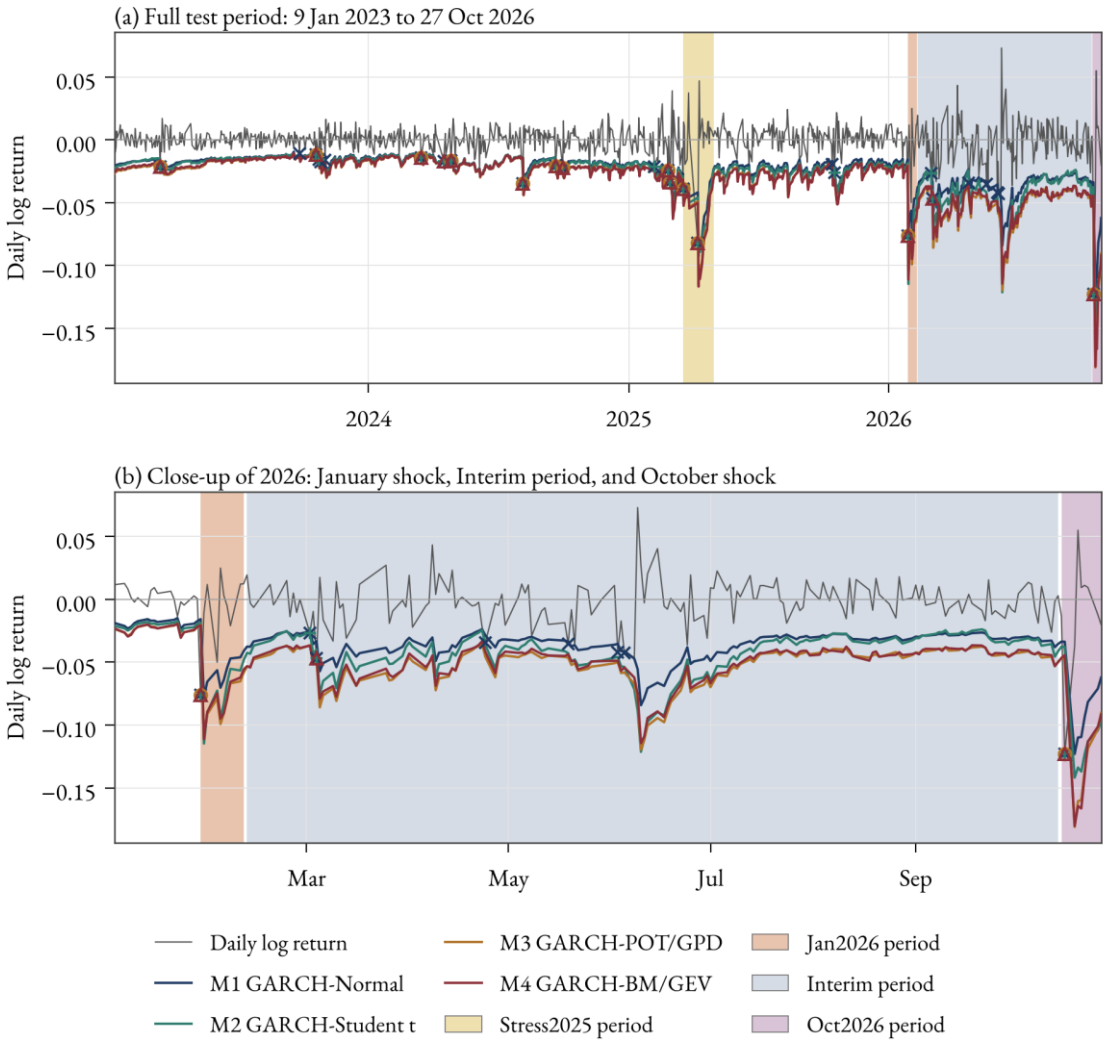


Figure 1: Realized log returns and 99% VaR of the four rolling models for the Jakarta Composite Index. Markers show violations (× for M1, + for M2, circle for M3, triangle for M4), and the unshaded area is the Baseline period.

4.1.3. Behavior within the event window

Within the event window, backtesting cannot be informative, and the reason lies in the data, not in the test statistic. The October 16 log returns (−12.30% to −16.06%, see the note to Table 1) are far below the 99% VaR of the same day (Table 4: −2.61% to −5.05% for the peaks over threshold model), and the same holds for M1, M2, and M4. All 20 asset-model combinations therefore record exactly one 99% VaR violation in the window, and that violation falls on October 16, like the JKSE markers at the right end of Figure 1. The five assets and four models consequently do not provide independent observations. At the 95% confidence level, 20 of the 23 violations in the window fall on October 16, and the other three on October 27.

This open access article is distributed under a Creative Commons Attribution (CC-BY-NC) 4.0 license.

https://doi.org/10.35877/jetech5606

ISSN : 2776-7914 (Print) / 2807-3045 (Online)

What the event window shows is model behavior on both sides of the shock (Table 4). On October 16, the 99% VaR covers only 18.7% to 38.2% of the realized loss, so the forecast was too narrow by a factor of about 2.6 to 5.4. For the four banks, the measured loss is itself truncated by the 15% price limit, so this figure is an upper bound on coverage. The widest VaR of the window appears afterward, on October 19 for 13 of the 20 combinations and on October 21 for the other 7, at 2.8 to 6.5 times the average pre-shock width, when realized losses were much smaller. On October 27, seven trading days after the day of the decline, 19 of the 20 combinations had not returned within 1.5 times their pre-shock width, and only BBRI with the Student-t model recovered, five trading days after its peak.

Table 4: Behavior of 99% VaR around the shock, peaks over threshold model. The range in the last row covers all four models.

Asset	VaR Oct 16	Share of loss covered	Widest VaR	Ratio to pre-shock	Recovered before Oct 27
JKSE	−0.0470	0.382	−0.1811 (Oct 19)	4.02	no
BBCA	−0.0505	0.315	−0.2751 (Oct 19)	6.09	no
BBRI	−0.0406	0.253	−0.1336 (Oct 21)	3.08	no
BBNI	−0.0261	0.189	−0.1475 (Oct 19)	5.08	no
BMRI	−0.0334	0.208	−0.1136 (Oct 21)	3.20	no
Range, all models	–	0.187 to 0.382	–	2.79 to 6.51	1 of 20 recovered

Note: Author's calculations. VaR and returns are in log units. The pre-shock base is the mean 99% VaR over October 1 to 14, 2026, and recovered means VaR returned within 1.5 times that base.

4.1.4. Instability of the GARCH likelihood optimum

Every daily GARCH estimation was run from ten starting values, and the estimate with the highest log-likelihood was kept. This procedure was added after two earlier estimation rounds produced different 99% VaR values for BBCA on October 19, 2026 from the same data window, namely −0.2226 and −0.0616. Re-estimation from scratch showed that the likelihood for that window is multimodal, and the single-start estimation stopped at a local optimum that still reported a successful convergence code. All results reported here come from the final ten-start estimation, and the single-start result is still recorded each day as a comparison.

The estimation window for October 19 ended on October 16, when the BBCA log return reached −16.03%. The single-start estimate gives an ARCH coefficient of exactly zero, persistence of 0.999, a log-likelihood of 584.95, and σ_t of 0.0260. The best of the ten starts gives an ARCH coefficient of 0.3443, persistence of 0.825, a log-likelihood of 600.93, and σ_t of 0.0948. An ARCH coefficient of zero means the variance equation does not respond to returns, so that estimate ignores the 16% decline inside its own window. The optimum also lies on the boundary of the parameter space (a_1=0), where β_1 is not identified (Zivot, 2009). Both estimates report the same convergence code.

Table 5 shows that this phenomenon is not rare. In 20.8% of asset-days, the best of ten starts beats the single-start estimate by a log-likelihood difference of more than one. The frequency is 16% to 22% in the Baseline, Jan2026, and Interim periods, 0% in Stress2025, and 40% in Oct2026. The last figure comes from 18 of 45 estimations, that is, 9 days and 5 correlated assets, so it cannot be read as independent observations. What differs markedly is the size of the consequence: the median relative range of σ_t across the ten estimates is 0.0018 in the Baseline period and 0.1802 in the Oct2026 period, one hundred times larger. A suboptimal single-start estimate therefore occurs in four of the five periods, but only in the Oct2026 period does it materially change the volatility estimate. That is the period when the volatility estimate is needed most.

Table 5: Instability of the GARCH likelihood optimum by period, from ten starting values in each daily estimation. The total number of estimations is 4,510.

This open access article is distributed under a Creative Commons Attribution (CC-BY-NC) 4.0 license.

https://doi.org/10.35877/jetech5606

ISSN : 2776-7914 (Print) / 2807-3045 (Online)

Period	Number of estimations	Share with log-likelihood gain above 1 (%)	Median relative σ range	90th percentile of relative σ range
Baseline	3,500	22.3	0.0018	0.1227
Stress2025	120	0.0	0.0014	0.2349
Jan2026	50	16.0	0.0018	0.1510
Interim	795	16.5	0.0001	0.1173
Oct2026	45	40.0	0.1802	0.5858
Full sample	4,510	20.8	0.0016	0.1253

Note: Author's calculations. The relative range of σ is the difference between the largest and smallest σ among the ten estimates, divided by the largest σ.

This phenomenon is already known in the literature. Zivot (2009) states that a global maximum is not guaranteed and recommends searching by perturbing starting values, and Doornik and Ooms (2003, 2008) report likelihood multimodality in GARCH regression models with dummy variables. The contribution here is quantification in rolling VaR applications. The consequence also propagates through the shared-estimation design, because an error in one estimation affects M1, M3, and M4 at once.

The influence of flagged days on the backtesting conclusions was checked by excluding them and recomputing. On average, 220 days were excluded from each test. Of the 40 tests, only one decision changes: M2 on JKSE at VaR 99% moves from rejected to not rejected. The main conclusions about the normal tail and heavy tails therefore do not change qualitatively when those days are excluded.

4.1.5. The tail index and its reliability

Averaged across all assets and days, the GPD shape parameter is −0.050 (58% of days negative) and the GEV shape parameter is 0.002 (50% of days negative), values that are small and mostly negative. Taken at face value, they suggest a tail that is close to exponential or even bounded.

The simulation in Section 3.5 shows that these values should not be taken at face value. With a true implied index of 0.200, the peaks over threshold estimator with 50 exceedances gives a mean of −0.025 (standard deviation 0.185), the block maxima estimator gives a mean of 0.000 (standard deviation 0.126), and the peaks over threshold estimator with 25 exceedances gives a mean of −0.041 (standard deviation 0.301). All three underestimate the true index by about 0.20 to 0.24, and 50% to 54% of estimates are negative even though the true tail is heavy. The expected difference of the block maxima estimator over the peaks over threshold estimator at the same true index is 0.025 with a standard deviation of 0.127, so the observed difference in the data (0.002 versus −0.050, that is 0.053) remains within the range of single-sample sampling variability.

The mapping from the true index to the estimated value in the second simulation allows a reverse reading. The mean peaks over threshold estimate of −0.050 lies between the simulation results for five and six degrees of freedom (−0.025 and −0.064), and linear interpolation gives a true index of about 0.18. The block maxima value of 0.002 is consistent with five degrees of freedom (0.000) and gives about 0.20. Both estimators therefore support a heavy tail with an index of about 0.18 to 0.20, equivalent to a Student-t distribution with about five degrees of freedom, and the conclusion that the tail is close to exponential is not supported. This mapping assumes Student-t innovations, so the result is a calibration under one family of distributions and not an assumption-free measurement of the tail index.

4.2. Discussion

The price pattern in this episode does not resemble a fundamental banking crisis, although this study does not test its cause. The loss is concentrated on one day, three of the four banks stopped near the lower price limit, and prices reversed sharply two trading days later. The pattern may be consistent with index-mandate-driven repricing, the mechanism that Chakrabarti et al. (2005) link to changes in MSCI index composition at the stock level. That

This open access article is distributed under a Creative Commons Attribution (CC-BY-NC) 4.0 license.

https://doi.org/10.35877/jetech5606

ISSN : 2776-7914 (Print) / 2807-3045 (Online)

study did not examine market-level announcements, so this interpretation is only an extrapolation and is not tested here, and the price limit means the uniform decline of the banks cannot serve as evidence for it.

This type of risk lies outside the information set of every model studied. Conditional volatility models learn from past prices and volatility, whereas an index provider's decision does not always reach the market through those two channels before it is announced. MSCI had already flagged November 2026 as a checkpoint (Kontan, 2026b), but the consultation opened earlier. A risk framework that uses only return history treats such an event as a draw from a heavy tail. Yet the process has a known decision timetable: the consultation is still running and the final decision falls no later than June 30, 2027 (MSCI, 2026).

Because M1, M3, and M4 share one conditional variance series, the results separate the contribution of the tail specification from that of the volatility specification. The normal quantile is inadequate for these data at the 99% level: it produces nearly twice the expected number of violations and is rejected for two of five assets (JKSE and BBRI), whereas at the 95% level it is not rejected for any asset. The Monte Carlo calibration separately indicates that the underlying tail is heavy, with an implied index of about 0.18 to 0.20. Replacing the normal quantile with a GPD or GEV quantile, while keeping the same variance series, removes more than half of the excess violations, from 86 to 62 and 64 against 45.1 expected.

The tail specification does not determine which heavy-tailed alternative should be used. The three differ by only a few violations over 902 days, and almost no Diebold-Mariano comparison separates them. The first modeling decision, abandoning the normal tail, therefore appears to matter more than the second, choosing among Student-t, peaks over threshold, and block maxima. This sample cannot resolve the second choice, and a claim that one option is superior on the basis of a single event should be treated with caution.

For risk management, post-shock behavior may matter more than the shock itself. The 99% VaR covers only 18.7% to 38.2% of the loss on October 16, the type of failure that backtesting is designed to detect. The VaR then widens to 2.8 to 6.5 times its pre-shock level on the following days and stays wide to the end of the window. This widening is a consequence of equation (2), which treats one extreme return as a persistent increase in volatility. It is a mismatch in the opposite direction, a VaR far wider than realized volatility (1.7 to 2.0 times the Interim level), and no coverage test reports it. A limit system driven by this number could mechanically force position reductions into a recovering market. That effect is not measured in this study and is not captured by any statistic in Table 2.

Several limitations constrain these results. The Stress2025, Jan2026, and Oct2026 periods contain 24, 10, and 9 test days respectively, so no test can establish calibration within them, and all statements about the event window are descriptive. Violations are concentrated on the same date across the five assets, so the five assets should not be counted as independent evidence. The GPD threshold was set at the 80th percentile without a threshold stability analysis. The move down from the 90th percentile was decided after the results were seen, and a threshold that low adds bias to the estimate of ξ in theory, although the simulation does not show it. The daily price limit also truncates the left tail of individual stock returns, so the tail estimates for the four banks may be biased toward negative values, and this is not corrected. The periods are date windows, not identified treatments, so no causal claim is made about the MSCI announcement. The MSCI consultation has not ended, and the post-shock window is only nine days. The ten-start search is also incomplete: in 66% of asset-days fewer than ten estimations converged, with a minimum of four, so a global optimum is not guaranteed even after this procedure. This study covers VaR only and does not report Expected Shortfall, which in the Basel framework for market risk replaces VaR as the measure for stressed conditions (BCBS, 2016).

5. Conclusion

This open access article is distributed under a Creative Commons Attribution (CC-BY-NC) 4.0 license.

https://doi.org/10.35877/jetech5606

ISSN : 2776-7914 (Print) / 2807-3045 (Online)

One day after MSCI announced a consultation on a proposed reclassification on October 15, 2026 (New York time), the Jakarta Composite Index fell 11.57% on October 16, 2026 and trading was halted, the third trading halt following an MSCI announcement under the current government. This study does not identify a causal relationship. In this episode, the tail specification appears to matter more at the 99% confidence level, whereas the choice among heavy-tailed alternatives does not. On the same conditional variance series, the normal quantile produces 86 violations against 45.1 expected over 902 test days and is rejected for two of five assets. The GPD and GEV quantiles on the same series produce 62 and 64 violations without rejection, and the Student-t model, with its own variance series, produces 58 violations with one rejection. No loss-based test separates the three. The evidence therefore supports abandoning the normal tail at the 99% confidence level but does not support favoring any one heavy-tailed treatment.

Three other findings go beyond model selection. First, no model anticipated the shock, and every model widened its VaR far beyond the loss that occurred afterward: the 99% VaR covered only 18.7% to 38.2% of the loss on the day of the event, then widened to several times its normal level until the end of the window. Second, estimated shape parameters cannot be read at face value: at this sample size the three estimators underestimate the true index by about 0.20 to 0.24, and the discrepancy between the two extreme value estimators is within what single-sample sampling variability would produce. Third, single-start GARCH estimation stops at a lower optimum than a ten-start search on about one-fifth of asset-days, with a frequency of 40% in the Oct2026 period against 16% to 22% in the other periods, and the median relative range of σ is one hundred times larger after the shock (45 correlated estimations).

The implication concerns how index classification risk is treated. An index provider's decision reaches the market as an announcement, not through the price and volatility history that models observe, and the final MSCI decision is still pending, due no later than June 30, 2027. Treating index provider decision dates as known risk events in stress tests may be more useful for risk management in Indonesia than further refinement of the tail, although this study has not tested that claim. Future research could extend the analysis to other markets undergoing reclassification review, add Expected Shortfall and its backtesting, and examine the behavior of risk limits in the weeks after a shock.

References

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

Basel Committee on Banking Supervision. (2009). *Revisions to the Basel II market risk framework*. Bank for International Settlements. Retrieved November 18, 2026, from https://www.bis.org/publ/bcbs158.htm

Basel Committee on Banking Supervision. (2016). *Minimum capital requirements for market risk*. Bank for International Settlements. Retrieved November 18, 2026, from https://www.bis.org/bcbs/publ/d352.pdf

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

Bollerslev, T. (1987). A conditionally heteroskedastic time series model for speculative prices and rates of return. *The Review of Economics and Statistics, 69*(3), 542–547. https://doi.org/10.2307/1925546

Bollerslev, T., & Wooldridge, J. M. (1992). Quasi-maximum likelihood estimation and inference in dynamic models with time-varying covariances. *Econometric Reviews, 11*(2), 143–172. https://doi.org/10.1080/07474939208800229

This open access article is distributed under a Creative Commons Attribution (CC-BY-NC) 4.0 license.

https://doi.org/10.35877/jetech5606

ISSN : 2776-7914 (Print) / 2807-3045 (Online)

Chakrabarti, R., Huang, W., Jayaraman, N., & Lee, J. (2005). Price and volume effects of changes in MSCI indices: Nature and causes. *Journal of Banking & Finance, 29*(5), 1237–1264. https://doi.org/10.1016/j.jbankfin.2004.04.002

Christoffersen, P. F. (1998). Evaluating interval forecasts. *International Economic Review, 39*(4), 841–862. https://doi.org/10.2307/2527341

Coles, S. (2001). *An introduction to statistical modeling of extreme values*. Springer. https://doi.org/10.1007/978-1-4471-3675-0

Detikcom. (2026, January 28). *Pengumuman MSCI picu keruntuhan IHSG 6%, ini isinya* [MSCI announcement triggers 6% JCI collapse, here is what it says]. Retrieved November 18, 2026, from https://finance.detik.com/bursa-dan-valas/d-8328839/pengumuman-msci-picu-keruntuhan-ihsg-6-ini-isinya

Diebold, F. X., & Mariano, R. S. (1995). Comparing predictive accuracy. *Journal of Business & Economic Statistics, 13*(3), 253–263. https://doi.org/10.1080/07350015.1995.10524599

Doornik, J. A., & Ooms, M. (2003). *Multimodality in the GARCH regression model* (Economics Working Papers). Nuffield College, University of Oxford. Retrieved November 18, 2026, from https://www.nuff.ox.ac.uk/economics/papers/2003/W20/GarchMode1.pdf

Doornik, J. A., & Ooms, M. (2008). Multimodality in GARCH regression models. *International Journal of Forecasting, 24*(3), 432–448. https://doi.org/10.1016/j.ijforecast.2008.06.002

Embrechts, P., Klüppelberg, C., & Mikosch, T. (1997). *Modelling extremal events for insurance and finance*. Springer. https://doi.org/10.1007/978-3-642-33483-2

Engle, R. F. (1982). Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation. *Econometrica, 50*(4), 987–1007. https://doi.org/10.2307/1912773

Engle, R. F., & González-Rivera, G. (1991). Semiparametric ARCH models. *Journal of Business & Economic Statistics, 9*(4), 345–359. https://doi.org/10.1080/07350015.1991.10509863

Giacomini, R., & White, H. (2006). Tests of conditional predictive ability. *Econometrica, 74*(6), 1545–1578. https://doi.org/10.1111/j.1468-0262.2006.00718.x

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

Harvey, D., Leybourne, S., & Newbold, P. (1997). Testing the equality of prediction mean squared errors. *International Journal of Forecasting, 13*(2), 281–291. https://doi.org/10.1016/S0169-2070(96)00719-4

Hendricks, D. (1996). Evaluation of value-at-risk models using historical data. *Economic Policy Review, 2*(1), 39–69. Retrieved November 18, 2026, from https://www.newyorkfed.org/medialibrary/media/research/epr/96v02n1/9604hend.html

Hukumonline. (2025, April 8). *BEI lakukan penyesuaian aturan perdagangan efek* [IDX makes adjustments to securities trading rules]. Retrieved November 18, 2026, from https://www.hukumonline.com/berita/a/bei-lakukan-penyesuaian-aturan-perdagangan-efek-lt67f4abb2a449a

Investor Trust. (2026a, October 16). *BEI langsung trading halt, dibuka anjlok 8,01%* [IDX immediately halts trading, opens down 8.01%]. Retrieved November 18, 2026, from https://investortrust.id/market/119480/bei-terapkang-trading-halt-dibuka-anjlok-8-01

Investor Trust. (2026b, October 16). *IHSG akhirnya ditutup jatuh 11,57% ke level terendah baru dalam 6 tahun terakhir, semua sektor anjlok* [JCI finally closes down 11.57% to a new six-year low, all sectors fall]. Retrieved November 18, 2026, from https://investortrust.id/market/119543/ihsg-akhirnya-ditutup-jatuh-11-57-ke-level-terendah-baru-dalam-6-tahun-terakhir-semua-sektor-anjlok

Jorion, P. (2007). *Value at risk: The new benchmark for managing financial risk* (3rd ed.). McGraw-Hill.

Kabar Bursa. (2026, October 17). *IHSG ambruk 14,91 persen sepekan, jualan asing tembus Rp88,19 triliun sepanjang 2026* [JCI collapses 14.91 percent in a week, foreign selling reaches Rp88.19 trillion throughout 2026]. Retrieved November 18, 2026, from https://www.kabarbursa.com/market-

This open access article is distributed under a Creative Commons Attribution (CC-BY-NC) 4.0 license.

https://doi.org/10.35877/jetech5606

ISSN : 2776-7914 (Print) / 2807-3045 (Online)

hari-ini/ihsg-ambruk-1491-persen-sepekan-jualan-asing-tembus-rp8819-triliun-sepanjang-2026

Katadata. (2026a, October 16). *IHSG dibuka trading halt 8,01% imbas MSCI, kapitalisasi pasar menguap Rp 760 T* [JCI opens with 8.01% trading halt due to MSCI, market capitalization evaporates Rp 760 T]. Retrieved November 18, 2026, from https://katadata.co.id/finansial/bursa/6ad1890554503/ihsg-dibuka-trading-halt-8-01-imbas-msci-kapitalisasi-pasar-menguap-rp-760-t

Katadata. (2026b, October 21). *2 tahun Prabowo-Gibran: IHSG ambruk 31,42% dan 5 kali trading halt* [Two years of Prabowo-Gibran: JCI collapses 31.42% and five trading halts]. Retrieved November 18, 2026, from https://katadata.co.id/finansial/bursa/6ad89444e4898/2-tahun-prabowo-gibran-ihsg-ambruk-31-42-dan-5-kali-trading-halt

Katadata Databoks. (2026, January 28). *The Jakarta Composite Index closed down 7.35% following the MSCI announcement (Wednesday, January 28, 2026)*. Retrieved November 18, 2026, from https://databoks.katadata.co.id/en/market/statistics/6979df85991ed/the-jakarta-composite-index-closed-down-735-following-the-msci-announcement-wednesday-january-28-2026

Kontan. (2026a, October 16). *Inilah isi lengkap pengumuman MSCI untuk pasar saham Indonesia* [Here is the full content of the MSCI announcement for the Indonesian stock market]. Retrieved November 18, 2026, from https://investasi.kontan.co.id/news/inilah-isi-lengkap-pengumuman-msci-untuk-pasar-saham-indonesia

Kontan. (2026b, June 24). *MSCI 2026 market classification review: Official announcement (23 June 2026)*. Retrieved November 18, 2026, from https://english.kontan.co.id/news/msci-2026-market-classification-review-official-announcement-23-june-2026

Kupiec, P. H. (1995). Techniques for verifying the accuracy of risk measurement models. *The Journal of Derivatives, 3*(2), 73–84. https://doi.org/10.3905/jod.1995.407942

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

MSCI. (2026, October 15). *MSCI launches a consultation on a proposal to reclassify Indonesia from Emerging Market to Standalone Market status* [Press release]. Retrieved November 18, 2026, from https://www.msci.com/downloads/documents/press-releases/media-room/Indonesia_PR_Oct_26.pdf

Pickands, J. (1975). Statistical inference using extreme order statistics. *The Annals of Statistics, 3*(1), 119–131. https://doi.org/10.1214/aos/1176343003

Zivot, E. (2009). Practical issues in the analysis of univariate GARCH models. In T. G. Andersen, R. A. Davis, J.-P. Kreiß, & T. Mikosch (Eds.), *Handbook of financial time series* (pp. 113–155). Springer. https://doi.org/10.1007/978-3-540-71297-8_5

This open access article is distributed under a Creative Commons Attribution (CC-BY-NC) 4.0 license.