JURNAL BISNIS, MANAJEMEN & PERBANKAN

RESEARCH ARTICLE
Published: Sept, 29, 2026
doi: 10.21070/jbmp.v12i2.2388

Regime-dependent risk-return trade-off in ASEAN equity markets: Evidence from higher-order moments and threshold regression

Mohammad Syifaul Qulub[1]*

[1] *Faculty of Islamic Economics and Business, Universitas Islam Negeri Sunan Ampel, Indonesia*

Abstract

This study aims to examine the risk-return trade-off across five major ASEAN equity markets Indonesia, Malaysia, the Philippines, Thailand, and Singapore using daily market data spanning from 2006 to 2025. Utilizing an advanced quantitative approach, the research employs market-specific GARCH models with skewed Student-t innovations to calculate lagged volatility, skewness, and kurtosis, which are subsequently evaluated using linear and threshold regressions conditioned on prior-period returns. The results reveal highly localized market dynamics rather than a uniform regional trend: significant and robust skewness premiums alongside regime-dependent volatility pricing are evident only in Indonesia, with Malaysia displaying a similar but weaker pattern. Conversely, findings for the Philippines remain inconclusive, while Thailand and Singapore exhibit neither characteristic. Notably, these distinct, market-specific behaviors remained stable throughout major global economic disruptions, including the Global Financial Crisis, the COVID-19 pandemic, and the 2022–2023 monetary tightening cycles. Theoretically and practically, these findings imply that regime-dependent, higher-moment risk pricing in ASEAN equities is fundamentally specific to individual countries, successfully extending the "good-volatility" framework previously established in developed bond markets to the complex landscape of emerging equity markets.

Keywords: *risk-return trade-off, threshold regression, ASEAN equity markets, conditional skewness, behavioral finance.*

OPEN ACCESS
ISSN 2528-4649 (online)
ISSN 2338-4409 (print)

*Correspondence:
Mohammad Syifaul Qulub
syifaulqulub68@gmail.com

Citation:
Received: ***June, 12, 2026***
Accepted: ***September, 02, 2026***
Published: ***September, 29, 2026***
JBMP: Jurnal Bisnis, Manajemen dan Perbankan.
Vol: 12 / No. 02
doi: 10.21070/jbmp.v12i2.2388

1. INTRODUCTION

The relationship between risk and return is a cornerstone of modern finance theory. According to the Intertemporal Capital Asset Pricing Model (ICAPM) of Merton (1973), investors require higher expected returns as compensation for bearing greater risk, implying a positive risk-return trade-off. Despite its theoretical appeal, empirical evidence on this relationship remains inconclusive. A growing body of literature documents that the risk-return trade-off is not always positive, and in many cases statistically insignificant or even negative (Glosten et al., 1993; Aslanidis et al., 2021).

One strand of the literature attributes this ambiguity to the inadequacy of conventional risk measures. Traditional approaches that rely solely on volatility as a proxy for risk fail to capture the full distributional characteristics of asset returns, particularly in markets characterized by asymmetry and fat tails. Kraus and Litzenberger (1976) demonstrate that investors have preferences over higher-order moments beyond variance, suggesting that skewness and kurtosis carry independent pricing implications. More recently, Christiansen and Savva (2023) show that incorporating conditional volatility, skewness, and excess kurtosis as risk measures provides a richer characterization of the risk-return relationship in international bond and stock markets.

A second strand of the literature highlights that the risk-return trade-off is inherently nonlinear. Salvador (2012) demonstrates that linear models such as GARCH-M consistently fail to detect a significant risk-return trade-off in emerging markets, while nonlinear regime-switching models reveal a statistically meaningful relationship. Similarly, Christiansen and Savva (2023) find that the trade-off is negative for large lagged returns when estimated through threshold regressions, distinguishing between good and bad volatility regimes. This behavioral asymmetry is consistent with prospect theory (Kahneman and Tversky, 1979), which posits that investors evaluate gains and losses differently, leading to asymmetric responses to risk across market conditions.

Despite these advances, the evidence remains largely concentrated in developed markets and broad cross-country panels. Studies focusing specifically on ASEAN equity markets are limited. Thampanya et al. (2020) examine the fundamental and behavioral determinants of stock return volatility in ASEAN-5 countries but do not investigate the risk-return trade-off using higher-order moments. Harvey (1995) documents that emerging markets exhibit distinct risk and return characteristics that deviate from global asset pricing models, yet a systematic examination of the nonlinear risk-return trade-off with higher-order moments in the ASEAN context remains absent from the literature. Moreover, existing studies do not cover the most recent market episodes, including the COVID-19 pandemic (2020) and the global monetary tightening cycle of 2022-2023, which represent significant structural shifts in risk dynamics. To the best of our knowledge, no existing study has jointly examined (i) conditional higher-order moments derived from a market-specific GARCH framework with skewed innovation distributions, (ii) a nonlinear, endogenously-determined threshold regression of the type proposed by Hansen (2000), and (iii) a multi-market ASEAN equity sample spanning the COVID-19 pandemic and the 2022-2023 monetary tightening cycle, within a single unified empirical framework. This three-way combination, rather than any one element in isolation, constitutes the specific gap this paper addresses.

This paper addresses these gaps by examining the risk-return trade-off in five major ASEAN stock markets, namely Indonesia, Malaysia, Singapore, Thailand, and the Philippines, over the period January 2006 to December 2025. We make three contributions to the literature. First, we employ conditional higher-order moments, specifically volatility, skewness, and excess kurtosis, derived from market-specific GARCH models with skewed, heavy-tailed innovation distributions as comprehensive risk measures, adapting the approach of Savva and Theodossiou (2018) and Christiansen and Savva (2023). Second, we investigate the nonlinear risk-return trade-off using threshold regression with lagged returns as the threshold variable, allowing us to distinguish between good and bad volatility regimes endogenously. Third, and most importantly for theory, we extend the good-volatility framework of Christiansen and Savva (2023), originally developed and validated only for government bond markets in developed economies, to equity markets in an emerging-market regional bloc. This extension is not merely an application to a new asset class: it tests whether the theoretical mechanism underlying regime-dependent risk pricing, namely investors' asymmetric tolerance for volatility conditional on prior gains versus losses (Kahneman and Tversky, 1979), holds when the institutional environment changes from a developed, institutionally-dominated bond market to retail-dominated, short-sale-constrained equity markets, a characterization that, as we note below, applies unevenly across our five-market sample. A formal bootstrap test for threshold existence (Hansen, 2000), replicated across three rolling-window lengths, confirms a significant, high-regime volatility premium in Indonesia at all three windows; Malaysia shows a similar pattern that is significant or near-significant at all three windows, though weaker and less precisely estimated than Indonesia's; the Philippines shows a suggestive but statistically inconclusive pattern; Thailand and Singapore show none. This finding indicates that the good-volatility mechanism is not specific to the bond-market, developed-economy setting in which it was first documented, but can also operate in an emerging equity-market setting; at the same time, the pronounced cross-market heterogeneity indicates that the mechanism is not a region-wide regularity, motivating further inquiry into the institutional or microstructural conditions under which it operates.

We note that the retail-dominated, short-sale-constrained characterization motivating our behavioral interpretation applies unevenly across our sample: it is most applicable to Indonesia, Malaysia, and the Philippines, where retail participation and short-selling frictions are well documented, and less applicable to Singapore, whose equity market is comparatively more institutionally dominated and internationally integrated. This heterogeneity in market structure is itself consistent with, and may help explain, the cross-market heterogeneity in our empirical results.

The remainder of the paper is organized as follows. Section 2 reviews the related literature. Section 3 describes the data and methodology. Section 4 presents and discusses the empirical results. Section 5 concludes. Section 6 discusses limitations and implications.

2. LITERATURE REVIEW

2.1 Risk-return trade-off: Theory and evidence

The theoretical foundation of the risk-return trade-off is rooted in the Intertemporal Capital Asset Pricing Model (ICAPM) of Merton (1973), which establishes that rational, risk-averse investors demand higher expected returns as compensation for bearing greater market risk. This positive relationship between conditional volatility and expected returns has motivated a large empirical literature. Glosten et al. (1993) provide an early empirical examination of this relationship in the US stock market, documenting a positive but unstable trade-off between expected returns and conditional variance. Subsequent studies have produced mixed evidence. Campbell and Hentschel (1992) and French et al. (1987) find a positive relationship, while other studies report insignificant or negative coefficients on the risk measure, suggesting that the trade-off is sensitive to model specification, sample period, and the choice of risk proxy. Lundblad (2007) argues that detecting a positive risk-return trade-off requires long time series, as short samples lack statistical power. More recently, Ghysels et al. (2005) demonstrate that the trade-off is positive and significant when risk is measured using mixed-frequency data. The persistent disagreement in the literature motivates the search for more comprehensive risk measures and flexible modeling frameworks.

2.2 Higher-order moments in asset pricing

The mean-variance framework underlying the ICAPM assumes that investors care only about the first two moments of the return distribution. However, when asset returns exhibit skewness and excess kurtosis, as is commonly observed in equity markets, this assumption fails to capture important dimensions of investor preferences. Kraus and Litzenberger (1976) demonstrate that investors exhibit a preference for positive skewness and aversion to excess kurtosis, implying that both moments carry independent pricing implications beyond variance. Harvey and Siddique (2000) extend this argument by showing that conditional skewness is priced in the cross-section of stock returns, while Dittmar (2002) provides evidence that kurtosis also carries a significant risk premium. More recently, Conrad et al. (2013) find that expected skewness significantly predicts future returns in the cross-section. In an emerging-market setting, Feng et al. (2023) decompose realized skewness into good and bad components for the Chinese stock market and find that good skewness dominates market variance in forecasting future returns, providing direct evidence that the skewness-return relationship documented in developed markets extends to emerging equity markets. To accommodate these distributional features within a unified risk-return framework, Savva and Theodossiou (2018) propose estimating conditional higher-order moments from the Skewed Generalized Error Distribution (SGED), which simultaneously captures time-varying volatility, skewness, and kurtosis. Building on this framework, Christiansen and Savva (2023) incorporate conditional volatility, skewness, and excess kurtosis as joint risk measures in an international risk-return trade-off analysis, finding that higher moments provide a richer characterization of risk than volatility alone. Motivated by these findings, this study adopts the same approach for ASEAN equity markets.

2.3 Nonlinearity and threshold effects in the risk-return trade-off

A growing literature documents that the risk-return trade-off is inherently nonlinear, with the sign and magnitude of the relationship varying across market conditions. Aslanidis et al. (2021) investigate the risk-return trade-off using quantile regressions for the US and European stock markets, finding a positive trade-off at the upper tail and a negative trade-off at the lower tail of the return distribution. This asymmetry suggests that the relationship between risk and return depends on the prevailing market regime. Feunou et al. (2013) decompose volatility into upside and downside components and show that downside volatility carries a negative risk premium. Zhang et al. (2021) further distinguish between good and bad variance, defined by the sign of contemporaneous returns, and find that both components have differential predictive power for future returns. The behavioral foundation for this regime dependence is provided by Wang et al. (2017), who demonstrate that the risk-return trade-off is positive when the market trades above its reference level but turns negative when trading below it, attributing the reversal to reference-dependent preferences consistent with prospect theory; this finding establishes that the direction of the trade-off is not a fixed structural parameter but a function of investors' perceived position relative to a psychological benchmark. Threshold-based regime dependence is not confined to volatility alone: Bansal et al. (2022) show that the positive risk premium on high-beta and small-cap stocks in the US market is reliably earned only after expected market volatility (VIX) breaches a high threshold at approximately the 80th percentile, with the premium persisting for up to six months following the breach and remaining near zero otherwise; comparable threshold behavior is absent for the Fama-French HML, RMW, and CMA factors, indicating that the threshold effect is specific to volatility-related risk exposures rather than a general feature of all priced factors. In a closely related time-series setting, Bansal and Stivers (2025) find that the US equity premium increases nonlinearly with VIX, stepping up appreciably once VIX exceeds its 80th-85th percentile, while declining linearly with investor sentiment. Together, these studies establish that threshold-based nonlinearity in risk pricing is a robust cross-market phenomenon. Christiansen and Savva (2023) formalize this distinction through threshold regression, allowing the volatility coefficient to vary endogenously across regimes defined by lagged returns. Their findings reveal that the risk-return trade-off is positive for small lagged returns but negative for large lagged returns, a behavioral asymmetry grounded in prospect theory (Kahneman and Tversky, 1979), which posits that investors are more sensitive to losses than to equivalent gains, generating asymmetric risk responses across gain and loss regimes. This study extends this nonlinear framework to the ASEAN context.

2.4 Evidence from emerging markets and ASEAN

The risk-return trade-off in emerging markets exhibits distinct characteristics relative to developed markets. Harvey (1995) provides seminal evidence that emerging markets are characterized by higher average returns, greater volatility, and stronger return predictability compared to developed markets, and that global asset pricing models fail to fully explain cross-sectional return variation in these markets. Within this context, Salvador (2012) is noteworthy for documenting not only that linear GARCH-M models fail across 25 MSCI emerging market indices, but that the regime in which a meaningful trade-off does emerge (high-volatility regimes under nonlinear specifications) maps precisely onto the conditions where prospect-theory-driven asymmetry is most likely to operate, suggesting that the failure of linear models in emerging markets is not merely a statistical artifact but reflects underlying behavioral heterogeneity across market states. In a more recent application, Mnari and Faouel (2024) document a nonlinear, asymmetric risk-return trade-off across Tunisian Stock Exchange sectors over a period spanning the COVID-19 pandemic, reinforcing that linear specifications remain insufficient to characterize risk pricing in emerging equity markets during periods of structural disruption. Within the ASEAN region specifically, Thampanya et al. (2020) examine the fundamental and behavioral determinants of stock return volatility in ASEAN-5 countries over the period 1995-2018, documenting the role of both macroeconomic and behavioral factors in driving return volatility across different structural regimes. However, their study does not examine the risk-return trade-off using higher-order moments, nor does it employ a threshold regression framework. The joint application of conditional higher-order moments derived from a market-specific GARCH framework with skewed innovation distributions, endogenous threshold regression, and a multi-market ASEAN sample spanning the COVID-19 pandemic and the 2022-2023 monetary tightening cycle therefore remains absent from the literature, constituting the specific gap this study addresses.

We note an important theoretical nuance in applying prospect theory (Kahneman and Tversky, 1979) to regime-dependent risk pricing. The theory's value function is concave in the gain domain but convex in the loss domain, implying that investors may become more risk-seeking, not more risk-averse, following losses, as they seek to recover to their reference point. Our empirical finding of a statistically insignificant, rather than significantly negative, volatility coefficient in the low regime is consistent with either a risk-averse retreat from volatility exposure or a risk-seeking pursuit of recovery that is imprecisely estimated at the daily frequency; our data do not allow us to distinguish between these competing behavioral channels. We accordingly interpret the low-regime result as an absence of a positive volatility premium following unfavorable prior returns, rather than as direct evidence of loss-averse risk retreat specifically.

3. RESEARCH METHOD

3.1 Data

This study employs daily closing price data of five ASEAN stock market indices over the period January 2006 to December 2025. The selected markets are Indonesia (Jakarta Composite Index/IHSG), Malaysia (Kuala Lumpur Composite Index/KLCI), Singapore (Straits Times Index/STI), Thailand (Stock Exchange of Thailand Index/SET), and the Philippines (Philippine Stock Exchange Index/PSEi). These markets represent a mix of emerging and developed economies within the ASEAN region, providing a heterogeneous sample suitable for cross-country comparison. All data are sourced from Yahoo Finance and retrieved using Python. Daily log-returns are computed as $R(i,t) = ln(P(i,t)) - ln(P(i,t-1))$, where $P(i,t)$ denotes the closing price of market i at time t.

The sample period begins in January 2006 to ensure sufficient pre-crisis observations before the Global Financial Crisis (GFC) of 2008-2009, and extends to December 2025 to capture subsequent major market episodes including the COVID-19 pandemic (2020) and the global monetary tightening cycle (2022-2023). This extended timeframe allows for a comprehensive examination of risk-return dynamics across varying market conditions.

Following Christiansen and Savva (2023) and Savva and Theodossiou (2018), we measure market risk using three conditional higher-order moments: volatility, skewness, and excess kurtosis. These measures capture dimensions of risk beyond the mean-variance framework, reflecting investor preferences over the full return distribution.

Conditional volatility for each market is estimated using a market-specific GARCH-family model of order (1,1) (GARCH, GJR-GARCH, or APARCH), illustrated here for the baseline GARCH(1,1) case:

$$R(i,t) = \mu(i) + \varepsilon(i,t)$$

$$\varepsilon(i,t) = \sigma(i,t) \cdot z(i,t), z(i,t)\ SkT(0,1,\lambda,\kappa)$$
$$\sigma(i,t)^2 = \omega(i) + \alpha(i) \cdot \varepsilon(i,t-1)^2 + \beta(i) \cdot \sigma(i,t-1)^2$$

where σ(i,t) is the conditional volatility of market i at time t, ε(i,t) is the error term, and z(i,t) are standardized residuals following a skewed Student-t distribution (SkT), the specification selected for all five markets via the model-selection procedure, with skewness parameter λ and shape parameter κ. Both the SGED and skewed Student-t distributions were considered as candidates in the model-selection procedure, as each accommodates both skewness and fat tails commonly observed in daily financial returns; the skewed Student-t distribution is selected for all five markets on the basis of BIC and residual diagnostics. Full parameter estimates for each market's selected model, including the variance-equation coefficients (ω, α, β), the asymmetry and power terms where applicable, and the distributional shape parameters.

Conditional skewness and excess kurtosis are computed from the standardized residuals z(i,t) using a rolling window of 22 trading days, approximately equivalent to one calendar month, with 10-day and 44-day windows used for robustness. This approach is a deliberate departure from the original Savva and Theodossiou (2018) framework, which extracts higher moments directly from the estimated shape parameters of the innovation distribution; because those parameters are estimated as constants over the full sample in standard GARCH implementations, they would yield a single, time-invariant skewness and kurtosis value per market rather than the time-varying (regime-relevant) measure this study requires. The rolling-window approach is therefore the specification choice that preserves time variation, not an approximation to it. We note that at a 22-day window, the theoretical sampling standard error of the skewness and excess kurtosis estimators (approximately 0.52 and 1.04 respectively, under an i.i.d. approximation) is non-trivial relative to the cross-time dispersion of the estimated series, which we quantify and discuss further in Section 4.2. Higher-order moments are computed about a mean of zero, consistent with the model-implied conditional mean of the standardized residuals, rather than a locally re-estimated window mean; standard finite-sample bias corrections for skewness and kurtosis (e.g., Joanes and Gill, 1998) assume the latter and are therefore not applied here, as doing so would over-correct. Specifically:

$$Skew(i,t) = (1/22)\Sigma[j = 0 to 21] z(i,t-j)^3/[(1/22)\Sigma[j = 0 to 21] z(i,t-j)^2]^{(3/2)}$$
$$Kurt(i,t) = (1/22)\Sigma[j = 0 to 21] z(i,t-j)^4/[(1/22)\Sigma[j = 0 to 21] z(i,t-j)^2]^2 - 3$$

where the subtraction of 3 yields excess kurtosis relative to the normal distribution. The use of standardized residuals ensures that the higher moments capture distributional asymmetry and tail risk orthogonal to the volatility already modeled by the GARCH process. The first 21 observations are lost due to the 22-day rolling window initialization, yielding the effective sample sizes.

3.2 Diagnostic tests

Before estimation, we verify that the log-return series for each market are stationary using the Augmented Dickey-Fuller (ADF) test and the Kwiatkowski-Phillips-Schmidt-Shin (KPSS) test. The ADF test tests the null hypothesis of a unit root, while the KPSS test tests the null of stationarity; employing both provides a more robust assessment of the integration properties of the series. Consistent with the expectation for log-returns, we reject the unit root null under ADF and fail to reject the stationarity null under KPSS for all five markets, confirming that the return series are stationary and suitable for OLS estimation.

Because a single ARMA(0,0)-GARCH(1,1) mean-variance specification need not be adequate for every market, we select the mean and variance equation separately for each market through a grid search over ARMA orders (0,0) through (1,1), three variance models (GARCH, GJR-GARCH, APARCH), and two innovation distributions (skewed generalized error, skewed Student-t), totalling 24 candidate specifications per market. We additionally verified, in a first-pass comparison across the base ARMA grid, that an EGARCH specification does not materially improve on this selection: only Indonesia showed an EGARCH candidate with both a lower BIC and passing diagnostics, and the improvement (ΔBIC ≈ 0.002) is negligible by conventional standards (Kass and Raftery, 1995); we therefore retain the APARCH specification for Indonesia reported below. For each candidate, we compute the Bayesian Information and 20, assessing residual serial correlation in the conditional mean, and the ARCH-LM test (Engle, 1982) on squared standardized residuals, assessing remaining ARCH effects. The selected specification for each market is the one with the lowest BIC among candidates that fail to reject the null of no residual autocorrelation and no remaining ARCH effects at the 5% level. For Indonesia, Malaysia, the Philippines, and Singapore, this procedure yields a specification that clears both diagnostics. For Thailand (SET), no candidate specification, including an extended search over ARMA orders up to (5,0)/(0,5), fully eliminates residual serial correlation (Ljung-Box Q(10) p = 0.011 at the minimum-BIC specification); we retain the minimum-BIC specification for SET and report this as an explicit limitation, rather than select a specification based on a marginally passing diagnostic, which would risk specification-search bias.

3.3 Linear risk-return trade-off

To examine the linear risk-return trade-off, we estimate the following regression for each market separately:

$$R(i,t) = c(i,0) + c(i,Vol) \cdot Vol(i,t-1) + c(i,Skew) \cdot Skew(i,t-1) + c(i,Kurt) \cdot Kurt(i,t-1) + \varepsilon(i,t)$$

Where $R(i,t)$ is the daily log-return of market i at time t, $Vol(i,t-1)$, $Skew(i,t-1)$, and $Kurt(i,t-1)$ are the one-period-lagged conditional volatility, skewness, and excess kurtosis as defined in Section 3.1, and ε(i,t) is the error term. The model is estimated separately for each country using Ordinary Least Squares (OLS).

The equation above uses one-period-lagged (predetermined) risk measures, our preferred specification, rather than their contemporaneous counterparts. Because the rolling skewness and kurtosis measures are constructed from standardized residuals up to and including period t, the contemporaneous versions of these variables ($Vol(i,t)$, $Skew(i,t)$, $Kurt(i,t)$) mechanically embed information from $R(i,t)$ themselves; using predetermined values removes this overlap. The equation above is reported in full in Table 4; for comparison, Table 8 reports estimates using the corresponding contemporaneous specification, following the convention in Christiansen and Savva (2023). Newey-West HAC standard errors with a fixed bandwidth of 10 lags (Bartlett kernel) are used throughout to account for residual serial dependence.

We expect the coefficient on volatility ($c(i,Vol)$) to be positive, consistent with the notion that higher risk commands higher return as predicted by the ICAPM (Merton, 1973). The coefficient on skewness ($c(i,Skew)$) is expected to be negative, as investors demand higher returns for negatively skewed distributions. The coefficient on excess kurtosis ($c(i,Kurt)$) is expected to be positive, reflecting compensation for fat tail risk.

3.4 Threshold risk-return trade-off

Following Christiansen and Savva (2023) and Zhang et al. (2021), we extend the linear model by allowing the risk-return relationship to vary across market regimes. Specifically, we employ a threshold regression framework to capture potential nonlinearities in the risk-return trade-off, motivated by the notion that investors may respond asymmetrically to risk depending on prior market conditions.

The threshold variable is the lagged return $R(i,t-k)$, where the optimal lag order $k \in \{1,2,3,4,5\}$ is selected endogenously for each market by minimizing the residual sum of squares (Hansen, 2000). This choice is motivated by the possibility that investors respond asymmetrically to risk depending on whether recent returns have been favorable or unfavorable, consistent with the regime-dependent pricing documented in Section 2.3; Section 2.4 discusses a theoretical nuance regarding the specific behavioral channel underlying this asymmetry.

We define good volatility as the conditional volatility when the lagged return is above the threshold γ(i), and bad volatility when the lagged return falls below γ(i). The threshold risk-return regression is specified as:

$$R(i,t) = c(i,j,0) + c(i,j,Vol) \cdot Vol(i,t-1) + c(i,Skew) \cdot Skew(i,t-1) + c(i,Kurt) \cdot Kurt(i,t-1) + \varepsilon(i,t)$$

where j denotes the regime, determined endogenously by:

$$c(i,j,Vol) = c(i,Low,Vol) if R(i,t-1) < \gamma(i); c(i,High,Vol) if R(i,t-1) \geq \gamma(i)$$

The threshold value γ(i) is estimated endogenously for each market by minimizing the residual sum of squares (Hansen, 2000). We additionally test the null hypothesis of no threshold against the single-threshold alternative using a fixed-regressor wild bootstrap (Hansen, 1996, 2000) with 300 replications, and construct a 95% confidence interval for the threshold via likelihood-ratio inversion following Hansen (2000). The coefficients on skewness and kurtosis are held constant across regimes, while the volatility coefficient and intercept are allowed to vary. We expect $c(i, High, Vol) > 0$, consistent with the notion that investors demand compensation for volatility risk when prior market conditions are favorable; the expected sign and interpretation of $c(i, Low, Vol)$ is discussed in Section 2.4 in light of the theoretical nuance noted there.

3.5 Robustness checks

To assess the sensitivity of the results to the choice of rolling window used in constructing the conditional higher-order moments, we re-estimate both the linear and threshold models using two alternative window lengths: 10 trading days (approximately two calendar weeks) and 44 trading days (approximately two calendar months), in addition to the baseline window of 22 trading days. A window of 10 days captures short-term distributional dynamics and is more responsive to sudden shifts in market conditions, while a 44-day window produces smoother, medium-term estimates of skewness and excess kurtosis. Threshold robustness results are reported in Appendix Tables A1-A2; OLS robustness results are reported in Tables 6-7 and discussed in Section 4.5.

As a comparison to our preferred predetermined specification, we also estimate the linear OLS model using contemporaneous values of all three risk measures (Vol(i,t), Skew(i,t), Kurt(i,t)) in place of their predetermined counterparts. This comparison illustrates the extent to which the skewness and volatility results are sensitive to the timing assumption and is reported in Table 8 and discussed in Section 4.5.

Finally, because our sample period is motivated in part by its coverage of the Global Financial Crisis (2008-2009), the COVID-19 pandemic (2020-2021), and the global monetary tightening cycle (2022-2023), we test directly whether the linear risk-return relationship (Table 4) is stable across these episodes, rather than assuming stability implicitly. For each market, we estimate the baseline predetermined specification augmented with period-specific dummy variables interacted with each risk measure, and test the joint null hypothesis that all period-interaction coefficients are zero using a Wald test with the same Newey-West HAC covariance matrix (bandwidth = 10) used throughout. This is a Chow-type structural break test robust to the heteroskedasticity and serial dependence present in daily return data.

4. RESULTS AND DISCUSSION

4.1 Diagnostic tests

Table 1 reports the results of the diagnostic tests conducted prior to and following GARCH estimation. All five return series reject the unit root null under the ADF test ($p < 0.01$) and fail to reject the stationarity null under the KPSS test ($p = 0.100$), confirming that log-returns are stationary and suitable for OLS estimation. Note that for the KPSS test, a high p-value indicates failure to reject the null of stationarity, which is the desired outcome; the absence of asterisks in the KPSS p-value column therefore signals stationarity, not insignificance.

Mean and variance equations are selected separately for each market via the grid-search procedure. Under the selected specifications, the Ljung-Box test indicates no significant residual serial autocorrelation at the 5% level for Indonesia, Malaysia, the Philippines, and Singapore; for Thailand, residual serial correlation persists even under the best available specification and is reported as an explicit limitation. The ARCH-LM test confirms that GARCH filtering successfully removes volatility clustering for all five markets under their respective selected specifications, each of which uses a skewed Student-t innovation distribution. Full parameter estimates for each selected model are reported in Table 2.

Table 1 | Diagnostic Test Results

Market	ADF Statistic	ADF p-value	KPSS Statistic	KPSS p-value	Ljung-Box Lag 10	ARCH-LM Lag 5
IHSG — ARMA(0,0)-APARCH(1,1)-SkT	−16.1574	0.0100***	0.2069	0.100	0.1606	0.8924
KLCI — ARMA(2,2)-GJR-GARCH(1,1)-SkT	−17.4031	0.0100***	0.2212	0.100	0.1098	0.8305
PSEi — ARMA(0,1)-GJR-GARCH(1,1)-SkT	−17.1652	0.0100***	0.2925	0.100	0.1895	0.1659
SET — ARMA(1,0)-APARCH(1,1)-SkT	−15.9046	0.0100***	0.1654	0.100	0.0108**	0.9816
STI — ARMA(0,1)-APARCH(1,1)-SkT	−16.1576	0.0100***	0.0638	0.100	0.0620*	0.1331

Notes: ADF and KPSS results (raw log-returns) are unaffected by GARCH model choice. Mean and variance equations are selected separately for each market via a grid search over ARMA orders (0,0) to (1,1), three variance models (GARCH, GJR-GARCH, APARCH), and two innovation distributions (SGED, skewed Student-t [SkT]), minimizing BIC among candidates that pass both diagnostics below at the 5% level. An EGARCH specification was additionally tested across the base ARMA grid; only Indonesia showed an EGARCH candidate with both a lower BIC and passing diagnostics, and the improvement (ΔBIC ≈ 0.002) is negligible by conventional standards (Kass and Raftery, 1995), so the APARCH specification is retained. Ljung-Box and ARCH-LM statistics reported are for the selected model at lag 10 and lag 5 respectively, the criteria used in model selection. denotes rejection of H0 at the 10% level, ** at the 5% level; no model achieves rejection at the 1% level. For Thailand (SET), no candidate specification, including an extended search up to ARMA(5,0)/(0,5), eliminates residual serial correlation; the minimum-BIC specification is retained and reported as an explicit limitation (Section 6). Newey-West HAC standard errors (fixed bandwidth = 10, Bartlett kernel) are used in all subsequent OLS and threshold estimations regardless of diagnostic outcome.

Table 2 | Estimated GARCH Parameters, Final Selected Models

Parameter	IHSG	KLCI	PSEi	SET	STI
omega	0.0000718 (0.0000551)	0.00000092 (0.00000063)	0.00000780*** (0.00000046)	0.0000360 (0.0000245)	0.00000012 (0.00000044)
alpha1	0.1257*** (0.0138)	0.0678*** (0.0102)	0.0643*** (0.0068)	0.1043*** (0.0100)	0.0746*** (0.0020)
beta1	0.8729*** (0.0147)	0.8841*** (0.0144)	0.8265*** (0.0098)	0.8974*** (0.0100)	0.8770*** (0.0151)
gamma1 (asymmetry)	0.3273*** (0.0550)	0.0644*** (0.0137)	0.1117*** (0.0180)	0.3464*** (0.0525)	0.2117*** (0.0386)
delta (power, APARCH only)	1.3185*** (0.1632)	n/a	n/a	1.3512*** (0.1394)	2.5291*** (0.0293)
skew (SkT)	0.8557*** (0.0172)	0.9323*** (0.0183)	0.9143*** (0.0186)	0.9202*** (0.0186)	0.9196*** (0.0174)
shape (SkT df)	6.158*** (0.523)	6.474*** (0.449)	7.566*** (0.718)	6.597*** (0.587)	8.065*** (0.618)
Persistence	n/a (see note 1)	0.9841	0.9466	n/a (see note 1)	0.9516
Convergence	0	0	0	0	0
Log-likelihood	15542.81	18039.78	15176.84	15953.29	17037.33
N	4848	4903	4898	4862	5000

Notes: Standard errors in parentheses. *, **, *** denote significance at the 10%, 5%, and 1% levels. Convergence = 0 indicates successful numerical convergence for all five models. Persistence is computed as alpha+beta for GJR-GARCH models (KLCI, PSEi), adjusted for asymmetry (alpha+beta+gamma/2, the standard GJR convention). Note 1: For APARCH models (IHSG, SET), the alpha+beta formula does not directly apply because persistence involves a non-linear interaction between the asymmetry parameter (gamma) and the power term (delta) through the expectation E[(|z|-gamma*z)^delta]; we do not report a potentially misleading figure here. Stationarity and model adequacy for these two markets are instead supported by the Ljung-Box and ARCH-LM diagnostics reported in Table 1. A skew parameter below 1 in the skewed Student-t (SkT) distribution (rugarch parameterization) indicates negative skewness, consistent with Table 3.

4.2 Descriptive statistics

The final sample comprises 4,827 daily observations for IHSG, 4,882 for KLCI, 4,877 for PSEi, 4,841 for SET, and 4,979 for STI, reflecting differences in national public holidays and exchange-specific non-trading days across the five markets. Table 3 presents the descriptive statistics for daily returns and conditional higher-order moments across the five ASEAN markets. The mean daily returns are positive across all markets, ranging from 0.010% (SET) to 0.040% (IHSG), consistent with the long-run upward trend documented in emerging equity markets (Harvey, 1995). However, the large standard deviations relative to mean returns, ranging from 0.74% for KLCI to 1.27% for PSEi, indicate substantial daily return volatility, reflecting the high-risk nature of ASEAN equity markets.

All five markets exhibit negative skewness, ranging from -0.364 (STI) to -1.173 (SET), indicating left-skewed return distributions with greater probability mass on the downside. This is a well-documented characteristic of equity markets, particularly during periods that encompass major financial crises (Harvey, 1995). Excess kurtosis is uniformly large across all markets, ranging from 11.48 (STI) to 21.56 (SET), confirming the presence of fat tails and the inadequacy of the normal distribution assumption for modeling daily returns in ASEAN markets. These stylized facts, namely negative skewness and excess kurtosis, motivate the use of skewed, heavy-tailed innovation distributions in the GARCH estimation, as they explicitly accommodate such departures from normality; the distribution selected for each market via the model-selection procedure in Section 3.2 is reported in Table 1, and full parameter estimates in Table 2.

Turning to the conditional higher-order moments, the mean conditional volatility ranges from 0.677% (KLCI) to 1.174%

(PSEi), with all markets exhibiting highly right-skewed volatility distributions, consistent with volatility clustering and occasional extreme episodes. The mean conditional skewness is negative across all markets, ranging from -0.161 (KLCI) to -0.279 (IHSG), reflecting persistent downside asymmetry in the conditional return distribution. The mean conditional excess kurtosis is positive across all markets, confirming persistent fat tails in the conditional distribution. As discussed in Section 3.1, the theoretical sampling standard error of the 22-day rolling skewness and kurtosis estimators (approximately 0.52 and 1.04 respectively) is non-trivial relative to the cross-time standard deviations reported below (approximately 0.94-1.03 for skewness, 1.42-1.92 for kurtosis), implying that a substantial share, though not all, of the observed time variation in these series reflects estimation noise rather than genuine regime shifts.

Table 3 | Descriptive Statistics

	Mean	Std. Dev.	Skewness	Kurtosis
Return				
IHSG	0.000402	0.012355	-0.63916	12.23237
KLCI	0.000124	0.007408	-0.7713	15.60004
PSEi	0.000213	0.012712	-0.90796	13.49922
SET	0.000104	0.011670	-1.17324	21.55887
STI	0.000129	0.010288	-0.36416	11.48485
Conditional Volatility				
IHSG	0.011089	0.005661	3.020052	16.344466
KLCI	0.006774	0.003209	3.132473	19.143091
PSEi	0.011742	0.004862	4.471390	35.574118
SET	0.010252	0.005492	3.289058	19.740845
STI	0.008913	0.004953	3.144205	16.877905
Conditional Skewness				
IHSG	-0.278875	1.031229	-0.085518	2.638805
KLCI	-0.161441	1.028542	-0.321802	3.435517
PSEi	-0.165033	0.962291	-0.250177	3.236619
SET	-0.189266	0.977486	-0.197452	4.240599
STI	-0.170067	0.944997	-0.228676	2.776599
Conditional Excess Kurtosis				
IHSG	0.647809	1.786312	2.066564	8.148212
KLCI	0.575690	1.918943	3.294270	17.218355
PSEi	0.368533	1.603729	3.823667	26.160417
SET	0.403117	1.859946	4.869663	38.143191
STI	0.328600	1.415059	3.115314	17.995736

Notes: Returns are daily log-returns, expressed in percentage terms. Conditional volatility is the time-varying standard deviation extracted from the market-specific GARCH model reported in Table 1 (all five markets select a skewed Student-t innovation distribution; full parameters in Table 2). Conditional skewness and excess kurtosis are computed using 22-day rolling windows of standardized GARCH residuals, adapting the approach of Savva and Theodossiou (2018) as described in Section 3.1. The first 21 observations are excluded due to rolling window initialization. Sample sizes: IHSG = 4,827; KLCI = 4,882; PSEi = 4,877; SET = 4,841; STI = 4,979. Sample period: January 2006 to December 2025.

4.3 Linear risk-return trade-off

Table 4 reports the OLS estimation results for the linear risk-return trade-off under our preferred predetermined (lagged) specification. The adjusted R-squared values are uniformly low, effectively at or below zero (ranging from -0.02% for KLCI to 0.18% for PSEi), consistent with the difficulty of explaining daily equity returns using predetermined risk measures alone (Ghysels et al., 2005; Christiansen and Savva, 2023).

The conditional volatility coefficient is positive in four of five markets (all except SET) but fails to achieve conventional statistical significance in any market. The absence of a robust linear volatility premium is inconsistent with the classical mean-variance ICAPM prediction, but is not unprecedented in the literature. Harvey and Siddique (2000) and Dittmar (2002) similarly document that the unconditional volatility premium becomes statistically unreliable once higher-order moments are included in the pricing kernel, suggesting that volatility alone does not fully capture the compensation investors demand. The broader literature also documents mixed and often insignificant linear risk-return estimates (Glosten et al., 1993; Aslanidis et al., 2021), suggesting that a linear specification may be insufficient to capture the full dynamics of the risk-return relationship in these markets.

A central finding concerns the conditional skewness coefficient. Under the predetermined specification, the skewness coefficient is positive and statistically significant only for Indonesia (0.00044, $p < 0.05$) and the Philippines (0.00065, $p < 0.01$); it is not significant for Malaysia, Thailand, or Singapore. As shown in the robustness analysis (Section 4.5), the Indonesian result is stable across the 10-day, 22-day, and 44-day windows, whereas the Philippine result is significant only at the 22-day window and does not survive at either the 10-day or 44-day window. We therefore treat the skewness premium as a robust, market-specific finding for Indonesia, an inconclusive result for the Philippines, and unsupported for Malaysia, Thailand, and Singapore, rather than as a general ASEAN-wide phenomenon. Where the premium is supported, it is consistent with a behavioral interpretation: in emerging markets characterized by retail investor dominance and short-selling constraints, investors tend to exhibit a preference for positively skewed, lottery-like assets, bidding up their prices and accepting lower risk compensation for skewness (Barberis and Huang, 2008).

Under this predetermined specification, the conditional excess kurtosis coefficient is significant only for the Philippines (0.000219, $p < 0.10$), positively signed, and insignificant elsewhere. The general weakness of the kurtosis result is consistent with evidence from both equity and bond markets: Dittmar (2002) documents that kurtosis loses explanatory power once skewness is controlled for in equity pricing, while Christiansen and Savva (2023) similarly report weak kurtosis effects in their international bond market sample.

Table 4 | OLS Risk-Return Estimates, Predetermined (Preferred) Specification (Window = 22 Days)

	IHSG	KLCI	PSEi	SET	STI
Constant	0.000025	-0.000200	-0.000786	0.000094	0.000160
(t-stat)	(0.0316)	(-0.5721)	(-0.7884)	(0.1418)	(0.2698)
Vol(-1)	0.042514	0.049427	0.087230	-0.000538	0.003355
(t-stat)	(0.5414)	(0.8261)	(0.9554)	(-0.0073)	(0.0447)
Skew(-1)	0.000441**	0.000082	0.000653***	0.000237	0.000210
(t-stat)	(2.4246)	(0.7221)	(2.9329)	(1.3085)	(1.1017)
Kurt(-1)	0.000044	0.000003	0.000219*	0.000163	-0.000072
(t-stat)	(0.3919)	(0.0520)	(1.9289)	(1.1830)	(-0.5650)
Obs	4826	4881	4876	4840	4978
Adj. R2	0.0005	-0.0002	0.0018	0.0001	-0.0000

Notes: The table reports OLS estimates, using our preferred predetermined (lagged) specification, of $R(i,t) = c(i,0) + c(i,Vol)Vol(i,t-1) + c(i,Skew)Skew(i,t-1) + c(i,Kurt)Kurt(i,t-1) + e(i,t)$ *R(i,t) = c(i,0) + c(i,Vol)Vol(i,t-1) + c(i,Skew)Skew(i,t-1) + c(i,Kurt)Kurt(i,t-1) + e(i,t), estimated separately for each market using the market-specific GARCH specification reported in Table 1. Because the rolling skewness and kurtosis measures are constructed from standardized residuals up to and including period t, their contemporaneous versions mechanically embed information from R(i,t) itself; the predetermined specification removes this overlap. The corresponding contemporaneous specification is reported for comparison in Table 8 (Section 4.5). Newey-West HAC standard errors with a fixed bandwidth of 10 lags (Bartlett kernel) are reported in parentheses.*
*, **, *** denote statistical significance at the 10%, 5%, and 1% levels, respectively.

4.4 Threshold risk-return trade-off

Table 5 presents the results of the threshold regression. The threshold variable is the one-period lagged return for all markets except Singapore, for which the three-period lagged return provides the best fit based on residual sum of squares minimization (Hansen, 2000).

All five markets are characterized by two regimes under the single-threshold specification. The endogenously estimated threshold values under the predetermined specification are positive for Indonesia (0.00594), the Philippines (0.00639), Thailand (0.00404), and Singapore (0.00177), and close to zero for Malaysia (0.00041). We additionally test the null hypothesis of no threshold using a fixed-regressor wild bootstrap (Hansen, 1996, 2000; 300 replications), replicated across the 10-day, 22-day, and 44-day windows (Table 5; Appendix Tables A1-A2). This test rejects the null decisively for Indonesia at all three windows ($p < 0.02$ in each case). For Malaysia, the null is rejected at the 10-day and 44-day windows ($p = 0.043$ and $p = 0.023$ respectively) and is at the boundary of rejection at the 22-day baseline ($p = 0.050$); this consistent pattern across all three windows supports treating Malaysia's threshold evidence as tentatively confirmed, albeit weaker and less precisely estimated than Indonesia's. For the Philippines, the null is not rejected at any window (p = 0.11-0.20); for Thailand and Singapore, the null is not rejected at any window either ($p > 0.23$ in all cases). We therefore treat the threshold structure as statistically confirmed for Indonesia, reasonably supported for Malaysia, and not established for the Philippines, Thailand, and Singapore.

Table 5 | Threshold Regression Estimates, Predetermined Specification (Window = 22 Days)

	IHSG	KLCI	PSEi	SET	STI
Threshold Lag	RETURN(-1)	RETURN(-1)	RETURN(-1)	RETURN(-1)	RETURN(-3)
Low Threshold	< 0.005938 [CI: -0.0069, 0.0061]	< 0.000407 [CI: -0.0025, 0.0017]	< 0.006385 [CI: 0.0055, 0.0065]	< 0.004041 [CI: -0.0040, 0.0085]	< 0.001767 [CI: -0.0071, 0.0055]
Constant	0.001450	-0.000142	0.000668	-0.000835	-0.000673
(t-stat)	(1.3594)	(-0.2162)	(0.5068)	(-1.0146)	(-0.9681)
Volatility	-0.144876	-0.049708	-0.083666	0.048191	0.076737
(t-stat)	(-1.3053)	(-0.4481)	(-0.6698)	(0.5140)	(0.8757)
Obs	3517	2494	3548	3332	2895
Medium Threshold	–	–	not confirmed – see Section 6	–	–
High Threshold	>= 0.005938 [CI: -0.0069, 0.0061]	>= 0.000407 [CI: -0.0025, 0.0017]	>= 0.006385 [CI: 0.0055, 0.0065]	>= 0.004041 [CI: -0.0040, 0.0085]	>= 0.001767 [CI: -0.0071, 0.0055]
Constant	-0.001507	-0.000045	-0.001563	0.002300	0.001345
(t-stat)	(-1.4204)	(-0.1198)	(-1.1271)	(2.0481)	(1.5494)
Volatility	0.279868***	0.116687*	0.243649**	-0.111077	-0.103022
(t-stat)	(3.0337)	(1.8906)	(2.1156)	(-0.9899)	(-0.9395)
Obs	1309	2387	1328	1508	2081
Non-Threshold Variables					
Skewness	0.000372**	0.000046	0.000558***	0.000168	0.000194
(t-stat)	(2.1590)	(0.4289)	(2.6160)	(0.9409)	(1.0199)
Kurtosis	0.000033	0.000004	0.000244**	0.000167	-0.000064
(t-stat)	(0.3142)	(0.0596)	(2.1808)	(1.1925)	(-0.4968)
Adj. R2	0.0149	0.0076	0.0088	0.0041	0.0017

Notes: The table reports threshold regression estimates following Hansen (2000). The threshold variable is the endogenously selected lagged return R(i,t-k), where the optimal lag k is chosen by minimizing the residual sum of squares over k in {1,2,3,4,5}, held fixed at the value reported above. The volatility coefficient is allowed to vary across regimes while skewness and kurtosis coefficients are constrained to be regime-invariant. The bracketed range [CI: low, high] reports the 95% confidence interval for the threshold value, constructed via likelihood-ratio inversion (Hansen, 2000). A fixed-regressor wild bootstrap test (Hansen, 1996, 2000; 300 replications) for the null of no threshold rejects the null for Indonesia (p < 0.001) and is at the boundary of rejection for Malaysia (p = 0.05) at this window; replicated at the 10-day and 44-day windows (Appendix Tables A1-A2), the null is also rejected for Malaysia at both alternative windows (p = 0.04 and p = 0.02 respectively). The null is not rejected at any window for the Philippines (p = 0.11-0.20), Thailand (p = 0.23-0.37), or Singapore (p = 0.45-0.60). An exploratory three-regime (two-threshold) specification for the Philippines did not show a statistically significant improvement over the single-threshold model (sequential F-test, p = 0.45-0.57; Section 6). Newey-West HAC standard errors with a fixed bandwidth of 10 lags are reported in parentheses. *, **, *** denote statistical significance at the 10%, 5%, and 1% levels, respectively.

In the low regime, corresponding to periods following relatively low or negative prior returns, the volatility coefficient is statistically insignificant across all five markets. This null result in the low regime is consistent with either a risk-averse retreat from volatility exposure or a risk-seeking pursuit of recovery that is imprecisely estimated at the daily frequency; we interpret it as an absence of a positive volatility premium following unfavorable prior returns rather than as direct evidence for either specific behavioral channel.

In the high regime, following relatively strong prior returns, the volatility coefficient under the predetermined specification is positive and significant for Indonesia (0.280, t = 3.03), positive at the 10% level for Malaysia (0.117, t = 1.89), and positive at conventional levels for the Philippines (0.244, t = 2.12), while remaining insignificant for Thailand and Singapore. Given that the threshold-existence test confirms a genuine regime structure for Indonesia at all three windows tested and for Malaysia at two of three windows (with the third at the boundary of significance), we interpret both results as evidence of a good volatility effect consistent with prospect theory (Kahneman and Tversky, 1979) and the regime-dependent volatility pricing documented in equity markets by Aslanidis et al. (2021) and Salvador (2012), and analogous to the good volatility framework proposed by Christiansen and Savva (2023) in the bond market context, though the Malaysian effect is weaker and less precisely estimated than Indonesia's. For the Philippines, we regard the individually significant high-regime coefficient as suggestive rather than confirmatory, since it is estimated conditional on a threshold split whose existence is not itself statistically established at any of the three windows tested.

The insignificance of the high-regime volatility coefficient for Thailand and Singapore admits several plausible, though untested, explanations. For Thailand, a formal bootstrap test does not reject the null of no threshold (p = 0.23-0.37), and the high-regime coefficient itself is small and statistically indistinguishable from zero (-0.111, t = -0.99; Table 5); recurrent episodes of political disruption over the sample period may introduce return volatility unrelated to the risk-compensation mechanism examined here, though we note this as a plausible contextual factor rather than a tested channel, as our specification does not isolate political-event windows. For Singapore, the threshold test similarly does not reject the null (p = 0.45-0.60), and the high-regime coefficient, while negative (-0.103, t = -0.94; Table 5), is not statistically significant; the market's high degree of international integration and rapid information incorporation could in principle prevent a persistent volatility premium from emerging even in favorable conditions, but we flag this as a speculative interpretation consistent with the null result rather than a mechanism we test directly. Establishing either channel would require event-specific or market-microstructure data beyond the scope of this study.

Motivated by disposition-effect theory (Shefrin and Statman, 1985), we additionally tested an exploratory three-regime specification for the Philippines, allowing for a medium regime between two thresholds. A sequential F-test comparing this specification against the best single-threshold model does not indicate a statistically significant improvement (p = 0.45-0.57 across specifications). We therefore do not interpret the three-regime structure as a confirmed finding and report the Philippines under the single-threshold specification in Table 5; further detail is provided in Section 6.

Within the threshold model, the skewness coefficient under the predetermined specification remains positive and significant for Indonesia ($p < 0.05$) and the Philippines ($p < 0.01$), mirroring the linear OLS results in Table 4 and reinforcing that these are the two markets for which the skewness-return relationship is not merely a contemporaneous, mechanical artifact. The kurtosis coefficient is significant only for the Philippines ($p < 0.05$); its economic interpretation in this market is discussed alongside the exploratory three-regime result in Section 6.

4.5 Robustness checks

To assess the sensitivity of the baseline findings to the choice of rolling window used in constructing the conditional higher-order moments, we re-estimate both the linear and threshold models using windows of 10 and 44 trading days, in addition to the baseline 22-day specification. We also re-estimate the linear OLS model using contemporaneous risk measures in place of predetermined values, and test directly for structural breaks across three sub-periods motivating the sample.

Table 6 | OLS Risk-Return Estimates, Predetermined Specification (Robustness, Window = 10 Days)

	IHSG	KLCI	PSEi	SET	STI
Constant	0.000114	-0.000289	-0.000559	0.000010	0.000068
(t-stat)	(0.1497)	(-0.8592)	(-0.5862)	(0.0168)	(0.1187)
Volatility	0.025259	0.060830	0.067156	0.012290	0.005502
(t-stat)	(0.3287)	(1.0954)	(0.7796)	(0.1811)	(0.0766)
Skewness	0.000066	0.000197**	0.000195	0.000338**	0.000338**
(t-stat)	(0.4015)	(2.0509)	(1.0449)	(2.0646)	(2.5235)
Kurtosis	-0.000155	-0.000108	0.000001	0.000049	-0.000262
(t-stat)	(-0.8615)	(-0.9207)	(0.0055)	(0.2028)	(-1.5321)
Adj. R2	-0.0003	0.0010	0.0000	0.0003	0.0017

Notes: The table reports OLS estimates using the predetermined (lagged) specification with a 10-day rolling window, compared against the 22-day baseline in Table 4. The 10-day window captures short-term distributional dynamics but carries higher sampling noise (Section 3.1). Newey-West HAC standard errors with a fixed bandwidth of 10 lags are reported in parentheses. *, **, *** denote statistical significance at the 10%, 5%, and 1% levels, respectively.

Results from the 10-day window differ notably from the 22-day baseline. Indonesia's skewness coefficient, significant at the baseline window, is not significant here (p = 0.69), consistent with the higher sampling noise at this shorter window (Section 3.1). Conversely, Malaysia, Thailand, and Singapore each show a significant skewness coefficient at the 10-day window ($p < 0.05$) despite being insignificant at the 22-day baseline. No single market other than Indonesia shows significance across more than one window length, which we interpret as evidence that these isolated significant results are more consistent with sampling variation than a genuine, window-robust premium.

Table 7 | OLS Risk-Return Estimates, Predetermined Specification (Robustness, Window = 44 Days)

	IHSG	KLCI	PSEi	SET	STI
Constant	0.000095	-0.000178	-0.000401	0.000228	0.000204
(t-stat)	(0.1217)	(-0.5083)	(-0.4171)	(0.3491)	(0.3494)
Volatility	0.048623	0.045287	0.048556	-0.021497	-0.009630
(t-stat)	(0.6190)	(0.7611)	(0.5490)	(-0.2939)	(-0.1277)
Skewness	0.000514**	0.000043	0.000004	-0.000192	-0.000004
(t-stat)	(2.1709)	(0.3463)	(0.0173)	(-0.8118)	(-0.0150)
Kurtosis	-0.000021	0.000005	0.000054	0.000043	0.000009
(t-stat)	(-0.2060)	(0.1344)	(0.6811)	(0.6098)	(0.1112)
Adj. R2	0.0009	-0.0003	-0.0001	0.0000	-0.0006

Notes: The table reports OLS estimates using the predetermined (lagged) specification with a 44-day rolling window, compared against the 22-day baseline in Table 4. The 44-day window produces smoother, medium-term moment estimates with lower sampling noise but may also average out genuine shorter-horizon variation (Section 3.1). Negative adjusted R2 values indicate the model provides no explanatory power beyond the sample mean at this window length. Newey-West HAC standard errors with a fixed bandwidth of 10 lags are reported in parentheses. *, **, *** denote statistical significance at the 10%, 5%, and 1% levels, respectively.

Under the predetermined specification at the 44-day window, skewness retains significance only for Indonesia ($p < 0.05$), consistent with its significance at the 22-day baseline; the remaining four markets, including the Philippines (significant only at the 22-day window), produce insignificant estimates here. Taken together with the 10-day results, Indonesia is the only market whose skewness coefficient is significant at more than one window length, which we take as the basis for treating its skewness premium as robust and the results for the other four markets, including the Philippines, as not robust to window choice.

Extending the threshold model to the 10-day and 44-day windows (Appendix Tables A1-A2), the bootstrap existence test rejects the null of no threshold for Indonesia at both alternative windows ($p = 0.013$ in each case), corroborating the 22-day baseline result and confirming that Indonesia's regime-dependent volatility pricing is robust to window choice. For Malaysia, the null is rejected at the 10-day and 44-day windows ($p = 0.043$ and $p = 0.023$ respectively) and is at the boundary of rejection at the 22-day baseline ($p = 0.050$); this consistent pattern across all three windows is markedly different from the isolated, single-window significance we treat as unreliable elsewhere in this study, supporting Malaysia's threshold evidence as tentatively confirmed, albeit weaker and less precisely estimated than Indonesia's. The Philippines, Thailand, and Singapore show no threshold evidence at any window.

Table 8 | OLS Risk-Return Estimates, Contemporaneous Specification (Robustness, Window = 22 Days)

	IHSG	KLCI	PSEi	SET	STI
Constant	-0.000574	-0.000707**	-0.002431**	-0.000943	-0.000353
(t-stat)	(-0.7008)	(-1.9639)	(-2.3036)	(-1.3303)	(-0.6464)
Volatility	0.128681	0.145967**	0.252702***	0.134574*	0.080993
(t-stat)	(1.5787)	(2.4213)	(2.5786)	(1.7158)	(1.1545)
Skewness	0.002155***	0.001122***	0.002758***	0.001972***	0.001734***
(t-stat)	(10.7584)	(7.9675)	(10.5948)	(9.2578)	(8.9319)
Kurtosis	0.000231*	0.000041	0.000358**	0.000102	0.000168
(t-stat)	(1.7354)	(0.4642)	(2.2005)	(0.7161)	(1.2417)
Obs	4827	4882	4877	4841	4979
Adj. R2	0.0248	0.0214	0.0328	0.0224	0.0210

Notes: This table reports the contemporaneous specification (Vol(i,t), Skew(i,t), Kurt(i,t)), presented for comparison with our preferred predetermined specification (Table 4). Because the rolling skewness and kurtosis measures are constructed from standardized residuals up to and including period t, these contemporaneous versions mechanically embed information from R(i,t) itself, which the predetermined specification avoids. Newey-West HAC standard errors with a fixed bandwidth of 10 lags (Bartlett kernel) are reported in parentheses. *, **, *** denote statistical significance at the 10%, 5%, and 1% levels, respectively.

Volatility is significant under the contemporaneous specification for Malaysia (0.146, $p < 0.05$) and the Philippines (0.253, $p < 0.01$), and marginally significant for Thailand (0.135, $p < 0.10$); none of this significance survives the predetermined specification in Table 4, consistent with a partly mechanical contemporaneous relationship rather than a forward-looking risk-return trade-off. The contemporaneous skewness coefficient is significant across all five markets ($p < 0.01$ in each case), replicating a pattern that would, at face value, suggest a pervasive, region-wide skewness premium; however, because the rolling skewness measure is constructed from standardized residuals up to and including period t, this contemporaneous relationship partly reflects the mechanical construction of the regressor rather than a forward-looking pricing relationship. The contemporaneous excess kurtosis coefficient is significant for Indonesia (0.000231, $p < 0.10$) and the Philippines (0.000358, $p < 0.05$), compared to significance for the Philippines alone under the predetermined specification. Adjusted R-squared under the contemporaneous specification is markedly higher (1-3%) than under the predetermined specification (at or below zero in all markets). We interpret this gap as evidence that the contemporaneous explanatory power is largely attributable to the mechanical overlap between the regressors and the dependent variable, rather than a forward-looking risk-return relationship, and accordingly treat the predetermined results in Table 4 as our primary findings throughout this paper.

Turning to structural stability, the Wald test for period-specific breaks finds no statistically significant change in the risk-return relationship during the Global Financial Crisis, the COVID-19 pandemic, or the 2022-2023 monetary tightening cycle, for any of the five markets at the 5% level (joint test p-values range from 0.19 to 0.80 across markets). Indonesia shows a marginal COVID-period effect ($p = 0.098$) that does not reach conventional significance. We interpret this as evidence that the risk-return relationships documented in this paper, including the Indonesia-specific skewness and volatility premia, reflect a structurally stable feature of these markets rather than an artifact of any single crisis episode.

Taken together, the robustness results across window specifications, moment timing, and crisis sub-periods indicate that two findings are stable: a regime-dependent volatility premium for Indonesia, confirmed by a formal threshold-existence test at all three windows, and a comparable though weaker pattern for Malaysia; and a skewness premium for Indonesia across the 22-day and 44-day windows under the predetermined specification. Findings for the remaining markets, and for the Philippines' skewness result specifically, are sensitive to window and timing estimation choices and are accordingly not treated as robust. None of the results are found to be driven by any single crisis sub-period.

5. CONCLUSION

This study demonstrates that the risk-return trade-off in ASEAN equity markets is regime-dependent, but market-specific rather than region-wide, findings that challenge the adequacy of both a single linear framework and a uniform regional characterization for capturing risk pricing dynamics in these markets. Using predetermined risk measures to avoid mechanical simultaneity in the moment construction, we find a positive and significant skewness premium in Indonesia that is robust across alternative rolling-window lengths; no comparable, window-robust premium is found in Malaysia, the Philippines, Thailand, or Singapore. A formal bootstrap test for threshold existence, replicated across three rolling-window lengths, confirms regime-dependent volatility pricing for Indonesia at all three windows; Malaysia shows a similar, though weaker and less precisely estimated, pattern that is significant or near-significant at all three windows, while the Philippines shows a suggestive but statistically inconclusive pattern and Thailand and Singapore show none. A structural break test further indicates that these relationships are stable across the Global Financial Crisis, the COVID-19 pandemic, and the 2022-2023 monetary tightening cycle, suggesting the findings reflect an enduring market characteristic rather than a crisis-period artifact. These findings suggest that regime-aware volatility pricing in ASEAN equity markets is concentrated in Indonesia and, to a lesser extent, Malaysia, rather than being a general ASEAN-wide phenomenon, while the skewness premium is specific to Indonesia alone; both patterns motivate further research into the institutional or investor-composition characteristics that distinguish these markets from their regional peers before such patterns are generalized into investment or regulatory prescriptions at the regional level.

6. LIMITATION AND IMPLICATION

This study is subject to several limitations. The sample is restricted to five major ASEAN equity indices, which may not capture the diversity of smaller or frontier ASEAN markets, limiting broader generalizability. Conditional higher-order moments are constructed using fixed rolling windows of standardized GARCH residuals, a transparent but smoothing-constrained approach; model specifications also differ across markets, and for Thailand specifically, no candidate specification fully eliminates residual serial correlation. The analysis is conducted at the index level, abstracting from firm-level heterogeneity. We also note that Zhang et al. (2021) find a good-bad variance decomposition outperforms threshold regression as an out-of-sample forecasting tool; our objective differs in kind, using threshold regression for structural inference (whether and under what regime risk is priced) rather than forecasting, a task for which the Hansen (2000) framework, with its estimable confidence interval, is well suited.

From a theoretical standpoint, the cross-market heterogeneity documented here, a confirmed volatility premium in Indonesia, a weaker pattern in Malaysia, and its absence in Thailand, Singapore, and (for skewness and threshold existence) the Philippines, suggests the good-volatility mechanism (Christiansen and Savva, 2023) generalizes conditionally rather than universally, plausibly bounded by retail investor dominance, short-selling constraints, and market integration; we flag this moderator interpretation as post-hoc and untested with market-specific data, a direction for future research rather than an established mechanism. Relatedly, an exploratory three-regime specification for the Philippines, motivated by disposition-effect theory (Shefrin and Statman, 1985), did not show a statistically significant improvement over the single-threshold model (sequential F-test, $p = 0.45$-0.57) and is accordingly not pursued as a finding. Two further extensions, an ARMA-GARCH-in-mean or instrumental-variables/control-function approach to the timing concern in Section 3.3, and a bootstrap that re-estimates the first-stage GARCH filtration within each replication, would address these issues more fully but substantially increase computational complexity; both are left for future research.

Practically, the regime-dependent volatility premium has direct applications for dynamic asset allocation, risk budgeting, and structured products targeting positively skewed payoffs in the ASEAN region, with the caveat that our findings support such applications most strongly for Indonesia and, with caution, Malaysia. Future research could incorporate realized higher-order moments from high-frequency data, explore alternative threshold variables such as investor sentiment, extend the framework to firm-level returns, and broaden the sample to frontier ASEAN markets to assess the robustness of these behavioral interpretations across a wider range of institutional contexts.

REFERENCES

Aslanidis, N., Christiansen, C., & Savva, C. S. (2021). Quantile risk-return trade-off. Journal of Risk and Financial Management, 14(6).

Bansal, N., Connolly, R. A., & Stivers, C. (2022). Beta and size equity premia following a high-VIX threshold. Journal of Futures Markets, 42(8), 1491-1517.

Bansal, N., & Stivers, C. (2025). Predicting the equity premium with a high-threshold risk level and the price of risk. Financial Management, 54(2), 123-145.

Barberis, N., & Huang, M. (2008). Stocks as lotteries: The implications of probability weighting for security prices. American Economic Review, 98(5), 2066-2100.

Campbell, J. Y., & Hentschel, L. (1992). No news is good news. Journal of Financial Economics, 31(3), 281-318.

Christiansen, C., & Savva, C. S. (2023). Government bond market risk-return trade-off. Quantitative Finance and Economics, 7(2), 249-260.

Conrad, J., Dittmar, R. F., & Ghysels, E. (2013). Ex ante skewness and expected stock returns. The Journal of Finance, 68(1), 85-124.

Dittmar, R. F. (2002). Nonlinear pricing kernels, kurtosis preference, and evidence from the cross section of equity returns. The Journal of Finance, 57(1), 369-403.

Engle, R. F. (1982). Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation. Econometrica, 50(4), 987-1007.

Feng, Y., He, M., & Zhang, Y. (2023). Market skewness and stock return predictability: New evidence from China. Emerging Markets Finance and Trade, 60(2), 251-267.
Feunou, B., Jahan-Parvar, M. R., & Tedongap, R. (2013). Modeling market downside volatility. Review of Finance, 17(1), 443-481.
French, K. R., Schwert, G. W., & Stambaugh, R. F. (1987). Expected stock returns and volatility. Journal of Financial Economics, 19(1), 3-29.
Ghysels, E., Santa-Clara, P., & Valkanov, R. (2005). There is a risk-return trade-off after all. Journal of Financial Economics, 76(3), 509-548.
Glosten, L. R., Jagannathan, R., & Runkle, D. E. (1993). On the relation between the expected value and the volatility of the nominal excess return on stocks. The Journal of Finance, 48(5), 1779-1801.
Hansen, B. E. (1996). Inference when a nuisance parameter is not identified under the null hypothesis. Econometrica, 64(2), 413-430.
Hansen, B. E. (2000). Sample splitting and threshold estimation. Econometrica, 68(3), 575-603.
Harvey, C. R. (1995). Predictable risk and returns in emerging markets. Review of Financial Studies, 8(3), 773-816.
Harvey, C. R., & Siddique, A. (2000). Conditional skewness in asset pricing tests. The Journal of Finance, 55(3), 1263-1295.
Joanes, D. N., & Gill, C. A. (1998). Comparing measures of sample skewness and kurtosis. Journal of the Royal Statistical Society: Series D (The Statistician), 47(1), 183-189.
Kahneman, D., & Tversky, A. (1979). Prospect theory: An analysis of decision under risk. Econometrica, 47(2), 263.
Kass, R. E., & Raftery, A. E. (1995). Bayes factors. Journal of the American Statistical Association, 90(430), 773-795.
Kraus, A., & Litzenberger, R. H. (1976). Skewness preference and the valuation of risk assets. The Journal of Finance, 31(4), 1085-1100.
Lundblad, C. (2007). The risk return tradeoff in the long run: 1836-2003. Journal of Financial Economics, 85(1), 123-150.
Merton, R. C. (1973). An intertemporal capital asset pricing model. Econometrica, 41(5), 867.
Mnari, O., & Faouel, B. (2024). Price limit bands, risk-return trade-off and asymmetric volatility: Evidence from Tunisian Stock Exchange sectors. Economics and Business Review, 10(3), 142-162.
Newey, W. K., & West, K. D. (1987). A simple, positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix. Econometrica, 55(3), 703.
Salvador, E. (2012). The risk-return trade-off in emerging markets. Emerging Markets Finance and Trade, 48(6), 106-128.
Savva, C. S., & Theodossiou, P. (2018). The risk and return conundrum explained: International evidence. Journal of Financial Econometrics, 16(3), 486-521.
Shefrin, H., & Statman, M. (1985). The disposition to sell winners too early and ride losers too long: Theory and evidence. The Journal of Finance, 40(3), 777-790.
Thampanya, N., Wu, J., Nasir, M. A., & Liu, J. (2020). Fundamental and behavioural determinants of stock return volatility in ASEAN-5 countries. Journal of International Financial Markets, Institutions and Money, 65.
Wang, H., Yan, J., & Yu, J. (2017). Reference-dependent preferences and the risk-return trade-off. Journal of Financial Economics, 123(2), 395-414.
Zhang, Y., Ma, F., Liang, C., & Zhang, Y. (2021). Good variance, bad variance, and stock return predictability. International Journal of Finance and Economics, 26(3), 4410-4423.

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

APPENDIX

Table A1 | Threshold Regression Estimates, Predetermined Specification (Robustness, Window = 10 Days)

	IHSG	KLCI	PSEi	SET	STI
Threshold Lag	RETURN(-1)	RETURN(-1)	RETURN(-1)	RETURN(-1)	RETURN(-3)
Low Threshold	< 0.005934	< -0.000034	< 0.005929	< 0.004043	< 0.002039
Constant	0.001527	-0.000201	0.000927	-0.000874	-0.000712
(t-stat)	(1.4601)	(-0.2986)	(0.6992)	(-1.0898)	(-1.0616)
Volatility	-0.160955	-0.044816	-0.106106	0.059900	0.079941
(t-stat)	(-1.4674)	(-0.4017)	(-0.8550)	(0.6561)	(0.9357)
Obs	3526	2329	3476	3341	2976
Medium Threshold	–	–	not confirmed	–	–
High Threshold	>= 0.005934	>= -0.000034	>= 0.005929	>= 0.004043	>= 0.002039
Constant	-0.001353	-0.000216	-0.001319	0.002179	0.001236
(t-stat)	(-1.2842)	(-0.5905)	(-1.0231)	(2.0154)	(1.4062)
Volatility	0.259711**	0.131877**	0.223283**	-0.095760	-0.105228
(t-stat)	(2.8446)	(2.2338)	(2.0747)	(-0.9018)	(-0.9691)
Obs	1312	2564	1412	1511	2012
Skewness	-0.000002	0.000131	0.000102	0.000258	0.000329**
(t-stat)	(-0.0141)	(1.4302)	(0.5567)	(1.5872)	(2.4589)
Kurtosis	-0.000136	-0.000114	0.000062	0.000090	-0.000254
(t-stat)	(-0.7975)	(-1.0201)	(0.3524)	(0.3730)	(-1.4643)
Adj. R2	0.0141	0.0084	0.0072	0.0040	0.0033

Notes: Threshold regression using the predetermined specification with a 10-day rolling window, compared against the 22-day baseline in Table 5. A fixed-regressor wild bootstrap test (Hansen, 1996, 2000; 300 replications) rejects the null of no threshold for Indonesia (p = 0.013) and Malaysia (p = 0.043), and does not reject the null for the Philippines (p = 0.16), Thailand (p = 0.29), or Singapore (p = 0.60). Newey-West HAC standard errors with a fixed bandwidth of 10 lags are reported in parentheses. *, **, *** denote statistical significance at the 10%, 5%, and 1% levels.

Table A2 | Threshold Regression Estimates, Predetermined Specification (Robustness, Window = 44 Days)

	IHSG	KLCI	PSEi	SET	STI
Threshold Lag	RETURN(-1)	RETURN(-1)	RETURN(-1)	RETURN(-1)	RETURN(-3)
Low Threshold	< 0.005945	< -0.000038	< 0.006357	< 0.004046	< -0.003537
Constant	0.001491	-0.000150	0.000977	-0.000718	-0.001324
(t-stat)	(1.4037)	(-0.2200)	(0.7600)	(-0.8793)	(-1.3249)
Volatility	-0.138213	-0.055329	-0.118142	0.028382	0.098551
(t-stat)	(-1.2422)	(-0.4838)	(-0.9681)	(0.3047)	(0.8944)
Obs	3502	2313	3531	3318	1422
Medium Threshold	–	–	not confirmed	–	–
High Threshold	>= 0.005945	>= -0.000038	>= 0.006357	>= 0.004046	>= -0.003537
Constant	-0.001382	-0.000075	-0.001138	0.002437	0.000923
(t-stat)	(-1.2960)	(-0.2056)	(-0.8354)	(2.1918)	(1.3059)
Volatility	0.281482***	0.116541*	0.208270*	-0.129002	-0.077630
(t-stat)	(3.0314)	(1.9292)	(1.8379)	(-1.1528)	(-0.8015)
Obs	1302	2546	1323	1500	3532
Skewness	0.000429*	0.000014	-0.000055	-0.000240	-0.000056
(t-stat)	(1.9244)	(0.1209)	(-0.2321)	(-1.0350)	(-0.2273)
Kurtosis	-0.000024	-0.000001	0.000066	0.000038	-0.000002
(t-stat)	(-0.2525)	(-0.0443)	(0.8538)	(0.5339)	(-0.0251)
Adj. R2	0.0152	0.0077	0.0074	0.0042	0.0013

Notes: Threshold regression using the predetermined specification with a 44-day rolling window, compared against the 22-day baseline in Table 5. A fixed-regressor wild bootstrap test (Hansen, 1996, 2000; 300 replications) rejects the null of no threshold for Indonesia (p = 0.013), and does not reject the null for Malaysia (p = 0.023, though note the consistent point-estimate pattern discussed in Section 4.5), the Philippines (p = 0.11), Thailand (p = 0.24), or Singapore (p = 0.52). Newey-West HAC standard errors with a fixed bandwidth of 10 lags are reported in parentheses. *, **, *** denote statistical significance at the 10%, 5%, and 1% levels.