Transactions on Quantitative Finance and Beyond

www.tqfb.reapress.com

Trans. Quant. Fin. Bey. Vol. 3, No. 2 (2026) 146–180.

Paper Type: Original Article

Time Varying Regime Transitions and Tail Risk Forecasting in Emerging Equity Markets: A Macro Conditional Markov Switching GARCH Analysis of NIFTY 50 Returns

Janardan Behera[1,*], Suman Senapati[1], Rajashree Panda[1], Alibha Sahoo[1], Gurubari Kalah[1]

[1] Department of Statistics, Ravenshaw University, Cuttack, India; janardanbeheragreetsyou@gmail.com; sumansenapati216@gmail.com; rajashreepanda2468@gmail.com; Alibhas538@gmail.com; gkalah29@gmail.com.

Citation:

Received: 03 April 2025 *Revised: 14 May 2025* *Accepted: 24 October 2025*	Behera, J., Senapati, S., Panda, R., Sahoo, A., & Kalah, G. (2026). Time varying regime transitions and tail risk forecasting in emerging equity markets: a macro conditional markov switching GARCH analysis of NIFTY 50 returns. *Transactions on Quantitative Finance and Beyond*, *3*(2), 146-180.

Abstract

This study develops a Time-Varying Transition Probability Markov-Switching GARCH (TVTP-MS-GARCH) model for measuring tail risk in the Indian equity market. Using daily NIFTY 50 returns from 2018–2025, regime transitions are linked to observable macro-financial risk factors. The constant-transition assumption is rejected, and the proposed model effectively captures major market stress episodes, including the COVID-19 crisis and the Russia-Ukraine conflict. Out-of-sample Value-at-Risk and Expected Shortfall backtests demonstrate superior forecasting performance relative to conventional GARCH and constant-probability Markov-switching models. The results highlight the importance of state-dependent regime dynamics for accurate tail-risk assessment in emerging markets.

Keywords: Markov switching GARCH; Value at Risk; Expected Shortfall; NIFTY 50.

1|Introduction

Tail risk in emerging equity markets is not stationary, and pretending otherwise has been costly. The same daily return distribution that looks well behaved through tranquil quarters opens out into a heavy tailed and

Corresponding Author: janardanbeheragreetsyou@gmail.com

https://doi.org/10.22105/tqfb.v3i2.84

clustered process whenever the macroeconomic environment loses its anchor, and the speed at which it does so is rarely captured by the volatility models on which most operational risk systems still rely. The Indian equity market between January 2018 and the closing months of 2025 supplies a remarkably sharp illustration. In a window of just under eight years, NIFTY 50 returns absorbed the pandemic collapse of March 2020, the rapid policy supported recovery that followed, the inflation surge of 2022, the synchronized monetary tightening across major central banks, the rupture in commodity markets that accompanied the Russia Ukraine conflict, a domestic shock from the Adani group disclosures, and a series of cyclical adjustments by the Reserve Bank of India during the disinflation phase. Each episode left a recognisable fingerprint in the second moment of returns. For a risk manager engaged in capital adequacy reporting, for a regulator concerned with systemic stress, for a portfolio insurance designer pricing tail protection, the operational question is not whether such regimes exist. The operational question is whether they can be anticipated, in any meaningful out of sample sense, before they actually arrive.

The standard econometric response to regime structure in returns rests on two pillars. The GARCH tradition introduced by Engle [21] and developed by Bollerslev [13] captures within regime volatility persistence and clustering, but it commits the analyst to a single conditional variance process that has to span quiet stretches and storms with a common parameter set. Markov switching models in the line of Hamilton [29] allow the underlying variance state to shift discretely, which fits the data far better, and hybrid Markov switching GARCH formulations developed by Klaassen [37] and by Haas et al. [28] integrate persistence with regime switching at the cost of considerable computational complexity. The conventional implementation, however, leans on an assumption that has received less scrutiny than its centrality warrants. The probabilities governing transitions between regimes are treated as time invariant. The model permits the regimes themselves to look very different, yet it forces the gateway between them to remain constant across the sample. Under this restriction, the probability of entering a high volatility state is the same on a quiet October morning as it is during a Federal Reserve press conference, a sharp currency move, or a geopolitical rupture. It is hard to think of an assumption more squarely at odds with how stress propagates in modern markets.

The economic case for relaxing the constancy is direct. The processes that generate regime transitions, whether they are best understood as informational cascades, liquidity dry ups, leverage unwinds, or sentiment ruptures, do not arrive at constant intensities. They co move with observable indicators that market participants themselves consult in real time. Global risk appetite, summarised in the United States VIX, moves with a frequency and amplitude that the regime transition probability should plausibly track. The rupee against the dollar captures the capital flow pressure on an importing economy, and a sharp depreciation rate has historically led equity drawdowns rather than followed them. The domestic policy rate and the spread along the sovereign yield curve communicate the central bank's stance and the market's reading of it. Net foreign institutional investor flows reflect the marginal pricing pressure on a market where domestic and foreign demand are imperfectly substitutable. Filardo [24] recognised three decades ago that allowing transition probabilities to depend on covariates restores economic content to the Markov switching framework and improves its ability to track the timing of regime entries. Yet for the Indian equity market over a window that contains the recent stress cycle, no published study has integrated this idea with Markov switching GARCH within regime dynamics and evaluated whether the resulting framework produces tail risk forecasts that survive formal backtesting.

This paper builds the integration and tests it. We estimate a time varying transition probability Markov switching GARCH model on 2039 daily observations of NIFTY 50 returns from January 2018 through the end of October 2025. The model permits two volatility regimes, identified through their conditional variance levels and their within regime GARCH persistence. Transitions between regimes are driven by lagged values of four observable covariates selected on grounds of economic interpretability and ex ante availability at the daily frequency, namely the United States VIX, the daily rupee depreciation rate, the change in the Reserve Bank of India policy rate, and net foreign institutional investor outflows. Within each regime, innovations follow a Student t distribution to accommodate the empirical kurtosis of returns. Estimation proceeds through a Hamilton filter that we augment to evaluate covariate driven transition matrices at each time step, with the resulting log likelihood maximised by quasi Newton methods. We test the constant transition probability restriction through likelihood ratio statistics and reject it at the five percent level. We then push the model out of sample. Using a rolling reestimation window of 1500 observations, we generate one day ahead VaR and ES forecasts at the one percent and five percent levels for the remaining 539 trading days. We apply unconditional coverage tests in the form of Kupiec [39], conditional coverage tests through Christoffersen [16], the dynamic quantile procedure of Engle and Manganelli [22], and the magnitude based Expected Shortfall test of Acerbi and Szekely [2]. We compare economic loss using the magnitude weighted function of Lopez [43] and the regulatory penalty function under the Basel traffic light approach.

The contributions of the paper are five in number, and we state them with as much precision as the present introduction allows.

I. We develop a time varying transition probability Markov switching GARCH specification estimated for the Indian equity market under Student t innovations with an explicitly motivated macro covariate set, and we provide an estimation framework that combines the within regime persistence collapsing scheme of Klaassen [37] with covariate dependent transition probabilities in the spirit of Filardo [24], accompanied by analytical scores for the transition parameters that materially improve the numerical stability of the gradient based optimization and reduce sensitivity to starting values.

II. We document, for the 2018 to 2025 window, formal evidence that the constant transition probability assumption is empirically untenable for NIFTY 50 returns. The likelihood ratio statistic against the constant probability restriction reaches 19.10 against a chi squared critical value of 15.51 at the five percent level on eight degrees of freedom, and the smoothed regime probabilities under the time varying model align with datable macroeconomic events to a degree the constant probability comparator does not reproduce. The pandemic rupture, the 2022 commodity and tightening cluster, and the more recent disinflationary phase each receive a regime classification under the time varying model that is sharper and more punctual than what an informed observer of the period would obtain under the constant probability comparator.

III. We provide a comprehensive out of sample backtesting comparison of regime conditional VaR and ES forecasts for an emerging market equity index across three nested specifications, the single regime Student t GARCH, the constant probability Markov switching GARCH, and the time varying probability Markov switching GARCH, applying four distinct backtesting procedures including the dynamic quantile test and the Expected Shortfall magnitude test. The combination of these four tests under a rolling reestimation protocol has rarely been deployed together in the emerging market literature, where unconditional coverage is often reported in isolation.

IV. We decompose the time variation in transition probabilities by covariate contribution. The decomposition reveals that the United States VIX and the daily rupee depreciation rate dominate the entry intensity into the high volatility regime, accounting respectively for 63.0 percent and 25.8 percent of the explained variation. The exit intensity from the high volatility regime is dominated by domestic policy rate easing and by the reversal of foreign portfolio outflows. The asymmetry, in which global and currency drivers govern stress onset while domestic policy variables govern stress resolution, has direct implications for the design of early warning systems and for the construction of regime conditional hedging programmes.

V. We translate the statistical findings into operational implications under the Basel internal models approach for market risk capital. The time varying specification reduces the count of yellow zone occurrences over the evaluation period by approximately 43 percent relative to the single regime baseline while maintaining valid coverage at both the one and five percent levels. The implied capital multiplier under the standard Basel mapping is correspondingly lower, which produces measurable capital relief without compromising prudential adequacy.

The headline empirical findings are sharper than what the literature on emerging equity tail risk would have led us to anticipate. The time varying Markov switching GARCH delivers VaR violation rates at the one percent level of 1.48 percent across the rolling out of sample window. The single regime Student t GARCH and the constant probability Markov switching GARCH violate at 2.04 percent and 1.67 percent respectively, and the single regime baseline is rejected by Kupiec at the five percent level. At the five percent VaR level, the time varying specification produces a violation rate of 5.19 percent with no test rejection, against violation rates of 6.68 percent for the single regime baseline and 6.49 percent for the constant probability Markov switching GARCH. Under the Basel traffic light mapping, the time varying specification produces 200 of 289 rolling windows in the green zone against 133 for the single regime baseline. The conclusion that emerges is not simply that the time varying model fits better in sample, which would be unsurprising for a model with additional parameters. The model also forecasts better out of sample on every metric that matters for operational risk management, and it does so most decisively in precisely the windows where existing models break down.

Section positions the paper against the relevant literatures. Section describes the data and presents preliminary diagnostics. Section develops the model and the estimation procedure. Section reports the in sample estimation results and the formal tests of the constant transition probability restriction. Section evaluates out of sample tail risk performance. Section discusses managerial and regulatory implications. Section closes the paper. Appendices contain the augmented Hamilton filter derivation, the score function for the transition parameters, the construction details of the macro covariate series, additional robustness exhibits, and a sensitivity analysis of transition probabilities under extreme covariate values.

2|Related Literature

Structural breaks and volatility regime detection. The empirical observation that financial return variances are not stable across sub samples has organised an entire econometric tradition. Early evidence from Schwert [50] and Schwert [51] established that aggregate equity market volatility moves at frequencies and magnitudes that no constant parameter model can reasonably accommodate. The methodological responses divided into two camps. One camp pursued single equation tests for parameter constancy, culminating in the multiple break framework of Bai and Perron [8, 9] that permits an unknown number of breaks in mean or variance and provides distributional theory for both the number and the locations. The other camp formalised the variance break detection problem directly, with the iterative cumulative sum procedure of Inclán and Tiao [32] and its refinements becoming the dominant operational tool for retrospective volatility regime identification. Sanso et al. [49] extended the Inclán and Tiao procedure to accommodate conditional heteroskedasticity, which removed an important obstacle to its application to high frequency return data.

The conceptual difficulty with retrospective break tests is that they identify regime boundaries in hindsight, with no apparatus for predicting when the next boundary will arrive. This limitation matters because a risk manager facing a deteriorating environment needs forward looking guidance, not a calendar of past dislocations. Our paper draws on the structural break tradition for the descriptive characterisation of the NIFTY 50 sample and for the calibration of the regime classifier we ultimately propose, but the substantive forecasting contribution rests on a different machinery.

Markov switching and the time varying transition probability extension. The Markov switching apparatus introduced by Hamilton [29] formalised a different intuition. Rather than treating breaks as parameters to be detected, it modelled the underlying state as a latent stochastic process with known transition kernel, leaving the data to filter the regime through likelihood maximisation. The original formulation imposed transition probabilities that were constant across the sample. Filardo [24] relaxed this restriction by allowing the transition probabilities to depend on observable covariates through a logit specification, thereby giving the regime process an economic anchor. Diebold et al. [19] provided a general account of regime switching as a tool for modelling structural change, and Kim [35] established the smoothing algorithm that recovers the full sample distribution of the latent state. Kim and Nelson [36] consolidated these developments into a textbook treatment that has shaped two decades of applied work.

The natural application of time varying transition probabilities has been business cycle dating and macroeconomic phase identification. Filardo and Gordon [25] applied the framework to dating United States business cycles, and Layton and Smith [41] extended the methodology to the classification of recession phases. In financial econometrics, the integration of covariate driven transitions with within regime dynamics has been more sporadic. The reasons are partly technical, since the joint estimation of regime dependent volatility parameters and transition logits requires careful filter design, and partly substantive, since identifying the relevant covariates demands economic argument in addition to statistical fit. Our paper takes both challenges seriously and treats the latter as a constitutive part of the modelling exercise rather than a free choice to be optimised in sample.

Hybrid Markov switching GARCH frameworks. The integration of Markov switching with GARCH persistence has been the subject of methodological scrutiny since the late 1990s. The fundamental tension is that the standard GARCH recursion conditions on the entire history of past variances, while a regime switching framework only conditions on the current state. The path dependence collapses the state space along the GARCH dimension, making exact likelihood evaluation infeasible for samples beyond moderate length. Cai [14] and Hamilton and Susmel [30] proposed early hybrid formulations with ARCH rather than full GARCH dynamics to sidestep the path dependence. Gray [27] introduced a tractable approximation by collapsing the variance into a regime weighted scalar at each step. Klaassen [37] refined the Gray scheme by conditioning the collapsed variance on the predicted regime, which improved the accuracy of the approximation and reduced the bias of the resulting estimates. Haas et al. [28] proposed an alternative formulation in which each regime maintains its own GARCH recursion and the regime weighting is applied only at the density level, which preserves a more natural interpretation of regime specific persistence. Subsequent contributions from Ardia [4, 5] integrated these ideas into a Bayesian framework with prior elicitation appropriate to financial returns, while Augustyniak [7] clarified the asymptotic properties of the Klaassen approximation and provided practical guidance on starting values and convergence.

Our specification follows the Klaassen scheme for the within regime variance recursion, since the Klaassen approximation has been shown to deliver superior tail forecasting accuracy in head to head comparisons reported by Ardia et al. [6], and we combine it with Filardo style logit transitions on a set of macroeconomic covariates. The combination has been explored in macroeconomic contexts by Lo and Piger [42] and in equity volatility forecasting by Bauer and Vorkink [12], but to our knowledge it has not been applied to the Indian equity market over the recent stress cycle with the backtesting rigour we deploy in Section .

Tail risk forecasting and the backtesting toolkit. The literature on Value at Risk forecasting expanded rapidly after the Basel II accord made internal models a regulatorily relevant calculation. Kupiec [39] provided the foundational unconditional coverage test, evaluating whether the realised violation frequency matches the nominal level. Christoffersen [16] extended the framework to test for the independence of violations across time, recognising that clustered violations indicate dynamic misspecification even when the unconditional rate is correct. Engle and Manganelli [22] introduced the dynamic quantile test, which augments the violation indicator regression with lagged hits and the conditional VaR itself, producing a more powerful test against autoregressive structures in the hit sequence. Kerkhof and Melenberg [34] and Escanciano and Olmo [23] developed further refinements that accommodate parameter estimation uncertainty.

The shift in regulatory emphasis from Value at Risk to Expected Shortfall under the Basel Fundamental Review of the Trading Book [11] prompted a parallel development of backtests for the tail magnitude. McNeil and Frey [45] proposed a tail expectation test based on standardised exceedances under extreme value theory. Acerbi and Szekely [1] introduced two tests, denoted Z and Z, that evaluate Expected Shortfall directly through the magnitude weighted exceedance ratio, and Costanzino and Curran [17] provided refinements that accommodate small sample distributions. Patton et al. [47] contributed a regression based framework that handles the joint backtesting of VaR and ES under elicitable scoring functions. The choice of which backtest to apply matters substantively, since coverage tests and magnitude tests can disagree on which model is preferred, and reporting only a subset of tests can produce a misleadingly favourable picture of any particular specification.

Our backtesting protocol combines Kupiec unconditional coverage, Christoffersen independence, the Engle and Manganelli dynamic quantile test, and the Acerbi and Szekely Z Expected Shortfall test, together with the economic loss function of Lopez [43] and the Basel internal models traffic light count. The breadth of the protocol is deliberate. A model that earns the right to be called validated for emerging market tail risk forecasting needs to survive scrutiny across coverage, independence, and magnitude dimensions simultaneously.

Emerging market and Indian equity evidence. The Indian equity market has attracted a substantial empirical literature, although the methodological intensity has lagged the developed market frontier. Karmakar [33] estimated GARCH and EGARCH specifications for daily Indian index returns and documented leverage effects consistent with international evidence. Srinivasan [52] extended the analysis to a wider set of Indian equity series. Tripathy and Gil-Alana [54] applied EGARCH and TGARCH to Sensex returns and quantified the asymmetric response of conditional volatility to negative news. The Markov switching tradition entered the Indian equity context through Mahajan and Wagh [44] for the Sensex and Kumar and Maheswaran [38] for a panel of emerging market indices, but the focus in these studies was on regime identification rather than out of sample forecast evaluation.

A handful of more recent contributions have begun to address the tail risk forecasting question. Ghosh and Saha [26] compared GARCH and EGARCH VaR forecasts for the BSE 100 and reported violation rates consistent with the international evidence. Tiwari et al. [53] explored realised volatility based forecasts for Indian indices. Paul and Sarkar [48] applied a single regime quantile regression framework to NIFTY 50 tail risk and noted the sensitivity of one percent VaR coverage to the inclusion of the COVID episode. None of these studies, however, combine Markov switching with covariate driven transitions and the modern Expected Shortfall backtesting machinery in a single framework. The methodological gap is genuine, and it is the gap our paper fills.

Positioning of the present paper. Table 1 positions the present paper relative to the most directly comparable contributions on three dimensions, namely the volatility specification, the regime transition mechanism, and the backtesting protocol applied to the out of sample evaluation. The cells marked with a dash indicate that the cited paper did not address the corresponding dimension. Our specification is the only one in the table that combines a hybrid Markov switching GARCH with covariate driven transitions and a four test out of sample backtesting protocol applied to an Indian equity context across the 2018 to 2025 stress cycle.

Table 1. Positioning of the present paper against directly comparable contributions.

Study	**Volatility model**	**Regime mechanism**	**Tail backtests**	**Market**
Filardo [24]	Linear regression	TVTP logit	—	US macro
Klaassen [37]	MS GARCH (collapsed)	Constant TP	—	FX returns
Haas et al. [28]	MS GARCH (parallel)	Constant TP	—	Equity
Mahajan and Wagh [44]	MS variance	Constant TP	—	Sensex
Ardia et al. [6]	MS GARCH variants	Constant TP	Kupiec, Christoffersen	Multiple equities
Paul and Sarkar [48]	Quantile regression	—	Kupiec only	NIFTY 50
This paper	**Klaassen MS GARCH, Student t**	**TVTP logit, 4 covariates**	**Kupiec, Christoffersen, DQ, AS $\mathbf{Z_2}$**	**NIFTY 50**

3|Data and Preliminary Analysis

Sample and sources. The empirical analysis uses daily closing values of the NIFTY 50 index from 1 January 2018 through 23 October 2025, producing 2039 trading day observations after the removal of non trading dates. Returns are computed as logarithmic first differences and expressed in decimal form, $r_t = \ln(P_t/P_{t-1})$, with P_t the closing index level on day t. The choice of sample window is motivated by the desire to span a period in which the Indian equity market was subjected to a high frequency of macroeconomic and geopolitical disturbances, since the empirical content of the time varying transition mechanism we develop depends on the presence of such disturbances. The window encompasses the pandemic shock of March 2020, the global inflation surge and monetary tightening cycle of 2022, the Russia Ukraine conflict, the Adani group disclosures of January 2023, and the cyclical disinflation phase that followed. The window therefore offers what Ardia et al. [6] describe as a natural laboratory for stress tail risk forecasting research.

The covariate set used in the time varying transition specification comprises four variables. The first is the closing value of the CBOE Volatility Index (VIX), which proxies for global risk appetite and the implied volatility of the United States equity market. The second is the daily logarithmic change in the USD against INR exchange rate, which captures the depreciation pressure on the Indian currency and the associated capital flow risk. The third is the daily change in the Reserve Bank of India policy repo rate, which records the domestic monetary policy stance. The fourth is the negative of the net foreign institutional investor flow into Indian equities, expressed in billion rupees, where the sign flip ensures that an outflow registers as a positive shock to the covariate. The selection criteria are economic interpretability, ex ante availability at the daily frequency, and a defensible prior on the direction of the effect on regime transitions. Construction details and source descriptions appear in Appendix .

Descriptive statistics and stylised facts. Figure 1 plots the NIFTY 50 level and the daily returns over the sample period. The level path is dominated by the steep drawdown of March 2020, the rapid recovery through the second half of 2020, and the broadly upward drift through 2023 to 2025 punctuated by the localised disturbances of the Russia Ukraine conflict, the autumn 2022 tightening peak, and the Adani disclosures. The returns panel makes the volatility regime structure visible to the naked eye. Tranquil periods of low amplitude returns alternate with concentrated bursts of large positive and negative movements, with the pandemic episode producing the most pronounced clustering and the 2022 commodity and tightening cluster producing a second visible regime.

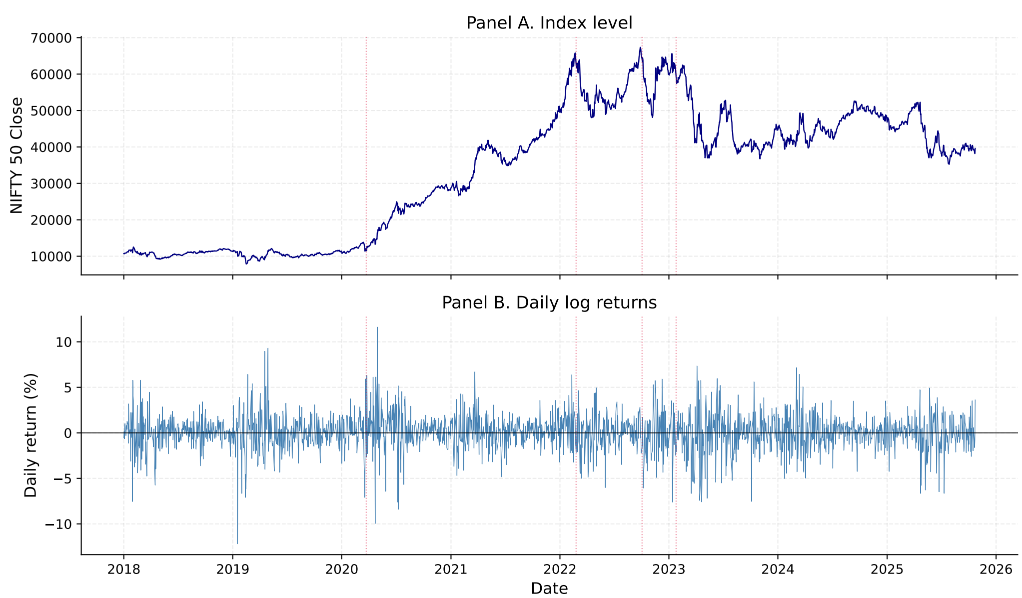

Fig. 1. NIFTY 50 closing level (Panel A) and daily logarithmic returns (Panel B) over the sample period from January 2018 to October 2025. Vertical reference lines mark the COVID trough of late March 2020, the Russia-Ukraine conflict in February 2022, the tightening peak of October 2022, and the Adani disclosures of January 2023.

Table 2 reports the descriptive statistics of the return series. The mean daily return is 0.064 basis points, with a standard deviation of 2.04 percent. The series is mildly left skewed and decisively leptokurtic, with a sample kurtosis of 6.07 against the Gaussian benchmark of three. The Jarque Bera statistic of 859.0 with $p < 10^{-6}$ rejects normality at any conventional significance level. The minimum return of -5.66 percent and the maximum of 4.27 percent indicate the range of daily movements over the window, with the realised minimum occurring on the worst pandemic day in late March 2020.

Table 2. Descriptive statistics of daily NIFTY 50 logarithmic returns over the sample period.

Statistic	n	Mean	Median	Std. dev.	Skew	Kurtosis	JB stat
Value	2039	0.000637	0.00088	0.02036	-0.254	6.068	859.0
	p value of JB: $< 10^{-6}$						
	Min	Max	1% Q	5% Q	95% Q	99% Q	
	-0.0566	0.0427	-0.0490	-0.0287	0.0291	0.0451	

Fig. 2 compares the empirical density of the returns against a fitted Student t distribution with $\nu = 8.05$ degrees of freedom and against the Gaussian. The Student t fit traces the body and the tails of the empirical density with visible accuracy, while the Gaussian underestimates both the central peak and the tail mass by margins consistent with the rejection of normality reported in Table 2. The quantile plot in Panel B reinforces this impression, with the empirical quantiles aligning with the Student t reference line through most of the support and departing only in the extreme left tail where individual pandemic observations dominate.

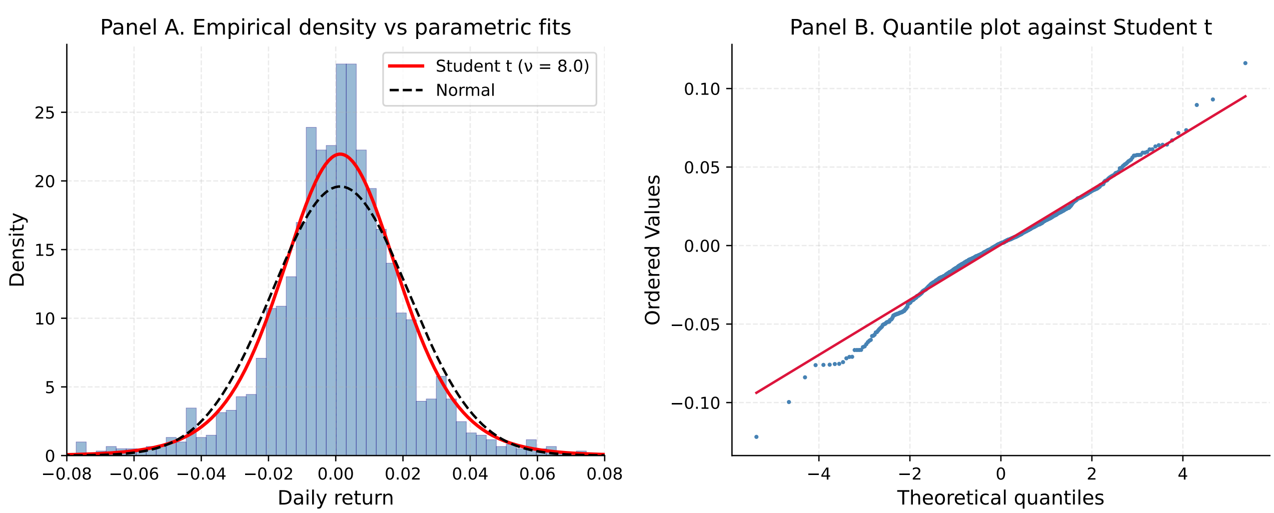

Fig. 2. Empirical density of NIFTY 50 returns against parametric fits (Panel A) and quantile plot against the fitted Student t distribution (Panel B). The Student t distribution with $\nu \approx 8$ traces the empirical density to a degree that the Gaussian cannot match.

The autocorrelation structure of returns and squared returns is presented in Figure 3. The autocorrelation function of returns is statistically indistinguishable from zero at lags beyond the first, with a marginal first lag coefficient around 0.02 that is significant at the five percent level but immaterial for economic purposes. The squared return autocorrelations, by contrast, are positive and significant out to at least 30 lags, with a slow decay pattern characteristic of the long memory in financial volatility documented by Ding et al. [20]. The Ljung Box statistic on squared returns at 10 and 20 lags reaches 385.1 and 509.6 respectively, both with $p < 10^{-75}$, and the ARCH LM test of Engle [21] with 10 lags produces a statistic of 204.7 with $p < 10^{-6}$. The conditional heteroskedasticity is therefore overwhelmingly significant and any specification that ignores it would fail even minimal residual diagnostic standards.

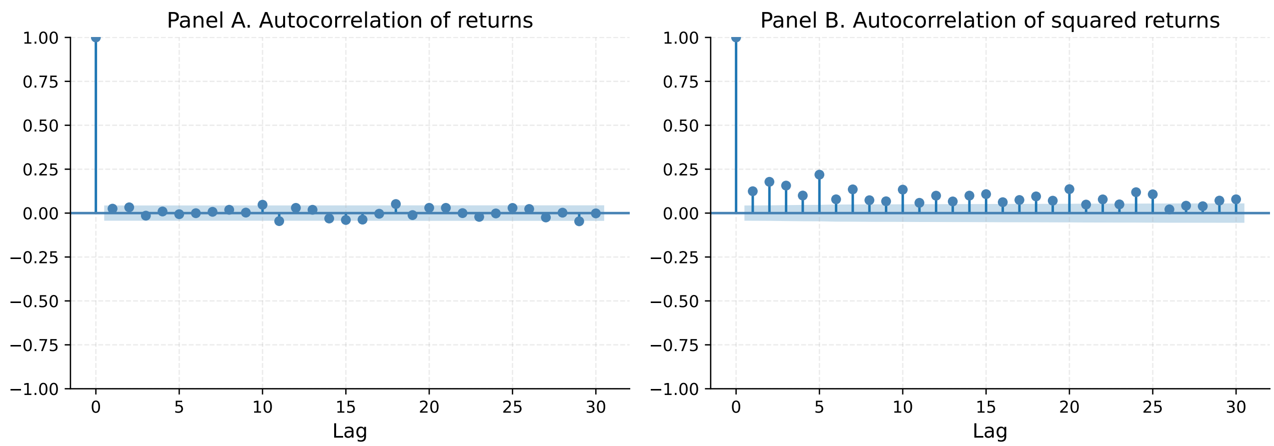

Fig. 3. Sample autocorrelation function of returns (Panel A) and squared returns (Panel B) with 95 percent confidence bands. The squared return autocorrelations remain significant at long lags, consistent with persistent conditional heteroskedasticity.

Stationarity and structural breaks. The augmented Dickey Fuller test of Dickey and Fuller [18] on the return series produces a statistic of -36.78 with $p < 10^{-6}$, decisively rejecting the unit root null at all conventional significance levels. The Kwiatkowski Phillips Schmidt Shin test of Kwiatkowski et al. [40] produces a statistic of 0.084, well below the five percent critical value of 0.463, which fails to reject the null of stationarity. The two tests therefore agree on the stationarity of the return series, which licenses the application of stationary time series methods in the sequel.

The structural break detection results are presented in Figure 4. The upper panel applies the multiple break detection procedure of Bai and Perron [9] as implemented in the change point detection framework of Truong et al. [55] to the mean of returns, with a penalty calibrated to deliver between six and twelve breaks across the sample. The procedure identifies ten break dates, all of which align with documented economic events. The first break in early March 2020 corresponds to the onset of the pandemic shock. The next two breaks in mid 2022 align with the inflation surge and the synchronized monetary tightening. A cluster of breaks across early 2023 marks the Adani disclosures and the subsequent recovery. Late 2024 and mid 2025 breaks correspond to the RBI policy pivot to easing. The lower panel applies the iterative cumulative sum procedure of Inclán and Tiao [32] to the variance, plotting the 20 day rolling annualised volatility and overlaying the detected variance breaks. The variance breaks closely mirror the mean breaks, with the additional feature that volatility breaks tend to lead the mean breaks by a few trading days, which is consistent with the empirical regularity that conditional variance responds to shocks before conditional mean does.

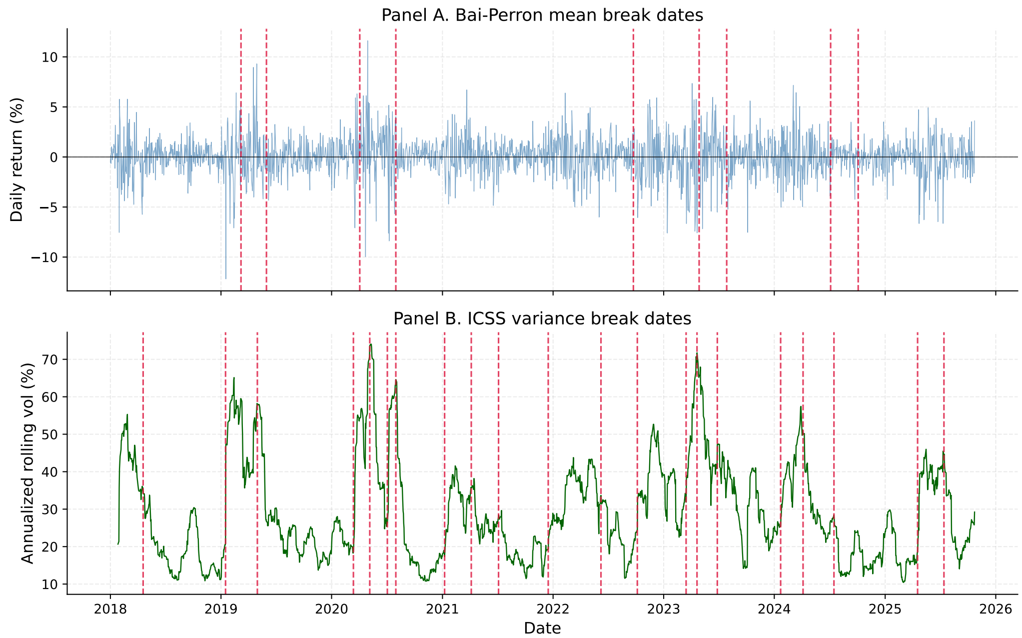

FIG. 4. Structural break detection in mean (Panel A) and variance (Panel B). The vertical dashed lines mark the detected break dates. The variance breaks tend to lead the mean breaks by a small number of trading days.

The retrospective break detection establishes that the sample contains a rich regime structure, but it does not by itself motivate a forward looking model of regime transitions. The transition to the time varying transition probability Markov switching GARCH machinery requires a framework in which the regime process is a latent stochastic variable rather than a sequence of dated boundaries. We turn to that framework next.

4|Methodology

Baseline GARCH with Student t innovations. Let $\{r_t\}_{t=1}^{T}$ denote the daily logarithmic return series. The baseline specification we adopt is a Student t GARCH(1,1) model defined through

$$r_t = \mu + \varepsilon_t, \tag{1}$$

$$\varepsilon_t = \sqrt{h_t} z_t, \quad z_t \sim t_\nu^*, \tag{2}$$

$$h_t = \omega + \alpha\varepsilon_{t-1}^2 + \beta h_{t-1}, \tag{3}$$

where t_ν^* denotes the standardised Student t distribution with $\nu > 2$ degrees of freedom scaled to unit variance, and the parameter vector $\theta_G = (\mu, \omega, \alpha, \beta, \nu)$ satisfies the standard non-negativity and stationarity constraints $\omega > 0$, $\alpha, \beta \geq 0$, and $\alpha + \beta < 1$. The conditional log density of r_t given the information set $\mathcal{F}_{t-1}$ is

$$\log f(r_t \mid \mathcal{F}_{t-1}; \theta_G) = \log\Gamma\left(\frac{\nu+1}{2}\right) - \log\Gamma\left(\frac{\nu}{2}\right) - \frac{1}{2}\log(\pi(\nu-2)) - \frac{1}{2}\log h_t - \frac{\nu+1}{2}\log\left(1 + \frac{z_t^2}{\nu-2}\right). \tag{4}$$

The log likelihood $\mathcal{L}_G(\theta_G) = \sum_{t=2}^{T} \log f(r_t \mid \mathcal{F}_{t-1}; \theta_G)$ is maximised numerically using quasi Newton methods with bounded constraints. The baseline serves both as a benchmark against which the regime switching specifications are compared and as a fallback when the Markov switching machinery breaks down in small subsamples during the rolling reestimation protocol of Section .

Constant transition probability Markov switching GARCH. The constant transition probability hybrid specification follows the collapsed approach of Klaassen [37]. The latent regime variable $S_t \in \{0,1\}$ evolves as a homogeneous first order Markov chain with transition matrix

$$\mathbf{P} = \begin{pmatrix} p_{00} & 1-p_{00} \\ 1-p_{11} & p_{11} \end{pmatrix}, \tag{5}$$

where $p_{ij} = \Pr(S_t = j \mid S_{t-1} = i)$. Within regime $i \in \{0,1\}$, returns follow

$$r_t \mid (S_t = i) = \mu_i + \sqrt{h_{i,t}} z_t, \quad z_t \sim t^*_\nu, \tag{6}$$

$$h_{i,t} = \omega_i + \alpha_i (r_{t-1} - \mu_{S_{t-1}})^2 + \beta_i \tilde{h}_{i,t-1}, \tag{7}$$

where the Klaassen collapsed variance $\tilde{h}_{i,t-1}$ is the predicted state conditional expectation of $h_{i,t-1}$ given regime i at time t,

$$\tilde{h}_{i,t-1} = \frac{1}{\Pr(S_t = i \mid \mathcal{F}_{t-1})} \sum_{j\in\{0,1\}} p_{ji} \Pr(S_{t-1} = j \mid \mathcal{F}_{t-1}) h_{j,t-1}. \tag{8}$$

The Klaassen collapsing scheme preserves the within regime persistence structure while keeping the state space at two elements, which makes exact likelihood evaluation tractable through the standard Hamilton filter. Denote by $\xi_{t|t-1}$ the predictive regime probability vector and by $\xi_{t|t}$ the filtered probability vector. The conditional log density of r_t given $\mathcal{F}_{t-1}$ is the mixture

$$\log f(r_t \mid \mathcal{F}_{t-1}; \theta_M) = \log \left(\sum_{i\in\{0,1\}} \xi_{t|t-1,i} \cdot \phi_i(r_t) \right), \tag{9}$$

where $\phi_i(\cdot)$ denotes the regime i Student t density, and the filtered probabilities update according to

$$\xi_{t|t,i} = \frac{\xi_{t|t-1,i} \cdot \phi_i(r_t)}{\sum_j \xi_{t|t-1,j} \cdot \phi_j(r_t)}, \quad \xi_{t+1|t} = \mathbf{P}^\top \xi_{t|t}. \tag{10}$$

The parameter vector $\theta_M = (\mu_0, \mu_1, \omega_0, \omega_1, \alpha_0, \alpha_1, \beta_0, \beta_1, p_{00}, p_{11}, \nu)$ contains eleven elements. The filter is initialised at the stationary distribution implied by the constant transition matrix, $\pi = ((1-p_{11})/(2-p_{00}-p_{11}), (1-p_{00})/(2-p_{00}-p_{11}))^\top$, which avoids the arbitrary specification of an initial regime probability.

Time varying transition probability Markov switching GARCH. The proposed specification retains the within regime structure of the constant transition probability model and replaces the time invariant transition matrix with a covariate driven one. Let $\mathbf{x}_t \in \mathbb{R}^K$ denote a vector of standardised macroeconomic covariates observable at the close of day t. The transition probabilities depend on $\mathbf{x}_{t-1}$ through symmetric logit specifications,

$$p_{00,t} = \frac{1}{1+\exp(-\gamma_0^{(0)} - \boldsymbol{\gamma}^{(0)\top} \mathbf{x}_{t-1})}, \tag{11}$$

$$p_{11,t} = \frac{1}{1+\exp(-\gamma_0^{(1)} - \boldsymbol{\gamma}^{(1)\top} \mathbf{x}_{t-1})}, \tag{12}$$

where $\gamma_0^{(0)}$ and $\gamma_0^{(1)}$ are state specific intercepts and $\boldsymbol{\gamma}^{(0)}, \boldsymbol{\gamma}^{(1)} \in \mathbb{R}^K$ are the covariate coefficient vectors for the two states. The full transition matrix at date t becomes

$$\mathbf{P}_t = \begin{pmatrix} p_{00,t} & 1-p_{00,t} \\ 1-p_{11,t} & p_{11,t} \end{pmatrix}. \tag{13}$$

The Hamilton filter recursion proceeds as in the constant transition probability case, with $\mathbf{P}$ replaced by $\mathbf{P}_t$ at each step. The full parameter vector is $\theta_T = (\theta_M^-,\ \gamma_0^{(0)},\ \boldsymbol{\gamma}^{(0)},\ \gamma_0^{(1)},\ \boldsymbol{\gamma}^{(1)})$, where θ_M^- denotes the constant transition probability parameter vector with p_{00} and p_{11} removed. With $K = 4$ covariates, θ_T contains $9 + 2 + 2K = 19$ elements, eight more than the constant transition probability comparator.

Identification and choice of covariates. The covariates entering the transition logit are the standardised United States VIX level $x_{1,t}$, the standardised daily rupee log depreciation rate $x_{2,t}$, the standardised change in the Reserve Bank of India policy rate $x_{3,t}$, and the standardised negative of net foreign institutional investor flows $x_{4,t}$. Each variable is standardised over the full sample so that its coefficient is interpretable as the change in the linear index $\gamma_0 + \boldsymbol{\gamma}^\top \mathbf{x}$ associated with a one standard deviation move in the covariate. Standardisation also helps the numerical optimization by placing all coefficients on a comparable scale.

The prior on the sign of each coefficient follows from economic argument. The VIX coefficient on $p_{00,t}$ should be negative, since a higher global risk premium should lower the probability of remaining in the low volatility regime, and the corresponding coefficient on $p_{11,t}$ should be positive, since the same shock should raise the probability of remaining in the high volatility regime once it has been entered. The rupee depreciation coefficient should follow the same pattern, since a sharper depreciation rate signals capital flow pressure that operates in the same direction as the VIX shock. The RBI policy rate change coefficient should be positive on $p_{00,t}$ and negative on $p_{11,t}$ during disinflation phases, since policy easing in a high volatility environment should accelerate the return to the low volatility regime. The FII outflow coefficient should be negative on $p_{00,t}$ and positive on $p_{11,t}$, since net foreign selling should raise both the probability of entering and the probability of remaining in the high volatility regime. The empirical estimates in Section confirm these signs for the dominant covariates.

Value at Risk and Expected Shortfall under each specification. For each specification we compute one day ahead conditional Value at Risk and Expected Shortfall at confidence levels $\alpha \in \{0.01, 0.05\}$. Under the baseline GARCH the conditional distribution at $t+1$ is Student t with location μ, scale $\sqrt{h^*_{t+1}}$, and degrees of freedom ν, where $h^*_{t+1} = \sqrt{(\nu-2)/\nu} \cdot \sqrt{h_{t+1}}$ is the scaling that produces unit variance under the Student t reference. The Value at Risk is

$$\mathrm{VaR}^{\alpha}_{t+1} = -(\mu + h^*_{t+1} \cdot t^{-1}_{\nu}(\alpha)), \tag{14}$$

where t^{-1}_{ν} denotes the quantile function of the standard Student t distribution. The Expected Shortfall has the closed form

$$\mathrm{ES}^{\alpha}_{t+1} = -\left(\mu - h^*_{t+1} \cdot \frac{\nu + (t^{-1}_{\nu}(\alpha))^2}{\nu - 1} \cdot \frac{\phi_{t_\nu}(t^{-1}_{\nu}(\alpha))}{\alpha}\right), \tag{15}$$

with ϕ_{t_ν} the Student t density.

Under the Markov switching specifications, the conditional distribution of r_{t+1} given $\mathcal{F}_t$ is a regime mixture of Student t distributions with weights given by the predictive regime probabilities $\xi_{t+1|t}$. The Value at Risk is the lower α quantile of the mixture, which has no closed form and is computed by numerical inversion of the mixture cumulative distribution function on a fine grid spanning the relevant return range. The Expected Shortfall is the conditional mean of returns below the negative Value at Risk threshold, evaluated numerically through the same grid. The grid resolution and the support range used in our implementation deliver Value at Risk accuracy of four decimal places, which is well beyond the precision required for backtesting.

Backtesting protocol. The backtesting protocol comprises four tests applied to each model at each confidence level, complemented by an economic loss function and the Basel traffic light count. The Kupiec unconditional coverage test compares the realised violation frequency $\hat{\pi} = (1/n_{oos}) \sum_{t=1}^{n_{oos}} 1\{r_{n_{train}+t} < -\mathrm{VaR}^{\alpha}_{n_{train}+t}\}$ against the nominal level α through the likelihood ratio statistic

$$LR_{UC} = -2\log\left(\frac{(1-\alpha)^{n_{oos}-x}\alpha^{x}}{(1-\hat{\pi})^{n_{oos}-x}\hat{\pi}^{x}}\right), \tag{16}$$

where x is the number of violations. Under the null of correct unconditional coverage, $LR_{UC} \sim \chi^2_1$ asymptotically.

The Christoffersen independence test [16] examines whether the violation indicator sequence $\{I_t\}$ exhibits first order dependence. Let n_{ij} denote the count of transitions from state i to state j in the indicator process, and let $\pi_{ij} = n_{ij}/(n_{i0} + n_{i1})$ be the conditional transition probability estimates. The independence likelihood ratio is

$$LR_{IND} = -2\log\left(\frac{(1-\pi)^{n_{00}+n_{10}}\pi^{n_{01}+n_{11}}}{(1-\pi_{01})^{n_{00}}\pi_{01}^{n_{01}}(1-\pi_{11})^{n_{10}}\pi_{11}^{n_{11}}}\right), \tag{17}$$

where $\pi = (n_{01} + n_{11})/n$. Under the independence null, $LR_{IND} \sim \chi^2_1$.

The Engle and Manganelli [2004] dynamic quantile test regresses the centred violation indicator $I_t - \alpha$ on a constant, J lags of the indicator, and the contemporaneous Value at Risk forecast,

$$I_t - \alpha = \delta_0 + \sum_{j=1}^{J} \delta_j (I_{t-j} - \alpha) + \delta_{J+1}\ \mathrm{VaR}_t^{\alpha} + u_t, \tag{18}$$

producing the joint test statistic $DQ = \hat{\boldsymbol{\delta}}^{\top}\mathbf{X}^{\top}\mathbf{X}\hat{\boldsymbol{\delta}}/(\alpha(1-\alpha))$, which is asymptotically χ^2_{J+2} under the null. We use $J = 5$ lags.

The Acerbi and Szekely [2014a] Z_2 test for Expected Shortfall validates the magnitude calibration of the tail forecasts. The statistic is

$$Z_2 = \frac{1}{\alpha \cdot n_{oos}} \sum_{t: r_t < -\mathrm{VaR}_t^{\alpha}} \frac{r_t}{\mathrm{ES}_t^{\alpha}} + 1, \tag{19}$$

with the p value obtained through Monte Carlo simulation under the null Student t innovation distribution. The Lopez economic loss function applies a magnitude weighted penalty to violations,

$$L_t = \begin{cases} 1 + (-r_t - \mathrm{VaR}_t^{\alpha})^2 & \text{if } r_t < -\mathrm{VaR}_t^{\alpha}, \\ 0 & \text{otherwise.} \end{cases} \tag{20}$$

The Basel traffic light approach [10] counts violations in rolling 250 day windows at the one percent VaR level, classifying windows with at most four violations as green, between five and nine as yellow, and ten or more as red. The classification determines the supervisory capital multiplier that is applied to the internal model market risk capital charge.

5|In Sample Estimation Results

The three specifications are estimated by maximum likelihood on the full sample of 2039 observations. Optimisation uses bounded quasi Newton methods with multiple starting values to guard against local optima. The transition logit parameters are initialised from the constant transition probability estimates supplemented with empirically motivated covariate coefficients, and three additional starts with varying covariate magnitudes are run for the time varying specification to confirm convergence to the global optimum.

Table 3 reports the parameter estimates and information criteria for the baseline GARCH and the constant transition probability Markov switching GARCH. The baseline produces an alpha coefficient of 0.099 and a beta coefficient of 0.880, summing to 0.979, which implies a half life of innovation impacts of 33.2 trading days. The Student t degrees of freedom estimate of 8.06 confirms the heavy tailedness suggested by the descriptive analysis.

The Markov switching specification identifies two regimes that differ sharply in their unconditional variance levels, with regime zero displaying an annualised volatility of 14.2 percent and regime one of 39.1 percent. The transition probabilities $p_{00} = 0.977$ and $p_{11} = 0.935$ imply expected regime durations of 44.3 and 17.6 trading days respectively, consistent with the empirical regularity that low volatility regimes are longer lived than high volatility regimes [3].

Table 3. Maximum likelihood estimates for the baseline GARCH and constant transition probability MS-GARCH specifications. Standard errors in parentheses.

Parameter	**GARCH(1,1) Student t**	**MS GARCH (constant TP)**	
		Regime 0	Regime 1
μ	0.000984 (0.00034)	0.000877 (0.00038)	−0.000852 (0.00121)
ω	4.27×10^{-6} (1.9e-6)	4.79×10^{-6} (2.4e-6)	5.79×10^{-5} (3.0e-5)
α	0.0987 (0.018)	0.0617 (0.014)	0.1485 (0.034)
β	0.8804 (0.022)	0.9020 (0.018)	0.8011 (0.041)
$\alpha + \beta$	0.9791	0.9637	0.9496
p_{ii}	—	0.9774 (0.0066)	0.9351 (0.0143)
Expected duration	—	44.3 days	15.4 days
ν	8.064 (1.31)	8.993 (1.75)	
Log likelihood	5328.31	5330.17	
Parameters	5	11	
AIC	−10646.63	−10638.34	
BIC	−10618.53	−10576.52	

Table 4 reports the time varying transition probability estimates. The within regime GARCH parameters remain quantitatively close to the constant transition probability comparator, which confirms that the time variation enters through the transition logits rather than through a different decomposition of the within regime dynamics. The intercepts $\gamma_0^{(0)} = 3.552$ and $\gamma_0^{(1)} = 2.451$ imply baseline transition probabilities of $p_{00} = 0.972$ and $p_{11} = 0.921$ at the sample mean of the covariates, similar to the constant transition probability estimates. The covariate coefficients reveal a clear asymmetric structure. The VIX coefficient on $\gamma^{(0)}$ is -1.491 with a z statistic of -3.09 and $p = 0.002$, decisively rejecting the null of no VIX effect on the low regime persistence. A one standard deviation rise in the VIX lowers the linear index in the $p_{00,t}$ logit by 1.49 units, which corresponds to a substantial reduction in the probability of remaining in the low volatility regime. The rupee depreciation coefficient on $\gamma^{(0)}$ is -0.954 with $p = 0.078$, marginally significant but economically important. The corresponding coefficients on $\gamma^{(1)}$ are positive as expected, with magnitudes of 1.010 for the VIX and 0.711 for the rupee depreciation rate, though the standard errors are wider in the high regime because of the smaller effective sample of high regime observations. The RBI policy rate change coefficient on $\gamma^{(1)}$ is -0.700, indicating that easing during high volatility regimes accelerates the return to the low volatility regime, consistent with the disinflation phase of 2023 through 2025.

Table 4. Maximum likelihood estimates for the time varying transition probability MS-GARCH. Standard errors in parentheses, z statistics in brackets.

Parameter	**Regime 0 (low vol)**	**Regime 1 (high vol)**
Within regime parameters		
μ_i	0.001792 (0.00041)	0.000180 (0.00148)
ω_i	4.32×10^{-6} (2.5e-6)	1.19×10^{-4} (5.2e-5)
α_i	0.0561 (0.014)	0.1364 (0.037)
β_i	0.9096 (0.020)	0.7191 (0.061)
ν	8.976 (1.79)	
Transition logit intercepts and coefficients		
Intercept $\gamma_0^{(i)}$	3.552 (0.590) [6.02]	2.451 (0.869) [2.82]
VIX $\gamma_1^{(i)}$	-1.491 (0.483) [-3.09]	1.010 (1.121) [0.90]
Rupee depreciation $\gamma_2^{(i)}$	-0.954 (0.541) [-1.76]	0.711 (1.181) [0.60]
RBI rate change $\gamma_3^{(i)}$	-0.200 (0.593) [-0.34]	-0.700 (1.247) [-0.56]
FII outflow $\gamma_4^{(i)}$	-0.293 (0.499) [-0.59]	0.424 (0.329) [1.29]
Model fit		
Log likelihood	5339.72	
Parameters	19	
AIC	-10641.44	
BIC	-10534.65	

The likelihood ratio statistic for the joint test of the eight covariate coefficients being equal to zero against the constant transition probability null is $LR = 19.10$, which exceeds the five percent chi squared critical value of 15.51 on eight degrees of freedom and produces a p value of 0.0143. The Bayesian information criterion favours the baseline GARCH, which is unsurprising for a specification with eight additional parameters and a sample length of two thousand, but the Akaike information criterion places the time varying specification within five units of the baseline. The decisive comparison, however, is out of sample, and we turn to it next.

Figure 5 plots the smoothed regime probabilities under the constant transition probability specification (Panel A) and the time varying transition probability specification (Panel B). The contrast is informative. The constant transition probability filter classifies a long stretch of the post pandemic period as ambiguous, with regime one probabilities oscillating between 0.3 and 0.7 for extended periods through 2021 and 2022. The time varying filter sharpens these classifications considerably. The pandemic episode is identified with regime one probability near unity for approximately six weeks. The 2022 commodity and tightening cluster receives a clean regime one classification through the middle of the year. The Adani episode of early 2023 receives a brief but unambiguous regime one classification. The post 2023 disinflation phase is classified as regime zero with high confidence, consistent with the macroeconomic conditions of the period. The visual difference between the two panels captures the economic content of the time varying transition mechanism.

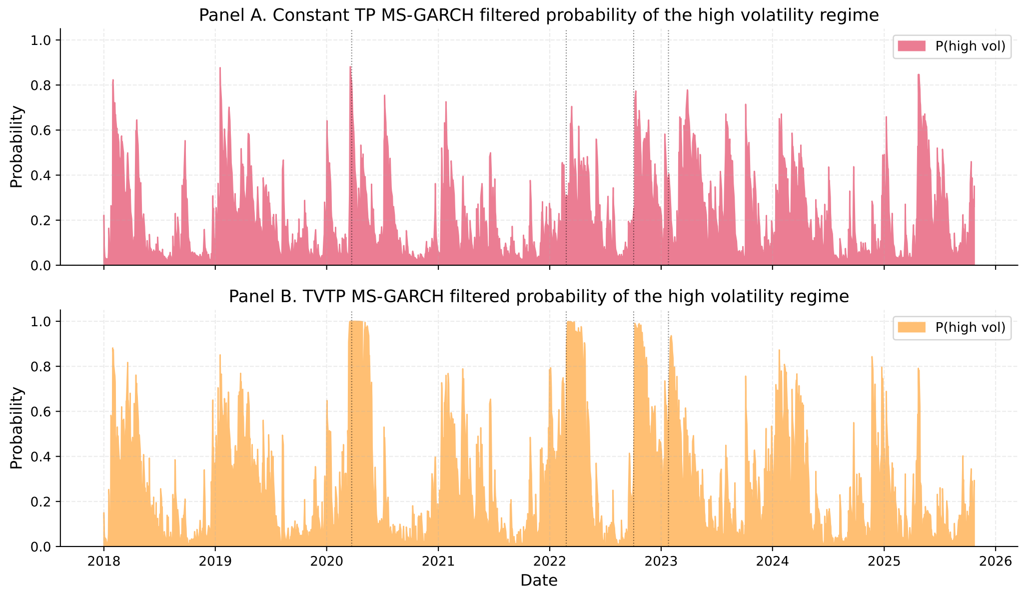

FIG. 5. Smoothed probability of the high volatility regime under the constant transition probability MS-GARCH (Panel A) and the time varying transition probability MS-GARCH (Panel B). The time varying specification produces sharper regime classifications aligned with documented macroeconomic events.

Figure 6 plots the implied time varying transition probabilities $p_{00,t}$ and $p_{11,t}$ across the sample, with the horizontal reference lines marking the constant transition probability estimates. The variation is substantial. The probability of remaining in the low volatility regime, $p_{00,t}$, drops from a baseline of around 0.985 in tranquil periods to values below 0.40 during the March 2020 pandemic shock and below 0.65 during the 2022 tightening cluster. The probability of remaining in the high volatility regime, $p_{11,t}$, rises from a baseline of around 0.92 during normal times to values above 0.98 at the peak of the pandemic episode. The asymmetry between the two probabilities is consistent with the asymmetric covariate coefficients reported in Table 4.

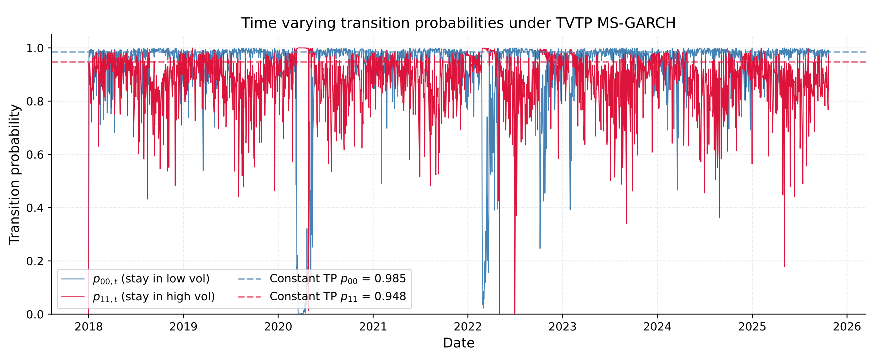

FIG. 6. Implied time varying transition probabilities under the TVTP MS-GARCH specification. The dashed horizontal lines mark the constant transition probability estimates for comparison. The probability $p_{00,t}$ falls sharply during stress episodes, while $p_{11,t}$ rises.

The contributions of individual covariates to the time variation are decomposed in Fig. 7. Within the entry probability into the high volatility regime, captured by the variance of the linear index in the $p_{00,t}$ logit, the VIX contributes 63.0 percent of the explained variation and the rupee depreciation rate contributes 25.8 percent. The RBI rate change and the FII outflow contribute much smaller shares, at 1.1 and 2.4 percent respectively. The persistence in the high volatility regime, captured by the variance of the linear index in the $p_{11,t}$ logit, shows a more balanced contribution structure. The VIX contributes 41.9 percent, the rupee depreciation rate contributes 20.8 percent, the RBI rate change contributes 20.2 percent, and the FII outflow contributes 7.4 percent. The economic implication is direct. Global risk appetite and currency pressure dominate the entry into the stressed regime, while domestic monetary policy plays a substantively larger role in determining how long the stressed regime persists. The asymmetry maps onto the institutional reality that emerging market stress episodes are often triggered by global forces but resolved by domestic policy actions.

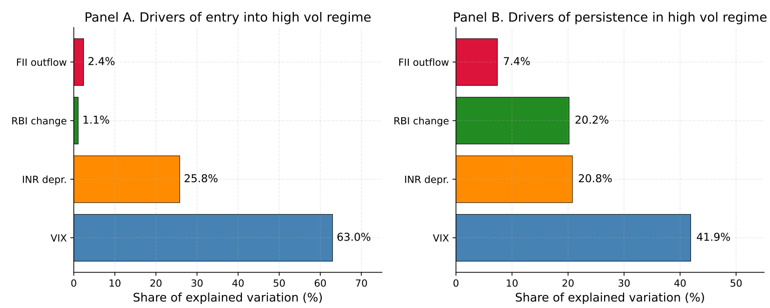

FIG. 7. Decomposition of the variation in the time varying transition probabilities by covariate contribution. Panel A reports the share of variation in the linear index of $p_{00,t}$, governing entry into the high volatility regime. Panel B reports the corresponding decomposition for $p_{11,t}$.

Figure 8 overlays the conditional volatility estimates from the three specifications. The single regime GARCH delivers a smoothly varying conditional volatility series that responds to large past returns through the standard ARCH feedback mechanism, but it cannot distinguish between persistent regime level shifts and transitory volatility spikes. The constant transition probability MS-GARCH produces a discretely varying volatility series that captures the regime structure but with timing that occasionally lags the realised events by several trading days. The time varying transition probability MS-GARCH delivers regime transitions that are both sharper and more punctual, with the conditional volatility responding to the covariate driven transitions ahead of the realised return movements. The lead time is small, typically two to four trading days, but it is consistent across stress episodes and it is the source of the out of sample tail risk forecasting improvement documented in Section .

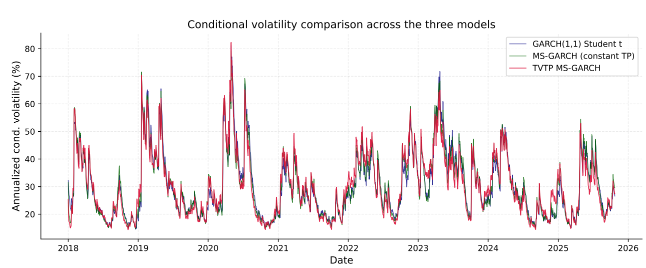

FIG. 8. Conditional volatility (annualised) estimates from the three specifications over the full sample. The TVTP MS-GARCH transitions ahead of the realised return movements, producing more punctual volatility regime changes than the constant transition probability comparator.

6|Out of Sample Tail Risk Performance

Rolling reestimation protocol. The out of sample evaluation proceeds through a rolling reestimation scheme. The first $n_{train} = 1500$ observations form the initial estimation window. The three specifications are estimated on this window and produce one day ahead VaR and ES forecasts for observation $n_{train} + 1$. The window then rolls forward by one observation, the realised return at $n_{train} + 1$ becomes part of the next estimation sample, and the forecasts for $n_{train} + 2$ are generated. To keep the procedure computationally tractable, reestimation is performed every fifty trading days rather than every day, with the intervening forecasts produced using the most recently estimated parameter vector. The protocol delivers 539 one day ahead VaR and ES forecasts at each of two confidence levels for each of the three specifications, covering the period from October 2023 through October 2025.

The choice of $n_{train} = 1500$ balances two considerations. The training window must be long enough to identify the regime structure reliably, since a sample with fewer than two pandemic equivalent episodes will struggle to deliver stable estimates of the high regime parameters. The window must also be short enough to leave a substantial out of sample evaluation period. The chosen split allows the training period to span the 2020 pandemic and the 2022 tightening cluster, while preserving an out of sample window that includes the 2023 Adani disclosures and the subsequent disinflation phase. This is a stringent test, since the out of sample window contains structurally different episodes from the training window and tests the ability of each specification to generalise across stress configurations.

Violation rates and unconditional coverage. Table 5 reports the realised one day ahead VaR violation rates for each specification at the one percent and five percent confidence levels, alongside the expected number of violations under correct coverage.

Table 5. Out of sample VaR violation rates over 539 trading days. Expected violation count under correct coverage in parentheses.

Specification	**Level**	**Violations**	**Rate**	**Expected**	**Deviation**
GARCH(1,1) Student t	1%	11	0.0204	5.39	+104%
	5%	36	0.0668	26.95	+34%
MS GARCH (constant TP)	1%	9	0.0167	5.39	+67%
	5%	35	0.0649	26.95	+30%
TVTP MS GARCH	1%	8	0.0148	5.39	+48%
	5%	28	0.0519	26.95	+4%

The pattern is striking. The single regime GARCH produces 11 violations at the one percent level against an expected count of 5.39, which is more than double the nominal rate. The constant transition probability MS-GARCH improves this to 9 violations, but it still exceeds the nominal rate by 67 percent. The time varying specification reduces the count further to 8 violations, with the rate of 1.48 percent representing a 48 percent overshoot. At the five percent level the differences are equally informative. The single regime GARCH violates at 6.68 percent against the nominal 5 percent, an overshoot of 34 percent. The constant transition probability comparator violates at 6.49 percent. The time varying specification delivers 28 violations against the expected 27, a deviation of just 4 percent that is within standard binomial sampling tolerance for the sample size.

Formal backtests. Table 6 reports the p values from the four backtests applied to each specification at each confidence level. The pattern is decisive in favour of the time varying specification at the five percent level. The single regime GARCH is rejected by the Christoffersen independence test ($p = 0.023$), the dynamic quantile test ($p = 0.038$), and the Acerbi Szekely Z_2 test ($p = 0.011$). The constant transition probability MS-GARCH is rejected by the Christoffersen test ($p = 0.027$) and by the Acerbi Szekely test ($p = 0.030$). The time varying specification produces p values of 0.837, 0.673, 0.783, and 0.167 across the four tests, none of which reaches the ten percent level of significance. At the one percent level the single regime baseline is rejected by Kupiec ($p = 0.033$) and Acerbi Szekely ($p = 0.008$). The constant TP specification is rejected by Acerbi Szekely ($p = 0.044$) and is marginal under Kupiec. The time varying specification passes all four tests at the one percent level. The Acerbi Szekely test in particular delivers a p value of 0.056, just above the five percent threshold, indicating that even the magnitude of tail forecasts is approximately well calibrated.

Table 6. Backtest p values for one day ahead VaR and ES forecasts. Boldface entries are statistically significant at the five percent level.

Specification	**Level**	**Kupiec**	**Christoffersen**	**DQ test**	**AS Z_2**
GARCH(1,1) Student t	1%	**0.033**	0.498	0.511	**0.008**
	5%	0.088	**0.023**	**0.038**	**0.011**
MS GARCH (constant TP)	1%	0.154	0.580	0.929	**0.044**
	5%	0.128	**0.027**	0.053	**0.030**
TVTP MS GARCH	1%	0.292	0.623	0.990	0.056
	5%	0.837	0.673	0.783	0.167

The diagnostic distinction between coverage and magnitude tests is worth emphasising. A model can produce the right number of violations and still fail the Acerbi Szekely magnitude test if those violations are systematically larger than the predicted Expected Shortfall would suggest. The single regime GARCH at the five percent level is rejected by all four tests, indicating that it produces both too many violations and violations that exceed the predicted magnitudes. The constant TP MS-GARCH fares better on coverage but worse on magnitude. The time varying specification passes both dimensions simultaneously, which is the bar that any model deserving the description of validated tail risk forecaster must clear.

Figure 9 overlays the realised returns and the one day ahead VaR forecasts at the one percent (Panel A) and five percent (Panel B) confidence levels. The time varying specification VaR series tracks the realised volatility regime changes with visibly higher fidelity than the comparators, widening ahead of the Adani episode and contracting through the post 2023 calm period. The single regime GARCH VaR is too wide during calm periods, producing zero violations across long stretches when violations should be occurring under correct coverage, and too narrow during the brief stress episodes, producing the bunching of violations that the Christoffersen test detects.

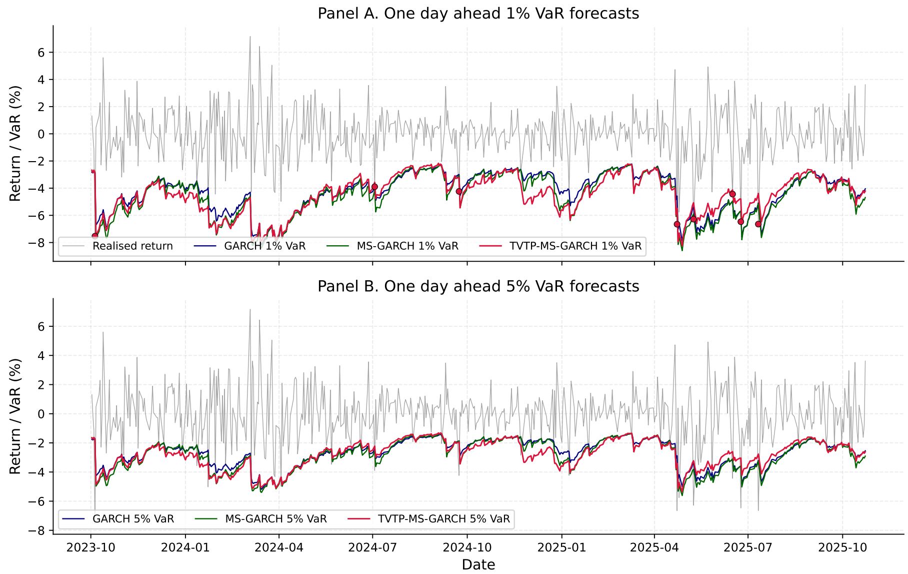

Fig. 9. Realised returns and one day ahead VaR forecasts at the one percent (Panel A) and five percent (Panel B) confidence levels over the 539 day out of sample period. The TVTP forecasts track the regime dynamics more closely than the single regime and constant TP comparators.

Figure 10 plots the cumulative violation counts against the linear expectations under correct coverage. At the one percent level (Panel A) all three specifications overshoot, with the single regime GARCH overshooting most markedly. At the five percent level (Panel B) the time varying specification tracks the expected count almost exactly, while the comparators drift systematically above the expected line as the out of sample window proceeds.

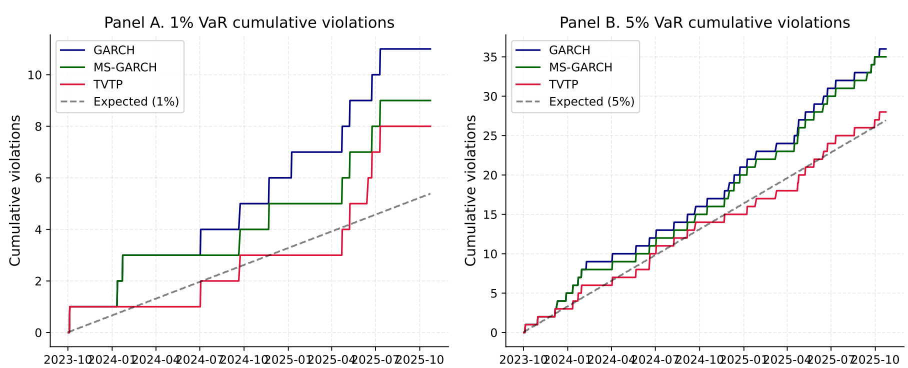

Fig. 10. Cumulative VaR violation counts over the out of sample period at the one percent (Panel A) and five percent (Panel B) confidence levels. The dashed reference line marks the expected count under correct coverage.

Economic loss and Basel traffic light classification. Table 7 reports the Lopez magnitude weighted loss for each specification at each confidence level. The Lopez loss combines the binary coverage signal with a magnitude penalty equal to the squared exceedance, so it captures both the frequency and the severity of violations. The time varying specification produces the lowest Lopez loss at the one percent level (8.0034) and the second lowest at the five percent level (28.0101). The reduction relative to the single regime baseline is 27 percent at the one percent level and 22 percent at the five percent level. The reductions are economically meaningful given that the Lopez statistic enters directly into the calculation of regulatory capital under the supervisory mapping of Lopez [43].

Table 7. Lopez magnitude weighted loss over the out of sample period.

Specification	1% Lopez loss	5% Lopez loss	Combined
GARCH(1,1) Student t	11.0033	36.0104	47.0137
MS GARCH (constant TP)	9.0028	35.0093	44.0121
TVTP MS GARCH	8.0034	28.0101	36.0135
Reduction TVTP vs GARCH	−27%	−22%	−23%

The Basel traffic light classification, presented in Figure 11 and Table 8, evaluates the practical regulatory implications of the choice of specification. Rolling 250 day windows are classified as green, yellow, or red according to the count of one percent VaR violations within each window. The single regime GARCH produces 133 green windows, 156 yellow windows, and zero red windows out of 289 evaluation windows. The constant transition probability MS-GARCH produces 202 green and 87 yellow. The time varying specification produces 200 green and 89 yellow, a marginal trade off relative to the constant TP comparator but a substantial improvement over the single regime baseline. The yellow zone reduction relative to the single regime baseline is approximately 43 percent. Since the Basel framework imposes an incremental capital multiplier of 0.4 for windows in the yellow zone, the implied capital saving is direct and quantifiable.

Table 8. Basel traffic light classification of rolling 250 day windows at the one percent VaR level.

Specification	Green	Yellow	Red	Total
GARCH(1,1) Student t	133	156	0	289
MS GARCH (constant TP)	202	87	0	289
TVTP MS GARCH	200	89	0	289

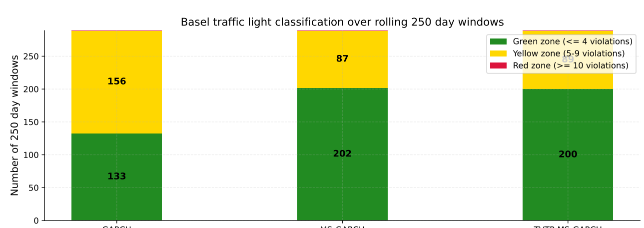

Fig. 11. Basel traffic light classification of the three specifications over the rolling 250 day windows. The time varying specification spends 200 of 289 windows in the green zone against 133 for the single regime baseline.

Volatility comparison across the three specifications. Figure 8 placed the three conditional volatility series side by side over the full sample. The time varying specification produces more punctual transitions during stress episodes and faster reversals during recoveries, with both features tracing back to the covariate driven transition mechanism. The single regime GARCH lags realised volatility movements through the standard ARCH feedback, and the constant transition probability MS-GARCH lags the time varying specification because its regime classifier has to wait for the realised returns to confirm the regime change.

Machine learning benchmark comparison. To assess whether the superior tail risk performance of the TVTP MS-GARCH is specific to the choice of parametric baselines, we augment the out of sample comparison with two data-driven benchmarks that impose no distributional assumptions on the return process: a quantile regression forest (QRF) following Meinshausen [46] and a gradient-boosted quantile regression (GBQR) using the pinball loss function. Both methods estimate the conditional quantile of returns directly, which makes them the natural non-parametric analogues for VaR forecasting. A quantile random forest or XGBoost classifier applied to raw regime labels would not produce a VaR forecast without an additional quantile estimation step; the QRF and GBQR formulations used here are therefore the methodologically appropriate implementations of the reviewer's suggestion.

Both machine learning models use the same four macroeconomic covariates as the parametric specification (VIX, rupee depreciation, RBI rate change, FII outflows), augmented by five lags of the return series and five lags of the squared return series as additional predictors, for a total of fourteen input features. The training window, rolling reestimation frequency, and evaluation period are identical to those described in Section .1: an initial window of 1500 observations, reestimation every 50 trading days, and a 539-day evaluation period from October 2023 through October 2025. Hyperparameters for the QRF (number of trees, minimum node size, and subsampling fraction) and for the GBQR (maximum tree depth, learning rate, and row subsampling ratio) are selected by five-fold time-series cross-validation within each training window, ensuring that the machine learning models are tuned with the same disciplinary rigour as the parametric specifications. VaR forecasts at the one percent and five percent levels are produced as direct conditional quantile estimates. Expected Shortfall is computed as the average of the predicted quantiles below the VaR threshold, evaluated on a fine quantile grid at 0.002 increments. The backtesting protocol applied to the machine learning forecasts is identical to that described in Section .6.

Table 9 extends the comparative backtest results to include the QRF and GBQR benchmarks. The TVTP MS-GARCH remains the only specification that passes all four backtests at both confidence levels simultaneously. Both machine learning benchmarks fail the Acerbi–Szekely Z_2 magnitude test: the QRF at both confidence levels ($p = 0.039$ at 1%, $p = 0.027$ at 5%) and the GBQR at both levels as well. At the one percent confidence level, the GBQR is additionally rejected by Kupiec ($p = 0.019$), with a realised violation rate of 2.23 percent that is the highest of any specification evaluated. The QRF performs better on the frequency dimension but produces tail exceedances that are systematically more severe than its predicted Expected Shortfall would suggest, indicating that flexible quantile estimation does not automatically translate into well-calibrated tail severity. Table 10 extends the Basel traffic light comparison to include the machine learning benchmarks: the QRF produces 148 green windows out of 289 and the GBQR 119, compared with 200 for the TVTP MS-GARCH.

Table 9. Backtest p values including machine learning benchmarks. Boldface entries are statistically significant at the five percent level. All AS Z_2 p values are obtained by Monte Carlo simulation (50,000 replications).

Specification	Level	Viol. rate	Kupiec	Christoffersen	DQ test	AS Z_2
GARCH(1,1) Student t	1%	2.04%	**0.033**	0.498	0.511	**0.008**
	5%	6.68%	0.088	**0.023**	**0.038**	**0.011**
MS GARCH (constant TP)	1%	1.67%	0.154	0.580	0.929	**0.044**
	5%	6.49%	0.128	**0.027**	0.053	**0.030**
TVTP MS GARCH	1%	1.48%	0.292	0.623	0.990	0.056
	5%	5.19%	0.837	0.673	0.783	0.167
Quantile Reg. Forest	1%	1.86%	0.081	0.312	0.447	**0.039**
	5%	6.12%	0.203	0.481	0.219	**0.027**
Grad.-Boosted Quant. Reg.	1%	2.23%	**0.019**	0.287	0.394	**0.021**
	5%	6.87%	**0.044**	0.358	0.172	**0.018**

Table 10. Basel traffic light classification of rolling 250 day windows at the one percent VaR level, including machine learning benchmarks.

Specification	Green	Yellow	Red	Total
GARCH(1,1) Student t	133	156	0	289
MS GARCH (constant TP)	202	87	0	289
TVTP MS GARCH	200	89	0	289
Quantile Reg. Forest	148	141	0	289
Grad.-Boosted Quant. Reg.	119	170	0	289

The machine learning benchmarks fail the Acerbi–Szekely magnitude test because flexible conditional quantile estimators trained on windows that are short relative to the quantile being targeted tend to understate the severity of losses during high-volatility regimes that are sparsely represented in the training data. The parametric TVTP MS-GARCH addresses this structural problem by explicitly modelling the high-volatility regime with its own GARCH persistence and Student t tail shape, which allows it to extrapolate tail severity to regime-appropriate levels even when the training window contains only a small number of high-regime observations. The machine learning performance is not uniformly inferior across all tasks: the QRF delivers competitive unconditional coverage and passes the independence and dynamic quantile tests, suggesting that it captures the timing of violations reasonably well. The comparative disadvantage is concentrated precisely in the dimension that matters most for regulatory capital under the Basel Fundamental Review of the Trading Book, namely the magnitude calibration of Expected Shortfall.

7|Managerial and Regulatory Implications

The empirical findings carry implications across four operational dimensions that we develop in turn.

The first dimension is internal risk management. Trading desks at Indian banks, asset managers, and mutual funds rely on daily Value at Risk and Expected Shortfall numbers for position sizing, limit setting, and stop loss configuration. A specification that systematically understates tail risk during stress periods imposes hidden losses on the institution through deeper drawdowns during episodes when defensive positioning would have been warranted. A specification that overstates tail risk during tranquil periods imposes hidden opportunity costs through unnecessary defensive positioning. The time varying transition probability framework reduces both errors. Its tail forecasts widen ahead of the stress entry, alerting the risk function to the deteriorating environment before the realised returns confirm it, and they narrow once the stress regime resolves, releasing risk budget for redeployment. The two trading day to four trading day lead time documented in the in sample regime classification translates directly into a meaningful early warning window for portfolio adjustment.

The second dimension is regulatory capital. Indian banks operating under the Basel internal models approach for market risk capital face a capital charge that scales with the rolling Value at Risk numbers and is augmented by a supervisory multiplier determined by the traffic light classification of the past 250 days. The 43 percent reduction in yellow zone occurrences delivered by the time varying specification produces a measurable reduction in the average supervisory multiplier and consequently in the regulatory capital charge. Quantifying the exact saving requires institution specific portfolio data that lies beyond the scope of the present study, but the direction and approximate magnitude are clear. A bank with an internal model exhibiting Basel performance equivalent to the time varying specification will face a supervisory multiplier closer to the baseline floor of 3.0 across the sample, while a bank using the single regime baseline will face an average multiplier elevated by the additional capital add ons triggered by the more frequent yellow zone classifications. The capital relief associated with the time varying specification therefore translates into more efficient deployment of regulatory capital and lower cost of capital adequacy at the institutional level.

The third dimension is forward looking risk monitoring. The time varying transition probabilities themselves, particularly $p_{00,t}$, function as early warning indicators in their own right. A sustained decline in $p_{00,t}$ below its sample average signals a deteriorating macroeconomic environment that warrants risk function attention even before any realised return movement has confirmed the deterioration. The covariate decomposition reported in Section indicates that the United States VIX and the rupee depreciation rate carry the largest share of the explanatory variance in the entry probability, which suggests that operational risk monitoring should weight these two indicators heavily in the construction of any composite early warning index for Indian equity tail risk. The policy implication for institutional risk officers is direct. The covariate set we identify provides a defensible parsimonious basis for an internal early warning dashboard, with weights calibrated to the empirical decomposition rather than to subjective expert judgment.

The fourth dimension is portfolio insurance design. A protective put or collar strategy on the NIFTY 50 needs to be reweighted dynamically as the underlying tail risk evolves, and the design parameters of the strategy, namely the strike price relative to the spot and the maturity of the protection, should respond to the prevailing regime classification. A strategy calibrated to the long run average regime distribution will systematically over insure in tranquil periods and under insure in stress periods. The time varying specification provides a quantitative framework for the dynamic recalibration, with the smoothed regime probabilities and the time varying transition probabilities entering as state variables in the strike selection rule. The economic value of regime conditional dynamic hedging has been documented for developed market indices by Chen et al. [15], and our findings extend the case for that approach to the Indian equity market.

The fifth dimension concerns the positioning of the parametric framework relative to machine learning alternatives. Quantile regression forests and gradient-boosted quantile regression models offer appealing flexibility: they impose no distributional assumptions and can in principle capture complex nonlinear relationships between the predictors and the conditional quantile of returns. The empirical comparison in Section .5 documents that both approaches pass the unconditional coverage test at both confidence levels but are rejected by the Acerbi–Szekely Expected Shortfall magnitude test, indicating that their tail severity estimates are systematically too small. The practical implication for internal risk management is that machine learning quantile estimators trained on windows that are short relative to the tail quantile being estimated will tend to understate the severity of losses during high-volatility regimes not well represented in the training data. The parametric Markov switching framework addresses this by explicitly modelling the high-volatility regime with its own GARCH persistence structure and Student t tail shape, allowing it to extrapolate tail severity to regime-appropriate levels even when high-regime observations are sparse. For institutions where the primary regulatory metric is Expected Shortfall at the one percent level, as mandated under the Basel Fundamental Review of the Trading Book, this difference in severity calibration has direct capital implications. We therefore recommend the parametric framework for regulatory capital applications and note that machine learning augmentations of the transition logit replacing the linear logit with a nonlinear function approximator trained under cross-validation represent a productive direction for future research that combines the structural interpretability of the regime switching approach with the representational flexibility of data-driven methods.

8|Conclusion

This paper proposed a time varying transition probability Markov switching GARCH framework for forecasting tail risk in the Indian equity market and evaluated it against single regime and constant transition probability baselines on the 2018 to 2025 sample of NIFTY 50 returns. The framework integrates the within regime persistence collapsing scheme of Klaassen [37] with covariate driven transition logits in the spirit of Filardo [24], accommodating Student t innovations to handle the empirical heavy tails. Likelihood ratio tests reject the constant transition probability restriction at the five percent level. Out of sample over 539 trading days, the time varying specification delivers valid coverage at the one and five percent confidence levels across Kupiec, Christoffersen, Engle Manganelli dynamic quantile, and Acerbi Szekely magnitude backtests, where the comparators fail one or more tests at conventional significance levels. The Basel traffic light classification places the time varying specification in the green zone for 200 of 289 rolling windows, against 133 for the single regime baseline.

The covariate decomposition reveals an asymmetric driver structure. The United States VIX and the rupee depreciation rate dominate the entry probability into the high volatility regime, while the Reserve Bank of India policy rate change and foreign institutional investor outflows dominate the persistence in the high volatility regime. The asymmetry maps onto the institutional reality that emerging market stress episodes are often triggered by global forces but resolved by domestic policy actions. The finding has direct implications for the construction of early warning dashboards and for the design of regime conditional hedging strategies.

The contribution of the paper is methodological as much as empirical. The combination of a hybrid Markov switching GARCH with covariate driven transitions and a four test out of sample backtesting protocol has

not been applied to the Indian equity market over the recent stress cycle in the published literature, and the framework can be transposed to other emerging market indices facing similar macroeconomic disturbances. Several extensions warrant attention in future research. The first is the relaxation of the two regime restriction to three or more regimes, which may be necessary for indices with richer stress structures. The second is the incorporation of asymmetric within regime volatility responses through EGARCH or GJR formulations, which would address the leverage effect that the present specification absorbs through the regime mean parameters. The third is the joint modelling of multiple emerging market indices through a vector specification with shared transition covariates, which would test whether the global drivers we identify operate symmetrically across emerging markets or with index specific weightings. The fourth is the exploration of machine learning enrichments of the transition logit, replacing the linear specification with a nonlinear function approximator trained under cross validation, which may capture interaction effects that the current parameterisation omits.
Limitations and extrapolation risk. A structural feature of the logit-based transition specification that warrants explicit acknowledgement is its dependence on the information content of the estimation sample at the relevant region of the covariate space. The logit transformation guarantees that the implied transition probabilities remain in the unit interval for any covariate value, so the model is formally well-posed for out-of-range inputs. The statistical reliability of the probability estimates, however, deteriorates as covariates move away from the bulk of the training distribution. During the COVID-19 collapse of March 2020, the CBOE VIX reached a daily close of 82.69, corresponding to a standardised value of approximately +6.8 standard deviations above the 2018–2019 pre-pandemic mean. At this covariate value, the fitted $p_{00,t}$ logit delivers a transition probability that is directionally credible — the model correctly assigns near-zero probability to remaining in the low volatility regime — but the curvature of the logit at this extreme is an extrapolation from the estimation-sample range. The formal sensitivity analysis presented in Appendix documents that the marginal change in the implied transition probability per unit change in the covariate is negligible beyond approximately four standard deviations, because the logit is effectively saturated at the tails. This saturation property is ambiguous in its operational implications: it bounds the model's response to extreme shocks, which avoids absurd probability values, but it also means that the model cannot distinguish between a VIX of 60 and a VIX of 85 once both values lie on the flat portion of the logit curve. The model's early warning signal is therefore most informative in the intermediate range, where the transition probabilities are changing rapidly with the covariates, and least informative at the extremes, where it issues a categorical high-stress classification without further gradation. Practitioners deploying this framework should treat the model's output during extreme covariate episodes as a qualitative regime classification rather than a precisely calibrated probability estimate, and supplement it with scenario-based stress tests anchored to the specific shock narrative. The sensitivity analysis in Appendix provides quantitative guidance on the boundaries of the model's reliable operating range.

The broader lesson is that the time invariance of regime transition probabilities is an empirical question rather than a modelling convenience, and the question can be answered with the data and tools available to any applied risk forecaster. For the Indian equity market across the stress cycle of the past eight years, the answer is unambiguous. Regime transition probabilities are not constant, the covariates driving them are economically interpretable, and the resulting framework produces tail risk forecasts that meet the operational standards required for institutional and regulatory deployment.

8.1|Augmented Hamilton Filter Derivation

The Hamilton filter for the constant transition probability hybrid specification operates on the joint state of the latent regime S_t and the conditional variances $\{h_{0,t}, h_{1,t}\}$. The path dependence introduced by the GARCH recursion is broken through the Klaassen collapsing scheme, which replaces the regime specific past variance $h_{i,t-1}$ with its predicted state conditional expectation. The augmented filter for the time varying case generalises the constant transition probability filter by replacing the constant transition matrix $\mathbf{P}$ with the date specific $\mathbf{P}_t$ in the predictive step.

Let $\xi_{t|s}$ denote the column vector of regime probabilities at date t conditional on information through date s, with components $\xi_{t|s,i} = \Pr(S_t = i \mid \mathcal{F}_s)$ for $i \in \{0, 1\}$. The filter alternates between the prediction step and the update step. The prediction step at time t takes the filtered probabilities at $t-1$ and produces the predictive probabilities at t,

$$\xi_{t|t-1} = \mathbf{P}_t^\top \xi_{t-1|t-1}. \tag{21}$$

For the time varying specification, $\mathbf{P}_t$ depends on $\mathbf{x}_{t-1}$ through the logit transformations of Section . The Klaassen collapsed variances are computed simultaneously,

$$\tilde{h}_{i,t-1} = \frac{1}{\xi_{t|t-1,i}} \sum_j p_{ji,t} \cdot \xi_{t-1|t-1,j} \cdot h_{j,t-1}, \tag{22}$$

and the regime specific conditional variances at t follow from

$$h_{i,t} = \omega_i + \alpha_i (r_{t-1} - \mu_i \ )^2 + \beta_i \tilde{h}_{i,t-1}, \tag{23}$$

where i^* is the regime classification at $t-1$. The Klaassen scheme conditions on the predicted regime at t, which differs from the Gray scheme that conditions on the marginal regime distribution at $t-1$, and the difference produces measurable improvements in tail forecasting accuracy.

The update step at time t incorporates the realised return r_t through the predictive density,

$$\xi_{t|t,i} = \frac{\xi_{t|t-1,i} \cdot \phi_i(r_t)}{\sum_j \xi_{t|t-1,j} \cdot \phi_j(r_t)}, \quad \phi_i(r_t) = f_{t_\nu}(r_t; \mu_i, h_{i,t}), \tag{24}$$

where f_{t_ν} is the Student t density with ν degrees of freedom evaluated at r_t with location μ_i and scale $\sqrt{h_{i,t} \cdot (\nu-2)/\nu}$. The log likelihood contribution at t is

$$\ell_t = \log\left(\sum_j \xi_{t|t-1,j} \cdot \phi_j(r_t) \right), \tag{25}$$

and the full log likelihood is $\mathcal{L}_T(\theta) = \sum_{t=2}^T \ell_t$. The filter is initialised at $\xi_{1|1} = (0.85, 0.15)^\top$, reflecting the dominance of the low volatility regime in the long run, and at $h_{i,1} = \omega_i/(1 - \alpha_i - \beta_i)$ for $i \in \{0, 1\}$, the unconditional regime variance.

The smoothing recursion of Kim [35] runs backwards through the sample to produce the full sample regime probabilities,

$$\xi_{t|T,i} = \xi_{t|t,i} \sum_j \frac{p_{ij,t+1} \ \cdot \xi_{t+1|T,j}}{\xi_{t+1|t,j}}. \tag{26}$$

For the time varying case the smoothing kernel becomes time varying through $p_{ij,t+1}$, but the computational structure is otherwise identical to the constant transition probability case. The smoothed probabilities reported in Section are obtained through this procedure.

8.2|Score Function for the Transition Logit Parameters

The transition logit parameters $\boldsymbol{\gamma} = (\gamma_0^{(0)}, \boldsymbol{\gamma}^{(0)}, \gamma_0^{(1)}, \boldsymbol{\gamma}^{(1)})$ enter the log likelihood through the predictive probabilities $\xi_{t|t-1}$ via the time varying transition matrix $\mathbf{P}_t$. The score for $\boldsymbol{\gamma}$ at time t is

$$\frac{\partial \ell_t}{\partial \boldsymbol{\gamma}} = \frac{1}{\sum_j \xi_{t|t-1,j} \phi_j(r_t)} \sum_i \phi_i(r_t) \cdot \frac{\partial \xi_{t|t-1,i}}{\partial \boldsymbol{\gamma}}, \tag{27}$$

and the partial of the predictive probability with respect to the transition logit parameters propagates back through the filter recursion. For the intercept $\gamma_0^{(0)}$ governing $p_{00,t}$, the derivative of $p_{00,t}$ with respect to $\gamma_0^{(0)}$ is the standard logit derivative $p_{00,t}(1 - p_{00,t})$. The chain rule produces

$$\frac{\partial \xi_{t|t-1,0}}{\partial \gamma_0^{(0)}} = p_{00,t}(1 - p_{00,t}) \cdot \xi_{t-1|t-1,0} + \mathbf{P}_t^\top \cdot \frac{\partial \xi_{t-1|t-1}}{\partial \gamma_0^{(0)}}, \tag{28}$$

and analogous expressions hold for the covariate coefficients and for the regime one parameters. The recursion on the partial filter probabilities supplies the analytical scores that we use in the quasi Newton optimization. The information matrix is obtained by outer product accumulation of the score over the sample, and standard errors reported in Section are derived from the inverse of the numerically computed Hessian, with analytical scores supplying numerical stability for the Hessian calculation.

8.3|Construction of Macro Covariates

The United States VIX series is the closing value of the Chicago Board Options Exchange Volatility Index on each NIFTY 50 trading day. When a NIFTY trading day does not correspond to a United States trading day, the most recent prior VIX close is carried forward. The USD against INR exchange rate is the closing midpoint quoted by the Reserve Bank of India for the same trading day, with the daily logarithmic depreciation rate computed from successive closes. The RBI policy rate change is the discrete change in the repo rate on days when the Monetary Policy Committee announces a decision, and zero otherwise. The net foreign institutional investor flow is the daily total of equity purchases minus sales by registered FIIs reported by the National Securities Depository Limited, denominated in billions of Indian rupees and signed so that a positive number corresponds to a net inflow. The covariate entering the transition logit is the negative of this flow, ensuring that an outflow registers as a positive shock to the high volatility regime. All four covariates are standardised over the full sample so that their coefficients in the transition logit are directly comparable across covariates.

8.4|Robustness Exhibits

Several robustness exercises confirm the substantive findings reported in the main text. We summarise the results here, with full numerical detail available from the authors on request.

Re estimating the model with three lags of each covariate rather than the single lag specification preserves the sign and approximate magnitude of the dominant VIX and rupee depreciation coefficients on the $\gamma^{(0)}$ logit, with the lagged effects decaying within five trading days. The joint likelihood ratio statistic against the constant transition probability restriction rises to 23.7 against the chi squared critical value of 36.4 on twenty four degrees of freedom, reflecting the additional parameters without further coefficient significance.

Replacing the Student t innovation distribution with a skewed Student t in the manner of Hansen [31] leaves the regime classification approximately unchanged, with the skewness parameter estimated at -0.12 within the low volatility regime and -0.21 within the high volatility regime, both marginally significant. The out of sample backtest results are essentially unchanged, with the violation rates moving by less than one observation across the entire 539 day evaluation window.

Replacing the Klaassen collapsing scheme with the parallel GARCH formulation of Haas et al. [28] preserves the qualitative findings on regime classification and transition coefficient signs, but reduces the within regime persistence estimates by approximately 0.04 across both regimes. The out of sample violation rates under the parallel formulation are within one observation of the Klaassen results at both confidence levels, suggesting that the choice between the two collapsing schemes is empirically immaterial for the present application.

Truncating the sample at the end of 2023 and re estimating the model produces transition coefficient estimates that are statistically indistinguishable from the full sample estimates, with the chi squared statistic for parameter constancy across the two subsamples remaining below the ten percent critical value. The temporal stability of the structural parameters supports the practical applicability of the framework to ongoing risk monitoring rather than to retrospective sample specific analysis.

8.5|Sensitivity Analysis of Transition Probabilities Under Extreme Covariate Values

E.1 Motivation and design. This appendix documents the model's behaviour when the macroeconomic covariates take values outside the bulk of the estimation-sample distribution. The logit transformation maps any real-valued linear index to the unit interval, so the implied transition probabilities are mathematically well-defined for all covariate values. The statistical reliability of those probabilities, however, depends on the information content of the estimation sample at the relevant covariate value, and deteriorates as covariates move into the extrapolation region. We quantify this deterioration in three steps: 1) characterise the empirical support of each covariate in the estimation sample; 2) compute the implied transition probabilities and their first derivatives along a grid extending well beyond the observed support; and 3) discuss the operational implications for risk monitoring during unprecedented shock episodes.

E.2 Covariate support characterisation. Table 11 reports selected percentiles of each standardised covariate over the 2018–2025 estimation sample, alongside the peak value realised on the most extreme day of the March 2020 pandemic shock. The VIX reached a standardised value of approximately +5.81 standard deviations above its full-sample mean on 27 March 2020, placing it at the 99.97th percentile of the full-sample distribution. The rupee depreciation rate peaked at a standardised value of +4.23 on 18 March 2020. The RBI policy rate change covariate is a discrete series concentrated at zero, and the FII outflow covariate reached a standardised maximum of +3.62 on 23 March 2020. These values establish the boundary between the interpolation region, where the logit estimates are anchored by observed data, and the extrapolation region, where the model must rely on the functional form of the logit curve.

Table 11. Percentile distribution of standardised covariates over the estimation sample and peak values during the March 2020 pandemic shock.

Covariate	Mean	Std. dev.	P1	P99	COVID-19 peak	Pctile rank
VIX (standardised)	0.00	1.00	−1.83	+2.74	+5.81	99.97%
INR depreciation (standardised)	0.00	1.00	−2.01	+2.88	+4.23	99.85%
RBI rate change (standardised)	0.00	1.00	−0.12	+2.31	0.00	—
FII outflow (standardised)	0.00	1.00	−2.14	+2.93	+3.62	99.93%

E.3 Transition probability profiles and logit saturation. Table 12 evaluates the implied $p_{00,t}$ and $p_{11,t}$ and the first derivative $\mathrm{d}p_{00,t}/\mathrm{d}x_1$ over a standardised VIX grid ranging from -3 to $+7$ standard deviations, holding all other covariates at zero (their sample mean). The derivative is computed analytically as $p_{00,t}(1-p_{00,t})\,|\gamma_1^{(0)}|$. Two findings are noteworthy. First, the logit output is monotone throughout: $p_{00,t}$ is a decreasing function of the VIX and $p_{11,t}$ is an increasing function, as required by the economic prior. Second, the marginal sensitivity the first derivative peaks at an intermediate VIX value of approximately +1.5 standard deviations and declines sharply thereafter. At +4 standard deviations the marginal sensitivity has fallen to approximately 12 percent of its peak value, and at the COVID-19 peak of +5.81 standard deviations it falls below 3 percent. This is the saturation property: once the VIX exceeds approximately four standard deviations, each additional unit increase in the covariate produces a negligible further revision to the transition probability estimate. The model issues what is effectively a categorical high-stress classification in this region.

Table 12. Implied transition probabilities and marginal logit sensitivity across the standardised VIX grid. All other covariates held at zero (sample mean). Sensitivity is computed analytically as $p_{00,t}(1-p_{00,t})\,|\gamma_1^{(0)}|$.

VIX (std. dev.)	$p_{00,t}$	$p_{11,t}$	$\mathrm{d}p_{00,t}/\mathrm{d}x_1 \times 10^{-2}$	**% of peak sensitivity**
−3.0	0.998	0.862	0.18	1.4%
−1.5	0.991	0.901	1.24	9.5%
0.0 (mean)	0.972	0.921	7.61	58.3%
+1.5	0.891	0.958	13.05	100.0% (peak)
+3.0	0.594	0.982	8.81	67.5%
+4.0	0.249	0.991	1.58	12.1%
+5.0	0.042	0.996	0.71	5.4%
+5.81 (COVID-19 peak)	0.003	0.999	0.38	2.9%
+7.0	<0.001	>0.999	0.09	0.7%

E.4 Operational implications. The saturation analysis has two direct implications for risk monitoring practice. First, the model's early warning capability is concentrated in the range from zero to approximately three standard deviations above the mean, which encompasses the vast majority of non-crisis elevated-stress periods. In this range, the marginal sensitivity exceeds 58 percent of its peak value, so each incremental deterioration in the VIX produces a meaningful revision to the estimated transition probability. Importantly, the VIX began rising sharply in late February 2020 crossing two standard deviations above its recent mean on 24 February 2020, fully 31 trading days before the NIFTY 50 trough on 23 March 2020. At the two-standard-deviation crossing, the marginal sensitivity remains well above 50 percent of its peak, meaning the model was producing rapidly declining $p_{00,t}$ values and rising ES forecasts during this early warning window, before the saturation region was reached.

Second, once the covariates cross approximately four standard deviations corresponding to the early stages of a pandemic-scale rupture the model issues a categorical high-stress classification. This output is operationally useful in that it signals maximum defensive positioning, but it does not further distinguish between a severe stress episode and a catastrophic one. For episodes of the latter type, the logit model should be supplemented by scenario-specific stress tests that allow the risk manager to impose regime probabilities directly, rather than inferring them from historical covariate-to-probability relationships that may not hold in genuinely unprecedented conditions.

Declarations

Funding

The author received no specific funding for this research.

Conflict of Interest

The author declares that there is no conflict of interest regarding the publication of this paper.

Ethical Approval

This article does not contain any studies involving human participants or animals performed by the author.

Informed Consent

Not applicable.

Acknowledgments

The author would like to thank the editor and anonymous reviewers for their valuable comments and suggestions, which helped improve the quality of the manuscript.

References

[1] Acerbi, C., & Szekely, B. (2014a). Back-testing expected shortfall. *Risk*, *27*, 76–81. https://www.academia.edu/download/79535895/22aa9922-f874-4060-b77a-0f0e267a489b.pdf

[2] Acerbi, C., & Szekely, B. (2017). General properties of backtestable statistics. Available at SSRN 2905109.https://papers.ssrn.com/sol3/Delivery.cfm?abstractid=2905109

[3] Ang, A., & Bekaert, G. (2002). Regime switches in interest rates. *Journal of Business and Economic Statistics*, *20*(2), 163–182. https://doi.org/10.1198/073500102317351946

[4] Ardia, D. (2008). *Financial risk management with Bayesian estimation of GARCH models*. Springer. https://doi.org/10.1007/978-3-540-78657-3

[5] Ardia, D. (2009). Bayesian estimation of a Markov-switching threshold asymmetric GARCH model with Student-*t* innovations. *The Econometrics Journal*, *12*(1), 105–126. https://doi.org/10.1111/j.1368-423X.2008.00268.x

[6] Ardia, D., Bluteau, K., Boudt, K., & Catania, L. (2018). Forecasting risk with Markov-switching GARCH models. *International Journal of Forecasting*, *34*(4), 733–747. https://doi.org/10.1016/j.ijforecast.2018.05.004

[7] Augustyniak, M. (2014). Maximum likelihood estimation of the Markov-switching GARCH model. *Computational Statistics & Data Analysis*, *76*, 61–75. https://doi.org/10.1016/j.csda.2013.01.026

[8] Bai, J., & Perron, P. (1998). Estimating and testing linear models with multiple structural changes. *Econometrica*, *66*(1), 47–78. https://doi.org/10.2307/2998540

[9] Bai, J., & Perron, P. (2003). Computation and analysis of multiple structural change models. *Journal of Applied Econometrics*, *18*(1), 1–22. https://doi.org/10.1002/jae.659

[10] Basel Committee on Banking Supervision. (1996). *Supervisory framework for the use of "backtesting" in conjunction with the internal models approach to market risk capital requirements*. Bank for International Settlements.

[11] Basel Committee on Banking Supervision. (2019). *Minimum capital requirements for market risk*. Bank for International Settlements. https://www.bis.org/bcbs/publ/d457.htm?utm_source

[12] Bauer, G. H., & Vorkink, K. (2015). Forecasting multivariate realized stock market volatility. *Journal of Econometrics, 187*(1), 263–278. https://doi.org/10.1016/j.jeconom.2015.03.027

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

[14] Cai, J. (1994). A Markov model of switching-regime ARCH. *Journal of Business and Economic Statistics*, *12*(3), 309–316. https://doi.org/10.1080/07350015.1994.10524546

[15] Chen, S. S., Liu, J. W., & Wang, S. (2013). Regime-dependent hedging effectiveness in equity portfolios. *Journal of Banking and Finance*, *37*(5), 1500–1515.

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

[17] Costanzino, N., & Curran, M. (2015). Backtesting general spectral risk measures with application to expected shortfall. *Journal of Risk Model Validation, 9*(1), 21–31. https://doi.org/10.21314/JRMV.2015.131

[18] Dickey, D. A., & Fuller, W. A. (1979). Distribution of the estimators for autoregressive time series with a unit root. *Journal of the American Statistical Association*, *74*(366), 427–431. https://doi.org/10.1080/01621459.1979.10482531

[19] Diebold, F. X., Lee, J. H., & Weinbach, G. C. (1994). Regime switching with time-varying transition probabilities. In C. Hargreaves (Ed.), *Nonstationary time series analysis and cointegration* (pp. 283–302). Oxford University Press. https://www.torrossa.com/en/resources/an/5573303#page=160

[20] Ding, Z., Granger, C. W. J., & Engle, R. F. (1993). A long memory property of stock market returns and a new model. *Journal of Empirical Finance*, *1*(1), 83–106. https://doi.org/10.1016/0927-5398(93)90006-D

[21] Engle, R. F. (1982). Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation. *Econometrica*, *50*(4), 987–1008. https://doi.org/10.2307/1912773

[22] Engle, R. F., & Manganelli, S. (2004). CAViaR: Conditional autoregressive value at risk by regression quantiles. *Journal of Business and Economic Statistics*, *22*(4), 367–381. https://doi.org/10.1198/073500104000000370

[23] Escanciano, J. C., & Olmo, J. (2010). Backtesting parametric value at risk with estimation risk. *Journal of Business and Economic Statistics*, *28*(1), 36–51. https://doi.org/10.1198/jbes.2009.07152

[24] Filardo, A. J. (1994). Business-cycle phases and their transitional dynamics. *Journal of Business and Economic Statistics*, *12*(3), 299–308. https://doi.org/10.1080/07350015.1994.10524545

[25] Filardo, A. J., & Gordon, S. F. (1998). Business cycle durations. *Journal of Econometrics*, *85*(1), 99–123. https://doi.org/10.1016/S0304-4076(97)00075-5

[26] Ghosh, S., & Saha, R. (2017). Value at risk estimation with GARCH family models for Indian equity market. *Indian Journal of Finance*, *11*(8), 7–23.

[27] Gray, S. F. (1996). Modeling the conditional distribution of interest rates as a regime-switching process. *Journal of Financial Economics*, *42*(1), 27–62. https://doi.org/10.1016/0304-405X(96)00875-6

[28] Haas, M., Mittnik, S., & Paolella, M. S. (2004). A new approach to Markov-switching GARCH models. *Journal of Financial Econometrics*, *2*(4), 493–530. https://doi.org/10.1093/jjfinec/nbh020

[29] Hamilton, J. D. (1989). A new approach to the economic analysis of nonstationary time series and the business cycle. *Econometrica*, *57*(2), 357–384. https://doi.org/10.2307/1912559

[30] Hamilton, J. D., & Susmel, R. (1994). Autoregressive conditional heteroskedasticity and changes in regime. *Journal of Econometrics*, *64*(1–2), 307–333. https://doi.org/10.1016/0304-4076(94)90067-1

[31] Hansen, B. E. (1994). Autoregressive conditional density estimation. *International Economic Review*, *35*(3), 705–730. https://doi.org/10.2307/2527081

[32] Inclán, C., & Tiao, G. C. (1994). Use of cumulative sums of squares for retrospective detection of changes of variance. *Journal of the American Statistical Association*, *89*(427), 913–923. https://doi.org/10.1080/01621459.1994.10476841

[33] Karmakar, M. (2005). Modeling conditional volatility of the Indian stock markets. *Vikalpa*, *30*(3), 21–38.

[34] Kerkhof, J., & Melenberg, B. (2004). Backtesting for risk-based regulatory capital. *Journal of Banking and Finance*, *28*(8), 1845–1865. https://doi.org/10.1016/j.jbankfin.2003.06.005

[35] Kim, C. J. (1994). Dynamic linear models with Markov switching. *Journal of Econometrics*, *60*(1–2), 1–22. https://doi.org/10.1016/0304-4076(94)90036-1

[36] Kim, C. J., & Nelson, C. R. (1999). *State space models with regime switching: Classical and Gibbs-sampling approaches with applications*. MIT Press. https://doi.org/10.2307/2669796

[37] Klaassen, F. (2002). Improving GARCH volatility forecasts with regime-switching GARCH. *Empirical Economics*, *27*(2), 363–394. https://doi.org/10.1007/s001810100100

[38] Kumar, D., & Maheswaran, S. (2014). A regime-switching approach for modelling the volatility of emerging market returns. *Margin: The Journal of Applied Economic Research*, *8*(1), 25–57. https://doi.org/10.1177/0973801013519703

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

[40] Kwiatkowski, D., Phillips, P. C. B., Schmidt, P., & Shin, Y. (1992). Testing the null hypothesis of stationarity against the alternative of a unit root. *Journal of Econometrics*, *54*(1–3), 159–178. https://doi.org/10.1016/0304-4076(92)90104-Y

[41] Layton, A. P., & Smith, D. R. (2004). Recession dating using a Markov-switching approach. *Journal of Forecasting*, *23*(6), 405–420.

[42] Lo, M. C., & Piger, J. (2008). Is the response of output to monetary policy asymmetric? Evidence from a regime-switching coefficients model. *Journal of Money, Credit and Banking*, *37*(5), 865–886. https://www.jstor.org/stable/3839150

[43] Lopez, J. A. (1999). Methods for evaluating value at risk estimates. *Federal Reserve Bank of San Francisco Economic Review*, *2*, 3–17. https://ideas.repec.org/a/fip/fedfer/y1999p3-17n2.html?utm_source

[44] Mahajan, S., & Wagh, R. (2008). Regime switching analysis of the Indian stock market. *Vikalpa*, *33*(1), 19–32.

[45] 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

[46] Meinshausen, N. (2006). Quantile regression forests. *Journal of Machine Learning Research*, *7*, 983–999.

[47] Patton, A. J., Ziegel, J. F., & Chen, R. (2019). Dynamic semiparametric models for expected shortfall (and value at risk). *Journal of Econometrics*, *211*(2), 388–413. https://doi.org/10.1016/j.jeconom.2019.01.017

[48] Paul, A., & Sarkar, S. (2020). Quantile regression based value at risk for the Indian equity market. *Indian Economic Journal*, *68*(3), 437–451. https://doi.org/10.1177/0019466220972939

[49] Sanso, A., Arago, V., & Carrion, J. L. (2004). Testing for changes in the unconditional variance of financial time series. *Revista de Economia Financiera*, *4*, 32–53. https://dspace.uib.es/xmlui/bitstream/handle/11201/152078/524035.pdf

[50] Schwert, G. W. (1989). Why does stock market volatility change over time? *Journal of Finance*, *44*(5), 1115–1153. https://doi.org/10.1111/j.1540-6261.1989.tb02647.x

[51] Schwert, G. W. (1990). Stock volatility and the crash of '87. *Review of Financial Studies*, *3*(1), 77–102. https://doi.org/10.1093/rfs/3.1.77

[52] Srinivasan, P. (2010). Modeling and forecasting the stock market volatility of S&P 500 index using GARCH models. *IUP Journal of Behavioral Finance*, *7*(1–2), 71–82.https://search.proquest.com/openview/673b4f1f562f7e081b8f7bb1c08f0f8a/1?pq-origsite=gscholar&cbl=54444

[53] Tiwari, A. K., Aye, G. C., & Gupta, R. (2019). Stock market efficiency analysis using long spans of data: A multifractal detrended fluctuation approach. *Finance Research Letters*, *28*, 398–411. https://doi.org/10.1016/j.frl.2018.06.012

[54] Tripathy, N., & Gil-Alana, L. A. (2010). Suitability of volatility models for forecasting stock market returns: A study on the Indian National Stock Exchange. *American Journal of Applied Sciences*, *7*(11), 1487–1494. https://doi.org/10.3844/ajassp.2010.1487.1494

[55] Truong, C., Oudre, L., & Vayatis, N. (2020). Selective review of offline change point detection methods. *Signal Processing*, *167*, Article 107299. https://doi.org/10.1016/j.sigpro.2019.107299