

Structural Break Detection and Regime-Dependent Financial Time-Series Forecasting Using Change-Point Models and Markov Switching: Evidence from Asia-Pacific Equity Markets

Abhinav Tiwari

abhinavkumartiwari78@gmail.com

Nirma University <https://orcid.org/0009-0006-7763-6930>

Mehul Mehta

North Carolina State University

Archit Arya

Birla Institute of Technology and Science, Pilani

Harsh Choudhary

Kalinga University

Varad Dahake

Indian Institute of Technology Bhubaneswar

Research Article

Keywords: Asia-Pacific equity markets, Bayesian online change-point detection, Markov switching, structural breaks, Value-at-Risk, volatility forecasting

Posted Date: October 5th, 2026

DOI: <https://doi.org/10.21203/rs.3.rs-11215014/v1>

License: © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Additional Declarations: The authors declare no competing interests.

Structural Break Detection and Regime-Dependent Financial Time-Series Forecasting Using Change-Point Models and Markov Switching: Evidence from Asia-Pacific Equity Markets

ARTICLE INFO

Keywords

Asia-Pacific equity markets; Bayesian online change-point detection; Markov switching; structural breaks; Value-at-Risk; volatility forecasting

JEL classification

C22; C53; C58; G17

Abstract

Financial return dynamics are inherently unstable, reflecting both recurring latent market regimes and structural changes in the underlying data-generating process. This study examines whether causal structural-break information improves regime-dependent financial forecasting across heterogeneous Asia-Pacific equity markets. The empirical analysis covers five major indices: the S&P BSE SENSEX, Nikkei 225, Hang Seng, KOSPI, and S&P/ASX 200. The framework combines retrospective Gaussian quasi-maximum-likelihood multiple-break segmentation, Bayesian Online Change-Point Detection (BOCPD), and a Markov-switching AR(1) model with regime-dependent parameters. The number of regimes is selected separately for each market using only pre-out-of-sample information, after which performance is evaluated through a strict 1,000-observation one-step-ahead walk-forward experiment against historical-mean, AR, ARMA, EWMA, GARCH, and conventional Markov-switching benchmarks. Controlled simulation with known regimes and break locations provides additional identification validation. The results reveal substantial cross-market heterogeneity in both structural-break evidence and selected regime structure. Break-aware adaptation does not consistently improve conditional-return or average-volatility forecasts, while conventional Markov switching without memory resetting remains highly competitive and, in some markets, significantly superior under QLIKE loss. At the 5% Value-at-Risk level, break-aware forecasts provide acceptable or improved calibration in several markets, although the benefit does not generalize uniformly across all indices or tail levels. Overall, the evidence indicates that structural-break information is market- and forecast-target dependent, offering selective value for adaptive downside-risk calibration rather than universal improvements in financial forecasting.

1. Introduction

1.1 Economic Motivation

Financial markets are unlikely to be governed by a single stable statistical process over long periods. Asset-return dynamics evolve as markets respond to changes in macroeconomic conditions, monetary policy, financial regulation, liquidity, investor behaviour, technological development, and episodes of economic or geopolitical stress. Periods of relatively stable returns may therefore be interrupted by abrupt increases in volatility, changes in persistence, altered tail behaviour, or more persistent shifts in the distribution of returns. These changes create an important challenge for financial forecasting because models estimated using long historical samples implicitly assume that older observations remain informative about current market conditions.

The problem is particularly important for volatility and downside-risk forecasting. During tranquil periods, a long estimation history may improve parameter precision. Following a financial crisis, liquidity disruption, or substantial policy shift, however, observations generated under an earlier environment may become less representative of the current market. A forecasting model that retains too much historical information may adapt slowly, while a model that reacts too aggressively to recent observations may discard information that remains economically useful. Financial forecasting therefore involves a trade-off between **parameter stability and adaptability**.

Asia-Pacific equity markets provide a useful setting in which to examine this problem. The region contains markets at different stages of financial development and with different institutional structures, investor compositions, monetary-policy environments, and exposure to domestic and global shocks. This study examines five major equity indices: the **S&P BSE SENSEX, Nikkei 225, Hang Seng, KOSPI, and S&P/ASX 200**. Together, these markets provide a heterogeneous empirical environment spanning India, Japan, Hong Kong, South Korea, and Australia. Such heterogeneity is useful because a forecasting method that appears effective in a single market may reflect market-specific dynamics rather than a general property of financial returns.

1.2 Regime Switching and Structural Breaks

Two related but conceptually different forms of instability are central to this study: **regime switching** and **structural breaks**.

A regime switch refers to movement among recurring market states. For example, an equity market may move from a relatively calm environment to a turbulent period and subsequently return to conditions resembling the earlier calm state:

Calm → Turbulent → Calm.

The statistical properties of returns depend on the prevailing state, but the states themselves can recur. Markov-switching models formalize this idea by allowing parameters such as the conditional mean, autoregressive persistence, and volatility to vary with an unobserved state that evolves probabilistically over time (Hamilton, 1989).

A structural break represents a different phenomenon. It occurs when the underlying parameter-generating environment itself changes. Conceptually,

Old statistical structure → New statistical structure.

The new environment need not correspond to a previously observed state, and observations generated before the break may become less informative for subsequent parameter estimation. Structural-break methods therefore focus on changes in the underlying statistical relationships, whereas regime-switching models focus on transitions among recurring latent states.

The distinction matters for forecasting. A Markov-switching model may correctly identify that a market has entered a high-volatility state while continuing to estimate the characteristics of that state using observations accumulated under an earlier structural environment. Conversely, a structural-break procedure can indicate that a change has occurred but does not naturally provide a probabilistic representation of recurring calm, intermediate, and crisis states. Combining information about both phenomena may therefore offer a more adaptive description of financial dynamics.

1.3 Limitations of Existing Model Classes

Different model classes address different parts of this problem. ARCH and GARCH models introduced by Engle (1982) and Bollerslev (1986) capture volatility clustering and conditional heteroskedasticity and remain central benchmarks in financial

forecasting. Standard specifications, however, generally assume stable parameters unless structural changes are modelled explicitly. Ignoring parameter instability can lead persistent volatility dynamics to be confused with changes in the underlying volatility process (Lamoureux & Lastrapes, 1990; Hillebrand, 2005).

Markov-switching models address another source of instability by allowing return dynamics to depend on recurrent latent states. They are particularly suitable for distinguishing relatively calm, turbulent, and crisis-like market environments. However, conventional Markov-switching models generally continue to use historical observations within their estimation window even after evidence suggests that the broader data-generating environment has changed.

Structural-break and change-point methods directly address parameter instability by identifying points at which the statistical properties of a series change. Retrospective procedures are useful for historical analysis, but full-sample break estimates cannot be used directly in genuine out-of-sample forecasting because their identification may depend on future observations.

Bayesian Online Change-Point Detection (BOCPD; Adams & MacKay, 2007) addresses this timing problem by updating the probability of a change sequentially as new observations arrive. BOCPD therefore provides a causal signal that can be used in real time. It is not, however, a complete financial forecasting model by itself. This creates a natural question: can an online change signal be used to make a regime-switching forecasting model adapt its historical memory more effectively?

1.4 Research Gap

Existing research has extensively examined structural breaks, regime-switching models, switching-volatility specifications, and combinations of recurrent and non-recurrent changes. The contribution of this paper is therefore not to claim that structural breaks and regime switching have never previously been combined.

The narrower question addressed here is whether a **causal online change-point signal can improve forecasting when it is used specifically to alter the estimation memory of an otherwise comparable Markov-switching model**. This distinction is important. Rather than merely documenting breaks ex post, the proposed framework uses BOCPD alarms available at the forecast origin to shorten the historical estimation window following evidence of change. The window subsequently expands as observations accumulate under the new environment.

The incremental value of this mechanism is evaluated against a conventional Markov-switching specification that does not reset its historical memory. This nested comparison helps isolate the effect of break-aware adaptation itself. Moreover, the analysis is conducted across five heterogeneous Asia-Pacific equity markets rather than relying on a single empirical application.

1.5 Research Questions

The analysis addresses three research questions:

RQ1: Can the proposed framework reliably identify persistent structural changes and recurring latent volatility regimes?

RQ2: Does structural-break-aware regime modelling improve one-day-ahead return and volatility forecasts relative to conventional benchmarks?

RQ3: Does break-aware adaptation improve downside-risk calibration, particularly Value-at-Risk and Expected Shortfall?

These questions are deliberately non-directional. Greater model complexity does not necessarily imply superior forecasting performance, and the value of structural-change information may differ depending on whether the forecasting objective concerns conditional returns, average volatility, or lower-tail risk.

1.6 Contributions

The study makes four contributions. First, it develops a **methodological integration** in which causal BOCPD information is used to adapt the estimation memory of a Markov-switching AR(1) forecasting model. The resulting break-aware specification is

evaluated directly against an otherwise comparable no-reset Markov model.

Second, the study combines **controlled simulation and strictly causal out-of-sample forecasting**. Simulation provides an environment in which the true latent regimes and structural breaks are known, allowing identification performance to be evaluated independently of real-market forecast accuracy. The empirical analysis then uses a 1,000-observation walk-forward design in which regime selection, filtered probabilities, and online change information are restricted to data available at each forecast origin.

Third, the study provides a **cross-market Asia-Pacific evaluation** using five major equity indices. This makes it possible to determine whether structural-break evidence, latent regime structure, and forecasting performance vary across markets rather than assuming a common process.

Fourth, the paper contributes empirically by showing that the usefulness of structural-break-aware adaptation is **both market dependent and forecast-target dependent**.

1.7 Headline Findings

The empirical results provide a nuanced rather than uniformly favourable assessment of break-aware forecasting. Structural-break evidence and the preferred number of latent regimes differ materially across the five markets, indicating that Asia-Pacific equity returns do not share a single common form of instability. The break-aware model does not consistently improve one-day-ahead return forecasts, and conventional Markov switching without memory resetting remains highly competitive for average-volatility forecasting under QLIKE loss. In some markets, retaining a longer historical information set produces significantly better volatility forecasts.

The downside-risk evidence is more favourable but remains heterogeneous. At the 5% Value-at-Risk level, break-aware forecasts provide acceptable or improved calibration in several markets, including the primary SENSEX application, while performance deteriorates in others. The benefits are also less consistent in the more extreme 1% tail. Overall, the findings suggest that structural-break-aware adaptation should be viewed as a **selective risk-management mechanism rather than a universally superior forecasting model**.

1.8 Structure of the Paper

The remainder of the paper is organized as follows. Section 2 reviews the literature on structural instability, regime switching, online change-point detection, volatility forecasting, and downside-risk evaluation, and positions the present contribution relative to prior work. Section 3 describes the five Asia-Pacific equity-market datasets and preliminary diagnostics. Section 4 presents the forecasting methodology and strict out-of-sample design. Section 5 reports the controlled simulation validation. Section 6 presents the main empirical results, while Section 7 examines robustness and cross-market sensitivity. Section 8 discusses the economic and forecasting implications of the findings. Section 9 outlines limitations and directions for future research, and Section 10 concludes.

2. Literature Review and Research Positioning

2.1 Structural Breaks and Parameter Instability

A substantial econometric literature recognizes that financial and economic time series may experience changes in their underlying parameters rather than being generated by a single stable process. Early work by Brown, Durbin, and Evans (1975) developed cumulative-sum procedures based on recursive residuals to assess whether regression relationships remain stable through time. The resulting CUSUM framework became an important diagnostic for parameter instability. Andrews (1993) subsequently developed tests for structural change when the timing of the potential break is unknown, emphasizing that inference about parameter stability must account for uncertainty surrounding the break date. These

contributions established structural stability as an empirical question rather than an assumption to be imposed automatically.

The literature later moved from single-break testing toward multiple structural changes. Bai and Perron (1998, 2003) developed influential methods for estimating and testing multiple breaks in linear models, permitting the data to be divided into statistically distinct segments subject to minimum-length restrictions. Liu, Wu, and Zidek (1997) similarly considered segmented regression and proposed an information criterion with a comparatively strong penalty for model complexity. This **Liu–Wu–Zidek (LWZ) criterion** is relevant when highly fragmented break structures are undesirable, because stronger penalization can favour a smaller number of persistent changes over numerous short-lived segments.

Structural instability is particularly relevant to financial volatility. Inclán and Tiao (1994) introduced a cumulative-sum-of-squares approach for identifying changes in unconditional variance, helping shift attention from breaks in conditional means toward breaks in second moments. Subsequent financial applications show that volatility changes can be mistaken for extreme persistence when parameter instability is ignored. Lamoureux and Lastrapes (1990), for example, demonstrate that omitted variance shifts can materially affect estimated GARCH persistence, while Hillebrand (2005) shows more generally that neglecting parameter changes can produce misleadingly persistent GARCH estimates. Andreou and Ghysels (2002) provide further evidence that financial-market volatility can contain multiple structural changes. Asian-market applications likewise document crisis-related structural breaks and changing cross-market relationships (Baek & Jun, 2011; Lee & Chou, 2020), while Yang, Chen, and Tian (2015) examine realized-volatility forecasting under structural breaks. Related applications consider structural breaks in volatility across financial series (Rapach & Strauss, 2008; Karanasos & Katsaklas, 2009).

This distinction is important for forecasting. A volatility model fitted over a long history may estimate substantial persistence partly because it is averaging observations generated under different parameter environments. Structural-break procedures provide a way of identifying such instability, but retrospective estimates alone do not solve the real-time forecasting problem: a break identified using the complete sample may depend on observations that were unavailable at the date when a historical forecast would actually have been formed.

2.2 Markov Switching and Regime-Dependent Volatility

A related but conceptually different literature represents instability through **recurrent latent regimes**. Hamilton (1989) introduced the modern Markov-switching framework into macroeconomic time-series analysis by allowing model parameters to depend on an unobserved state evolving according to a finite-state Markov chain. The central appeal of this approach is that different statistical environments can recur. A high-volatility state, for example, does not need to represent a permanent break; the market can move from a calm state to a turbulent state and subsequently return to a state resembling the earlier calm environment. For a broader review of regime changes in financial markets, see Ang and Timmermann (2012).

This framework was extended extensively to financial volatility. Hamilton and Susmel (1994) and Cai (1994) introduced switching-regime ARCH specifications in which volatility dynamics differ across latent states. Gray (1996) developed a regime-switching approach to the conditional distribution of financial variables, while Kim (1994) formalized filtering and smoothing techniques that are central to probabilistic inference in Markov-switching models. Together, these studies established the basis for distinguishing latent market states while accounting for uncertainty about which state is active.

Later research integrated regime switching more directly with GARCH dynamics. Klaassen (2002) showed how regime information could be incorporated into volatility forecasting, while Haas, Mittnik, and Paolella (2004) developed an alternative Markov-switching GARCH formulation designed to address path-dependence problems. Marcucci (2005) evaluated regime-switching

GARCH models specifically for stock-market volatility forecasting, and Bauwens, Preminger, and Rombouts (2010) further developed the theory and inference of Markov-switching GARCH specifications. These contributions illustrate that recurrent regimes can materially alter estimates of conditional volatility and its persistence. Related evidence for Asian equity markets documents regime-switching volatility and cross-market transmission (Worthington & Higgs, 2004; Chiang et al., 2007; Gallo & Otranto, 2008; Wang & Theobald, 2008; Mateus et al., 2024).

The forecasting implications extend beyond variance. Ardia, Bluteau, Boudt, and Catania (2018), in a large-scale comparison of single-regime and Markov-switching GARCH models, find that regime-switching specifications can improve Value-at-Risk, Expected Shortfall, and left-tail distribution forecasts across several equity-return horizons. Their results reinforce the idea that regime modelling may be particularly valuable when the objective concerns the predictive distribution and downside risk rather than only the conditional mean.

Nevertheless, conventional Markov-switching models address recurrent state dependence rather than permanent parameter instability. A model may successfully infer that the market is currently in a turbulent regime while continuing to estimate the characteristics of that regime using observations accumulated before an economically important structural change. This creates the possibility that regime inference is appropriate but the historical information used to estimate regime parameters is partly stale.

2.3 Online Change-Point Detection

Traditional structural-break procedures are predominantly retrospective: they identify changes after examining a sample containing observations both before and after the potential break. This is useful for historical characterization but creates a fundamental difficulty for genuine out-of-sample forecasting. If future observations contribute to the identification of a break, the estimated break cannot legitimately be used to construct forecasts that supposedly preceded those observations. The broader distinction between offline and online change-point methods is reviewed by Truong, Oudre, and Vayatis (2020).

Bayesian Online Change-Point Detection (BOCPD), developed by Adams and MacKay (2007), addresses this timing problem by estimating sequentially the posterior distribution of the **run length**, defined as the number of observations since the most recent change point. Rather than producing only a final segmentation of the sample, BOCPD updates the probability that a new run has begun as each observation arrives. The approach therefore provides information about potential changes using only data available at the current date.

The offline–online distinction is especially important in financial forecasting. Retrospective methods can be used to describe historical structural instability, while online procedures can provide causal adaptation signals. However, an online change detector is not itself a complete return, volatility, or risk-forecasting model. It indicates that recent observations may be inconsistent with the prevailing run, but an additional modelling mechanism is required to convert that signal into a predictive distribution. In the present study, BOCPD therefore plays a deliberately limited role: it supplies **online change alarms** that affect the amount of historical information used by a Markov-switching forecasting model rather than replacing regime modelling altogether.

2.4 Forecast Evaluation and Downside Risk

The appropriate evaluation of financial forecasts depends on the variable being forecast. Point-return forecasts are commonly compared using measures such as RMSE and MAE, but numerical rankings alone are insufficient to establish whether differences in predictive accuracy are statistically meaningful. Diebold and Mariano (1995) developed a formal framework for testing equal predictive accuracy using differences between model loss sequences. This approach is particularly important when forecast-performance differences are economically small, as is often the case for daily equity returns.

Volatility and density forecasts require additional criteria. QLIKE loss is widely used for variance forecasting because of its favourable properties when volatility itself is latent and must be evaluated using an imperfect proxy. Research on volatility evaluation emphasizes that conclusions can depend materially on the quality of the volatility proxy and on the loss function used (Andersen & Bollerslev, 1998; Hansen & Lunde, 2005; Patton, 2011). Predictive distributions can additionally be evaluated with log scores and other proper scoring rules, which reward forecasts for assigning high probability to outcomes that subsequently occur. Gneiting and Raftery (2007) provide a general treatment of strictly proper scoring rules, while Diebold, Gunther, and Tay (1998) demonstrate the importance of evaluating entire predictive densities rather than only conditional means or variances.

Downside-risk forecasting introduces a further distinction between the probability of crossing a risk threshold and the severity of losses beyond that threshold. Kupiec (1995) developed an unconditional-coverage test for Value-at-Risk, examining whether the observed frequency of exceedances is consistent with the nominal VaR level. Christoffersen (1998) extended this framework by testing whether exceedances are independent over time and by combining frequency and independence into a conditional-coverage assessment. These tests are particularly useful because a model can produce approximately the correct number of breaches while still generating them in undesirable clusters. Alternative VaR forecasting and validation approaches include Engle and Manganelli (2004), Kuester, Mittnik, and Paolella (2006), and Berkowitz, Christoffersen, and Pelletier (2011).

Expected Shortfall addresses information beyond the VaR threshold by measuring average loss conditional on being in the tail. Theoretical work by Artzner et al. (1999) and Acerbi and Tasche (2002) helped establish Expected Shortfall as a coherent tail-risk measure. More recent work has focused on formal evaluation. Fissler and Ziegel (2016) show that VaR and Expected Shortfall can be evaluated jointly using suitable scoring functions, while Patton, Ziegel, and Chen (2019) develop dynamic semiparametric approaches for forecasting and evaluating both quantities. This literature motivates evaluating the proposed model separately for point returns, volatility, predictive distributions, VaR, and Expected Shortfall rather than assuming that a model superior under one loss function must also be superior for another risk objective. Asia-Pacific applications also document extreme downside-risk spillovers and high-frequency volatility/VaR forecasting (Liu, 2014; So & Xu, 2013).

2.5 Closest Literature and Research Positioning

The present study lies at the intersection of these literatures, but the intersection itself is not entirely new. Bauwens, Dufays, and Rombouts (2014) explicitly compare Markov-switching and change-point GARCH models and develop marginal-likelihood methods for determining regime or change-point structures. Bauwens, De Backer, and Dufays (2014) go further by developing a Bayesian GARCH framework that permits both **recurrent regimes and non-recurrent change-point regimes**, using simulation and several financial return series to examine structural changes in volatility dynamics. These studies are particularly important because they demonstrate that recurrent regimes and structural breaks need not be treated as mutually exclusive explanations of financial instability.

Related work also connects structural instability directly to downside-risk forecasting. Hood and Malik (2018) incorporate volatility breaks into stock-return risk modelling and report improved out-of-sample VaR performance relative to several benchmarks. In parallel, Ardia et al. (2018) show that recurrent-

regime GARCH models can improve VaR, Expected Shortfall, and left-tail density forecasts compared with single-regime alternatives. Together, these studies indicate that both structural changes and regime dependence can matter for tail-risk estimation.

The contribution of the present study is therefore **not merely to combine structural breaks and regime switching**. Instead, it tests a more specific forecasting mechanism: whether **causal change-point information improves regime-dependent forecasts when it is used dynamically to adjust the historical estimation memory of a Markov-switching model**. The proposed break-aware specification is compared with an otherwise similar Markov-switching model that retains its conventional estimation window, allowing the incremental effect of memory adaptation to be isolated more clearly. The comparison is conducted under a strict no-look-ahead protocol, with regime-count selection completed before the test period and only filtered state probabilities and previously observed BOCPD alarms entering forecasts.

A second distinction is empirical breadth. Rather than drawing conclusions from a single return series, the framework is evaluated across the **S&P BSE SENSEX, Nikkei 225, Hang Seng, KOSPI, and S&P/ASX 200**, allowing structural instability and forecasting performance to vary across heterogeneous Asia-Pacific equity markets. This cross-market design permits the analysis to distinguish a general forecasting effect from one that is market specific.

This positioning maps directly onto the paper's three research questions. **RQ1** evaluates whether persistent structural changes and recurring latent regimes can be identified reliably, including under controlled simulation. **RQ2** tests whether causal break-aware adaptation improves one-day-ahead return and volatility forecasting relative to conventional benchmarks and, critically, relative to an otherwise comparable no-reset Markov specification. **RQ3** evaluates whether any incremental value is concentrated instead in downside-risk calibration through VaR and Expected Shortfall. The resulting framework therefore treats structural-break information not as an assumed source of universal forecasting improvement, but as an empirical signal whose usefulness must be established separately across markets and forecast objectives.

3. Data and Preliminary Analysis

3.1 Market Universe

The empirical analysis uses daily observations from five major Asia-Pacific equity indices: the **S&P BSE SENSEX (India), Nikkei 225 (Japan), Hang Seng Index (Hong Kong), KOSPI (South Korea), and S&P/ASX 200 (Australia)**. The sample begins on 15 September 2006 and extends through 26 August 2026, providing approximately two decades of observations encompassing substantially different economic and financial environments, including the Global Financial Crisis, the European sovereign-debt episode, the COVID-19 shock, subsequent inflationary pressures, and periods of substantial monetary-policy adjustment.

The selection of these markets provides a heterogeneous setting for examining structural instability and regime-dependent forecasting. The sample includes both developed and emerging-market characteristics, different monetary and regulatory environments, varying degrees of international capital-market integration, and distinct exposures to domestic and global economic shocks. This cross-market variation is important because structural changes and latent volatility regimes need not occur at the same frequency or with the same persistence across markets. A forecasting mechanism that performs well for one market may therefore not generalize automatically to another.

Table 1. Asia-Pacific Market Sample and Data Coverage

Market	Country/region	Yahoo ticker	Exchange calendar	Price sample start	Sample end	Return observations	OOS observations
S&P BSE SENSEX	India	^BSESN	XBOM	15 Sep 2006	26 Aug 2026	4,898	1,000
Nikkei 225	Japan	^N225	XTKS	15 Sep 2006	26 Aug 2026	4,874	1,000
Hang Seng	Hong Kong	^HSI	XHKG	15 Sep 2006	26 Aug 2026	4,908	1,000
KOSPI	South Korea	^KS11	XKRX	15 Sep 2006	26 Aug 2026	4,909	1,000
S&P/ASX 200	Australia	^AXJO	XASX	15 Sep 2006	26 Aug 2026	5,035	1,000

Note. OOS denotes the out-of-sample forecasting period. Price data begin on 15 September 2006; the first return is calculated from the next available trading session. Each market is evaluated over its final 1,000 available return observations; consequently, the calendar date of the first OOS forecast differs slightly across markets because national trading calendars differ.

Daily market data are obtained from Yahoo Finance through the yfinance interface. Trading sessions are independently checked against the relevant exchange calendars. Because national holidays

and exchange closures differ across markets, the number of available observations varies despite the common calendar start and end dates.

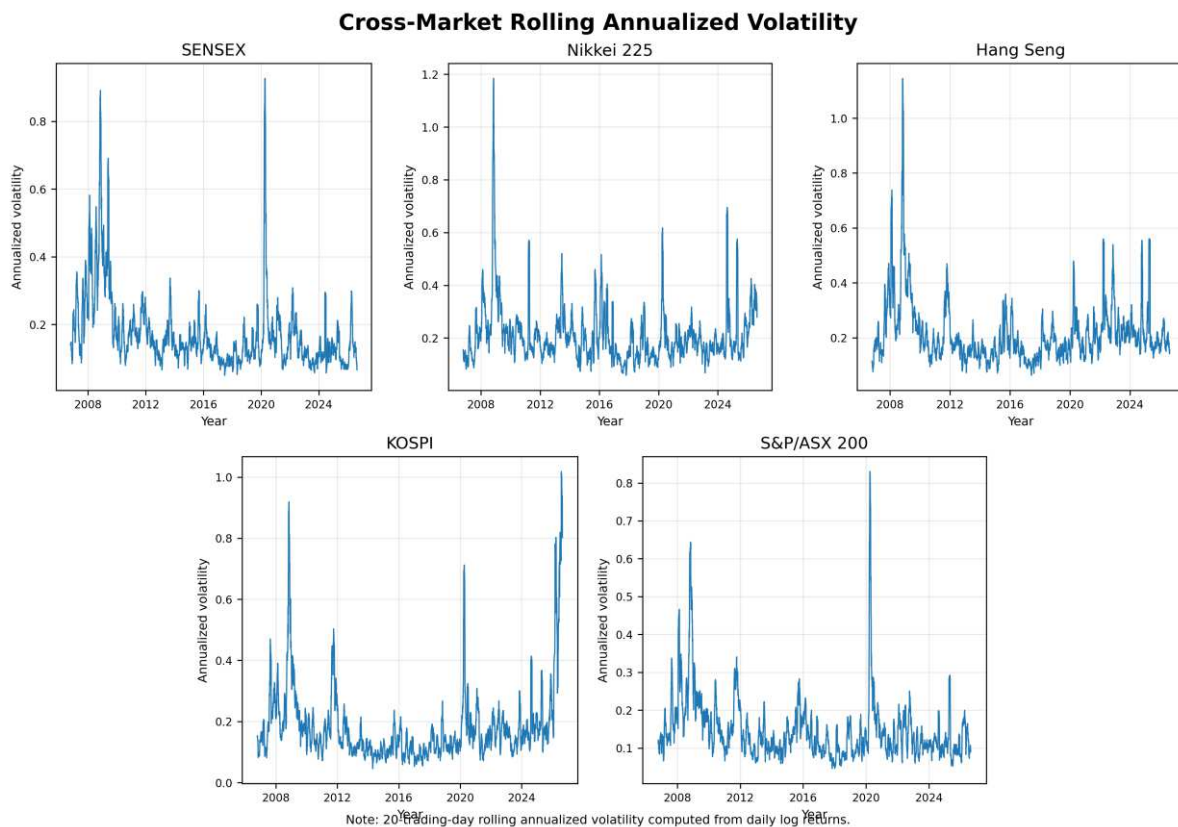


Fig. 1. Cross-Market Rolling Volatility in Asia-Pacific Equity Indices, September 2006–August 2026

Note. Each panel plots 20-trading-day annualized rolling volatility for the S&P BSE SENSEX, Nikkei 225, Hang Seng, KOSPI, and S&P/ASX 200. Annualized volatility equals the 20-day rolling standard deviation of daily log returns multiplied by $\sqrt{252}$. Individual market price, return, and volatility panels are reported in the supplementary appendix.

The cross-market visualization is useful because volatility episodes are neither identical in magnitude nor perfectly synchronized. Common global shocks generate increases in volatility across several markets, but the scale and persistence of those increases differ, providing an initial indication that a single common regime or break structure is unlikely to describe all five return series adequately.

3.2 Return Construction and Preprocessing

Daily continuously compounded returns for market i are constructed from adjusted index prices as

$$r_{i,t} = \ln(P_{i,t}) - \ln(P_{i,t-1}),$$

where $P_{i,t}$ denotes the adjusted closing level of index i on trading day t . Logarithmic returns are used because they are additive through time and are standard in financial time-series analysis.

For each market, observations are sorted chronologically and checked for duplicate timestamps, non-positive prices, and inconsistencies relative to the corresponding exchange calendar.

Returns are calculated only from observed consecutive price records; missing observations are **not mechanically forward-filled or interpolated**. This treatment avoids generating artificial zero returns that could distort volatility, structural-break, and regime estimates.

Extreme returns are retained unless there is evidence that they result from a data error. This is particularly important in the present context because unusually large observations during crisis periods are economically meaningful for identifying changes in volatility and downside risk.

3.3 Data Validation

Calendar matching indicates very high data coverage across all markets. Coverage is **99.53% for SENSEX**, **99.94% for the Nikkei 225**, **99.94% for Hang Seng**, **99.84% for KOSPI**, and **99.76% for the ASX 200**. The corresponding numbers of expected exchange sessions for which no Yahoo Finance observation is available are 23, 3, 3, 8, and 12, respectively.

These small vendor gaps are left unfilled rather than being replaced with mechanically generated prices. All retained prices are positive, timestamps are strictly increasing, and no duplicate timestamps remain in the validated datasets. A robust median-absolute-deviation procedure is also used to flag unusually large

returns for inspection, but observations are not removed simply because they are extreme. Complete data-integrity checks, flagged-observation counts, price-reconstruction checks, and market-specific SHA-256 hashes are reported in the reproducibility appendix.

3.4 Descriptive Statistics

Table 2. Descriptive Statistics of Daily Log Returns Across Markets

Statistic	SENSEX	Nikkei 225	Hang Seng	KOSPI	ASX 200
Observations	4,898	4,874	4,908	4,909	5,035
Mean daily return	0.0381%	0.0293%	0.0081%	0.0328%	0.0118%
Daily standard deviation	1.3004%	1.4965%	1.5276%	1.4210%	1.0589%
Minimum return	-14.1017%	-13.2341%	-14.1833%	-12.8489%	-10.2030%
Maximum return	15.9900%	13.2346%	13.4068%	16.4764%	6.7665%
Skewness	-0.226	-0.449	-0.120	-0.434	-0.675
Excess kurtosis	13.547	8.320	8.280	14.061	7.774
Annualized volatility	20.64%	23.76%	24.25%	22.56%	16.81%
Jarque–Bera statistic	37,412.69	14,187.55	13,999.70	40,502.15	13,030.52

Note. Annualized volatility is calculated using 252 trading days per year. The Jarque–Bera null hypothesis of normality is rejected at $p < 0.001$ for every market.

The descriptive statistics reveal substantial cross-market heterogeneity. Hang Seng exhibits the highest annualized volatility at approximately 24.25%, followed by the Nikkei 225 and KOSPI, whereas the ASX 200 has the lowest at approximately 16.81%. All five distributions are negatively skewed, although the magnitude differs considerably. More importantly, every market exhibits pronounced excess kurtosis, ranging from approximately 7.77 for the ASX 200 to more than 14 for KOSPI. Normality is therefore decisively rejected in every case.

3.5 Preliminary Diagnostics

The preliminary diagnostics reveal several common features across the five markets. Return series are strongly stationary under Augmented Dickey–Fuller tests (Dickey & Fuller, 1979), while KPSS tests (Kwiatkowski et al., 1992) do not reject return stationarity. Price-level evidence, by contrast, is broadly consistent with non-stationarity. The Hang Seng constant-only ADF specification provides somewhat different evidence, but its trend specification and KPSS results do not support treating the price

level as a stationary process. The subsequent analysis therefore focuses on logarithmic returns rather than index levels.

More importantly for the forecasting framework, all five markets display strong second-moment dependence. ARCH-LM tests reject the absence of conditional heteroskedasticity at extremely small significance levels in every market. Ljung–Box tests on squared returns (Ljung & Box, 1978) reject the absence of serial dependence across all 20 examined lags, providing clear evidence of volatility clustering. Hill tail-index diagnostics (Hill, 1975) are consistent with the large excess-kurtosis estimates reported in Table 2. These common characteristics justify modelling time-varying volatility and tail risk, while the variation in their magnitude across markets motivates allowing structural breaks and latent regime structures to be determined separately for each index. Normality is assessed using the Jarque–Bera statistic (Jarque & Bera, 1980). Full market-specific ADF, KPSS, ARCH, Ljung–Box, and tail-index results are reported in Appendix A.

4. Methodology

This section develops the empirical framework used to distinguish persistent structural changes from recurring latent market regimes and to assess whether causal structural-break information improves financial forecasting. The methodology is applied independently to each of the five Asia-Pacific equity indices described in Section 3. A central design principle is the separation between **retrospective analysis**, which uses the full historical sample for descriptive identification of structural instability, and **real-time forecasting**, in which only information available at a given forecast origin is permitted to influence the next prediction. This distinction prevents ex-post break dates or smoothed regime classifications from contaminating the out-of-sample experiment.

4.1 Research Framework

The empirical framework consists of six linked stages:

Research Framework

Structural-Break Detection and Regime-Dependent Forecasting in Asia-Pacific Equity Markets

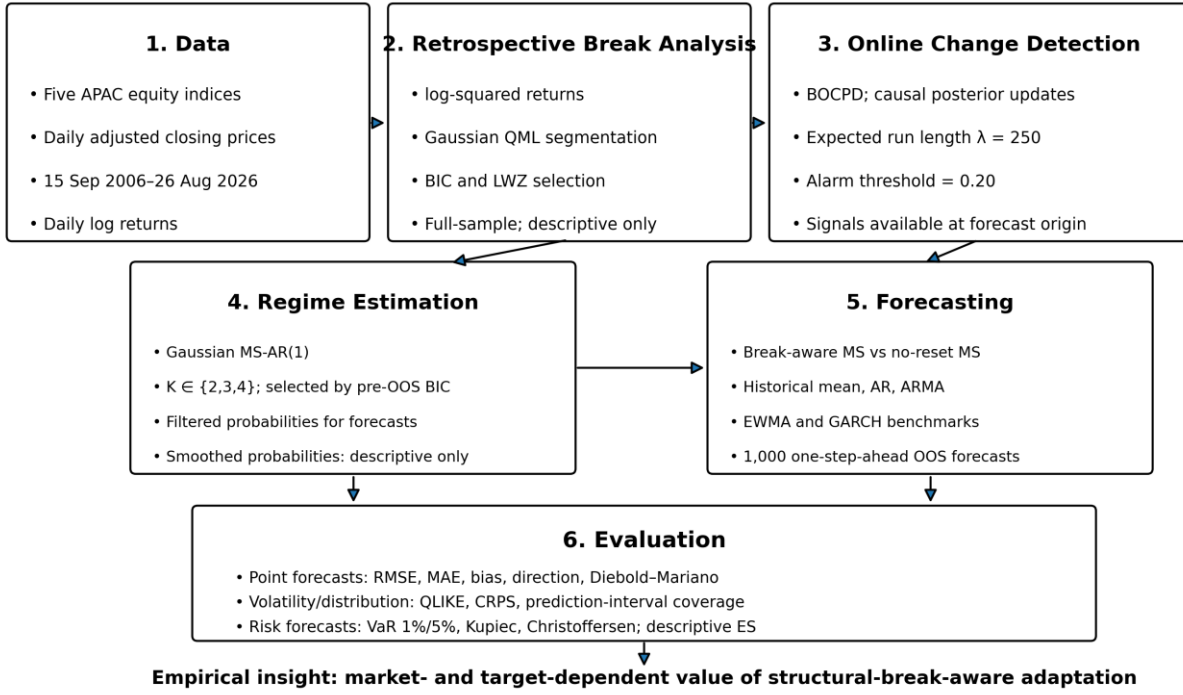


Fig. 2. Research Framework

Note. The framework separates retrospective historical analysis from causal forecasting. Retrospective structural-break estimates and full-sample smoothed regime probabilities are used only for interpretation. Out-of-sample forecasts use only information available before each target date, including filtered Markov probabilities and previously observed BOCPD alarms.

The first stage uses daily logarithmic returns for each Asia-Pacific equity index. Retrospective structural-break analysis then identifies historically important shifts in the volatility-related distribution of returns. In parallel, Bayesian Online Change-Point Detection generates sequential change signals using only information available through the current date. A Markov-switching autoregressive model captures recurrent latent market states. The proposed break-aware specification uses online change alarms to adjust the amount of historical information used in regime estimation, while an otherwise comparable no-reset Markov model retains a conventional estimation window. Forecasts are subsequently compared across point-return, volatility, predictive-distribution, and downside-risk criteria.

4.2 Retrospective Structural-Break Detection

Historical structural instability is examined using a multiple-break segmentation procedure applied to a volatility-oriented transformation of returns. Because financial instability is typically more visible in second moments than in daily conditional means, the segmentation variable for market i is

$$x_{i,t} = \log(r_{i,t}^2 + \varepsilon), \quad (1)$$

where

$$\varepsilon = 10^{-10}$$

is included to ensure numerical stability when daily returns are close to zero.

The procedure is best described as a **Bai–Perron-style segmentation framework**, rather than as the canonical Bai–Perron hypothesis-testing procedure. For a candidate segment containing observations $t = a, \dots, b$, a Gaussian quasi-maximum-likelihood segment cost is calculated from the estimated within-segment mean and variance. Dynamic programming then determines the partition that minimizes the total segment cost subject to restrictions on minimum segment length.

For the baseline specification, a minimum segment length of **126 trading observations** is imposed, corresponding approximately to half a trading year. Up to **12 structural breaks** are considered, and

candidate break positions are evaluated at five-observation increments. These restrictions prevent the procedure from identifying extremely short-lived episodes as independent persistent structural environments.

The number of breaks is evaluated using two information criteria. The Bayesian Information Criterion is defined generally as

$$BIC = -2 \log L + k \log n, \quad (2)$$

where L is the maximized quasi-likelihood, k is the number of estimated parameters, and n is the sample size. The second criterion is the Liu–Wu–Zidek information criterion (Liu, Wu, & Zidek, 1997), which imposes a stronger complexity penalty and therefore tends to favour more parsimonious segmentations.

Both criteria are retained because they represent different attitudes toward structural complexity. BIC is useful for identifying finer segmentation, whereas the **LWZ criterion is designated as the primary conservative criterion** because it is less likely to interpret short-lived volatility episodes as persistent structural breaks. The controlled simulation in Section 5 provides additional justification for this choice by assessing how the two criteria behave when the true number of breaks is known.

The resulting break dates are strictly retrospective. They are used to characterize historical structural instability and to compare break patterns across markets, but they are **not inserted into the forecasting engine**. This restriction is necessary because a break identified with the full sample may depend on observations that were unavailable at an earlier forecast origin.

4.3 Bayesian Online Change-Point Detection

To incorporate structural-change information into real-time forecasting, the analysis employs Bayesian Online Change-Point Detection (BOCPD; Adams & MacKay, 2007). Unlike retrospective segmentation, BOCPD processes observations sequentially and maintains a posterior distribution over the **run length**

$$R_t,$$

defined as the number of observations elapsed since the most recent change point.

At every date t , the procedure considers two possibilities. The current statistical run may continue, in which case the run length increases, or a new run may begin, in which case

$$R_t = 0.$$

The probability of a new change is therefore summarized by

$$P(R_t = 0 \mid x_{1:t}), \quad (3)$$

where $x_{1:t}$ denotes the volatility-related signal observed through date t .

BOCPD is applied to the same transformed squared-return series defined in Equation (1). A Normal–Inverse-Gamma conjugate specification is used for the unknown within-run mean and variance, generating a Student- t posterior predictive distribution. The hazard function is constant and defined as

$$H = \frac{1}{\lambda}, \quad (4)$$

with an expected run length of

$$\lambda = 250$$

trading days in the baseline specification.

An observation is classified as an **online change alarm** when the posterior probability of a new change reaches **0.20**, subject to a minimum separation of **20 trading days** between alarms. The posterior run-length distribution is truncated at a maximum of **1,200 observations** for computational efficiency.

The term *online change alarm* is used deliberately. A BOCPD alarm is not interpreted automatically as evidence of a permanent economic structural break. Instead, it represents a causal statistical signal that the recent volatility-generating environment may have become less consistent with the preceding run.

The information timing is central. At date t , the BOCPD posterior is computed using only observations through t . Therefore, when producing a forecast for $t + 1$, only alarms that have already occurred may affect the model. Future observations cannot alter historical alarm decisions.

4.4 Markov-Switching AR(1) Model

Recurring market conditions are represented using a Gaussian Markov-switching AR(1) model. For market i ,

$$r_{i,t} = c_{S_{i,t}} + \phi_{S_{i,t}} r_{i,t-1} + \sigma_{S_{i,t}} \varepsilon_{i,t}, \quad (5)$$

where

$$\varepsilon_{i,t} \sim N(0, 1),$$

and $S_{i,t} \in \{1, \dots, K_i\}$ denotes the unobserved regime at date t .

Each regime has its own intercept c_j , autoregressive coefficient ϕ_j , and conditional standard deviation σ_j . The latent state follows a first-order homogeneous Markov chain with transition probabilities $p_{jk} = P(S_t = k \mid S_{t-1} = j)$. (6)

The probability of remaining in regime j is p_{jj} . The corresponding expected duration is

$$D_j = \frac{1}{1 - p_{jj}}. \quad (7)$$

The model is estimated by maximum likelihood using an Expectation–Maximization procedure (Dempster, Laird, & Rubin, 1977). The expectation step applies the Hamilton forward filter together with a backward smoother to obtain state probabilities and expected transition counts. The maximization step updates regime-specific autoregressive parameters using probability-weighted estimation and updates the transition matrix from expected state transitions. Fitted states are subsequently ordered by estimated volatility to ensure economically consistent labeling across repeated estimations.

For numerical stability, returns are internally expressed in percentage units during estimation and converted back to decimal form for forecast evaluation.

Forecasts incorporate uncertainty over the next regime rather than assigning the observation to a single state. If

$$\pi_{t|t}$$

denotes the filtered regime-probability vector at date t , then the one-step-ahead state probabilities are

$$\pi_{t+1|t} = \pi_{t|t} P, \quad (8)$$

where P is the estimated transition matrix.

Conditional on state j ,

$$r_{t+1} \mid S_{t+1} = j, \mathcal{F}_t \sim N(\mu_{j,t+1}, \sigma_j^2),$$

with

$$\mu_{j,t+1} = c_j + \phi_j r_t. \quad (9)$$

The predictive distribution is therefore a finite Gaussian mixture. The one-step-ahead predictive mean is

$$\hat{\mu}_{t+1|t} = \sum_{j=1}^K \pi_{j,t+1|t} \mu_{j,t+1}, \quad (10)$$

while predictive variance accounts for both within-regime volatility and uncertainty regarding the next latent state.

4.5 Market-Specific Selection of the Number of Regimes

The number of regimes is not imposed uniformly across markets. Different Asia-Pacific equity markets may exhibit different latent-state structures, and selecting K using the full sample would introduce specification leakage.

For each market independently, candidate models with

$$K \in \{2, 3, 4\}$$

are estimated using only observations available before the first out-of-sample target. The preferred specification is the valid converged model with the lowest BIC (Schwarz, 1978).

A candidate that fails to converge is treated as invalid and is excluded from selection rather than being forced into the comparison. This avoids allowing a numerically unstable specification to influence the forecasting experiment.

Once the preferred K_i is selected for market i , it is **frozen for the entire 1,000-observation OOS period**. No subsequent OOS performance information is allowed to alter the number of states. This design ensures that market-specific regime complexity is selected causally rather than retrospectively.

4.6 Break-Aware Forecasting Mechanism

The proposed model differs from conventional Markov switching through the way historical information is used after an online change alarm.

The conventional no-reset Markov model uses a rolling estimation history of up to **1,500 observations**. Under the break-aware specification, the effective estimation window contracts following a causal BOCPD alarm. Immediately after an alarm, the usable window is reduced toward a minimum of **252 trading observations**, approximately one trading year. As additional observations accumulate following the detected change, the available estimation window expands progressively until the 1,500-observation maximum is reached.

Conceptually,

$$W_t = \min(W_{\max}, \max(W_{\min}, A_t)), \quad (11)$$

where $W_{\min} = 252$, $W_{\max} = 1,500$, and A_t represents the amount of post-alarm information available at the relevant forecast origin, subject to the implementation's causal dating convention.

The purpose of the mechanism is not to delete older observations permanently, but to **temporarily reduce their influence when recent evidence suggests that the statistical environment may have changed**. The usable historical sample subsequently expands as evidence accumulates under the new environment.

A new BOCPD alarm triggers an immediate refit of the break-aware Markov model. In the absence of a new alarm, formally estimated models are refitted at regular intervals of **21 trading observations**.

This structure creates a nested empirical comparison. Break-aware MS and no-reset MS share the same regime specification and

estimation machinery; the key difference is whether causal change information alters historical memory.

4.7 Filtered versus Smoothed Regime Probabilities

A critical safeguard against look-ahead bias is the distinction between filtered and smoothed regime probabilities.

Filtered probabilities are

$$P(S_t = j | r_1 \dots r_t), \quad (12)$$

and depend only on information available through date t . They are therefore appropriate for real-time forecasting.

4.8 Benchmark Forecasting Models

The proposed break-aware specification is compared against a set of conventional benchmarks covering simple conditional-mean models, volatility models, and regime-switching alternatives.

Table 3. Forecasting Models and Benchmark Specifications

Model	Mean dynamics	Variance / distribution dynamics	Regime-aware	Break-aware
Zero return	$E[r_{t+1}] = 0$	—	No	No
Historical mean	Historical sample mean	Historical sample variance	No	No
AR(1)	Autoregressive mean	—	No	No
ARMA(1,1)	AR and MA mean dynamics	—	No	No
EWMA	Historical mean	Recursive EWMA variance, $\lambda = 0.94$	No	No
GARCH(1,1)-t	Constant conditional mean	GARCH variance; Student- t innovations	No	No
MS without reset	Regime-specific AR(1)	Gaussian regime mixture	Yes	No
Break-aware MS	Regime-specific AR(1)	Gaussian regime mixture	Yes	Yes

The ARMA(1,1) benchmark is restricted to stationary and invertible solutions. Invalid or numerically implausible fits are rejected rather than permitted to generate unstable forecasts. The EWMA specification uses the standard decay factor

$$\lambda = 0.94.$$

GARCH(1,1)- t uses a constant conditional mean, recursive conditional variance, and Student- t innovations to accommodate heavy tails. The most important benchmark is MS without reset because it provides a direct test of whether BOCPD-driven memory adaptation adds value beyond the regime-switching structure itself.

4.9 Out-of-Sample Forecasting Design

Forecast performance is evaluated using the **final 1,000 return observations of each market**. Because national trading calendars differ, the first OOS calendar date is market specific, but the evaluation horizon is identical in number of forecasts.

For a forecast target $t + 1$, the information set is restricted to

$$\mathcal{F}_t.$$

A valid forecast therefore satisfies

$$\hat{r}_{t+1|t} = E(r_{t+1} | \mathcal{F}_t). \quad (14)$$

No observation dated $t + 1$ or later may affect model parameters, regime probabilities, BOCPD alarms, or recursive variance estimates used for that forecast.

All markets follow the same forecasting protocol. Periodically estimated models are refitted every **21 trading observations**, while the break-aware Markov model is additionally refitted when a new causal BOCPD alarm becomes available. Realized target returns are incorporated into recursive states only **after** forecasts for that target have been stored.

This strict walk-forward procedure ensures that forecasting performance reflects information that would actually have been available in real time.

4.10 Forecast Evaluation

Forecast evaluation is conducted separately for conditional returns, volatility and predictive distributions, and downside risk.

Point-return forecasting

Point forecasts are assessed using root mean squared error,

Smoothed probabilities are

$$P(S_t = j | r_1 \dots r_T), \quad T > t, \quad (13)$$

and incorporate observations occurring after date t .

Smoothed probabilities generally provide clearer historical regime classification, but using them in a historical forecasting experiment would introduce future information. Consequently, **every OOS forecast uses filtered probabilities only**. Smoothed probabilities are restricted to descriptive historical figures and ex-post regime interpretation.

$$RMSE = \sqrt{\frac{1}{T} \sum_{t=1}^T (r_t - \hat{r}_t)^2}, \quad (15)$$

mean absolute error,

$$MAE = \frac{1}{T} \sum_{t=1}^T |r_t - \hat{r}_t|, \quad (16)$$

mean forecast bias, and directional accuracy. Numerical differences in squared-error loss are evaluated statistically using two-sided **Diebold–Mariano tests** of equal predictive accuracy.

Volatility and predictive-distribution forecasting

Because latent daily variance is unobservable, squared daily returns are used as the common variance proxy:

$$RV_t = r_t^2. \quad (17)$$

The principal volatility loss is QLIKE,

$$QLIKE_t = \log(\hat{\sigma}_t^2) + \frac{r_t^2}{\hat{\sigma}_t^2}, \quad (18)$$

where lower average loss indicates better performance. Variance mean squared error is reported as an additional metric.

Predictive distributions are evaluated using the mean negative log score, the Continuous Ranked Probability Score (CRPS), and empirical coverage of nominal **80% and 95% prediction intervals**. For the Markov specifications, these quantities are computed from the full Gaussian-mixture predictive distribution. Pairwise Diebold–Mariano tests based on QLIKE differences assess whether volatility-performance differences are statistically meaningful.

Tail-risk forecasting

Downside-risk forecasts are evaluated at the

1% and 5%

levels.

Value-at-Risk forecasts for the Markov models are obtained from the full regime-mixture distribution. GARCH VaR uses the fitted Student- t distribution, while EWMA uses its corresponding conditional distributional specification.

VaR calibration is evaluated using three complementary tests. The **Kupiec unconditional-coverage test** examines whether the observed frequency of exceedances matches the nominal probability. The **Christoffersen independence test** evaluates whether exceedances are temporally independent. The

Christoffersen conditional-coverage test combines frequency and independence into a joint assessment of VaR adequacy.

Expected Shortfall is also calculated at the 1% and 5% levels. The empirical ES analysis is intentionally treated as **descriptive calibration rather than a formal ES hypothesis test**. Average realized returns on VaR-breach dates are compared with average predicted Expected Shortfall on those dates. Formal joint VaR–ES scoring is therefore left as an extension rather than being presented as part of the baseline inference.

Taken together, the evaluation framework allows the proposed model to be judged across multiple forecasting objectives. This distinction is essential because a model that performs well for point returns need not provide the best volatility forecast, and a model that minimizes average variance loss need not produce the most accurately calibrated downside-risk thresholds.

5. Controlled Simulation Validation

Before evaluating the forecasting framework on observed Asia-Pacific equity returns, a controlled simulation exercise is used to examine whether the modelling procedures can recover known latent regimes and persistent structural changes. In empirical financial data, the true regime sequence and exact locations of structural breaks are unobservable. Simulation therefore provides an environment in which estimated states and break locations can be compared directly with known ground truth. The exercise is designed as an **internal validation of identification performance**, rather than as evidence that the proposed break-aware model must necessarily improve real-market forecasts.

5.1 Simulation Design

The validation separates regime recovery from structural-break recovery through two distinct synthetic data-generating processes (DGPs). This separation avoids imposing structural breaks and recurrent Markov regimes simultaneously and makes the source of identification performance easier to interpret.

DGP A evaluates the ability of the Markov-switching component to recover recurring latent regimes. A sample of 4,200 business-day observations is generated from a three-state Markov-switching AR(1) process representing low-, medium-, and high-volatility environments. The daily innovation standard deviations are set to 0.5%, 1.5%, and 5.5%, respectively. The corresponding

autoregressive coefficients are 0.03, -0.04, and 0.08, while regime-specific daily intercepts are 0.035%, 0.005%, and -0.070%.

The states evolve according to a known first-order transition process. The probabilities of remaining in the low-, medium-, and high-volatility states are 97.6%, 92.0%, and 90.0%, respectively, generating persistent but recurrent regimes. Innovations follow standardized Student-*t* distributions with 10 degrees of freedom. The estimated Markov model nevertheless retains the Gaussian regime-emission specification used in the empirical analysis. The simulation therefore contains deliberate distributional misspecification rather than creating an artificially favourable Gaussian environment.

DGP B separately evaluates structural-break recovery. It also contains 4,200 observations, but the return process follows three persistent volatility segments rather than recurrent latent states. Daily volatility equals 0.7% for observations 1–1,400, increases to 2.2% after observation 1,400, and falls to 1.0% after observation 2,800. The true persistent break locations are therefore

$$\tau_1 = 1400, \tau_2 = 2800.$$

Innovations follow a standardized Student-*t* distribution with eight degrees of freedom. In addition, an isolated **-10% return shock** is introduced at observation 3,200 without changing subsequent parameters. This shock is deliberately included to test whether the structural-break procedure distinguishes a temporary extreme observation from a persistent change in the underlying volatility process.

The same retrospective segmentation methodology employed in the empirical analysis is applied to DGP B, with LWZ retained as the primary break-selection criterion and BIC reported for comparison. BOCPD is also applied using the baseline online-detection parameters.

5.2 Regime Recovery

The three-state Markov-switching AR(1) model is estimated on DGP A without revealing the true state sequence. Because numerical regime labels are arbitrary, estimated states are aligned ex post with the true states using the assignment that maximizes classification agreement. This alignment is used only to measure simulation accuracy and has no role in the empirical forecasting procedure.

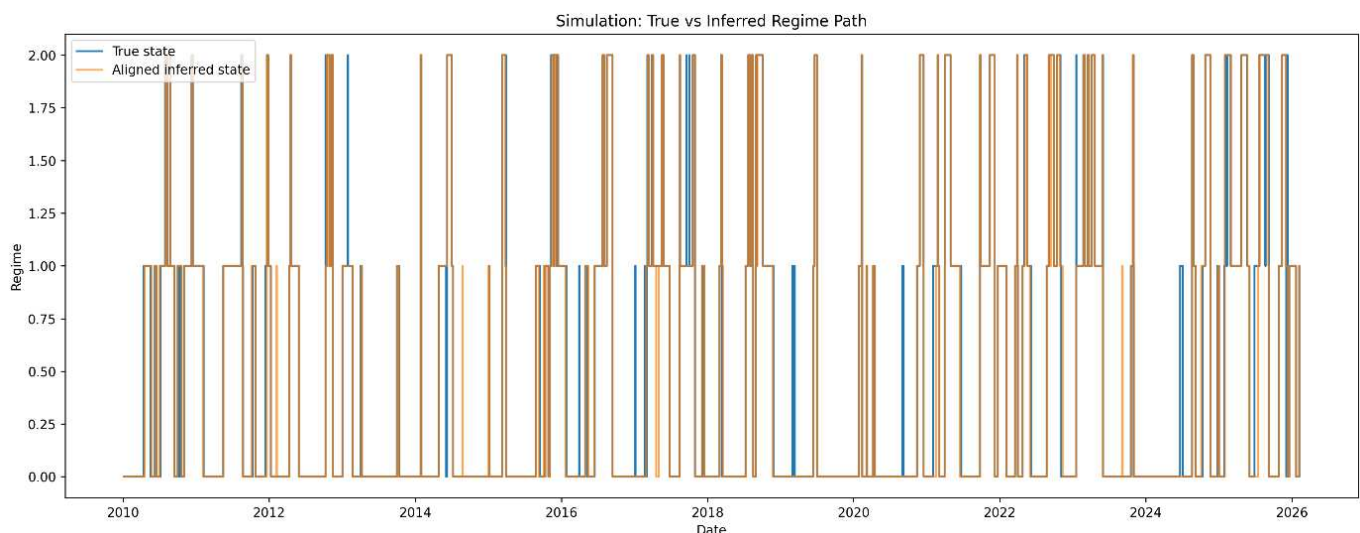


Fig. 3. True and Inferred Regime Paths in the Controlled Simulation

Note. The figure compares the known latent-state sequence generated by DGP A with the state sequence inferred by the estimated three-state Markov-switching AR(1) model. Estimated labels are aligned ex post with the true states solely for classification evaluation.

The inferred state sequence closely follows the simulated regime process. Overall aligned classification accuracy is **93.64%**. Recovery is strongest for the low-volatility state, which has recall of **97.68%**. Recall equals **87.51%** for the medium-volatility regime

and **84.39%** for the high-volatility regime. The lower recall for the highest-volatility state is consistent with its smaller representation and shorter average duration in the simulated sample.

5.3 Simulation Results

Table 4. Controlled Simulation Recovery Performance

Validation dimension	Metric	Result
Regime recovery	Simulated observations	4,200
	Overall aligned regime accuracy	93.64%
	Low-volatility state recall	97.68%
	Medium-volatility state recall	87.51%
	High-volatility state recall	84.39%
Structural-break recovery	Simulated observations	4,200
	True persistent structural breaks	2
	Primary criterion	LWZ
	Breaks selected by LWZ	2
	Breaks selected by BIC	3
	True breaks recovered within ± 30 trading days	2 of 2
	First true break: 1,400	Detected at 1,400
	Second true break: 2,800	Detected at 2,790
	Absolute localization error, first break	0 days
	Absolute localization error, second break	10 days
	BOCPD alarms in break DGP	29

Under the primary LWZ criterion, the structural-break procedure correctly selects **two persistent breaks**, exactly matching the true break count in DGP B. The first estimated break coincides with the true break at observation 1,400, while the second is identified at observation 2,790, only 10 trading days before the true location. Both breaks are therefore recovered within the pre-specified ± 30 -trading-day tolerance. BIC selects three breaks and consequently overstates the true number of persistent changes by one, supporting the use of the more conservative LWZ criterion in the empirical analysis.

Importantly, the isolated -10% shock at observation 3,200 does not cause LWZ to create an additional persistent structural segment. BOCPD generates 29 online alarms, which need not coincide with the two permanent breaks because the online detector is designed to signal short-run evidence of distributional change rather than to classify permanent structural shifts.

Overall, the simulation provides evidence of **internal identification validity**: the Markov component recovers most known latent states, while the primary structural-break procedure identifies both known persistent breaks with small localization error. These results do not establish that break-aware adaptation improves real-market forecasting performance. That question is evaluated

separately using the strictly out-of-sample Asia-Pacific experiments reported in Section 6.

6. Empirical Results

This section reports the empirical evidence from the five Asia-Pacific equity markets. The analysis proceeds from structural-break identification and online change-point evidence to market-specific regime selection, detailed regime interpretation for the primary SENSEX application, and strictly out-of-sample forecasting comparisons. Each market is evaluated over its final 1,000 return observations using the common forecasting protocol described in Section 4. Because trading calendars differ, the first OOS target varies slightly across markets, but the number of forecasts and evaluation procedure are identical.

6.1 Structural-Break Evidence Across Markets

The retrospective analysis reveals substantial heterogeneity in the degree of persistent structural instability across Asia-Pacific equity markets. Table 5 reports the number of breaks selected by the conservative LWZ criterion and by BIC, together with the LWZ-selected dates and the number of BOCPD online alarms.

Table 5. Structural-Break Evidence Across Asia-Pacific Equity Markets

Market	LWZ breaks	BIC breaks	LWZ-selected break dates	BOCPD alarms
S&P BSE SENSEX	1	10	9 Apr 2012	24
Nikkei 225	0	6	None	33
Hang Seng	2	6	5 Jul 2007; 22 Dec 2009	41
KOSPI	3	6	16 Aug 2012; 23 Dec 2019; 22 Jan 2026	32
S&P/ASX 200	1	7	5 Jan 2012	32

Note. Retrospective break detection is applied to log-squared returns using Gaussian QML dynamic programming. LWZ is the pre-designated conservative primary selection criterion. BOCPD alarms are causal online adaptation signals and need not correspond one-for-one with persistent retrospective structural breaks.

The results do not support a common break structure across the five markets. Under LWZ, the Nikkei 225 contains no persistent break, whereas the SENSEX and ASX 200 each contain one, Hang Seng contains two, and KOSPI contains three. The result is economically important because it suggests that structural instability cannot be represented adequately by imposing the same number or timing of breaks across regional equity markets.

The contrast between LWZ and BIC is also pronounced. BIC selects between six and ten breaks in every market, compared with zero to three under LWZ. This indicates that the volatility-related return series contain numerous shorter-lived shifts that can be represented through finer segmentation, but only a smaller subset is classified as persistent structural change when a stronger complexity penalty is imposed. The empirical pattern therefore reinforces the

rationale for retaining LWZ as the primary criterion while reporting BIC as evidence of more granular volatility variation.

For the primary SENSEX application, LWZ identifies a single break on **9 April 2012**. Segment-level estimates show that

annualized return volatility is approximately **29.3% before the break and 16.1% thereafter**, indicating a substantial change in the unconditional scale of daily fluctuations.

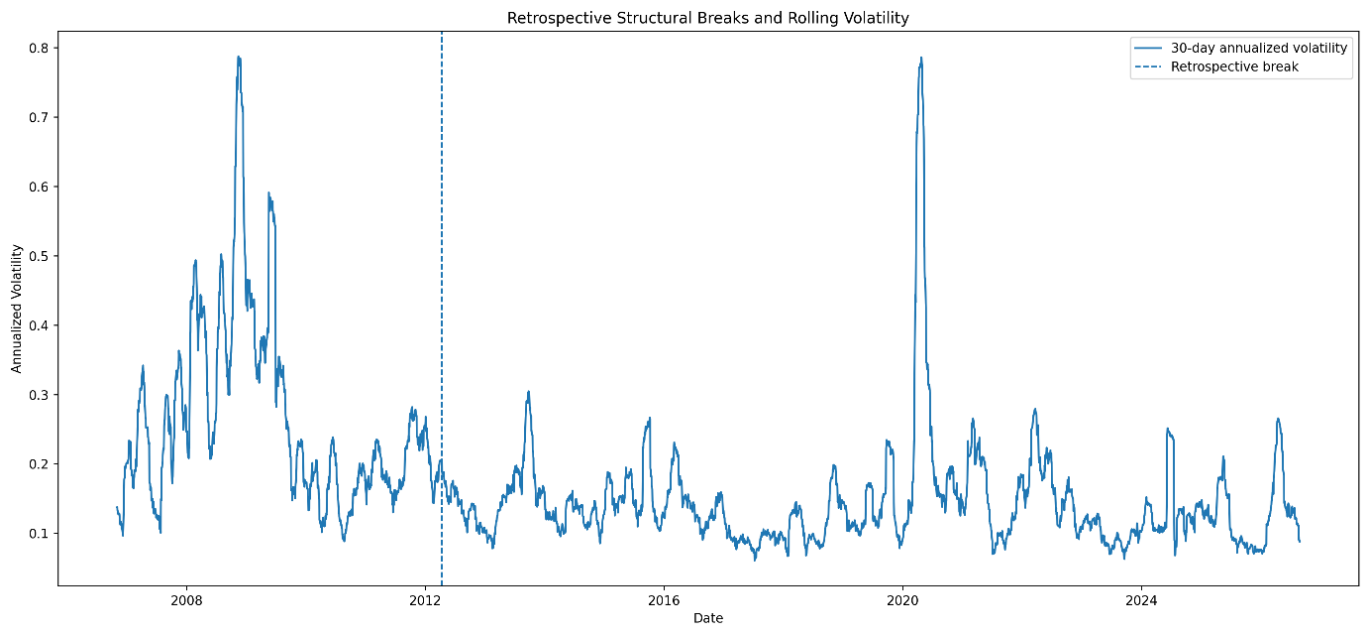


Fig. 4. Retrospective Structural Break and SENSEX Volatility

Note. The figure plots rolling annualized SENSEX volatility over the full sample. The vertical dashed line identifies the structural break selected by the primary LWZ criterion on 9 April 2012. Retrospective break dates are descriptive and are not used directly in the OOS forecasting process.

The cross-market results also illustrate why a structural break should not be equated automatically with any period of high volatility. Large volatility episodes can occur without being classified as permanent breaks, and the Nikkei result shows that a market can experience substantial volatility variation without LWZ identifying a persistent structural segmentation. This distinction becomes particularly important when retrospective breaks are compared with the much more frequent online BOCPD signals.

6.2 Online Change-Point Evidence

The forecasting mechanism relies on BOCPD rather than on the retrospective LWZ dates. In the SENSEX series, the baseline BOCPD specification generates **24 online change alarms** over the full sample. These signals are distributed across the observation

period rather than being concentrated around the single LWZ break. The highest individual alarm probability is approximately **0.944**, illustrating that some dates produce strong evidence of a change in the recent volatility-generating process even when the retrospective procedure does not classify the episode as a permanent structural segment.

Importantly, BOCPD remains active during the OOS forecasting period. For SENSEX, online alarms are observed during 2023, 2024, 2025, and 2026, so the break-aware forecasting mechanism actually contracts and subsequently rebuilds its estimation memory during the evaluation period rather than remaining observationally equivalent to the no-reset specification.

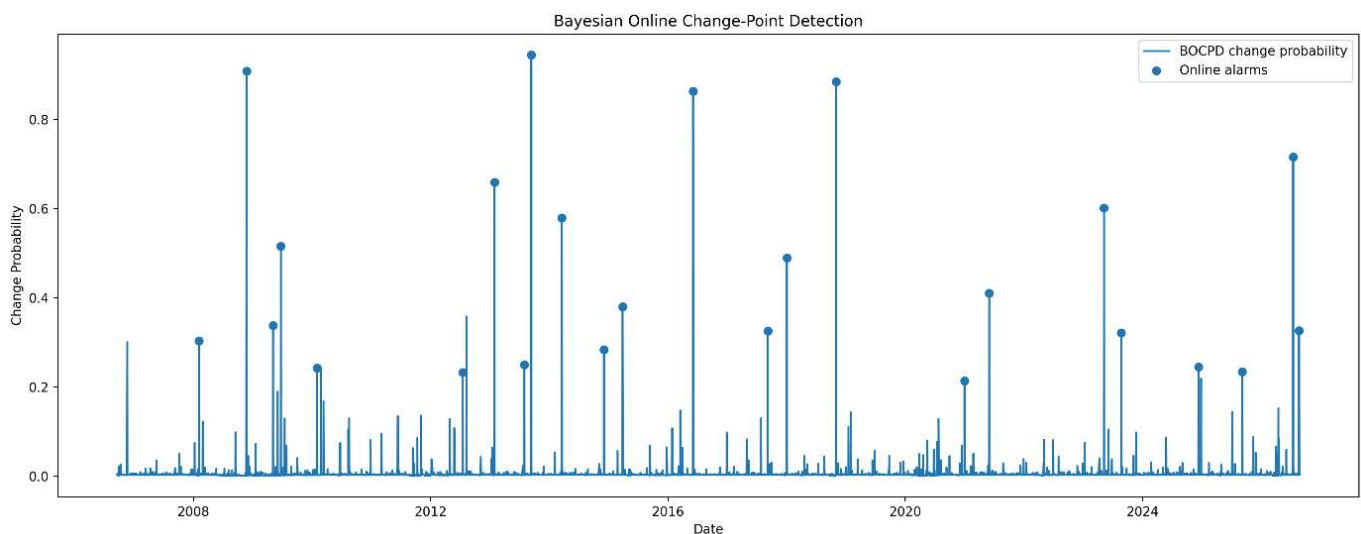


Fig. 5. SENSEX Bayesian Online Change-Point Probabilities

Note. The line shows the sequential posterior probability that a new run begins at each date. Markers identify observations satisfying the baseline online-alarm rule. The probabilities use information available only through the corresponding date. An online alarm is an adaptation signal rather than definitive evidence of a permanent structural break.

The frequency of online alarms also differs across markets. BOCPD generates 33 alarms for the Nikkei 225, 41 for Hang Seng, and 32 each for KOSPI and the ASX 200. In all markets, the number of online alarms materially exceeds the number of LWZ

structural breaks. This difference is expected rather than contradictory. LWZ is intended to identify persistent full-sample segmentation, whereas BOCPD is designed to react sequentially to evidence that the current run may have changed. Temporary shocks,

volatility transitions, or shorter-lived changes can therefore trigger BOCPD without constituting permanent retrospective breaks.

The distinction between the two mechanisms is important for interpreting the forecasting results that follow. The break-aware model is not reacting only to a small number of ex-post crisis dates; it is responding to a substantially richer stream of real-time statistical change signals.

Table 6. Market-Specific Pre-Out-of-Sample Markov Regime Selection

Market	K=2 BIC	K=3 BIC	K=4 BIC	Selected K	Convergence note
SENSEX	12,137.99	11,970.00	12,028.81	3	All candidates converged
Nikkei 225	13,106.74	12,937.31	12,968.48	3	All candidates converged
Hang Seng	13,197.35	13,091.51	13,087.65	4	All candidates converged
KOSPI	11,457.16	11,253.08	11,263.94*	3	K = 4 did not converge
ASX 200	11,112.18	10,891.94	10,888.76	4	All candidates converged

Note. Selection is based exclusively on each market's pre-OOS sample. An asterisk denotes a non-converged candidate, which is excluded from valid model selection. The selected number of regimes is frozen throughout the subsequent 1,000-observation forecasting experiment.

Three regimes are selected for SENSEX, the Nikkei 225, and KOSPI, while four are selected for Hang Seng and the ASX 200. The evidence therefore rejects the assumption that a common number of latent states is appropriate for all five markets. Hang Seng and the ASX 200 provide sufficient pre-OOS likelihood improvement to offset the additional parameters associated with a fourth state, whereas SENSEX and the Nikkei favour the more parsimonious three-state representation.

For KOSPI, the four-state candidate reaches the iteration limit without satisfying the convergence criterion. It is therefore excluded rather than treated as a valid competitor. The selected three-state

6.3 Pre-Out-of-Sample Regime Selection

The number of Markov regimes is selected independently for each market using only observations preceding the first OOS target. Table 6 reports BIC values for the two-, three-, and four-state candidates.

model is fully converged and has a substantially lower BIC than the valid two-state alternative. Reporting this case explicitly is important because it prevents numerical instability from being mistaken for evidence supporting greater regime complexity.

6.4 Detailed Regime Characteristics: SENSEX

The SENSEX is used as the primary market for detailed regime interpretation. The selected three-state model produces economically distinct low-, medium-, and high-volatility environments.

Table 7. SENSEX Markov-Switching Regime Characteristics and Transition Dynamics

Panel A. Regime characteristics

State	Economic label	Daily intercept	AR(1)	Daily volatility	Annualized volatility	Persistence	Expected duration	Occupancy
0	Low volatility	0.0776%	0.0950	0.6520%	10.35%	98.16%	54.3 days	47.97%
1	Medium volatility	0.0182%	0.0557	1.1724%	18.61%	97.13%	34.8 days	41.48%
2	High volatility	-0.1061%	0.0162	2.9334%	46.57%	96.97%	33.1 days	10.55%

Panel B. Transition matrix

Current state	Next: Low	Next: Medium	Next: High
Low	98.159%	1.783%	0.058%
Medium	2.168%	97.130%	0.702%
High	0.004%	3.021%	96.975%

Note. Regime labels are assigned after estimation according to increasing conditional volatility. Occupancy uses full-sample smoothed probabilities and is descriptive only.

The three regimes are clearly differentiated by volatility. Annualized volatility rises from approximately **10.4%** in the low-volatility state to **18.6%** in the medium state and **46.6%** in the high-volatility state. The high-volatility regime also has a negative conditional intercept, whereas the lower-volatility states have positive intercepts. The intercepts should not be interpreted as unconditional state returns because they enter the regime-specific AR(1) specification.

All three regimes are highly persistent. The probability of remaining in the low-volatility state exceeds 98%, while the medium- and high-volatility states have persistence of approximately 97%. Expected durations range from approximately

33 to 54 trading days. The low- and medium-volatility regimes jointly account for nearly 90% of the sample, while the high-volatility state represents about 10.6%.

The transition matrix also reveals an ordered structure. A direct transition from the low-volatility state to the high-volatility state has a probability of only about **0.06%**, while a high-to-low transition is effectively negligible. Movement between extreme states therefore occurs predominantly through the medium-volatility regime. This pattern is consistent with a market in which volatility conditions typically adjust through intermediate states rather than repeatedly jumping directly between calm and stressed environments.

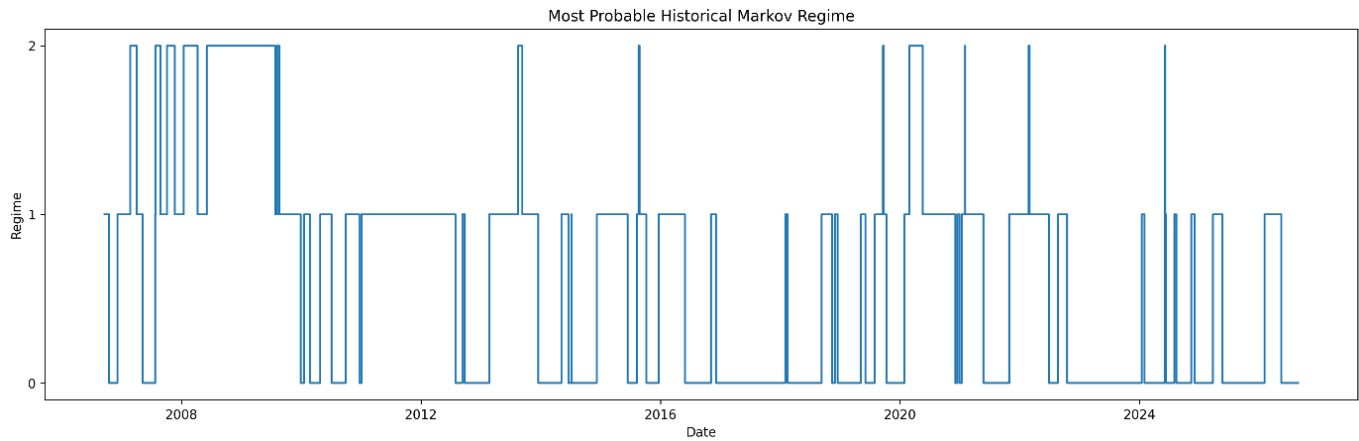


Fig. 6. Historical SENSEX Markov Regime Timeline

Note. The figure displays the most probable historical regime inferred from full-sample smoothed state probabilities. Smoothed probabilities are used only for descriptive historical interpretation; all OOS forecasts use filtered probabilities.

Complete regime parameters, transition matrices, and probability plots for the other four markets are reported in the Appendix.

6.5 Cross-Market Point-Return Forecasting

Table 8 summarizes the principal point-forecast results. Because daily return predictability is weak, the central question is whether additional structural and regime complexity produces a robust reduction in forecast error relative to simple benchmarks.

Table 8. Cross-Market Point-Forecast Performance

Market	Best non-break-aware model	Best comparator RMSE	Break-aware MS RMSE	Break-aware MS MAE	MS without reset RMSE	Break-aware directional accuracy	DM p-value, break-aware vs best comparator
SENSEX	ARMA(1,1)	0.8031%	0.7957%	0.5806%	0.8046%	55.2%	0.357
Nikkei 225	GARCH mean	1.4987%	1.5679%	1.0675%	1.5138%	51.4%	0.011
Hang Seng	Zero return	1.5540%	1.5627%	1.1401%	1.5600%	52.0%	0.110
KOSPI	GARCH mean	1.9534%	1.9638%	1.2003%	1.9556%	51.7%	0.575
ASX 200	Historical mean	0.7779%	0.7841%	0.5785%	0.7866%	50.2%	0.327

Note. RMSE and MAE are reported in percentage-return units. The final column reports the two-sided Diebold–Mariano p -value comparing the break-aware model with the lowest-RMSE non-break-aware comparator. Complete model-by-model performance and pairwise DM matrices are reported in the Appendix.

SENSEX is the only market in which the break-aware MS achieves the lowest numerical RMSE. Its RMSE of 0.7957% is below the ARMA, AR, conventional Markov, historical-mean, zero-return, and GARCH-mean forecasts. However, none of the corresponding DM comparisons is significant. The hybrid-versus-ARMA p -value is 0.357 and the hybrid-versus-no-reset MS p -value is 0.204. The SENSEX evidence therefore suggests a small numerical gain, but not statistically robust superior conditional-mean forecasting.

For the other four markets, simple models remain at least as competitive. In Hang Seng, a zero-return forecast has the lowest RMSE; in the ASX 200, the historical mean performs best; and GARCH conditional-mean forecasts achieve the lowest RMSE for the Nikkei and KOSPI. Most differences remain statistically insignificant.

The notable exception is the **Nikkei 225**, where the break-aware MS is materially weaker. Its RMSE of 1.5679% exceeds the GARCH-mean RMSE of 1.4987%, and the DM comparison rejects

equal predictive accuracy at approximately the 1% level. The hybrid is also significantly worse than the historical mean, AR(1), ARMA(1,1), and zero-return forecasts at conventional significance levels.

Overall, the cross-market evidence provides no general support for break-aware adaptation as a method for predicting next-day equity returns. Even where the hybrid delivers the best numerical result, as in SENSEX, statistical evidence of superiority is absent. The result is consistent with the low signal-to-noise ratio of daily equity returns and illustrates that successful identification of regimes and structural instability does not automatically create economically meaningful predictability in the conditional mean.

6.6 Cross-Market Volatility and Predictive-Distribution Forecasting

The conclusions become sharper when evaluation shifts to volatility. Table 9 focuses on the four models that produce explicit conditional variance forecasts.

Table 9. Cross-Market Volatility and Distribution Forecast Performance

Market	Break-aware QLIKE	MS no-reset QLIKE	GARCH QLIKE	EWMA QLIKE	Break-aware CRPS	Break-aware 95% coverage	DM p-value: Break-aware vs MS
SENSEX	-8.6542	-8.7280	-8.7465	-8.7291	0.004287	94.8%	0.040
Nikkei 225	-7.4469	-7.6322	-7.6304	-7.5755	0.007854	93.4%	0.076
Hang Seng	-7.3226	-7.3857	-7.3845	-7.3768	0.008311	96.1%	0.159
KOSPI	-7.3284	-7.5036	-7.4946	-7.4800	0.008883	91.2%	0.037
ASX 200	-8.7922	-8.7981	-8.8088	-8.7610	0.004199	94.2%	0.754

Note. Lower QLIKE and CRPS indicate better performance. The final column reports the two-sided Diebold–Mariano test comparing QLIKE loss for break-aware MS and MS without reset. Full CRPS, negative-log-score, variance-MSE, and interval-coverage results for all four models are reported in the Appendix.

The most consistent finding is that the break-aware specification **does not attain the lowest QLIKE in any of the five markets**. Conventional Markov switching produces the lowest QLIKE for the Nikkei, Hang Seng, and KOSPI, while GARCH performs best for SENSEX and the ASX 200.

Relative to the nested no-reset MS benchmark, the break-aware model has higher QLIKE in all five markets. The difference is

statistically significant for SENSEX ($p = 0.040$) and KOSPI ($p = 0.037$). In SENSEX, the hybrid is also significantly inferior to GARCH under QLIKE ($p = 0.006$) and to EWMA ($p = 0.029$). For the Nikkei and Hang Seng, the no-reset model's numerical advantage is not significant at the 5% level, while the ASX difference is economically small and statistically negligible.

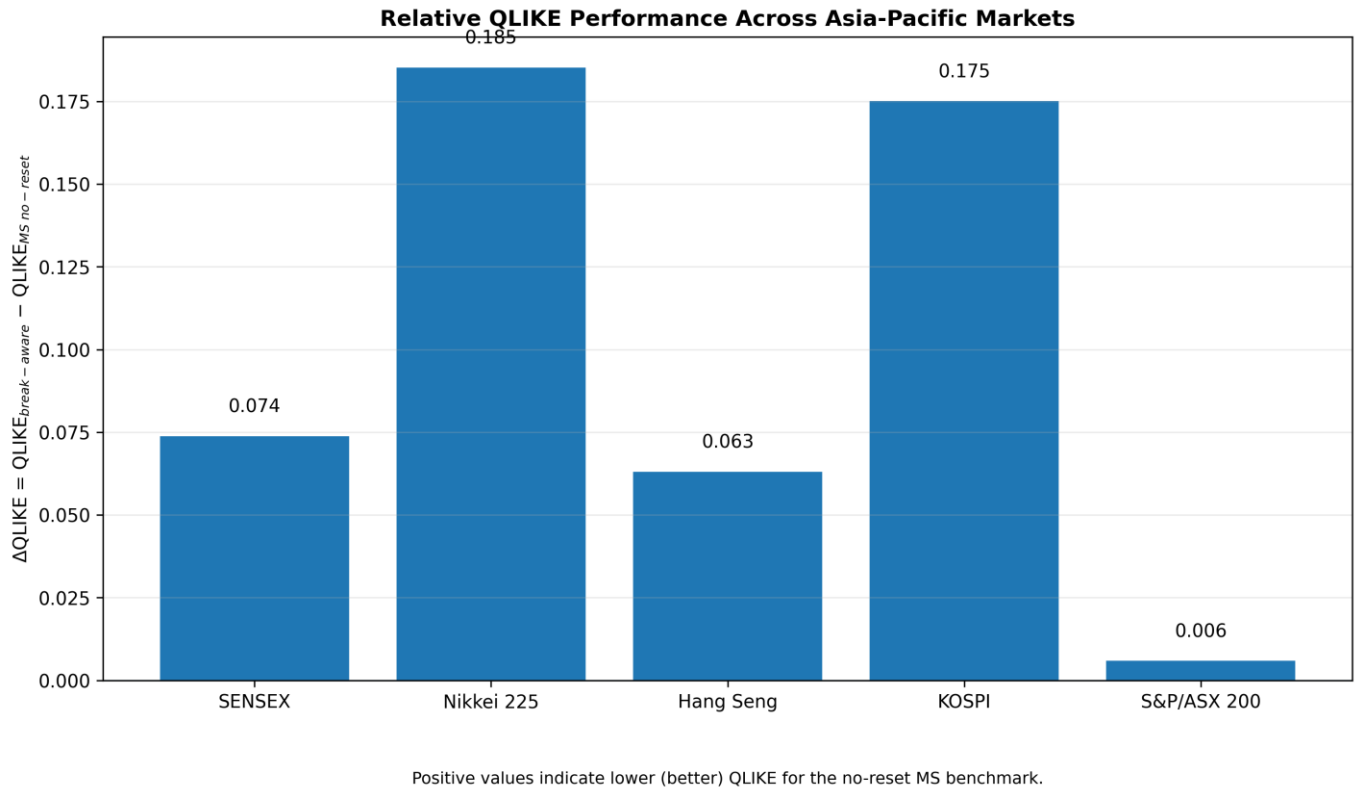


Fig. 7. Relative QLIKE Performance of Break-Aware and No-Reset Markov Models

Note. The figure plots $\Delta QLIKE = QLIKE_{break-aware} - QLIKE_{MS no-reset}$ for each market. Positive values indicate worse QLIKE performance for break-aware adaptation. The difference is positive in all five markets and statistically significant at the 5% level for SENSEX and KOSPI.

Predictive-distribution measures provide a similar message. The break-aware model does not consistently minimize CRPS, although its 95% interval coverage is reasonably close to the nominal rate in SENSEX, Hang Seng, and ASX 200. Coverage is lower for the Nikkei at 93.4% and particularly KOSPI at 91.2%, indicating undercoverage of realized outcomes in those markets.

The volatility results suggest that reducing historical memory after BOCPD alarms can impose an estimation cost. Longer samples may contain information that remains useful for estimating state transition probabilities and regime-specific variances, even when the recent environment has changed. The conventional Markov structure is already capable of responding to new observations through changing filtered state probabilities; additional memory contraction therefore need not improve the average conditional variance forecast.

6.7 Value-at-Risk Across Markets

The relative ranking changes when the objective shifts from average variance forecasting to downside-risk calibration. Table 10 reports 1% and 5% VaR backtests. Each OOS sample contains 1,000 observations, so the expected numbers of exceedances are 10 and 50, respectively.

Table 10. Cross-Market Value-at-Risk Backtesting

Panel A. 1% VaR

Market	Model	Expected	Actual	Kupiec p-value	Conditional-coverage p-value
SENSEX	Break-aware MS	10	15	0.139	0.266
	MS without reset	10	10	1.000	0.904
	GARCH(1,1)-t	10	8	0.510	0.755
	EWMA	10	18	0.022	0.053
Nikkei 225	Break-aware MS	10	10	1.000	0.226
	MS without reset	10	14	0.231	0.204
	GARCH(1,1)-t	10	15	0.139	0.157
	EWMA	10	19	0.011	0.026
Hang Seng	Break-aware MS	10	6	0.170	0.376
	MS without reset	10	7	0.314	0.573
	GARCH(1,1)-t	10	6	0.170	0.376
	EWMA	10	6	0.170	0.376

Market	Model	Expected	Actual	Kupiec p-value	Conditional-coverage p-value
KOSPI	Break-aware MS	10	26	<0.001	<0.001
	MS without reset	10	21	0.002	0.002
	GARCH(1,1)-t	10	27	<0.001	<0.001
	EWMA	10	28	<0.001	<0.001
ASX 200	Break-aware MS	10	17	0.043	0.074
	MS without reset	10	12	0.538	0.715
	GARCH(1,1)-t	10	17	0.043	0.074
	EWMA	10	23	<0.001	<0.001

Panel B. 5% VaR

Market	Model	Expected	Actual	Kupiec p-value	Conditional-coverage p-value
SENSEX	Break-aware MS	50	49	0.884	0.601
	MS without reset	50	48	0.770	0.251
	GARCH(1,1)-t	50	60	0.159	0.172
	EWMA	50	65	0.037	0.105
Nikkei 225	Break-aware MS	50	57	0.320	0.092
	MS without reset	50	63	0.069	0.064
	GARCH(1,1)-t	50	67	0.019	0.003
	EWMA	50	57	0.320	0.034
Hang Seng	Break-aware MS	50	21	<0.001	<0.001
	MS without reset	50	33	0.009	0.023
	GARCH(1,1)-t	50	31	0.003	0.013
	EWMA	50	28	<0.001	0.001
KOSPI	Break-aware MS	50	60	0.159	0.172
	MS without reset	50	57	0.320	0.034
	GARCH(1,1)-t	50	68	0.013	0.014
	EWMA	50	60	0.159	0.086
ASX 200	Break-aware MS	50	50	1.000	0.652
	MS without reset	50	56	0.393	0.406
	GARCH(1,1)-t	50	56	0.393	0.614
	EWMA	50	68	0.013	0.025

The 1% results show substantial market variation. Break-aware MS performs well for the Nikkei, where it records exactly ten exceedances, and remains statistically acceptable for SENSEX and Hang Seng. In SENSEX, however, conventional MS and GARCH are more closely aligned with the expected breach count. In the ASX 200, the break-aware model records 17 breaches; unconditional coverage is rejected at the 5% level, although the joint conditional-coverage test remains above the 5% threshold.

The most difficult market is KOSPI. Every model generates considerably more than the ten expected 1% exceedances, ranging from 21 under no-reset MS to 28 under EWMA. All four models fail conditional coverage decisively. This suggests that the extreme lower tail of KOSPI during the OOS period is not adequately represented by any of the baseline conditional distributions.

At the **5% level**, break-aware adaptation performs more favourably. SENSEX records **49 exceedances**, almost exactly matching the expected 50, and comfortably passes both coverage criteria. The ASX 200 produces exactly **50 exceedances**, with a conditional-coverage *p*-value of 0.652. The Nikkei break-aware model records 57 exceedances and is not rejected at 5%, whereas GARCH fails both unconditional and conditional coverage. For KOSPI, the break-aware model records 60 exceedances and passes conditional coverage, while conventional MS and GARCH are

rejected. EWMA also remains acceptable for KOSPI, showing that the hybrid's performance is not unique.

Hang Seng provides the clearest counterexample to any claim of universal improvement. The break-aware model records only **21** 5% exceedances against 50 expected, indicating an excessively conservative VaR threshold, and fails coverage decisively. All other models also underproduce breaches, although the discrepancy is largest for the break-aware specification.

The VaR evidence therefore supports a **market- and tail-dependent interpretation**. Break-aware adaptation provides strong 5% calibration for SENSEX and ASX 200 and acceptable calibration for the Nikkei and KOSPI, but fails materially for Hang Seng. Its relative benefit is also less consistent at the 1% level. Structural-change information can therefore help with moderate-tail threshold calibration in some markets without delivering universal tail-risk dominance.

6.8 Expected Shortfall

Expected Shortfall provides additional information about the severity of losses conditional on crossing the VaR threshold. Table 11 presents the calibration gap, defined as the average realized return on breach dates minus the average predicted ES. Negative gaps indicate that realized tail losses were more severe than forecast.

Table 11. Descriptive Expected Shortfall Calibration Across Markets

Market	Break-aware 1% ES gap	Closest non-hybrid comparator, 1%	Break-aware 5% ES gap	Closest non-hybrid comparator, 5%
SENSEX	-0.439 pp	GARCH: -0.108 pp	-0.138 pp	MS no-reset: +0.055 pp
Nikkei 225	-0.659 pp	MS no-reset: -0.051 pp	-0.096 pp	MS no-reset: -0.038 pp
Hang Seng	-3.002 pp	MS no-reset: -1.634 pp	-0.828 pp	GARCH: -0.194 pp
KOSPI	-0.807 pp	GARCH: -0.193 pp	-0.826 pp	MS no-reset: -0.496 pp
ASX 200	-0.091 pp	GARCH: +0.201 pp	-0.129 pp	MS no-reset: -0.006 pp

Note. ES calibration gap equals realized mean return on VaR-breach dates minus mean predicted ES on those dates. Negative values indicate that realized losses were more severe than predicted. The comparison is descriptive and does **not** constitute a formal ES hypothesis test. Detailed model-level results are reported in the Appendix.

The ES results do not provide evidence of broad superiority for break-aware adaptation. The break-aware calibration gap is negative at both tail levels in every market, implying that realized losses on breach dates are generally more severe than its predicted ES. The discrepancy is relatively small for the ASX 200 and for the Nikkei at the 5% level, but substantially larger for Hang Seng and KOSPI.

Conventional MS frequently produces a smaller absolute calibration gap, particularly at the 5% level for SENSEX, the Nikkei, KOSPI, and the ASX 200. GARCH is closest at the 1% level for SENSEX and KOSPI and at the 5% level for Hang Seng. The most pronounced discrepancy occurs for Hang Seng, where the break-aware 1% ES gap exceeds three percentage points in absolute value.

These results reinforce the distinction between **VaR threshold calibration and tail-loss severity**. A model may generate an acceptable number of 5% VaR breaches without necessarily predicting the average magnitude of losses beyond that threshold accurately. The break-aware specification's relatively favourable 5% VaR results in several markets therefore should not be interpreted as evidence of corresponding superiority in Expected Shortfall.

Taken together, Sections 6.5–6.8 provide a consistent central conclusion. Structural-break-aware adaptation does not universally improve financial forecasts. Its strongest numerical evidence appears in selected moderate-tail VaR applications, while simple models remain highly competitive for daily-return forecasting and conventional Markov switching or GARCH generally perform

better for average volatility. The usefulness of causal change-point information is therefore conditional on both the **market being modelled and the forecasting objective being evaluated**, rather than representing a uniform improvement over established forecasting approaches.

7. Robustness and Sensitivity Analysis

The robustness analysis examines whether the principal empirical conclusions depend materially on key modelling choices or numerical implementation. Because the S&P BSE SENSEX is used as the primary detailed application, the most computationally intensive sensitivity exercises are conducted for this market, while the remaining four Asia-Pacific indices provide external cross-market validation. The purpose of these tests is not to search for alternative specifications that improve the break-aware model, but to determine which findings are stable and which remain dependent on modelling assumptions.

7.1 Structural-Break Sensitivity

The baseline retrospective segmentation imposes a minimum segment length of 126 trading observations. To evaluate whether the primary structural-break result depends on this restriction, the analysis is repeated using minimum segment lengths of 63 and 252 observations, corresponding approximately to one quarter and one trading year.

Table 12. SENSEX Structural-Break Robustness to Minimum Segment Length

Minimum segment length	LWZ breaks	BIC breaks	LWZ-selected break date	BIC boundary reached	LWZ boundary reached
63 trading days	1	10	9 Apr 2012	No	No
126 trading days	1	10	9 Apr 2012	No	No
252 trading days	1	5	9 Apr 2012	No	No

The primary LWZ result is unchanged across all three specifications. Irrespective of whether the minimum permissible segment contains 63, 126, or 252 observations, LWZ selects a single persistent structural break on **9 April 2012**. This stability is notable because changing the segment-length restriction materially affects the degree of segmentation allowed by the procedure.

BIC is more sensitive to the same restriction. It selects ten breaks under both the 63- and 126-observation specifications but only five when the minimum segment is increased to 252 observations. This difference reinforces the interpretation developed in Section 6: BIC captures considerably finer volatility segmentation, whereas the more strongly penalized LWZ criterion identifies a smaller set of persistent structural shifts. Importantly, neither criterion reaches the maximum-break search boundary under any of the three SENSEX specifications. The primary one-break result is therefore not an artifact of either the baseline minimum segment length or the maximum-break constraint.

7.2 BOCPD Hazard Sensitivity

BOCPD requires an assumption about the expected duration of a statistical run. The baseline specification uses an expected run length of $\lambda = 250$ trading observations. Sensitivity is examined using $\lambda = 125$ and $\lambda = 500$.

As expected, the number of online alarms declines as the assumed run duration increases. With $\lambda = 125$, the SENSEX series generates 35 alarms, compared with 24 alarms under the baseline $\lambda = 250$ specification and 13 alarms when $\lambda = 500$. Mean posterior change probabilities also decline from approximately 0.0121 to 0.0070 and 0.0042, respectively, while the maximum posterior change probability remains high in all three cases, ranging from approximately 0.916 to 0.960.

This sensitivity is mechanically consistent with the role of the hazard function: a model that expects shorter regimes assigns greater prior probability to a new run and consequently generates more alarms. The result reinforces the need to interpret BOCPD signals as **model-dependent online adaptation indicators rather than immutable economic break dates**.

The current robustness exercise varies the BOCPD hazard at the detection stage only; the complete 1,000-observation forecasting experiment is not rerun under every alternative hazard. Accordingly, the analysis does not claim that forecast rankings are invariant to the hazard parameter. Forecast-performance sensitivity to alternative BOCPD hazards remains a useful extension.

7.3 Regime-Count Robustness

Regime complexity is controlled through the pre-OOS selection procedure described in Section 4. For SENSEX, the two-state model produces a BIC of **12,137.99**, the three-state model **11,970.00**, and the four-state model **12,028.81**. All three specifications converge successfully.

The three-state specification therefore remains preferred despite the four-state model achieving a somewhat higher likelihood, because the additional parameters are not sufficient to offset the BIC penalty. The BIC difference between the selected three-state model and the four-state alternative is approximately 58.8 points, providing substantial support for the more parsimonious specification.

More importantly, this comparison is completed using the **3,898 observations available before the first SENSEX OOS forecast on 3 August 2022**. The selected $K = 3$ is subsequently frozen and is never reconsidered in response to test-period forecast performance. The regime-count exercise therefore functions both as a robustness comparison and as a safeguard against specification leakage.

7.4 Estimation-Window Sensitivity

The break-aware specification is additionally evaluated using maximum estimation-window limits of **1,000, 1,500, and 2,000 observations**. All three settings generate identical SENSEX results: RMSE remains 0.7957%, QLIKE remains -8.6542, 95% interval coverage remains 94.8%, and the model records 15 and 49 breaches at the 1% and 5% VaR levels, respectively.

These identical results should **not** be interpreted as evidence of unusually strong robustness. During the SENSEX OOS experiment, the actual break-aware estimation window ranges from **252 to 731 observations**, with a mean of approximately **416 observations**. Consequently, even the smallest tested maximum of 1,000 observations never binds. The exercise therefore confirms only that the selected maximum-window ceiling is irrelevant over this particular OOS period. Detailed results are retained in the Appendix as a reproducibility check rather than presented as substantive evidence that forecast performance is insensitive to estimation-memory design.

7.5 Benchmark Stability

Numerical stability of the ARMA(1,1) benchmark is separately audited because repeated walk-forward estimation can occasionally generate non-stationary, non-invertible, or implausible solutions. Across all **1,000 SENSEX OOS forecasts**, the final implementation records **zero fallback forecasts**, all estimated states satisfy the validity restrictions, and every forecast is finite. The maximum absolute ARMA forecast is approximately **0.38%**.

These results indicate that the ARMA benchmark's performance is generated by valid estimated models rather than by fallback rules or numerical failures. The complete stability audit is reported in the Appendix.

7.6 Cross-Market External Validation

The strongest robustness check is the application of the same forecasting architecture to four additional Asia-Pacific markets. The Nikkei 225, Hang Seng, KOSPI, and S&P/ASX 200 differ materially from SENSEX in both their retrospective break structures and their selected numbers of Markov regimes. This heterogeneity provides a meaningful test of whether the primary empirical interpretation is specific to a single Indian equity index.

Several conclusions generalize. First, break-aware adaptation does **not** consistently improve point-return prediction. SENSEX is the only market in which it achieves the lowest numerical RMSE, and that advantage is not statistically significant; in the Nikkei 225 it performs significantly worse than several simpler alternatives. Second, conventional Markov switching without memory resetting achieves lower QLIKE than the break-aware model in **all five markets**, with the difference statistically significant for SENSEX and KOSPI. The conclusion that memory contraction does not systematically improve average-volatility forecasting is therefore not a SENSEX-specific result.

The downside-risk evidence is less uniform. At the 5% VaR level, the break-aware model passes conditional-coverage tests for SENSEX, the Nikkei 225, KOSPI, and the ASX 200, but fails decisively for Hang Seng. At the 1% level, performance is more fragile, particularly for KOSPI. Consequently, the cross-market evidence supports the central interpretation of the paper: **structural-break-aware adaptation has selective rather than universal forecasting value, and its effectiveness depends jointly on the market and the forecast objective.**

8. Discussion

The empirical results show that both structural instability and recurrent volatility regimes are important features of Asia-Pacific equity returns, but their forecasting implications are more nuanced than a simple preference for greater model complexity. Across the five markets, the evidence varies substantially in terms of structural-break frequency, preferred regime count, point-forecast performance, volatility loss, and downside-risk calibration. Taken together, the findings suggest that structural-break-aware adaptation is most useful when treated as a selective mechanism whose value depends on the market and forecasting objective rather than as a universally superior alternative to conventional volatility and regime-switching models.

8.1 Market Regimes Are Economically Meaningful but Heterogeneous

The first important finding is that latent market regimes and structural changes are clearly present, but their form differs across markets. Pre-out-of-sample model selection identifies three Markov regimes for the SENSEX, Nikkei 225, and KOSPI, compared with four regimes for Hang Seng and the S&P/ASX 200. The retrospective structural-break evidence is even more heterogeneous. The conservative LWZ criterion identifies no persistent break in the Nikkei 225, one in the SENSEX and ASX 200, two in Hang Seng, and three in KOSPI.

This variation argues against imposing a common structural representation across Asia-Pacific equity markets. Although these markets are exposed to many common global shocks, their domestic financial structures, monetary-policy environments, investor bases, market microstructures, and exposure to regional and global capital flows differ materially. A global event may therefore appear as a temporary high-volatility regime in one market but as part of a more persistent statistical change in another.

The SENSEX results provide an intuitive example. Its three-state model separates low-, medium-, and high-volatility environments, with annualized volatility increasing from approximately 10% in the calmest state to more than 46% in the highest-volatility state. All three states are highly persistent, and direct transitions between the lowest- and highest-volatility regimes are rare. This suggests that volatility conditions generally evolve through intermediate states rather than repeatedly jumping directly between extremes.

At the same time, the difference between retrospective LWZ breaks and more frequent BOCPD alarms illustrates an important conceptual distinction. Persistent structural breaks are relatively rare, whereas sequential evidence of local distributional change occurs much more frequently. The fact that BOCPD generates substantially more alarms than LWZ identifies permanent breaks is therefore not a contradiction. Rather, it reflects the difference between detecting possible real-time changes and concluding ex post that the statistical process entered a new persistent segment.

8.2 Greater Model Complexity Does Not Reliably Improve Return Prediction

The second finding is that successful identification of regimes and breaks does not translate automatically into better one-day-ahead return forecasts. Across most markets, simple conditional-mean benchmarks remain highly competitive. The break-aware Markov specification achieves the lowest numerical RMSE only for SENSEX, and even there the improvement is not statistically significant. In other markets, historical means, zero-return forecasts,

ARMA specifications, or GARCH conditional means perform as well as or better than the more elaborate regime-dependent model.

The Nikkei 225 provides the clearest example of the cost of additional complexity. The break-aware model produces significantly worse point forecasts than several simpler alternatives. This result emphasizes the difference between modelling the **distribution of returns** and predicting the **conditional mean of tomorrow's return**. A model can correctly distinguish calm and turbulent environments while still offering little information about the sign or magnitude of the next daily return.

This pattern is consistent with the well-known low signal-to-noise ratio of daily equity returns. Regime probabilities, structural-break signals, and time-varying volatility may contain meaningful information about risk without generating correspondingly strong predictability in expected returns. Additional parameters also increase estimation uncertainty. When the predictable component of daily returns is small, greater flexibility may fit changes in historical dynamics without producing sufficiently stable gains in conditional-mean forecasting.

The implication is methodological as well as empirical. Model complexity should not be justified solely by better in-sample characterization of market states. Its value must ultimately be evaluated against the specific out-of-sample objective for which the model is intended.

8.3 Break-Aware Memory Reduction Is Not Generally Superior for Average Volatility Forecasting

The third finding is more direct: shortening historical estimation memory after online change alarms does not systematically improve average volatility forecasts. The conventional Markov-switching model without memory resetting records a lower QLIKE value than the break-aware specification in all five markets. The difference is statistically significant for SENSEX and KOSPI, while in the remaining markets the conventional model is numerically superior but the difference is not statistically significant.

Several mechanisms may explain this result. First, older observations may continue to contain useful information about regime-specific variances and transition probabilities even after recent evidence of change. Removing too much history can therefore reduce estimation precision.

Second, BOCPD is intentionally sensitive to sequential evidence of distributional change and can respond to temporary volatility episodes that do not represent lasting structural shifts. Contracting the estimation window after such episodes may cause the model to overreact to short-lived disturbances.

Third, the Markov-switching structure already provides an endogenous form of adaptation. As new returns arrive, filtered state probabilities change and the model can rapidly shift probability toward a higher- or lower-volatility regime. A second adaptation mechanism based on reducing historical memory may therefore duplicate part of the flexibility that the latent-state model already provides.

Finally, shorter estimation samples are particularly costly for parameters associated with less frequently observed high-volatility regimes. Transition probabilities and crisis-state variances are difficult to estimate precisely when only a limited number of observations are available. The empirical results consequently reveal a trade-off between **responsiveness to structural change and statistical precision**. More rapid forgetting is not necessarily beneficial when recurring regimes themselves already absorb substantial time variation.

8.4 Tail-Risk Benefits Are Market Dependent

The downside-risk results provide the most nuanced—and economically relevant—evidence. At the 5% VaR level, the break-aware specification performs well in several markets. It records 49 exceedances in SENSEX against 50 expected and exactly 50 in the ASX 200, while passing conditional-coverage tests in both cases. It also remains acceptable at conventional significance levels for the Nikkei 225 and KOSPI.

These results suggest that structural adaptation can improve the positioning of a moderate-tail risk threshold even when it does not minimize average volatility loss. KOSPI is particularly informative: the break-aware model passes 5% conditional coverage whereas conventional Markov switching and GARCH are rejected. This indicates that the criteria used to evaluate average conditional variance and downside-risk thresholds need not produce the same model ranking.

However, the Hang Seng results demonstrate clearly that this benefit is not universal. The break-aware model generates only 21 breaches at the 5% level, substantially below the expected 50, implying an excessively conservative VaR threshold and a decisive rejection of correct coverage. At the more extreme 1% level, performance is also less stable across markets. KOSPI poses a particular challenge, with all baseline models producing too many extreme-tail exceedances.

The distinction between QLIKE and VaR performance is economically important. QLIKE evaluates how accurately a model predicts conditional variance on average across all observations. VaR evaluation instead concentrates on the location and frequency of lower-tail outcomes. A model can therefore perform relatively poorly in average variance forecasting while still positioning a particular tail quantile more effectively. Conversely, accurate average volatility forecasts do not guarantee well-calibrated tail probabilities.

Expected Shortfall adds a further qualification. The break-aware model frequently understates the average severity of losses once the VaR threshold is breached, and conventional Markov switching or GARCH often produces smaller descriptive ES calibration gaps. Therefore, successful 5% VaR calibration should not be interpreted as evidence that the entire tail distribution is estimated correctly.

The overall implication is that **structural-break-aware adaptation appears to be a target- and market-dependent risk-management tool rather than a universally superior forecasting mechanism**. Its value is most evident in selected moderate-tail applications, while conventional Markov switching and GARCH remain highly competitive for average volatility and, in several cases, extreme-tail forecasting. The broader contribution is therefore not the identification of one dominant model, but evidence that the usefulness of structural-change information depends jointly on the market, the horizon of adaptation, and the specific forecasting objective being evaluated.

9. Limitations and Future Research

The findings should be interpreted in light of several limitations concerning the empirical universe, volatility measurement, change-point specification, structural-break selection, regime modelling, and risk evaluation. These limitations do not invalidate the results, but they define the boundaries within which the conclusions should be understood and indicate several directions for future research.

First, the empirical analysis spans five major Asia-Pacific markets, but all of the underlying assets are **broad equity indices**. This provides useful cross-market heterogeneity, yet it does not establish that the same conclusions would hold for other asset classes. Structural instability and regime persistence may operate differently in fixed-income, foreign-exchange, commodity, or sector-specific markets because their information flows, trading structures, and exposures to macroeconomic shocks differ. Future research could therefore apply the same framework to **government and corporate bonds, exchange rates, commodities, and sectoral equity indices**, allowing the role of structural-break-aware adaptation to be assessed across a wider range of financial environments.

Second, volatility evaluation relies on daily squared returns as a proxy for latent daily variance. This measure enables a consistent comparison across all five markets but is inherently noisy because a single daily return provides only one realization from the underlying conditional distribution. Future work could employ high-frequency intraday data to construct realized variance or realized volatility

(Andersen et al., 2003), providing a more informative target for QLIKE and related volatility-loss functions.

Third, the BOCPD mechanism depends on tuning choices, particularly the **hazard rate and alarm threshold**. The sensitivity analysis demonstrates that alternative expected run lengths materially affect the number of detected online alarms. Although this behaviour is consistent with the design of BOCPD, it means that the timing and frequency of memory contraction remain partly model dependent. Future extensions could use **time-varying or state-dependent hazard functions, data-driven alarm thresholds, or learned change-detection mechanisms**, and could compare BOCPD with alternative sequential procedures such as CUSUM-based monitoring or other online change-point algorithms.

Fourth, retrospective structural-break inference depends on the model-selection penalty. Across the empirical markets, BIC generally identifies substantially more breaks than the more conservative LWZ criterion. The present study reports both and uses LWZ as the primary criterion, supported by the controlled simulation in which BIC over-selects the true break count. Nevertheless, future work could compare these results with a **verified canonical Bai–Perron testing framework**, sequential break tests, or **Bayesian multiple-change-point models** that place explicit prior distributions on both the number and location of structural changes.

Fifth, the Markov-switching specification imposes several simplifying assumptions. The number of states is finite and discrete, transition probabilities are homogeneous through time, and regime-specific innovations are Gaussian. Although the model captures economically distinct volatility states, financial returns exhibit substantial excess kurtosis and may involve transition dynamics that depend on market or macroeconomic conditions. Extensions could therefore consider **Student-*t* or other heavy-tailed regime distributions, time-varying transition probabilities, hidden semi-Markov models, or duration-dependent state processes**.

Sixth, Expected Shortfall evaluation in the present study is intentionally **descriptive rather than inferential**. Comparing realized losses on VaR-breach dates with predicted ES provides useful information about tail severity, but it does not constitute a formal test of joint VaR–ES forecast adequacy. Future research should employ **strictly consistent joint VaR–ES scoring functions and formal backtesting procedures**, building on the elicibility literature developed by Fissler and Ziegel and subsequent dynamic risk-forecasting work.

Finally, the study evaluates **statistical forecast performance rather than trading profitability**. Transaction costs, bid–ask spreads, market impact, turnover, portfolio constraints, and implementation delays do not affect the forecast comparisons presented here. These considerations would become essential only if regime probabilities, volatility forecasts, or VaR estimates were translated into dynamic trading, hedging, or asset-allocation strategies. A natural extension would therefore examine whether any statistical gains from break-aware adaptation remain economically meaningful after realistic implementation costs and portfolio constraints are introduced.

10. Conclusion

This study examines whether explicitly incorporating causal structural-break information into a regime-switching framework improves financial time-series forecasting. The central motivation is that equity markets may experience both recurring latent regimes and more persistent changes in their underlying data-generating processes. A model that captures only recurring states may continue to rely excessively on observations generated under an earlier structural environment, while a pure change-point model does not naturally represent recurring calm, turbulent, and crisis-like conditions. The paper therefore investigates whether online evidence of structural change can improve forecasting when it is used specifically to adjust the historical estimation memory of a Markov-switching model.

The empirical framework combines retrospective structural-break analysis, Bayesian Online Change-Point Detection, and a Markov-switching AR(1) specification with regime-dependent intercepts, autoregressive coefficients, and variances. The analysis is conducted across five major Asia-Pacific equity markets: the S&P BSE SENSEX, Nikkei 225, Hang Seng, KOSPI, and S&P/ASX 200. For each market, the number of latent regimes is selected independently using only information available before the first out-of-sample forecast and is then fixed throughout a 1,000-observation walk-forward experiment. The break-aware specification is evaluated against zero-return, historical-mean, AR(1), ARMA(1,1), EWMA, GARCH(1,1)-*t*, and conventional no-reset Markov-switching benchmarks. A controlled simulation with known regimes and structural breaks provides additional validation of the identification procedures.

The results show substantial heterogeneity in both structural instability and latent regime structure. The conservative LWZ criterion identifies no persistent break for the Nikkei 225, one for SENSEX and the ASX 200, two for Hang Seng, and three for KOSPI. Similarly, pre-out-of-sample model selection favours three Markov regimes for SENSEX, the Nikkei 225, and KOSPI, but four regimes for Hang Seng and the ASX 200. The detailed SENSEX application reveals economically distinct low-, medium-, and high-volatility states with high persistence and limited direct movement between the calmest and most volatile regimes. These findings indicate that Asia-Pacific equity markets do not share a uniform structural or regime configuration and that market-specific modelling is preferable to imposing a common state structure.

Identification of structural change and recurrent regimes, however, does not translate into general forecasting superiority. Break-aware Markov switching does not consistently outperform simpler benchmarks for one-day-ahead return prediction. SENSEX is the only market in which the hybrid model achieves the lowest numerical RMSE, and even there the improvement is not statistically significant. For volatility forecasting, the evidence is less favourable to memory resetting. Conventional Markov switching without break-induced contraction produces lower QLIKE loss than the break-aware specification in all five markets, with statistically significant advantages in SENSEX and KOSPI. The results suggest that shortening estimation memory can discard observations that remain informative for regime variances and transition probabilities, particularly when the latent-state model already adapts through changing filtered probabilities.

The downside-risk evidence is more nuanced. At the 5% Value-at-Risk level, break-aware forecasts are well calibrated for several markets. SENSEX records 49 breaches against 50 expected, while the ASX 200 records exactly 50, and the hybrid also passes conditional-coverage tests for the Nikkei 225 and KOSPI. Hang Seng provides a clear counterexample, with substantially too few breaches and a decisive rejection of correct coverage. Performance at the 1% level is also less stable, particularly for KOSPI. Expected Shortfall results further qualify the VaR evidence, as the break-aware specification frequently understates the severity of realized tail losses relative to conventional Markov switching or GARCH. Structural adaptation therefore appears more promising for selected moderate-tail calibration tasks than for the entire lower-tail distribution.

The principal contribution of the study is consequently not the identification of a universally dominant forecasting model, but a clearer characterization of when structural-change information adds value. By combining causal online change detection, regime-dependent forecasting, strict no-look-ahead evaluation, controlled simulation, and cross-market evidence, the paper shows that the usefulness of structural adaptation must be assessed against the specific market and forecasting objective. **Structural instability and recurring regimes are empirically relevant across Asia-Pacific equity markets, but identifying those features does not automatically improve forecast accuracy. The incremental value of structural-break-aware adaptation varies across markets and forecast objectives, with its greatest potential appearing in**

selected downside-risk calibration tasks rather than universal point or volatility forecasting.

Data and Code Availability

Daily index data are publicly available from Yahoo Finance and were accessed through the yfinance interface. The accompanying reproducibility archive contains the analysis code, market-specific diagnostics, model outputs, simulation results, robustness files, data-integrity hashes, and the no-look-ahead audit used to generate the reported results.

References

- Acerbi, C., & Tasche, D. (2002). On the coherence of expected shortfall. *Journal of Banking & Finance*, 26(7), 1487–1503. doi:10.1016/S0378-4266(02)00283-2
- Adams, R. P., & MacKay, D. J. C. (2007). Bayesian online changepoint detection. *arXiv preprint arXiv:0710.3742*.
- Andersen, T. G., & Bollerslev, T. (1998). Answering the skeptics: Yes, standard volatility models do provide accurate forecasts. *International Economic Review*, 39(4), 885–905. doi:10.2307/2527343
- Andersen, T. G., Bollerslev, T., Diebold, F. X., & Labys, P. (2003). Modeling and forecasting realized volatility. *Econometrica*, 71(2), 579–625. doi:10.1111/1468-0262.00418
- Andreou, E., & Ghysels, E. (2002). Detecting multiple breaks in financial market volatility dynamics. *Journal of Applied Econometrics*, 17(5), 579–600. doi:10.1002/jae.684
- Andrews, D. W. K. (1993). Tests for parameter instability and structural change with unknown change point. *Econometrica*, 61(4), 821–856. doi:10.2307/2951764
- Ang, A., & Timmermann, A. (2012). Regime changes and financial markets. *Annual Review of Financial Economics*, 4, 313–337. doi:10.1146/annurev-financial-110311-101808
- Ardia, D., Bluteau, K., Boudt, K., & Catania, L. (2018). Forecasting risk with Markov-switching GARCH models: A large-scale performance study. *International Journal of Forecasting*, 34(4), 733–747. doi:10.1016/j.ijforecast.2018.05.004
- Artzner, P., Delbaen, F., Eber, J.-M., & Heath, D. (1999). Coherent measures of risk. *Mathematical Finance*, 9(3), 203–228. doi:10.1111/1467-9965.00068
- Back, I.-M., & Jun, J. (2011). Testing contagion of the 1997–98 crisis in Asian stock markets with structural breaks and incubation periods. *Journal of Asian Economics*, 22(5), 356–368. doi:10.1016/j.asieco.2011.05.005
- Bai, J., & Perron, P. (1998). Estimating and testing linear models with multiple structural changes. *Econometrica*, 66(1), 47–78. doi:10.2307/2998540
- Bai, J., & Perron, P. (2003). Computation and analysis of multiple structural change models. *Journal of Applied Econometrics*, 18(1), 1–22. doi:10.1002/jae.659
- Bauwens, L., De Backer, B., & Dufays, A. (2014). A Bayesian method of change-point estimation with recurrent regimes: Application to GARCH models. *Journal of Empirical Finance*, 29, 207–229. doi:10.1016/j.jempfin.2014.06.008
- Bauwens, L., Dufays, A., & Rombouts, J. V. K. (2014). Marginal likelihood for Markov-switching and change-point GARCH models. *Journal of Econometrics*, 178, 508–522. doi:10.1016/j.jeconom.2013.08.017
- Bauwens, L., Preminger, A., & Rombouts, J. V. K. (2010). Theory and inference for a Markov switching GARCH model. *The Econometrics Journal*, 13(2), 218–244. doi:10.1111/j.1368-423X.2009.00307.x
- Berkowitz, J., Christoffersen, P., & Pelletier, D. (2011). Evaluating Value-at-Risk models with desk-level data. *Management Science*, 57(12), 2213–2227. doi:10.1287/mnsc.1080.0964
- Bollerslev, T. (1986). Generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics*, 31(3), 307–327. doi:10.1016/0304-4076(86)90063-1
- Brown, R. L., Durbin, J., & Evans, J. M. (1975). Techniques for testing the constancy of regression relationships over time. *Journal of the Royal Statistical Society: Series B (Methodological)*, 37(2), 149–163. doi:10.1111/j.2517-6161.1975.tb01532.x
- Cai, J. (1994). A Markov model of switching-regime ARCH. *Journal of Business & Economic Statistics*, 12(3), 309–316. doi:10.1080/07350015.1994.10524546
- Chiang, T. C., Jeon, B. N., & Li, H. (2007). Dynamic correlation analysis of financial contagion: Evidence from Asian markets. *Journal of International Money and Finance*, 26(7), 1206–1228. doi:10.1016/j.jimonfin.2007.06.005
- Christoffersen, P. F. (1998). Evaluating interval forecasts. *International Economic Review*, 39(4), 841–862. doi:10.2307/2527341
- Dempster, A. P., Laird, N. M., & Rubin, D. B. (1977). Maximum likelihood from incomplete data via the EM algorithm. *Journal of the Royal Statistical Society: Series B (Methodological)*, 39(1), 1–22. doi:10.1111/j.2517-6161.1977.tb01600.x
- 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. doi:10.1080/01621459.1979.10482531
- Diebold, F. X., & Mariano, R. S. (1995). Comparing predictive accuracy. *Journal of Business & Economic Statistics*, 13(3), 253–263. doi:10.1080/07350015.1995.10524599
- Diebold, F. X., Gunther, T. A., & Tay, A. S. (1998). Evaluating density forecasts with applications to financial risk management. *International Economic Review*, 39(4), 863–883. doi:10.2307/2527342
- Engle, R. F. (1982). Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation. *Econometrica*, 50(4), 987–1007. doi:10.2307/1912773
- Engle, R. F., & Manganelli, S. (2004). CAViaR: Conditional autoregressive Value at Risk by regression quantiles. *Journal of Business & Economic Statistics*, 22(4), 367–381. doi:10.1198/073500104000000370
- Fissler, T., & Ziegel, J. F. (2016). Higher order elicibility and Osband's principle. *The Annals of Statistics*, 44(4), 1680–1707. doi:10.1214/16-AOS1439
- Gallo, G. M., & Otranto, E. (2008). Volatility spillovers, interdependence and comovements: A Markov switching approach. *Computational Statistics & Data Analysis*, 52(6), 3011–3026. doi:10.1016/j.csda.2007.09.016
- Gneiting, T., & Raftery, A. E. (2007). Strictly proper scoring rules, prediction, and estimation. *Journal of the American Statistical Association*, 102(477), 359–378. doi:10.1198/016214506000001437
- Gray, S. F. (1996). Modeling the conditional distribution of interest rates as a regime-switching process. *Journal of Financial Economics*, 42(1), 27–62. doi:10.1016/0304-405X(96)00875-6
- Haas, M., Mittnik, S., & Paolella, M. S. (2004). A new approach to Markov-switching GARCH models. *Journal of Financial Econometrics*, 2(4), 493–530. doi:10.1093/jjfinec/nbh020
- Hamilton, J. D. (1989). A new approach to the economic analysis of nonstationary time series and the business cycle. *Econometrica*, 57(2), 357–384. doi:10.2307/1912559
- Hamilton, J. D., & Susmel, R. (1994). Autoregressive conditional heteroskedasticity and changes in regime. *Journal of Econometrics*, 64(1–2), 307–333. doi:10.1016/0304-4076(94)90067-1
- Hansen, P. R., & Lunde, A. (2005). A forecast comparison of volatility models: Does anything beat a GARCH(1,1)? *Journal of Applied Econometrics*, 20(7), 873–889. doi:10.1002/jae.800
- Hill, B. M. (1975). A simple general approach to inference about the tail of a distribution. *The Annals of Statistics*, 3(5), 1163–1174. doi:10.1214/aos/1176343247
- Hillebrand, E. (2005). Neglecting parameter changes in GARCH models. *Journal of Econometrics*, 129(1–2), 121–138. doi:10.1016/j.jeconom.2004.09.005
- Hood, M., & Malik, F. (2018). Estimating downside risk in stock returns under structural breaks. *International Review of Economics & Finance*, 58, 102–112. doi:10.1016/j.iref.2018.03.002
- 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. doi:10.1080/01621459.1994.10476824
- Jarque, C. M., & Bera, A. K. (1980). Efficient tests for normality, homoscedasticity and serial independence of regression residuals. *Economics Letters*, 6(3), 255–259. doi:10.1016/0165-1765(80)90024-5
- Karanasos, M., & Kartsaklas, A. (2009). Dual long-memory, structural breaks and the link between turnover and the range-based volatility.

- Journal of Empirical Finance*, 16(5), 838–851. doi:10.1016/j.jempfin.2009.06.001
- Kim, C.-J. (1994). Dynamic linear models with Markov-switching. *Journal of Econometrics*, 60(1–2), 1–22. doi:10.1016/0304-4076(94)90036-1
- Klaassen, F. (2002). Improving GARCH volatility forecasts with regime-switching GARCH. *Empirical Economics*, 27, 363–394. doi:10.1007/s001810100100
- Kuester, K., Mittnik, S., & Paoletta, M. S. (2006). Value-at-Risk prediction: A comparison of alternative strategies. *Journal of Financial Econometrics*, 4(1), 53–89. doi:10.1093/jjfinec/nbj002
- Kupiec, P. H. (1995). Techniques for verifying the accuracy of risk measurement models. *The Journal of Derivatives*, 3(2), 73–84. doi:10.3905/jod.1995.407942
- Kwiatkowski, D., Phillips, P. C. B., Schmidt, P., & Shin, Y. (1992). Testing the null hypothesis of stationarity against the alternative of a unit root: How sure are we that economic time series have a unit root? *Journal of Econometrics*, 54(1–3), 159–178. doi:10.1016/0304-4076(92)90104-Y
- Lamoureux, C. G., & Lastrapes, W. D. (1990). Persistence in variance, structural change, and the GARCH model. *Journal of Business & Economic Statistics*, 8(2), 225–234. doi:10.1080/07350015.1990.10509794
- Lee, C.-H., & Chou, P.-I. (2020). Structural breaks in the correlations between Asian and US stock markets. *The North American Journal of Economics and Finance*, 51, 101087. doi:10.1016/j.najef.2019.101087
- Liu, J., Wu, S., & Zidek, J. V. (1997). On segmented multivariate regression. *Statistica Sinica*, 7(2), 497–525.
- Liu, L. (2014). Extreme downside risk spillover from the United States and Japan to Asia-Pacific stock markets. *International Review of Financial Analysis*, 33, 39–48. doi:10.1016/j.irfa.2013.07.009
- Ljung, G. M., & Box, G. E. P. (1978). On a measure of lack of fit in time series models. *Biometrika*, 65(2), 297–303. doi:10.1093/biomet/65.2.297
- Marcucci, J. (2005). Forecasting stock market volatility with regime-switching GARCH models. *Studies in Nonlinear Dynamics & Econometrics*, 9(4). doi:10.2202/1558-3708.1145
- Mateus, C., Bagirov, M., & Mateus, I. (2024). Return and volatility connectedness and net directional patterns in spillover transmissions: East and Southeast Asian equity markets. *International Review of Finance*, 24(1), 83–103. doi:10.1111/irfi.12435
- Patton, A. J. (2011). Volatility forecast comparison using imperfect volatility proxies. *Journal of Econometrics*, 160(1), 246–256. doi:10.1016/j.jeconom.2010.03.034
- 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. doi:10.1016/j.jeconom.2018.10.008
- Perron, P. (1989). The Great Crash, the oil price shock, and the unit root hypothesis. *Econometrica*, 57(6), 1361–1401. doi:10.2307/1913712
- Rapach, D. E., & Strauss, J. K. (2008). Structural breaks and GARCH models of exchange rate volatility. *Journal of Applied Econometrics*, 23(1), 65–90. doi:10.1002/jae.976
- Schwarz, G. (1978). Estimating the dimension of a model. *The Annals of Statistics*, 6(2), 461–464. doi:10.1214/aos/1176344136
- So, M. K. P., & Xu, R. (2013). Forecasting intraday volatility and Value-at-Risk with high-frequency data. *Asia-Pacific Financial Markets*, 20(1), 83–111. doi:10.1007/s10690-012-9160-1
- Truong, C., Oudre, L., & Vayatis, N. (2020). Selective review of offline change point detection methods. *Signal Processing*, 167, 107299. doi:10.1016/j.sigpro.2019.107299
- Wang, P., & Theobald, M. (2008). Regime-switching volatility of six East Asian emerging markets. *Research in International Business and Finance*, 22(3), 267–283. doi:10.1016/j.ribaf.2007.07.001
- Worthington, A. C., & Higgs, H. (2004). Transmission of equity returns and volatility in Asian developed and emerging markets: A multivariate GARCH analysis. *International Journal of Finance & Economics*, 9(1), 71–80. doi:10.1002/ijfe.222
- Yang, K., Chen, L., & Tian, F. (2015). Realized volatility forecast of stock index under structural breaks. *Journal of Forecasting*, 34(1), 57–82. doi:10.1002/for.2318

Supplementary Appendices

Appendix A. Market-Specific Statistical Diagnostics

This appendix reports the complete preliminary statistical diagnostics underlying Section 3. Diagnostics are conducted independently for the S&P BSE SENSEX, Nikkei 225, Hang Seng, KOSPI, and S&P/ASX 200.

A.1 Unit-Root and Stationarity Tests

Table A1. Full ADF and KPSS Diagnostics Across Markets

Market	Series	Test	Specification	Statistic	p-value	Lag / bandwidth
SENSEX	Log price	ADF	constant	-0.793	0.821	12
SENSEX	Log price	ADF	constant + trend	-3.654	0.026	17
SENSEX	Returns	ADF	constant	-19.025	<0.001	11
SENSEX	Returns	ADF	constant + trend	-19.023	<0.001	11
SENSEX	Log price	KPSS	constant	11.025	0.010	42
SENSEX	Log price	KPSS	constant + trend	0.741	0.010	42
SENSEX	Returns	KPSS	constant	0.025	0.100	17
SENSEX	Returns	KPSS	constant + trend	0.025	0.100	17
Nikkei 225	Log price	ADF	constant	0.596	0.988	3
Nikkei 225	Log price	ADF	constant + trend	-2.296	0.436	3
Nikkei 225	Returns	ADF	constant	-41.957	<0.001	2
Nikkei 225	Returns	ADF	constant + trend	-42.012	<0.001	2
Nikkei 225	Log price	KPSS	constant	9.498	0.010	42
Nikkei 225	Log price	KPSS	constant + trend	1.006	0.010	42
Nikkei 225	Returns	KPSS	constant	0.378	0.087	17
Nikkei 225	Returns	KPSS	constant + trend	0.053	0.100	18
Hang Seng	Log price	ADF	constant	-3.000	0.035	10
Hang Seng	Log price	ADF	constant + trend	-3.019	0.127	10
Hang Seng	Returns	ADF	constant	-23.366	<0.001	9
Hang Seng	Returns	ADF	constant + trend	-23.364	<0.001	9
Hang Seng	Log price	KPSS	constant	1.323	0.010	42
Hang Seng	Log price	KPSS	constant + trend	0.779	0.010	42
Hang Seng	Returns	KPSS	constant	0.038	0.100	15
Hang Seng	Returns	KPSS	constant + trend	0.031	0.100	15
KOSPI	Log price	ADF	constant	0.241	0.974	22
KOSPI	Log price	ADF	constant + trend	-1.213	0.908	22
KOSPI	Returns	ADF	constant	-14.717	<0.001	21
KOSPI	Returns	ADF	constant + trend	-14.766	<0.001	21
KOSPI	Log price	KPSS	constant	7.248	0.010	42
KOSPI	Log price	KPSS	constant + trend	0.334	0.010	42
KOSPI	Returns	KPSS	constant	0.218	0.100	13
KOSPI	Returns	KPSS	constant + trend	0.101	0.100	14
ASX 200	Log price	ADF	constant	-0.982	0.760	12
ASX 200	Log price	ADF	constant + trend	-2.857	0.177	12
ASX 200	Returns	ADF	constant	-20.812	<0.001	11
ASX 200	Returns	ADF	constant + trend	-20.827	<0.001	11
ASX 200	Log price	KPSS	constant	8.724	0.010	42
ASX 200	Log price	KPSS	constant + trend	1.010	0.010	42
ASX 200	Returns	KPSS	constant	0.076	0.100	7
ASX 200	Returns	KPSS	constant + trend	0.022	0.100	7

ADF and KPSS results jointly support treating daily returns as stationary while price levels are generally non-stationary. The Hang Seng constant-only ADF result is an exception to the general price-level ADF pattern, but its trend-augmented ADF and KPSS results do not support treating the price level as stationary.

A.2 Conditional Heteroskedasticity and Tail Behaviour

Table A2. Distributional, ARCH, and Tail Diagnostics

Market	Jarque–Bera	ARCH-LM	ARCH-LM p	Hill upper	Hill lower
SENSEX	37,412.69	803.56	<0.001	2.639	2.643
Nikkei 225	14,187.55	1,024.03	<0.001	3.071	2.845
Hang Seng	13,999.70	980.78	<0.001	3.162	2.943
KOSPI	40,502.15	1,004.27	<0.001	2.613	2.274
ASX 200	13,030.52	1,468.82	<0.001	2.798	2.894

Normality is rejected for every market, and the ARCH-LM tests provide strong evidence of conditional heteroskedasticity. The Hill estimates are consistent with heavy-tailed return distributions.

A.3 Ljung–Box Diagnostics

Table A3a. Ljung–Box Tests for Returns — SENSEX

Lag	LB statistic	p-value
1	8.695	0.003
2	10.090	0.006
3	13.329	0.004
4	17.103	0.002
5	17.138	0.004
6	22.953	<0.001
7	35.365	<0.001
8	36.188	<0.001
9	36.381	<0.001
10	41.546	<0.001
11	43.682	<0.001
12	54.118	<0.001
13	55.109	<0.001
14	55.366	<0.001
15	56.733	<0.001
16	56.833	<0.001
17	65.659	<0.001
18	66.062	<0.001
19	69.486	<0.001
20	69.772	<0.001

Table A3b. Ljung–Box Tests for Returns — Nikkei 225

Lag	LB statistic	p-value
1	9.904	0.002
2	10.346	0.006
3	15.318	0.002
4	15.337	0.004
5	18.312	0.003
6	19.966	0.003
7	20.587	0.004
8	20.626	0.008
9	20.646	0.014
10	20.810	0.022
11	21.071	0.033
12	21.182	0.048
13	22.091	0.054
14	23.292	0.056
15	24.894	0.051
16	27.576	0.036
17	28.300	0.042
18	28.500	0.055
19	28.613	0.072
20	30.238	0.066

Table A3c. Ljung–Box Tests for Returns — Hang Seng

Lag	LB statistic	p-value
1	0.911	0.340
2	0.971	0.615
3	4.112	0.250
4	4.460	0.347
5	4.852	0.434
6	6.727	0.347
7	8.059	0.327
8	14.128	0.079
9	18.047	0.035
10	25.346	0.005
11	25.650	0.007
12	25.745	0.012
13	27.727	0.010
14	28.496	0.012
15	28.835	0.017
16	31.031	0.013
17	31.713	0.016
18	31.914	0.023
19	40.747	0.003
20	41.046	0.004

Table A3d. Ljung–Box Tests for Returns — KOSPI

Lag	LB statistic	p-value
1	3.570	0.059
2	6.807	0.033
3	14.374	0.002
4	16.417	0.003
5	21.291	<0.001
6	21.755	0.001
7	24.434	<0.001
8	24.947	0.002
9	25.798	0.002
10	28.190	0.002
11	40.005	<0.001
12	42.712	<0.001
13	42.714	<0.001
14	67.276	<0.001
15	67.396	<0.001
16	71.021	<0.001
17	71.232	<0.001
18	78.641	<0.001
19	83.116	<0.001
20	92.803	<0.001

Table A3e. Ljung–Box Tests for Returns — ASX 200

Lag	LB statistic	p-value
1	23.756	<0.001
2	27.578	<0.001
3	30.276	<0.001
4	31.304	<0.001
5	37.516	<0.001
6	40.229	<0.001
7	41.176	<0.001
8	42.372	<0.001
9	45.268	<0.001
10	46.293	<0.001

Lag	LB statistic	p-value
11	50.069	<0.001
12	61.937	<0.001
13	61.997	<0.001
14	62.071	<0.001
15	62.712	<0.001
16	64.484	<0.001
17	64.561	<0.001
18	64.688	<0.001
19	66.311	<0.001
20	66.435	<0.001

Table A4a. Ljung–Box Tests for Squared Returns — SENSEX

Lag	LB statistic	p-value
1	139.625	<0.001
2	399.846	<0.001
3	645.339	<0.001
4	841.991	<0.001
5	1044.067	<0.001
6	1203.460	<0.001
7	1502.177	<0.001
8	1597.714	<0.001
9	1890.520	<0.001
10	2116.010	<0.001
11	2194.271	<0.001
12	2280.355	<0.001
13	2348.454	<0.001
14	2438.587	<0.001
15	2517.343	<0.001
16	2615.935	<0.001
17	2673.914	<0.001
18	2729.600	<0.001
19	2766.743	<0.001
20	2815.747	<0.001

Table A4b. Ljung–Box Tests for Squared Returns — Nikkei 225

Lag	LB statistic	p-value
1	580.631	<0.001
2	1085.012	<0.001
3	1471.674	<0.001
4	1628.918	<0.001
5	1821.273	<0.001
6	2038.180	<0.001
7	2155.975	<0.001
8	2364.453	<0.001
9	2566.489	<0.001
10	2709.557	<0.001
11	2854.968	<0.001
12	3042.602	<0.001
13	3123.658	<0.001
14	3206.431	<0.001
15	3321.591	<0.001
16	3390.147	<0.001
17	3450.609	<0.001
18	3491.388	<0.001
19	3528.937	<0.001
20	3559.242	<0.001

Table A4c. Ljung–Box Tests for Squared Returns — Hang Seng

Lag	LB statistic	p-value
1	462.256	<0.001
2	895.232	<0.001
3	1222.594	<0.001
4	1403.045	<0.001
5	1528.249	<0.001
6	1622.388	<0.001
7	1733.417	<0.001
8	1919.786	<0.001
9	2018.232	<0.001
10	2227.429	<0.001
11	2436.609	<0.001
12	2530.431	<0.001
13	2763.470	<0.001
14	2941.257	<0.001
15	2987.730	<0.001
16	3070.542	<0.001
17	3111.858	<0.001
18	3166.301	<0.001
19	3220.274	<0.001
20	3236.305	<0.001

Table A4d. Ljung–Box Tests for Squared Returns — KOSPI

Lag	LB statistic	p-value
1	400.304	<0.001
2	647.746	<0.001
3	1086.573	<0.001
4	1353.878	<0.001
5	1558.390	<0.001
6	1724.482	<0.001
7	1855.573	<0.001
8	1984.603	<0.001
9	2106.715	<0.001
10	2472.906	<0.001
11	2660.610	<0.001
12	2803.986	<0.001
13	3050.144	<0.001
14	3208.002	<0.001
15	3304.936	<0.001
16	3438.277	<0.001
17	3616.494	<0.001
18	3729.273	<0.001
19	3838.591	<0.001
20	4000.786	<0.001

Table A4e. Ljung–Box Tests for Squared Returns — ASX 200

Lag	LB statistic	p-value
1	586.451	<0.001
2	1497.747	<0.001
3	2199.274	<0.001
4	2683.690	<0.001
5	3378.575	<0.001
6	3748.905	<0.001
7	4180.223	<0.001
8	4617.012	<0.001
9	4927.800	<0.001
10	5294.548	<0.001

Lag	LB statistic	p-value
11	5484.433	<0.001
12	5706.025	<0.001
13	5877.220	<0.001
14	6037.655	<0.001
15	6247.930	<0.001
16	6383.776	<0.001
17	6459.415	<0.001
18	6566.895	<0.001
19	6614.027	<0.001
20	6703.185	<0.001

Squared-return autocorrelation is highly significant across all 20 examined lags in every market, providing strong evidence of volatility clustering.

Appendix B. Structural-Break Diagnostics

This appendix provides the detailed retrospective evidence supporting the structural-break results reported in Section 6.

B.1 Break-Count Selection

Table B1. Cross-Market Break-Selection Summary

Market	LWZ breaks	BIC breaks	Maximum searched	Minimum segment	BIC boundary	LWZ boundary
S&P BSE SENSEX	1	10	12	126	No	No
Nikkei 225	0	6	12	126	No	No
Hang Seng Index	2	6	12	126	No	No
KOSPI Composite Index	3	6	12	126	No	No
S&P/ASX 200	1	7	12	126	No	No

Table B2a. BIC/LWZ Break-Selection Path — SENSEX

Breaks	Segments	-2 log L	Parameters	BIC	LWZ	BIC selected	LWZ selected
0	1	22507.628	2	22524.621	22561.098	No	No
1	2	22248.344	5	22290.827	22382.020	No	Yes
2	3	22187.655	8	22255.628	22401.537	No	No
3	4	22116.616	11	22210.079	22410.704	No	No
4	5	22090.946	14	22209.898	22465.239	No	No
5	6	22044.389	17	22188.831	22498.888	No	No
6	7	22010.951	20	22180.883	22545.656	No	No
7	8	21985.281	23	22180.702	22600.191	No	No
8	9	21960.585	26	22181.497	22655.702	No	No
9	10	21933.635	29	22180.036	22708.957	No	No
10	11	21907.965	32	22179.855	22763.493	Yes	No
11	12	21892.917	35	22190.298	22828.651	No	No
12	13	21871.616	38	22194.486	22887.555	No	No

Table B2b. BIC/LWZ Break-Selection Path — Nikkei 225

Breaks	Segments	-2 log L	Parameters	BIC	LWZ	BIC selected	LWZ selected
0	1	22595.955	2	22612.938	22649.361	No	Yes
1	2	22553.179	5	22595.637	22686.693	No	No
2	3	22469.484	8	22537.418	22683.107	No	No
3	4	22435.791	11	22529.199	22729.521	No	No
4	5	22381.331	14	22500.215	22755.170	No	No
5	6	22344.015	17	22488.374	22797.963	No	No
6	7	22313.260	20	22483.093	22847.316	Yes	No
7	8	22296.196	23	22491.504	22910.360	No	No
8	9	22265.440	26	22486.224	22959.713	No	No
9	10	22248.969	29	22495.227	23023.350	No	No
10	11	22230.842	32	22502.576	23085.332	No	No
11	12	22218.459	35	22515.668	23153.057	No	No
12	13	22200.333	38	22523.016	23215.039	No	No

Table B2c. BIC/LWZ Break-Selection Path — Hang Seng

Breaks	Segments	-2 log L	Parameters	BIC	LWZ	BIC selected	LWZ selected
0	1	22752.211	2	22769.209	22805.709	No	No
1	2	22638.606	5	22681.099	22772.350	No	No
2	3	22539.978	8	22607.967	22753.968	No	Yes
3	4	22476.904	11	22570.389	22771.140	No	No
4	5	22455.449	14	22574.430	22829.931	No	No
5	6	22428.800	17	22573.277	22883.529	No	No
6	7	22397.192	20	22567.165	22932.167	Yes	No
7	8	22375.708	23	22571.176	22990.929	No	No
8	9	22357.686	26	22578.650	23053.153	No	No
9	10	22336.202	29	22582.662	23111.915	No	No
10	11	22315.320	32	22587.276	23171.279	No	No
11	12	22300.319	35	22597.771	23236.524	No	No
12	13	22285.234	38	22608.182	23301.686	No	No

Table B2d. BIC/LWZ Break-Selection Path — KOSPI

Breaks	Segments	-2 log L	Parameters	BIC	LWZ	BIC selected	LWZ selected
0	1	23033.768	2	23050.765	23087.268	No	No
1	2	22872.134	5	22914.628	23005.885	No	No
2	3	22750.715	8	22818.706	22964.716	No	No
3	4	22625.780	11	22719.267	22920.031	No	Yes
4	5	22595.432	14	22714.416	22969.933	No	No
5	6	22541.928	17	22686.408	22996.679	No	No
6	7	22515.733	20	22685.709	23050.734	Yes	No
7	8	22491.075	23	22686.548	23106.327	No	No
8	9	22471.297	26	22692.267	23166.799	No	No
9	10	22450.753	29	22697.219	23226.505	No	No
10	11	22436.071	32	22708.033	23292.073	No	No
11	12	22422.161	35	22719.620	23358.413	No	No
12	13	22409.266	38	22732.222	23425.769	No	No

Table B2e. BIC/LWZ Break-Selection Path — ASX 200

Breaks	Segments	-2 log L	Parameters	BIC	LWZ	BIC selected	LWZ selected
0	1	23068.589	2	23085.638	23122.425	No	No
1	2	22926.242	5	22968.863	23060.832	No	Yes
2	3	22875.728	8	22943.922	23091.071	No	No
3	4	22834.651	11	22928.417	23130.747	No	No
4	5	22793.048	14	22912.386	23169.898	No	No
5	6	22739.642	17	22884.553	23197.246	No	No
6	7	22715.373	20	22885.856	23253.730	No	No
7	8	22682.549	23	22878.605	23301.660	Yes	No
8	9	22662.502	26	22884.130	23362.366	No	No
9	10	22636.949	29	22884.150	23417.567	No	No
10	11	22616.902	32	22889.675	23478.273	No	No
11	12	22602.192	35	22900.538	23544.317	No	No
12	13	22581.573	38	22905.491	23604.451	No	No

B.2 Segment Statistics

Table B3. Segment Statistics Under the Primary LWZ Specifications

Market	Segment	Start	End	N	Mean return	Daily volatility	Annualized volatility
SENSEX	1	18 Sep 2006	09 Apr 2012	1365	0.0264%	1.8457%	29.30%
SENSEX	2	10 Apr 2012	26 Aug 2026	3533	0.0426%	1.0143%	16.10%
Nikkei 225	1	19 Sep 2006	26 Aug 2026	4874	0.0293%	1.4965%	23.76%
Hang Seng	1	18 Sep 2006	05 Jul 2007	195	0.1310%	1.0257%	16.28%
Hang Seng	2	06 Jul 2007	22 Dec 2009	610	-0.0088%	2.6074%	41.39%
Hang Seng	3	23 Dec 2009	26 Aug 2026	4103	0.0048%	1.3159%	20.89%
KOSPI	1	18 Sep 2006	16 Aug 2012	1470	0.0247%	1.6064%	25.50%
KOSPI	2	17 Aug 2012	23 Dec 2019	1805	0.0066%	0.7620%	12.10%
KOSPI	3	24 Dec 2019	22 Jan 2026	1490	0.0543%	1.2913%	20.50%
KOSPI	4	23 Jan 2026	26 Aug 2026	144	0.2210%	4.2499%	67.46%

Market	Segment	Start	End	N	Mean return	Daily volatility	Annualized volatility
ASX 200	1	18 Sep 2006	05 Jan 2012	1335	-0.0146%	1.3948%	22.14%
ASX 200	2	06 Jan 2012	26 Aug 2026	3700	0.0214%	0.9078%	14.41%

B.3 SENSEX Minimum-Segment Sensitivity

Table B4. SENSEX Structural-Break Sensitivity

Minimum segment	LWZ breaks	BIC breaks	LWZ date
63	1	10	2012-04-09
126	1	10	2012-04-09
252	1	5	2012-04-09

The primary LWZ break remains unchanged across all three minimum-segment restrictions.

Appendix C. BOCPD Diagnostics

BOCPD is interpreted as a sequential adaptation mechanism rather than a retrospective classifier of permanent economic breaks.

C.1 Complete Baseline Alarm Lists

Table C1. SENSEX BOCPD Alarms

Date	Posterior change probability	Run-length mode	Posterior mean run length
05 Feb 2008	0.3031	0	142.85
24 Nov 2008	0.9075	0	1.96
05 May 2009	0.3378	0	65.73
23 Jun 2009	0.5147	0	59.19
01 Feb 2010	0.2424	0	44.04
17 Jul 2012	0.2323	0	93.93
28 Jan 2013	0.6580	0	5.09
31 Jul 2013	0.2494	0	99.05
11 Sep 2013	0.9435	0	1.26
19 Mar 2014	0.5781	0	12.69
03 Dec 2014	0.2828	0	217.94
27 Mar 2015	0.3793	0	266.77
03 Jun 2016	0.8624	0	5.40
07 Sep 2017	0.3254	0	187.64
02 Jan 2018	0.4888	0	169.18
01 Nov 2018	0.8839	0	1.88
31 Dec 2020	0.2136	0	100.56
01 Jun 2021	0.4097	0	115.88
09 May 2023	0.6007	0	91.93
22 Aug 2023	0.3207	0	144.05
10 Dec 2024	0.2448	0	131.27
05 Sep 2025	0.2341	0	167.00
16 Jul 2026	0.7149	0	4.02
21 Aug 2026	0.3258	88	52.52

Table C2. Nikkei 225 BOCPD Alarms

Date	Posterior change probability	Run-length mode	Posterior mean run length
08 Feb 2007	0.7566	0	16.62
20 Dec 2007	0.8500	0	5.98
12 Feb 2008	0.5123	0	31.46
24 Mar 2008	0.2014	28	33.20
17 Jun 2008	0.5282	0	25.63
05 Dec 2008	0.4493	0	14.65
26 Feb 2009	0.2379	0	45.31
28 Jul 2009	0.5285	0	55.37
11 Nov 2009	0.5067	0	52.68
15 Mar 2010	0.4182	0	46.82
30 Sep 2011	0.2011	0	372.71
23 Jan 2012	0.9311	0	21.45
29 Feb 2012	0.2569	0	176.21
15 Jun 2012	0.4999	0	164.77
06 Sep 2012	0.2151	0	381.20

Date	Posterior change probability	Run-length mode	Posterior mean run length
09 Apr 2013	0.9714	0	0.04
11 Sep 2013	0.6577	0	27.37
18 Nov 2013	0.3192	0	68.78
10 Apr 2014	0.8715	0	12.31
02 Oct 2015	0.8459	0	4.34
05 Aug 2016	0.9360	0	4.43
03 Feb 2017	0.3969	0	62.20
26 Jun 2018	0.2063	58	44.00
03 Apr 2020	0.6127	0	10.22
01 Sep 2020	0.3244	0	37.22
16 Aug 2022	0.7267	0	55.24
30 Dec 2022	0.9047	0	8.72
15 Jun 2023	0.2283	0	191.08
18 Oct 2023	0.5378	0	181.90
05 Jul 2024	0.7403	0	106.95
16 May 2025	0.5314	0	216.30
02 Dec 2025	0.8898	0	10.47
27 May 2026	0.4816	0	106.53

Table C3. Hang Seng BOC PD Alarms

Date	Posterior change probability	Run-length mode	Posterior mean run length
23 Mar 2007	0.3236	0	56.43
05 Mar 2008	0.4344	0	64.53
22 Jul 2008	0.2277	0	150.89
17 Nov 2008	0.4374	0	15.55
11 Jun 2009	0.4179	0	88.35
21 Jul 2009	0.9718	0	0.42
09 Oct 2009	0.2412	0	139.45
31 May 2010	0.8608	0	8.14
29 Jun 2011	0.4889	0	30.31
10 Oct 2011	0.4214	0	38.41
08 Nov 2011	0.9504	0	0.44
26 Mar 2012	0.9235	0	0.52
07 Jan 2013	0.3674	0	76.13
18 Jun 2013	0.9247	0	4.39
06 Nov 2013	0.5085	0	133.79
18 Aug 2014	0.3166	0	45.29
05 Jan 2015	0.9550	0	0.10
10 Feb 2015	0.2151	25	13.66
05 Nov 2015	0.3653	0	102.38
07 Mar 2016	0.2248	0	46.37
23 Aug 2016	0.7961	0	5.66
23 Nov 2016	0.3912	0	24.50
14 Mar 2017	0.4745	0	85.67
05 Sep 2017	0.4997	0	81.65
13 Mar 2018	0.8780	0	4.22
10 Jul 2018	0.3859	0	58.31
07 Sep 2018	0.9198	0	4.44
23 Jan 2019	0.3775	0	134.79
05 Mar 2019	0.2483	0	84.10
23 Apr 2019	0.3701	2	3.88
29 Mar 2021	0.4599	0	85.28
14 Jun 2022	0.9461	0	2.44
07 Aug 2023	0.6581	0	110.29
22 Nov 2023	0.5291	0	66.30
04 Jan 2024	0.3799	0	61.90
05 Apr 2024	0.2037	0	217.96
24 Jun 2024	0.2690	0	179.94
06 Aug 2025	0.2071	0	120.21

Date	Posterior change probability	Run-length mode	Posterior mean run length
19 Sep 2025	0.5005	0	2.32
20 Nov 2025	0.2698	0	173.63
21 Jul 2026	0.5779	0	26.51

Table C4. KOSPI BOCPD Alarms

Date	Posterior change probability	Run-length mode	Posterior mean run length
26 Feb 2008	0.9684	0	0.08
26 Jun 2008	0.6679	0	16.56
29 Aug 2008	0.5714	0	44.69
14 Nov 2008	0.8799	0	1.55
02 Jul 2009	0.2842	0	54.88
20 Nov 2009	0.6808	0	38.63
06 Aug 2010	0.6901	0	14.00
01 Sep 2011	0.6100	0	27.64
01 Nov 2011	0.2021	0	44.66
02 Dec 2011	0.8784	0	6.32
19 Jun 2012	0.5748	0	50.72
20 Jul 2012	0.5368	0	51.25
10 Dec 2012	0.7216	0	40.31
28 Mar 2013	0.3059	0	183.08
04 Jun 2013	0.3009	0	136.66
16 Apr 2014	0.2233	2	34.66
19 Nov 2014	0.2969	0	172.23
19 Oct 2015	0.7892	0	16.10
22 Feb 2016	0.2092	0	255.91
16 Oct 2018	0.9013	0	3.96
17 Sep 2019	0.8360	0	22.46
30 Mar 2020	0.7668	0	3.99
11 Sep 2020	0.3123	0	71.16
19 Apr 2021	0.4876	0	6.83
02 Jul 2021	0.2325	0	47.20
25 Mar 2022	0.2418	0	107.91
08 Jun 2022	0.2013	0	80.90
24 May 2023	0.6073	0	27.40
28 May 2024	0.2473	0	135.00
06 Aug 2025	0.6061	0	300.52
06 Mar 2026	0.2330	0	54.43
24 Apr 2026	0.9009	0	2.24

Table C5. ASX 200 BOCPD Alarms

Date	Posterior change probability	Run-length mode	Posterior mean run length
26 Sep 2007	0.3916	0	39.26
03 Dec 2007	0.5945	0	63.95
05 Mar 2008	0.2731	0	20.22
15 Aug 2008	0.2387	0	124.52
21 Jul 2009	0.2961	0	60.63
10 Mar 2010	0.8656	0	2.69
15 Jun 2010	0.2681	0	173.49
06 Aug 2010	0.2863	0	224.07
28 Apr 2011	0.6954	0	50.43
30 Sep 2011	0.6735	0	64.92
31 Aug 2012	0.2676	0	98.67
23 Nov 2012	0.9171	0	6.35
03 May 2013	0.4138	0	168.62
15 Jul 2014	0.8364	0	49.17
18 Mar 2015	0.7115	0	123.68
09 Jul 2015	0.2004	0	368.10
22 Mar 2016	0.8932	0	1.88
26 Sep 2016	0.8630	0	14.97

Date	Posterior change probability	Run-length mode	Posterior mean run length
07 Jun 2017	0.3059	0	170.90
07 Sep 2017	0.3200	0	113.19
27 Dec 2017	0.2621	0	71.42
06 Apr 2018	0.6149	0	81.02
13 Jul 2018	0.5823	0	111.51
22 Apr 2020	0.9853	0	0.03
29 Sep 2020	0.7566	0	7.86
15 Jan 2021	0.7865	0	14.37
18 Feb 2021	0.3368	0	120.70
26 Jul 2021	0.5609	0	43.04
24 Apr 2024	0.3551	0	213.66
20 Jun 2024	0.4476	0	72.72
28 Aug 2024	0.5977	0	35.36
26 May 2025	0.8079	0	12.75

C.2 Hazard Sensitivity

Table C6. SENSEX BOCPD Hazard Sensitivity

Expected run length λ	Alarms	Mean change probability	Maximum probability
125	35	0.0121	0.9602
250	24	0.0070	0.9435
500	13	0.0042	0.9162

Alarm frequency falls monotonically as the assumed expected run length increases.

Appendix D. Markov-Switching Diagnostics

This appendix reports the complete Markov-model evidence underlying the selected market-specific regime structures.

D.1 Full Pre-OOS Regime Selection

Table D1a. Pre-OOS Regime Selection — SENSEX

K	Log likelihood	AIC	BIC	Parameters	EM iterations	Converged	Selection N	Selection end	First OOS target	Selected
2	-6031.789	12081.578	12137.990	9	20	Yes	3898	02 Aug 2022	03 Aug 2022	No
3	-5914.721	11863.443	11969.998	17	45	Yes	3898	02 Aug 2022	03 Aug 2022	Yes
4	-5902.787	11859.573	12028.808	27	140	Yes	3898	02 Aug 2022	03 Aug 2022	No

Table D1b. Pre-OOS Regime Selection — Nikkei 225

K	Log likelihood	AIC	BIC	Parameters	EM iterations	Converged	Selection N	Selection end	First OOS target	Selected
2	-6516.192	13050.383	13106.739	9	30	Yes	3874	25 Jul 2022	26 Jul 2022	No
3	-6398.431	12830.863	12937.313	17	49	Yes	3874	25 Jul 2022	26 Jul 2022	Yes
4	-6372.708	12799.416	12968.484	27	175	Yes	3874	25 Jul 2022	26 Jul 2022	No

Table D1c. Pre-OOS Regime Selection — Hang Seng

K	Log likelihood	AIC	BIC	Parameters	EM iterations	Converged	Selection N	Selection end	First OOS target	Selected
2	-6561.457	13140.913	13197.348	9	22	Yes	3908	29 Jul 2022	01 Aug 2022	No
3	-6475.458	12984.915	13091.514	17	91	Yes	3908	29 Jul 2022	01 Aug 2022	No
4	-6432.172	12918.343	13087.648	27	273	Yes	3908	29 Jul 2022	01 Aug 2022	Yes

Table D1d. Pre-OOS Regime Selection — KOSPI

K	Log likelihood	AIC	BIC	Parameters	EM iterations	Converged	Selection N	Selection end	First OOS target	Selected
2	-5691.362	11400.725	11457.162	9	24	Yes	3909	21 Jul 2022	22 Jul 2022	No
3	-5556.240	11146.480	11253.083	17	59	Yes	3909	21 Jul 2022	22 Jul 2022	Yes
4	-5520.316	11094.633	11263.944	27	300	No	3909	21 Jul 2022	22 Jul 2022	No

Table D1e. Pre-OOS Regime Selection — ASX 200

K	Log likelihood	AIC	BIC	Parameters	EM iterations	Converged	Selection N	Selection end	First OOS target	Selected
2	-5518.730	11055.461	11112.183	9	25	Yes	4035	12 Sep 2022	13 Sep 2022	No
3	-5375.398	10784.795	10891.938	17	69	Yes	4035	12 Sep 2022	13 Sep 2022	No
4	-5332.296	10718.592	10888.760	27	236	Yes	4035	12 Sep 2022	13 Sep 2022	Yes

D.2 Full-Sample Regime Parameters**Table D2. Market-Specific Regime Parameters**

Market	State	Daily intercept	AR(1)	Daily vol.	Annualized vol.	Persistence	Duration	Occupancy
SENSEX	0	0.0776%	0.0950	0.6520%	10.35%	98.16%	54.3	47.97%
SENSEX	1	0.0182%	0.0557	1.1724%	18.61%	97.13%	34.8	41.48%
SENSEX	2	-0.1061%	0.0162	2.9334%	46.57%	96.97%	33.1	10.55%
Nikkei 225	0	0.1215%	-0.0022	0.8426%	13.38%	97.60%	41.6	46.15%
Nikkei 225	1	0.0026%	-0.0422	1.4569%	23.13%	96.49%	28.5	47.03%
Nikkei 225	2	-0.4483%	-0.0916	3.6014%	57.17%	92.13%	12.7	6.82%
Hang Seng	0	0.1664%	-0.1168	0.4371%	6.94%	10.29%	1.1	14.65%
Hang Seng	1	0.0095%	0.1315	1.0757%	17.08%	59.00%	2.4	32.98%
Hang Seng	2	-0.0249%	-0.0471	1.4330%	22.75%	98.38%	61.6	42.87%
Hang Seng	3	-0.1152%	-0.0075	3.2925%	52.27%	96.01%	25.0	9.50%
KOSPI	0	0.0692%	-0.0039	0.6817%	10.82%	97.88%	47.2	49.98%
KOSPI	1	0.0347%	0.0175	1.2766%	20.27%	96.65%	29.9	42.35%
KOSPI	2	-0.2413%	-0.0698	3.7624%	59.73%	95.24%	21.0	7.66%
ASX 200	0	0.0875%	-0.0086	0.5643%	8.96%	96.72%	30.4	48.84%
ASX 200	1	-0.0364%	-0.0417	1.0132%	16.08%	94.93%	19.7	36.53%
ASX 200	2	-0.0534%	-0.0855	1.5745%	24.99%	97.63%	42.2	12.66%
ASX 200	3	-0.7045%	-0.1757	3.5860%	56.93%	96.50%	28.6	1.98%

D.3 Transition Matrices and Probability Plots**Table D3a. Transition Matrix — SENSEX**

Current state	Next_0	Next_1	Next_2
Current_0	98.159%	1.783%	0.058%
Current_1	2.168%	97.130%	0.702%
Current_2	0.004%	3.021%	96.975%

Table D3b. Transition Matrix — Nikkei 225

Current state	Next_0	Next_1	Next_2
Current_0	97.595%	2.404%	0.001%
Current_1	2.366%	96.491%	1.143%
Current_2	0.001%	7.867%	92.132%

Table D3c. Transition Matrix — Hang Seng

Current state	Next_0	Next_1	Next_2	Next_3
Current_0	10.290%	89.708%	0.001%	0.001%
Current_1	39.722%	59.002%	0.985%	0.290%
Current_2	0.008%	0.954%	98.377%	0.661%
Current_3	0.001%	0.005%	3.987%	96.006%

Table D3d. Transition Matrix — KOSPI

Current state	Next_0	Next_1	Next_2
Current_0	97.879%	2.120%	0.001%
Current_1	2.457%	96.654%	0.889%
Current_2	0.001%	4.759%	95.240%

Table D3e. Transition Matrix — ASX 200

Current state	Next_0	Next_1	Next_2	Next_3
Current_0	96.716%	3.244%	0.001%	0.039%
Current_1	4.435%	94.930%	0.633%	0.002%
Current_2	0.008%	1.973%	97.631%	0.388%
Current_3	0.001%	0.001%	3.498%	96.500%

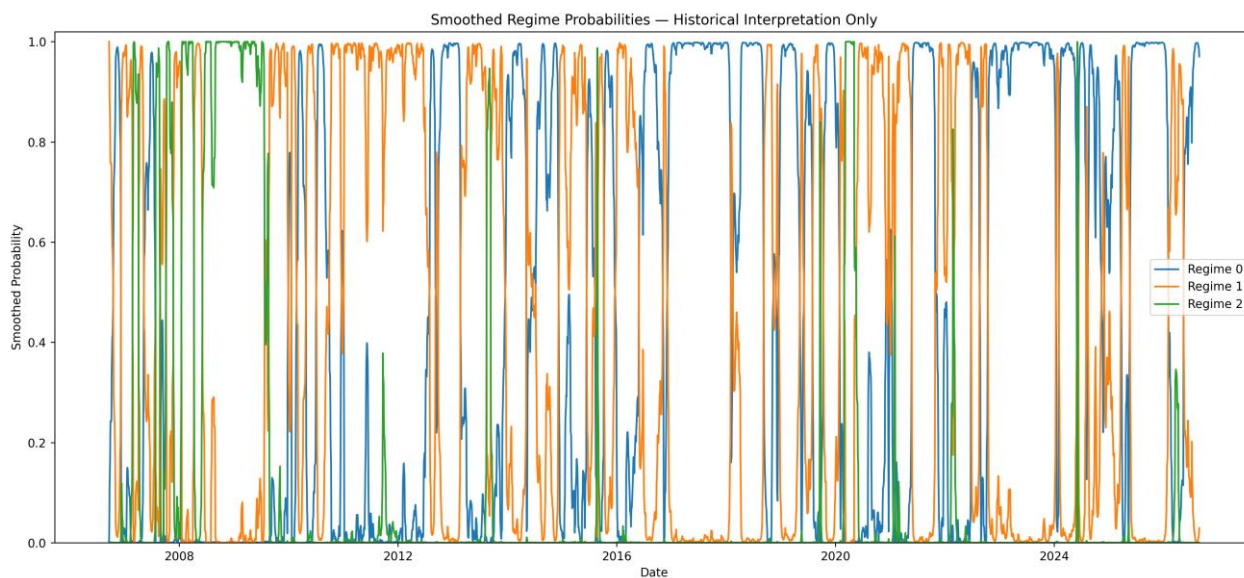


Fig. D1. Smoothed Regime Probabilities — SENSEX

Note. Full-sample smoothed probabilities are descriptive only.

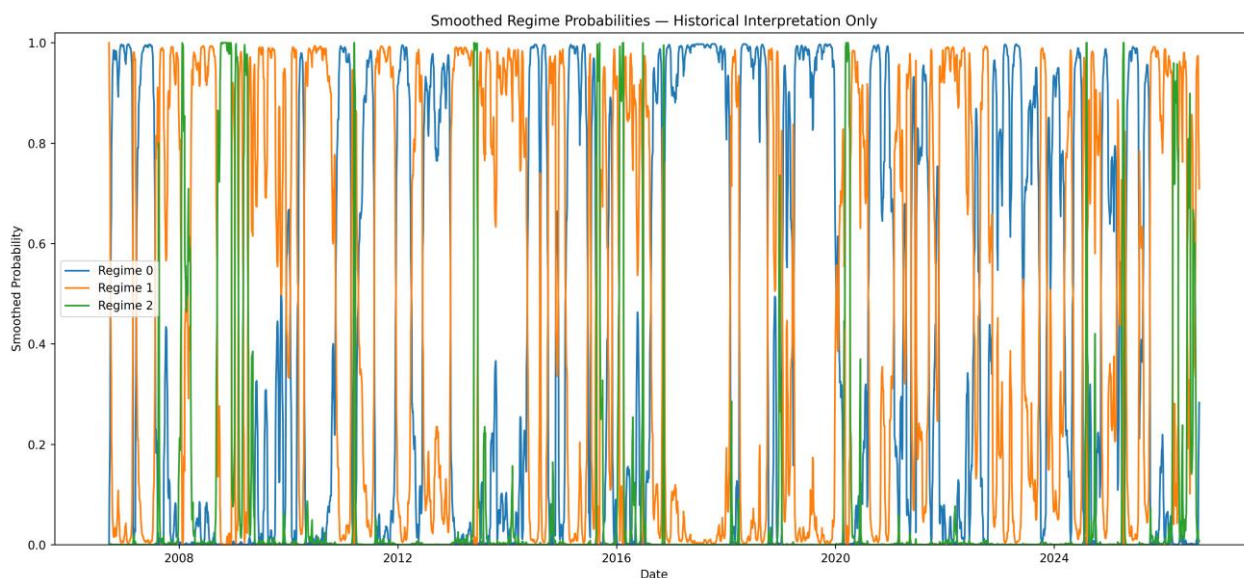


Fig. D2. Smoothed Regime Probabilities — Nikkei 225

Note. Full-sample smoothed probabilities are descriptive only.

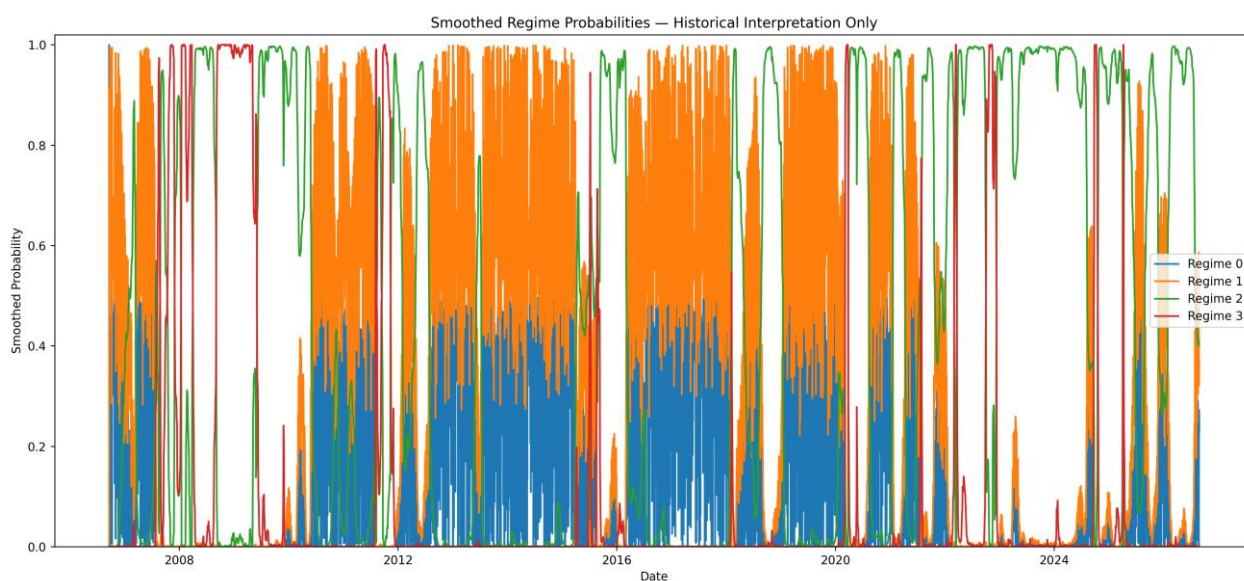


Fig. D3. Smoothed Regime Probabilities — Hang Seng

Note. Full-sample smoothed probabilities are descriptive only.

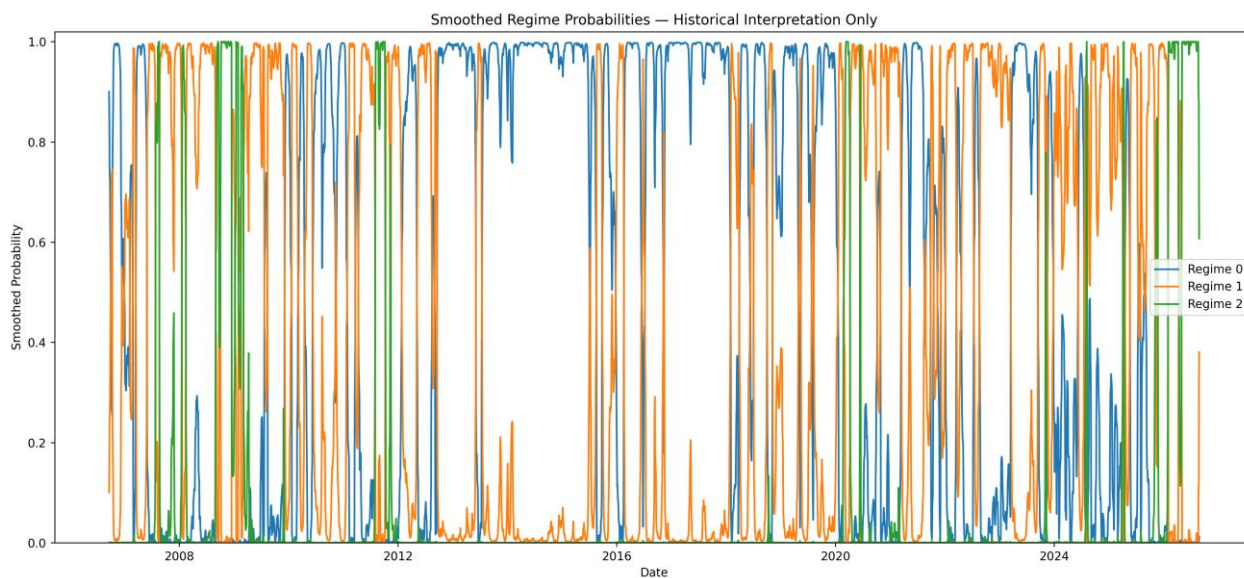


Fig. D4. Smoothed Regime Probabilities — KOSPI

Note. Full-sample smoothed probabilities are descriptive only.

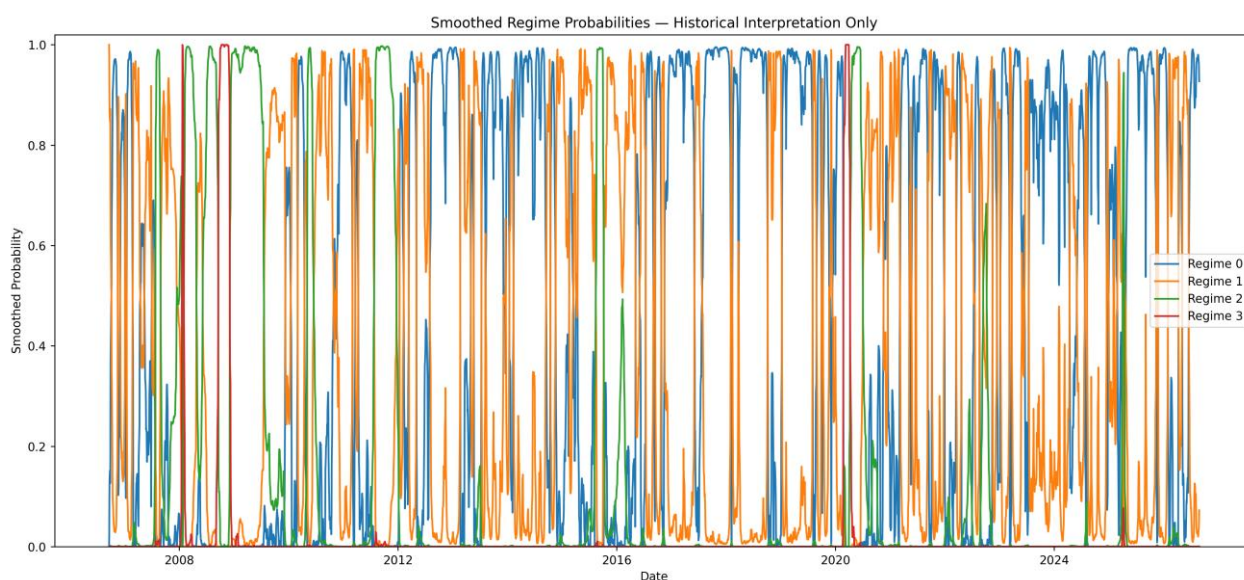


Fig. D5. Smoothed Regime Probabilities — ASX 200

Note. Full-sample smoothed probabilities are descriptive only.

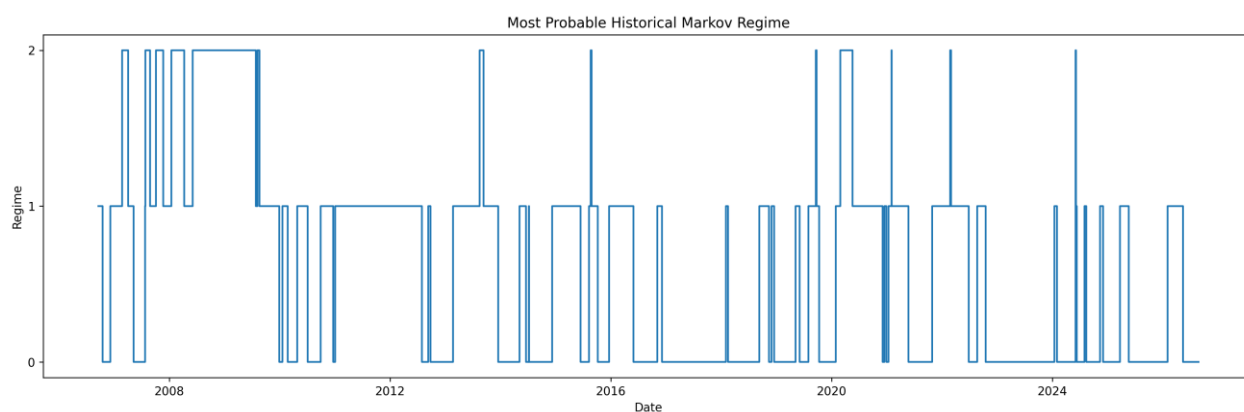


Fig. D6. Most-Probable Historical Regime Timeline — SENSEX

Note. Full-sample smoothed classifications are descriptive only.

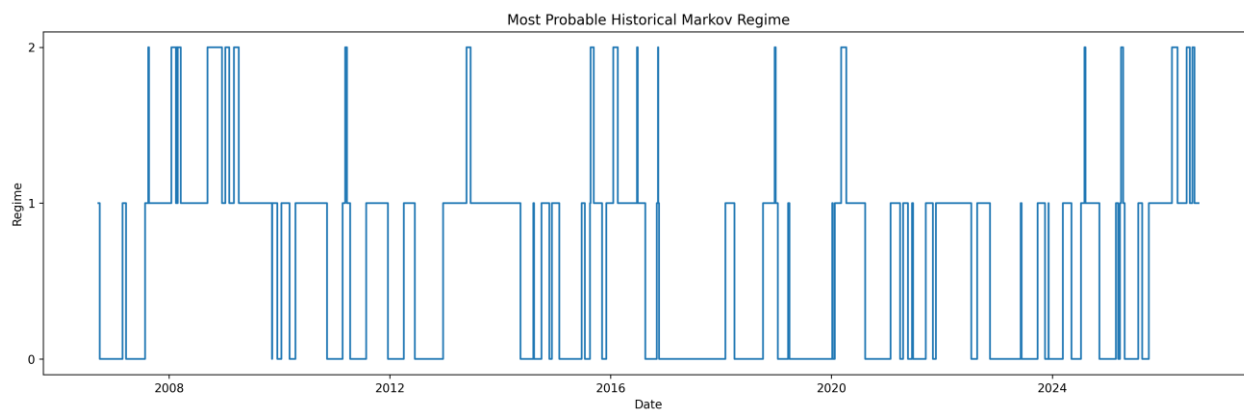


Fig. D7. Most-Probable Historical Regime Timeline — Nikkei 225

Note. Full-sample smoothed classifications are descriptive only.

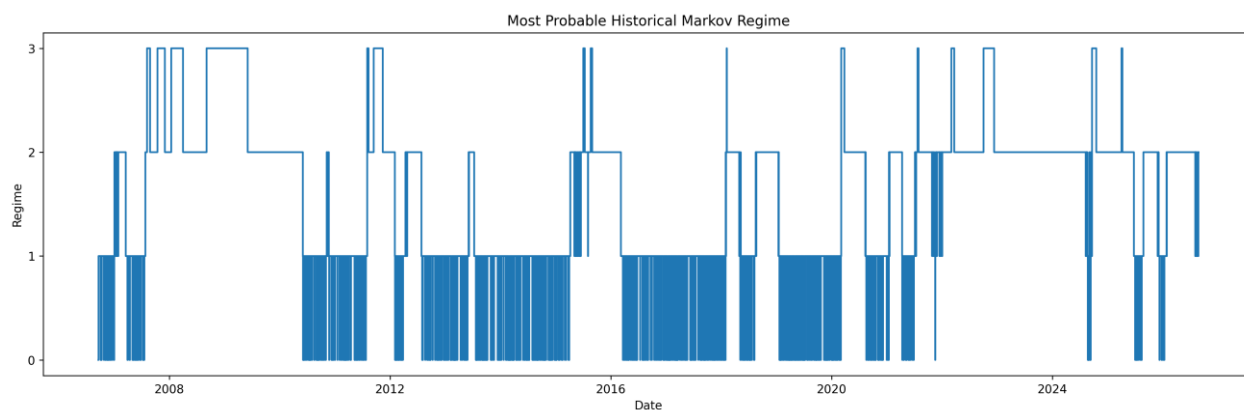


Fig. D8. Most-Probable Historical Regime Timeline — Hang Seng

Note. Full-sample smoothed classifications are descriptive only.

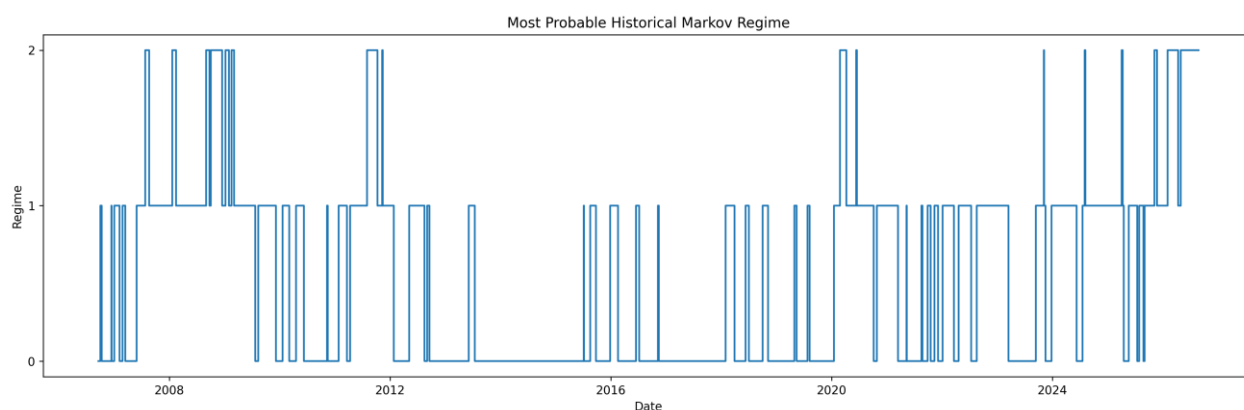


Fig. D9. Most-Probable Historical Regime Timeline — KOSPI

Note. Full-sample smoothed classifications are descriptive only.

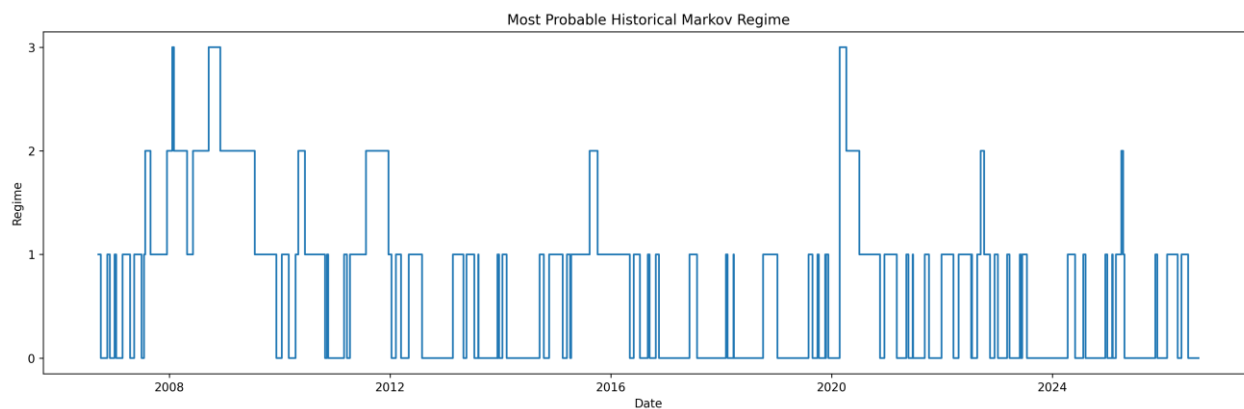


Fig. D10. Most-Probable Historical Regime Timeline — ASX 200

Note. Full-sample smoothed classifications are descriptive only.

Appendix E. Forecast Evaluation

This appendix contains the complete forecasting output underlying the selective comparisons reported in Section 6.

E.1 Full Point-Forecast Metrics

Table E1. Complete Point-Forecast Metrics Across Markets

Market	Model	N	RMSE	MAE	Bias	Direction accuracy
S&P BSE SENSEX	hybrid	1000	0.7957%	0.5806%	0.0099%	55.2%
S&P BSE SENSEX	arma	1000	0.8031%	0.5860%	0.0229%	53.0%
S&P BSE SENSEX	ar1	1000	0.8037%	0.5866%	0.0234%	52.6%
S&P BSE SENSEX	ms_noreset	1000	0.8046%	0.5868%	0.0383%	52.1%
S&P BSE SENSEX	zero	1000	0.8050%	0.5874%	-0.0287%	0.0%
S&P BSE SENSEX	hist_mean	1000	0.8050%	0.5861%	0.0223%	53.1%
S&P BSE SENSEX	garch_mean	1000	0.8067%	0.5864%	0.0610%	53.1%
Nikkei 225	garch_mean	1000	1.4987%	1.0275%	-0.0075%	54.3%
Nikkei 225	hist_mean	1000	1.4997%	1.0292%	-0.0460%	54.3%
Nikkei 225	arma	1000	1.5002%	1.0300%	-0.0475%	50.9%
Nikkei 225	zero	1000	1.5012%	1.0313%	-0.0872%	0.0%
Nikkei 225	ar1	1000	1.5020%	1.0291%	-0.0486%	53.6%
Nikkei 225	ms_noreset	1000	1.5138%	1.0339%	-0.0582%	50.6%
Nikkei 225	hybrid	1000	1.5679%	1.0675%	-0.1033%	51.4%
Hang Seng Index	zero	1000	1.5540%	1.1369%	-0.0241%	0.0%
Hang Seng Index	hist_mean	1000	1.5549%	1.1371%	-0.0405%	49.7%
Hang Seng Index	arma	1000	1.5553%	1.1373%	-0.0423%	49.7%
Hang Seng Index	garch_mean	1000	1.5554%	1.1389%	-0.0204%	47.6%
Hang Seng Index	ar1	1000	1.5557%	1.1379%	-0.0402%	50.9%
Hang Seng Index	ms_noreset	1000	1.5600%	1.1378%	-0.0526%	51.6%
Hang Seng Index	hybrid	1000	1.5627%	1.1401%	-0.0259%	52.0%
KOSPI Composite Index	garch_mean	1000	1.9534%	1.1953%	-0.0436%	56.0%
KOSPI Composite Index	hist_mean	1000	1.9547%	1.1980%	-0.0812%	54.7%
KOSPI Composite Index	ar1	1000	1.9550%	1.1982%	-0.0834%	54.0%
KOSPI Composite Index	ms_noreset	1000	1.9556%	1.2039%	-0.1035%	50.3%
KOSPI Composite Index	arma	1000	1.9557%	1.1999%	-0.0853%	53.4%
KOSPI Composite Index	zero	1000	1.9563%	1.2025%	-0.1039%	0.0%
KOSPI Composite Index	hybrid	1000	1.9638%	1.2003%	-0.0771%	51.7%
S&P/ASX 200	hist_mean	1000	0.7779%	0.5737%	-0.0087%	52.6%
S&P/ASX 200	zero	1000	0.7780%	0.5740%	-0.0270%	0.1%
S&P/ASX 200	garch_mean	1000	0.7785%	0.5730%	0.0382%	52.6%
S&P/ASX 200	ar1	1000	0.7831%	0.5776%	-0.0103%	50.4%
S&P/ASX 200	arma	1000	0.7836%	0.5783%	-0.0098%	49.9%
S&P/ASX 200	hybrid	1000	0.7841%	0.5785%	0.0007%	50.2%
S&P/ASX 200	ms_noreset	1000	0.7866%	0.5797%	-0.0021%	50.4%

E.2 Point-Forecast Diebold–Mariano Tests**Table E2a. Point-Forecast DM Statistics — SENSEX**

Model	hybrid	ms_noreset	arma	garch_mean	ar1	hist_mean	zero
hybrid	0.000	-1.272	-0.921	-1.143	-0.996	-0.996	-1.004
ms_noreset	1.272	0.000	0.407	-0.517	0.245	-0.107	-0.091
arma	0.921	-0.407	0.000	-1.367	-1.461	-0.813	-0.647
garch_mean	1.143	0.517	1.367	0.000	1.130	1.311	0.594
ar1	0.996	-0.245	1.461	-1.130	0.000	-0.550	-0.433
hist_mean	0.996	0.107	0.813	-1.311	0.550	0.000	0.023
zero	1.004	0.091	0.647	-0.594	0.433	-0.023	0.000

Table E3a. Point-Forecast DM p-values — SENSEX

Model	hybrid	ms_noreset	arma	garch_mean	ar1	hist_mean	zero
hybrid	1.000	0.204	0.357	0.253	0.319	0.320	0.316
ms_noreset	0.204	1.000	0.684	0.605	0.807	0.915	0.927
arma	0.357	0.684	1.000	0.172	0.144	0.416	0.518
garch_mean	0.253	0.605	0.172	1.000	0.259	0.190	0.553
ar1	0.319	0.807	0.144	0.259	1.000	0.582	0.665
hist_mean	0.320	0.915	0.416	0.190	0.582	1.000	0.982
zero	0.316	0.927	0.518	0.553	0.665	0.982	1.000

Table E2b. Point-Forecast DM Statistics — Nikkei 225

Model	hybrid	ms_noreset	arma	garch_mean	ar1	hist_mean	zero
hybrid	0.000	1.563	2.621	2.562	2.368	2.529	2.470
ms_noreset	-1.563	0.000	0.682	0.839	0.823	0.812	0.737
arma	-2.621	-0.682	0.000	0.441	-0.271	0.143	-0.248
garch_mean	-2.562	-0.839	-0.441	0.000	-0.682	-0.817	-0.984
ar1	-2.368	-0.823	0.271	0.682	0.000	0.524	0.170
hist_mean	-2.529	-0.812	-0.143	0.817	-0.524	0.000	-1.029
zero	-2.470	-0.737	0.248	0.984	-0.170	1.029	0.000

Table E3b. Point-Forecast DM p-values — Nikkei 225

Model	hybrid	ms_noreset	arma	garch_mean	ar1	hist_mean	zero
hybrid	1.000	0.118	0.009	0.011	0.018	0.012	0.014
ms_noreset	0.118	1.000	0.495	0.402	0.411	0.417	0.461
arma	0.009	0.495	1.000	0.659	0.786	0.887	0.804
garch_mean	0.011	0.402	0.659	1.000	0.495	0.414	0.326
ar1	0.018	0.411	0.786	0.495	1.000	0.601	0.865
hist_mean	0.012	0.417	0.887	0.414	0.601	1.000	0.304
zero	0.014	0.461	0.804	0.326	0.865	0.304	1.000

Table E2c. Point-Forecast DM Statistics — Hang Seng

Model	hybrid	ms_noreset	arma	garch_mean	ar1	hist_mean	zero
hybrid	0.000	0.550	1.392	1.241	1.349	1.479	1.598
ms_noreset	-0.550	0.000	1.153	0.977	1.098	1.260	1.402
arma	-1.392	-1.153	0.000	-0.027	-0.555	0.747	1.407
garch_mean	-1.241	-0.977	0.027	0.000	-0.217	0.416	1.593
ar1	-1.349	-1.098	0.555	0.217	0.000	1.755	2.018
hist_mean	-1.479	-1.260	-0.747	-0.416	-1.755	0.000	1.334
zero	-1.598	-1.402	-1.407	-1.593	-2.018	-1.334	0.000

Table E3c. Point-Forecast DM p-values — Hang Seng

Model	hybrid	ms_noreset	arma	garch_mean	ar1	hist_mean	zero
hybrid	1.000	0.583	0.164	0.215	0.178	0.139	0.110
ms_noreset	0.583	1.000	0.249	0.329	0.273	0.208	0.161
arma	0.164	0.249	1.000	0.978	0.579	0.455	0.160
garch_mean	0.215	0.329	0.978	1.000	0.828	0.677	0.111
ar1	0.178	0.273	0.579	0.828	1.000	0.080	0.044
hist_mean	0.139	0.208	0.455	0.677	0.080	1.000	0.183
zero	0.110	0.161	0.160	0.111	0.044	0.183	1.000

Table E2d. Point-Forecast DM Statistics — KOSPI

Model	hybrid	ms_noreset	arma	garch_mean	ar1	hist_mean	zero
hybrid	0.000	0.651	0.480	0.561	0.564	0.488	0.389
ms_noreset	-0.651	0.000	-0.012	0.180	0.067	0.071	-0.059
arma	-0.480	0.012	0.000	0.488	0.148	0.206	-0.119
garch_mean	-0.561	-0.180	-0.488	0.000	-0.285	-1.449	-1.112
ar1	-0.564	-0.067	-0.148	0.285	0.000	0.052	-0.217
hist_mean	-0.488	-0.071	-0.206	1.449	-0.052	0.000	-0.759
zero	-0.389	0.059	0.119	1.112	0.217	0.759	0.000

Table E3d. Point-Forecast DM p-values — KOSPI

Model	hybrid	ms_noreset	arma	garch_mean	ar1	hist_mean	zero
hybrid	1.000	0.515	0.631	0.575	0.573	0.626	0.697
ms_noreset	0.515	1.000	0.991	0.857	0.947	0.943	0.953
arma	0.631	0.991	1.000	0.626	0.882	0.837	0.905
garch_mean	0.575	0.857	0.626	1.000	0.776	0.148	0.267
ar1	0.573	0.947	0.882	0.776	1.000	0.958	0.828
hist_mean	0.626	0.943	0.837	0.148	0.958	1.000	0.448
zero	0.697	0.953	0.905	0.267	0.828	0.448	1.000

Table E2e. Point-Forecast DM Statistics — ASX 200

Model	hybrid	ms_noreset	arma	garch_mean	ar1	hist_mean	zero
hybrid	0.000	-0.309	0.056	0.851	0.124	0.981	0.963
ms_noreset	0.309	0.000	0.436	1.497	0.530	1.712	1.688
arma	-0.056	-0.436	0.000	1.002	0.690	1.083	1.036
garch_mean	-0.851	-1.497	-1.002	0.000	-0.914	0.418	0.263
ar1	-0.124	-0.530	-0.690	0.914	0.000	1.003	0.960
hist_mean	-0.981	-1.712	-1.083	-0.418	-1.003	0.000	-0.144
zero	-0.963	-1.688	-1.036	-0.263	-0.960	0.144	0.000

Table E3e. Point-Forecast DM p-values — ASX 200

Model	hybrid	ms_noreset	arma	garch_mean	ar1	hist_mean	zero
hybrid	1.000	0.757	0.956	0.395	0.901	0.327	0.336
ms_noreset	0.757	1.000	0.663	0.135	0.596	0.087	0.092
arma	0.956	0.663	1.000	0.317	0.490	0.279	0.300
garch_mean	0.395	0.135	0.317	1.000	0.361	0.676	0.793
ar1	0.901	0.596	0.490	0.361	1.000	0.316	0.337
hist_mean	0.327	0.087	0.279	0.676	0.316	1.000	0.885
zero	0.336	0.092	0.300	0.793	0.337	0.885	1.000

E.3 Volatility and Distribution Metrics

Table E4. Complete Volatility and Predictive-Distribution Metrics

Market	Model	N	QLIKE	Variance MSE	Neg. log score	CRPS	80% coverage	95% coverage
S&P BSE SENSEX	garch	1000	-8.7465	0.0000000	-3.4512	0.004290	82.6%	95.8%
S&P BSE SENSEX	ewma	1000	-8.7291	0.0000000	-3.4449	0.004289	81.1%	94.7%
S&P BSE SENSEX	ms_noreset	1000	-8.7280	0.0000000	-3.4750	0.004266	83.8%	95.6%
S&P BSE SENSEX	hybrid	1000	-8.6542	0.0000000	-3.4154	0.004287	81.5%	94.8%
Nikkei 225	ms_noreset	1000	-7.6322	0.0000006	-2.9121	0.007600	77.4%	93.3%
Nikkei 225	garch	1000	-7.6304	0.0000006	-2.8983	0.007584	77.2%	92.9%
Nikkei 225	ewma	1000	-7.5755	0.0000006	-2.8702	0.007670	80.4%	93.2%
Nikkei 225	hybrid	1000	-7.4469	0.0000007	-2.7825	0.007854	78.7%	93.4%
Hang Seng Index	ms_noreset	1000	-7.3857	0.0000006	-2.8160	0.008234	78.0%	95.2%
Hang Seng Index	garch	1000	-7.3845	0.0000006	-2.7719	0.008260	77.8%	96.3%
Hang Seng Index	ewma	1000	-7.3768	0.0000007	-2.7689	0.008281	79.2%	95.5%
Hang Seng Index	hybrid	1000	-7.3226	0.0000007	-2.7458	0.008311	82.8%	96.1%

Market	Model	N	QLIKE	Variance MSE	Neg. log score	CRPS	80% coverage	95% coverage
KOSPI Composite Index	ms_noreset	1000	-7.5036	0.0000018	-2.8668	0.008859	75.5%	93.2%
KOSPI Composite Index	garch	1000	-7.4946	0.0000018	-2.8305	0.008821	77.0%	92.8%
KOSPI Composite Index	ewma	1000	-7.4800	0.0000018	-2.8222	0.008848	80.7%	93.4%
KOSPI Composite Index	hybrid	1000	-7.3284	0.0000018	-2.8137	0.008883	74.8%	91.2%
S&P/ASX 200	garch	1000	-8.8088	0.0000000	-3.4825	0.004187	81.0%	95.6%
S&P/ASX 200	ms_noreset	1000	-8.7981	0.0000000	-3.4910	0.004197	80.7%	95.1%
S&P/ASX 200	hybrid	1000	-8.7922	0.0000000	-3.4342	0.004199	80.5%	94.2%
S&P/ASX 200	ewma	1000	-8.7610	0.0000000	-3.4610	0.004208	80.2%	94.2%

E.4 QLIKE Diebold–Mariano Tests

Table E5a. QLIKE DM Statistics — SENSEX

Model	hybrid	ms_noreset	garch	ewma
hybrid	0.000	2.051	2.772	2.189
ms_noreset	-2.051	0.000	1.309	0.042
garch	-2.772	-1.309	0.000	-1.031
ewma	-2.189	-0.042	1.031	0.000

Table E6a. QLIKE DM p-values — SENSEX

Model	hybrid	ms_noreset	garch	ewma
hybrid	1.000	0.040	0.006	0.029
ms_noreset	0.040	1.000	0.191	0.967
garch	0.006	0.191	1.000	0.303
ewma	0.029	0.967	0.303	1.000

Table E5b. QLIKE DM Statistics — Nikkei 225

Model	hybrid	ms_noreset	garch	ewma
hybrid	0.000	1.778	1.777	1.343
ms_noreset	-1.778	0.000	-0.097	-1.438
garch	-1.777	0.097	0.000	-1.946
ewma	-1.343	1.438	1.946	0.000

Table E6b. QLIKE DM p-values — Nikkei 225

Model	hybrid	ms_noreset	garch	ewma
hybrid	1.000	0.076	0.076	0.180
ms_noreset	0.076	1.000	0.923	0.151
garch	0.076	0.923	1.000	0.052
ewma	0.180	0.151	0.052	1.000

Table E5c. QLIKE DM Statistics — Hang Seng

Model	hybrid	ms_noreset	garch	ewma
hybrid	0.000	1.409	1.647	1.425
ms_noreset	-1.409	0.000	-0.074	-0.383
garch	-1.647	0.074	0.000	-0.527
ewma	-1.425	0.383	0.527	0.000

Table E6c. QLIKE DM p-values — Hang Seng

Model	hybrid	ms_noreset	garch	ewma
hybrid	1.000	0.159	0.100	0.154
ms_noreset	0.159	1.000	0.941	0.701
garch	0.100	0.941	1.000	0.598
ewma	0.154	0.701	0.598	1.000

Table E5d. QLIKE DM Statistics — KOSPI

Model	hybrid	ms_noreset	garch	ewma
hybrid	0.000	2.083	1.960	1.805
ms_noreset	-2.083	0.000	-0.537	-0.613
garch	-1.960	0.537	0.000	-0.476
ewma	-1.805	0.613	0.476	0.000

Table E6d. QLIKE DM p-values — KOSPI

Model	hybrid	ms_noreset	garch	ewma
hybrid	1.000	0.037	0.050	0.071
ms_noreset	0.037	1.000	0.592	0.540
garch	0.050	0.592	1.000	0.634
ewma	0.071	0.540	0.634	1.000

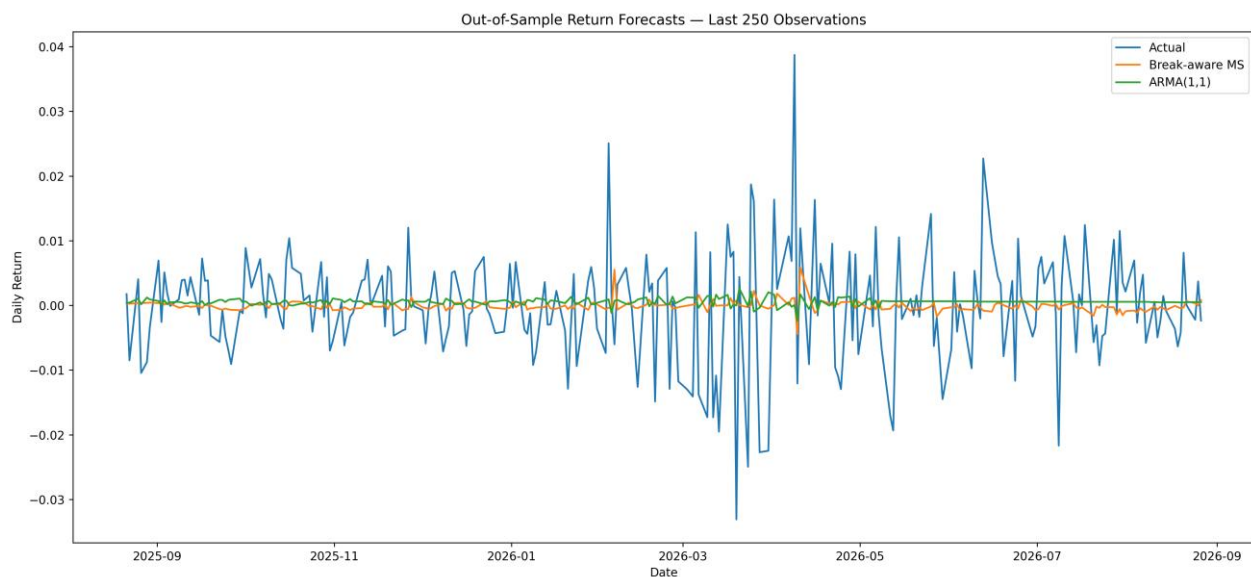
Table E5e. QLIKE DM Statistics — ASX 200

Model	hybrid	ms_noreset	garch	ewma
hybrid	0.000	0.313	0.926	-1.357
ms_noreset	-0.313	0.000	1.398	-1.910
garch	-0.926	-1.398	0.000	-2.914
ewma	1.357	1.910	2.914	0.000

Table E6e. QLIKE DM p-values — ASX 200

Model	hybrid	ms_noreset	garch	ewma
hybrid	1.000	0.754	0.355	0.175
ms_noreset	0.754	1.000	0.162	0.056
garch	0.355	0.162	1.000	0.004
ewma	0.175	0.056	0.004	1.000

E.5 OOS Visualizations

**Fig. E1. OOS Point Forecasts — SENSEX**

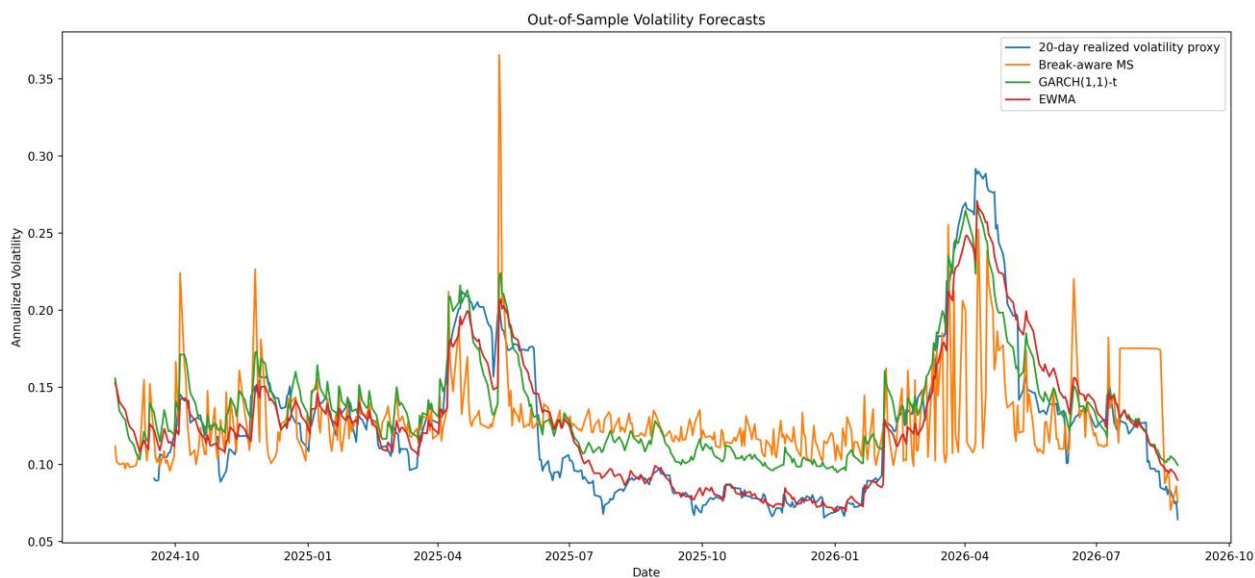


Fig. E2. OOS Volatility Forecasts — SENSEX

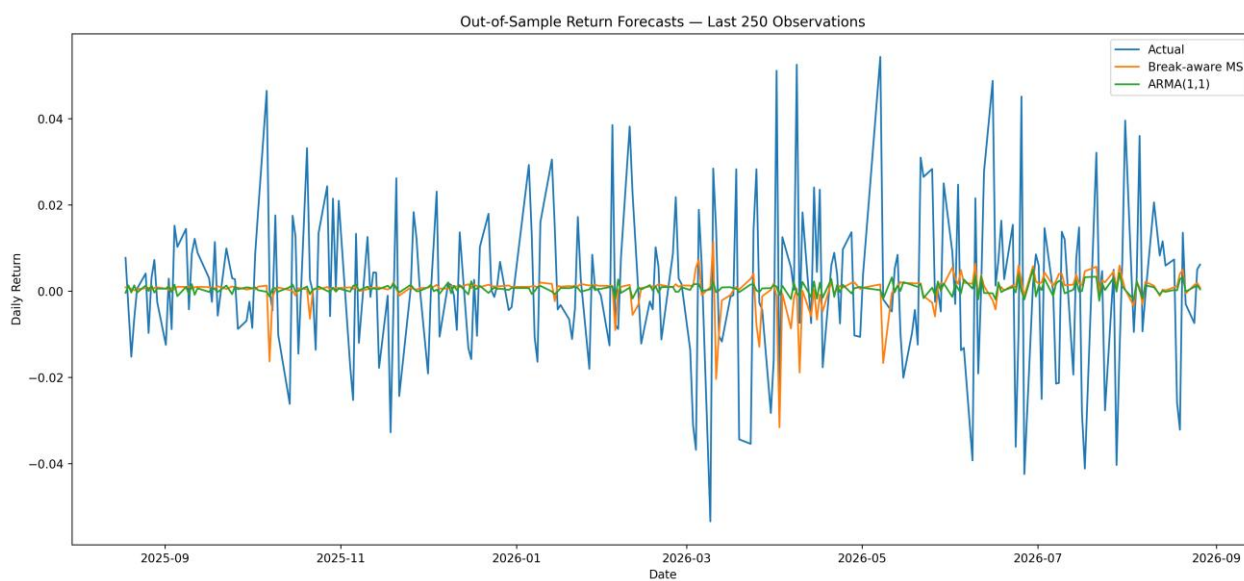


Fig. E3. OOS Point Forecasts — Nikkei 225

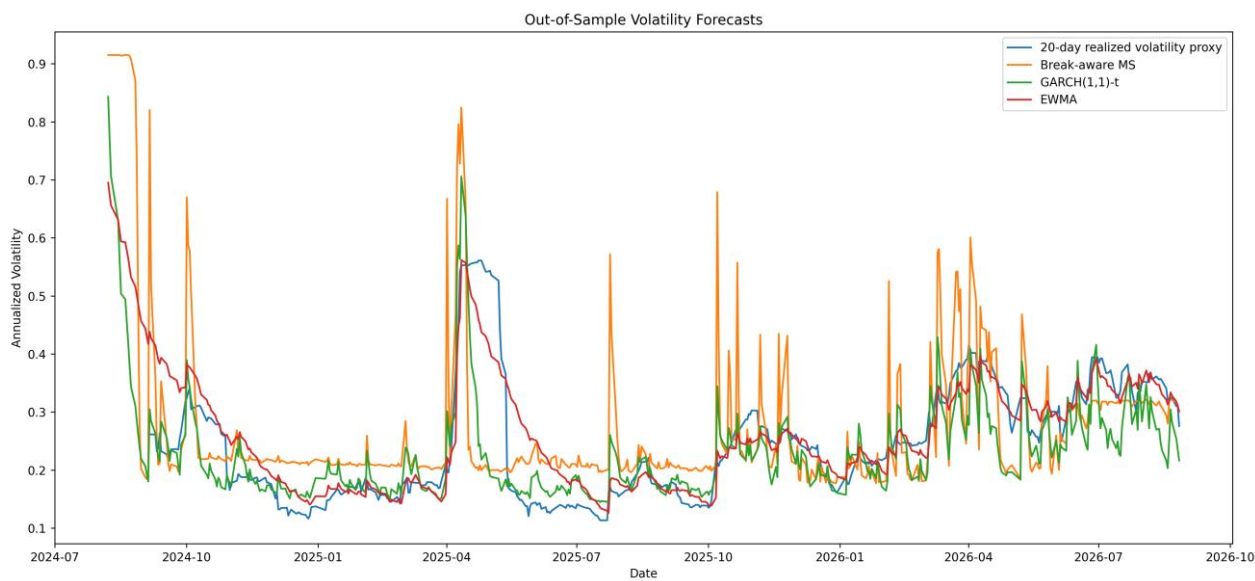


Fig. E4. OOS Volatility Forecasts — Nikkei 225

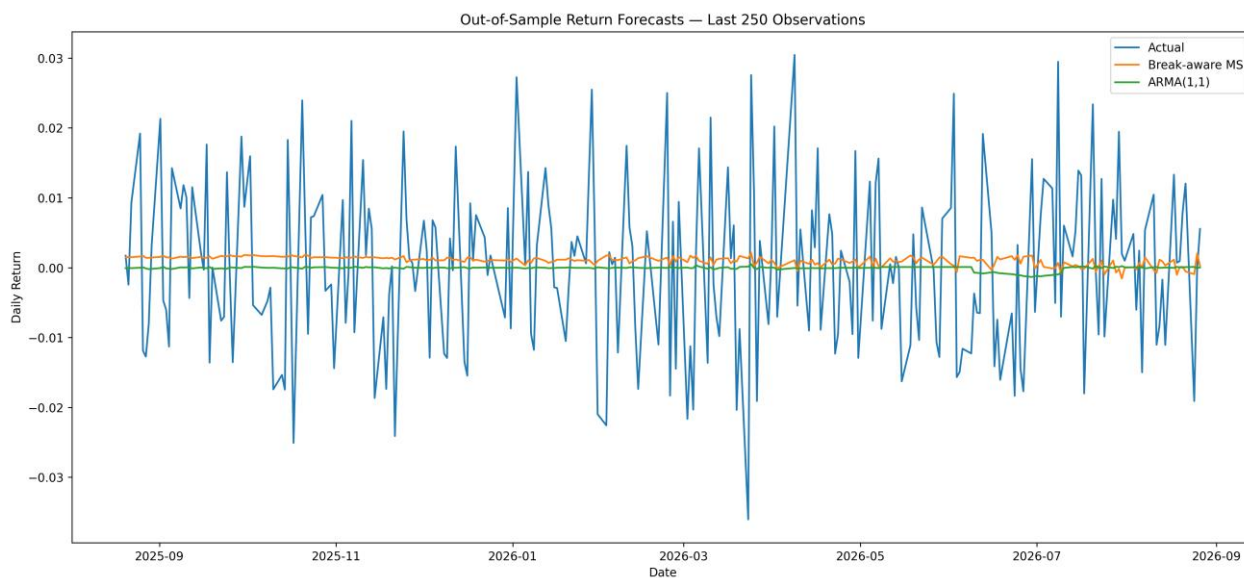


Fig. E5. OOS Point Forecasts — Hang Seng

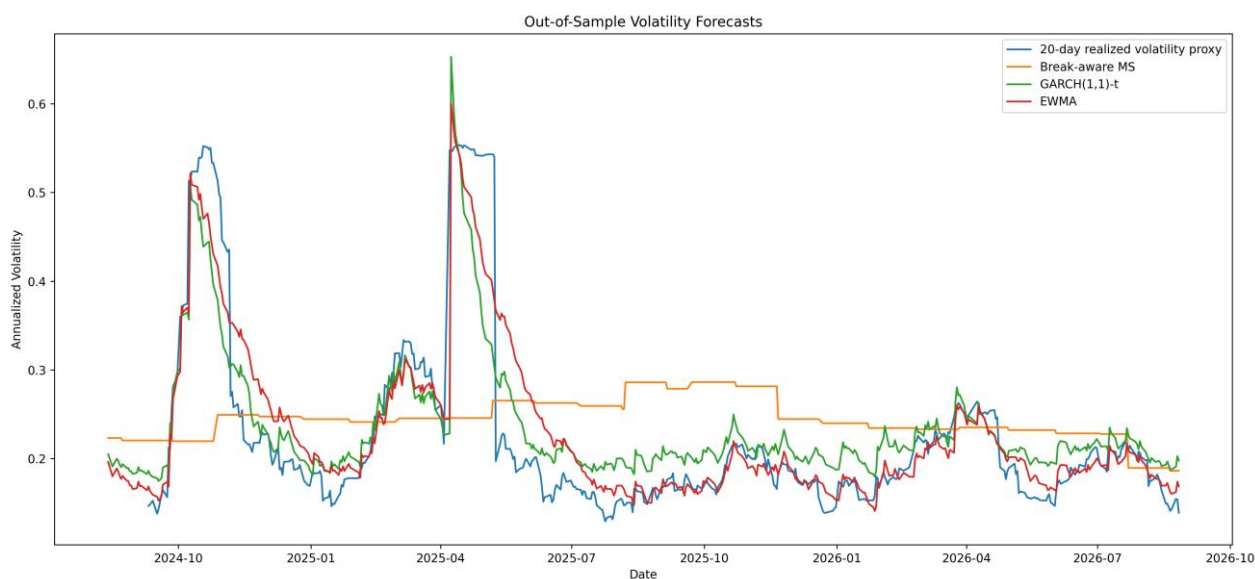


Fig. E6. OOS Volatility Forecasts — Hang Seng

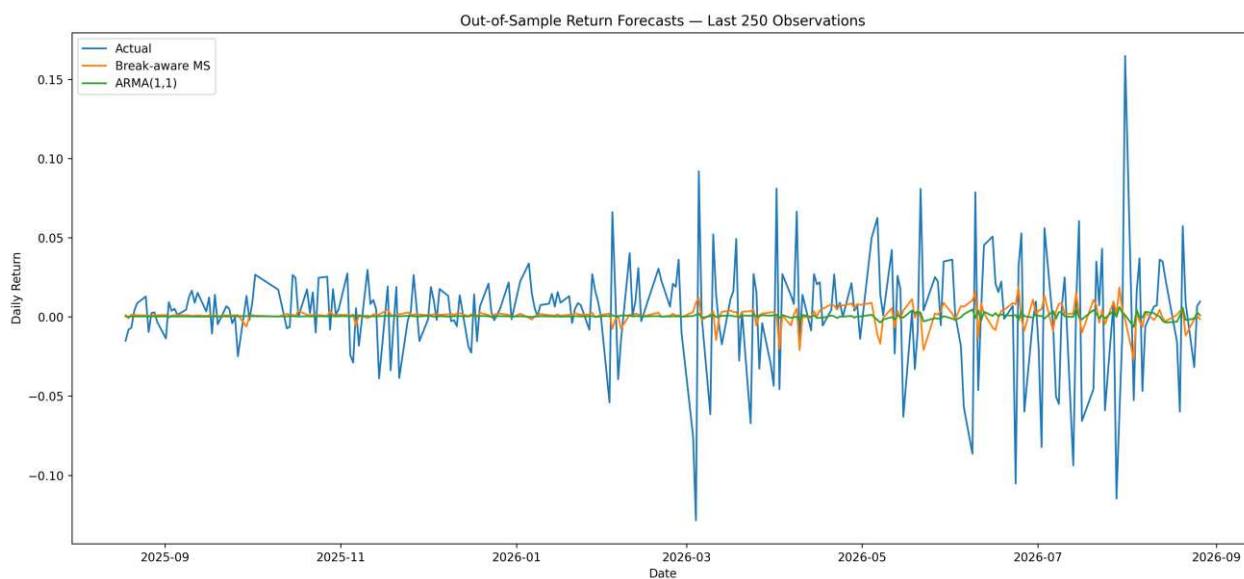


Fig. E7. OOS Point Forecasts — KOSPI

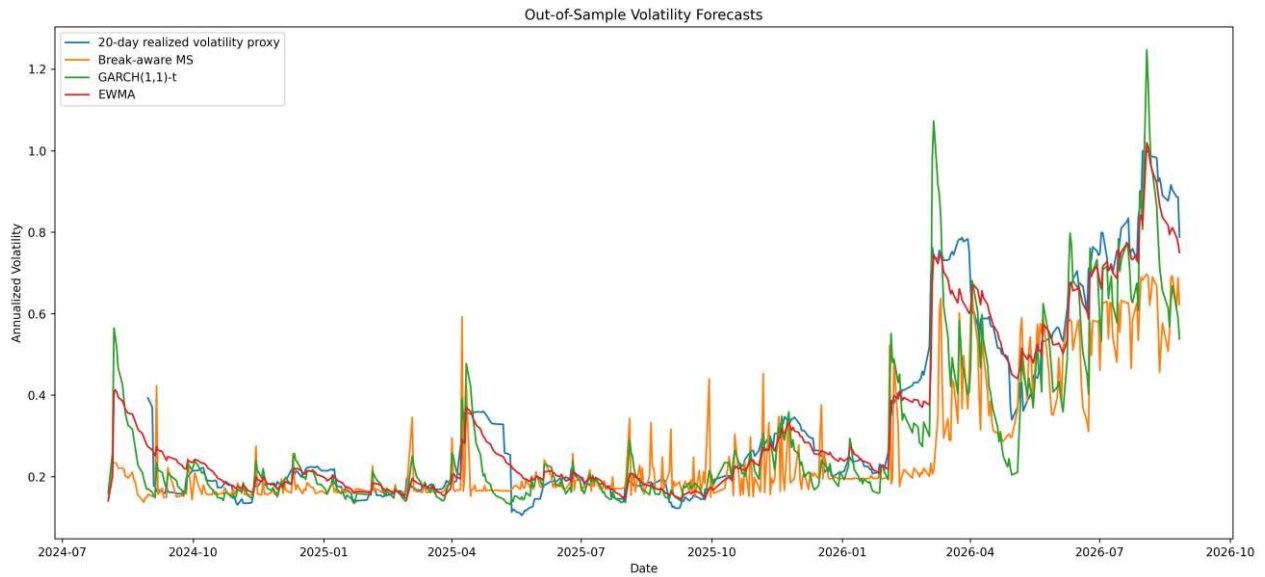


Fig. E8. OOS Volatility Forecasts — KOSPI

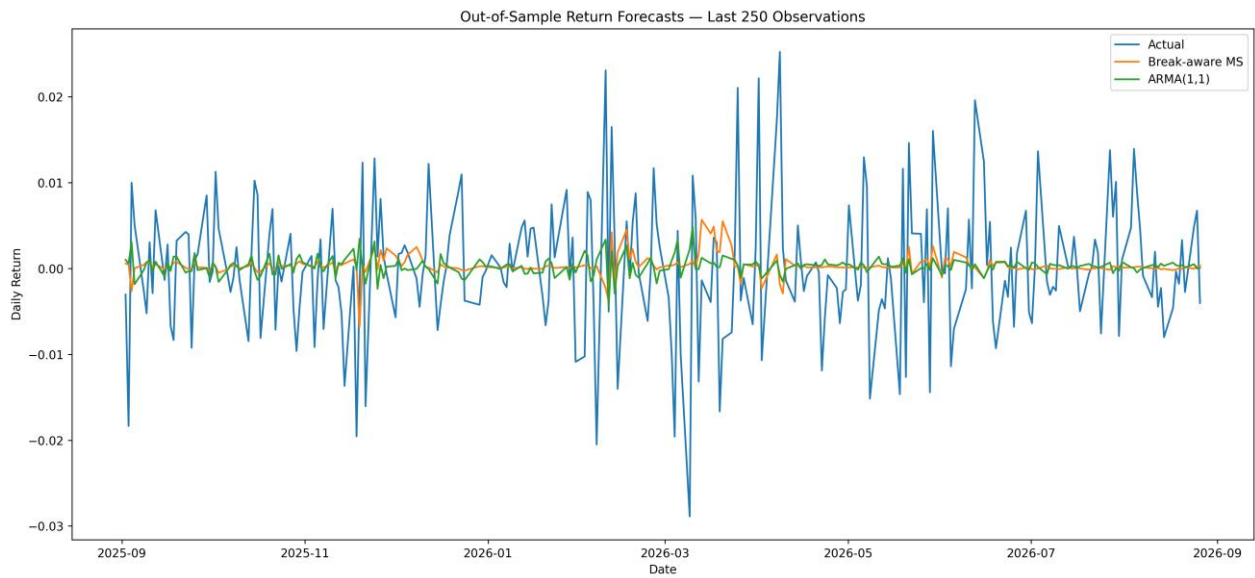


Fig. E9. OOS Point Forecasts — ASX 200

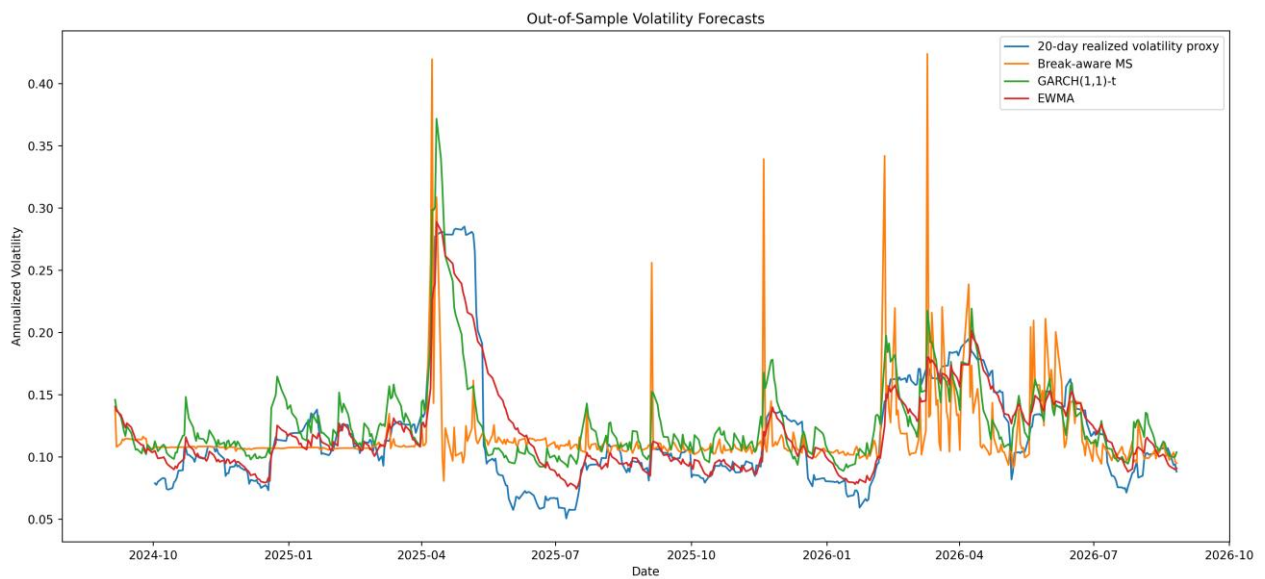


Fig. E10. OOS Volatility Forecasts — ASX 200

Appendix F. Value-at-Risk and Expected Shortfall Detail

This appendix contains the complete downside-risk results.

F.1 1% VaR**Table F1a. 1% VaR Backtest — SENSEX**

Model	Expected	Actual	Rate	Kupiec LR	Kupiec p	Independence LR	Independence p	CC LR	CC p
hybrid	10	15	1.50%	2.189	0.139	0.457	0.499	2.647	0.266
ms_noreset	10	10	1.00%	-0.000	1.000	0.202	0.653	0.202	0.904
garch	10	8	0.80%	0.434	0.510	0.129	0.719	0.563	0.755
ewma	10	18	1.80%	5.225	0.022	0.661	0.416	5.886	0.053

Table F1b. 1% VaR Backtest — Nikkei 225

Model	Expected	Actual	Rate	Kupiec LR	Kupiec p	Independence LR	Independence p	CC LR	CC p
hybrid	10	10	1.00%	-0.000	1.000	2.973	0.085	2.973	0.226
ms_noreset	10	14	1.40%	1.437	0.231	1.746	0.186	3.183	0.204
garch	10	15	1.50%	2.189	0.139	1.515	0.218	3.704	0.157
ewma	10	19	1.90%	6.473	0.011	0.803	0.370	7.276	0.026

Table F1c. 1% VaR Backtest — Hang Seng

Model	Expected	Actual	Rate	Kupiec LR	Kupiec p	Independence LR	Independence p	CC LR	CC p
hybrid	10	6	0.60%	1.886	0.170	0.073	0.788	1.959	0.376
ms_noreset	10	7	0.70%	1.016	0.314	0.099	0.753	1.114	0.573
garch	10	6	0.60%	1.886	0.170	0.073	0.788	1.959	0.376
ewma	10	6	0.60%	1.886	0.170	0.073	0.788	1.959	0.376

Table F1d. 1% VaR Backtest — KOSPI

Model	Expected	Actual	Rate	Kupiec LR	Kupiec p	Independence LR	Independence p	CC LR	CC p
hybrid	10	26	2.60%	17.947	<0.001	1.831	0.176	19.777	<0.001
ms_noreset	10	21	2.10%	9.284	0.002	3.171	0.075	12.455	0.002
garch	10	27	2.70%	19.929	<0.001	4.351	0.037	24.280	<0.001
ewma	10	28	2.80%	21.988	<0.001	1.423	0.233	23.411	<0.001

Table F1e. 1% VaR Backtest — ASX 200

Model	Expected	Actual	Rate	Kupiec LR	Kupiec p	Independence LR	Independence p	CC LR	CC p
hybrid	10	17	1.70%	4.091	0.043	1.121	0.290	5.212	0.074
ms_noreset	10	12	1.20%	0.380	0.538	0.292	0.589	0.672	0.715
garch	10	17	1.70%	4.091	0.043	1.121	0.290	5.212	0.074
ewma	10	23	2.30%	12.485	<0.001	2.574	0.109	15.059	<0.001

F.2 5% VaR**Table F2a. 5% VaR Backtest — SENSEX**

Model	Expected	Actual	Rate	Kupiec LR	Kupiec p	Independence LR	Independence p	CC LR	CC p
hybrid	50	49	4.90%	0.021	0.884	0.996	0.318	1.017	0.601
ms_noreset	50	48	4.80%	0.085	0.770	2.683	0.101	2.768	0.251
garch	50	60	6.00%	1.984	0.159	1.538	0.215	3.522	0.172
ewma	50	65	6.50%	4.345	0.037	0.153	0.696	4.498	0.105

Table F2b. 5% VaR Backtest — Nikkei 225

Model	Expected	Actual	Rate	Kupiec LR	Kupiec p	Independence LR	Independence p	CC LR	CC p
hybrid	50	57	5.70%	0.989	0.320	3.787	0.052	4.776	0.092
ms_noreset	50	63	6.30%	3.299	0.069	2.202	0.138	5.501	0.064
garch	50	67	6.70%	5.524	0.019	6.021	0.014	11.545	0.003
ewma	50	57	5.70%	0.989	0.320	5.796	0.016	6.785	0.034

Table F2c. 5% VaR Backtest — Hang Seng

Model	Expected	Actual	Rate	Kupiec LR	Kupiec p	Independence LR	Independence p	CC LR	CC p
hybrid	50	21	2.10%	22.441	<0.001	0.902	0.342	23.343	<0.001
ms_noreset	50	33	3.30%	6.878	0.009	0.661	0.416	7.539	0.023
garch	50	31	3.10%	8.739	0.003	0.002	0.968	8.741	0.013
ewma	50	28	2.80%	12.036	<0.001	1.423	0.233	13.459	0.001

Table F2d. 5% VaR Backtest — KOSPI

Model	Expected	Actual	Rate	Kupiec LR	Kupiec p	Independence LR	Independence p	CC LR	CC p
hybrid	50	60	6.00%	1.984	0.159	1.538	0.215	3.522	0.172
ms_noreset	50	57	5.70%	0.989	0.320	5.796	0.016	6.785	0.034
garch	50	68	6.80%	6.161	0.013	2.390	0.122	8.552	0.014
ewma	50	60	6.00%	1.984	0.159	2.933	0.087	4.918	0.086

Table F2e. 5% VaR Backtest — ASX 200

Model	Expected	Actual	Rate	Kupiec LR	Kupiec p	Independence LR	Independence p	CC LR	CC p
hybrid	50	50	5.00%	-0.000	1.000	0.855	0.355	0.855	0.652
ms_noreset	50	56	5.60%	0.731	0.393	1.070	0.301	1.800	0.406
garch	50	56	5.60%	0.731	0.393	0.246	0.620	0.977	0.614
ewma	50	68	6.80%	6.161	0.013	1.235	0.267	7.396	0.025

F.3 Expected Shortfall**Table F3. Complete Descriptive ES Calibration**

Market	α	Model	Breaches	Realized mean on breaches	Mean predicted ES	Calibration gap
S&P BSE SENSEX	1%	hybrid	15	-2.400%	-1.960%	-0.439 pp
S&P BSE SENSEX	1%	ms_noreset	10	-2.231%	-2.556%	0.324 pp
S&P BSE SENSEX	1%	garch	8	-2.932%	-2.825%	-0.108 pp
S&P BSE SENSEX	1%	ewma	18	-2.099%	-1.815%	-0.284 pp
S&P BSE SENSEX	5%	hybrid	49	-1.717%	-1.579%	-0.138 pp
S&P BSE SENSEX	5%	ms_noreset	48	-1.686%	-1.741%	0.055 pp
S&P BSE SENSEX	5%	garch	60	-1.629%	-1.721%	0.092 pp
S&P BSE SENSEX	5%	ewma	65	-1.565%	-1.466%	-0.098 pp
Nikkei 225	1%	hybrid	10	-4.917%	-4.258%	-0.659 pp
Nikkei 225	1%	ms_noreset	14	-4.782%	-4.730%	-0.051 pp
Nikkei 225	1%	garch	15	-4.725%	-4.343%	-0.382 pp
Nikkei 225	1%	ewma	19	-3.926%	-3.087%	-0.839 pp
Nikkei 225	5%	hybrid	57	-3.061%	-2.966%	-0.096 pp
Nikkei 225	5%	ms_noreset	63	-2.972%	-2.934%	-0.038 pp
Nikkei 225	5%	garch	67	-2.904%	-2.723%	-0.181 pp
Nikkei 225	5%	ewma	57	-3.019%	-2.612%	-0.408 pp
Hang Seng Index	1%	hybrid	6	-6.959%	-3.958%	-3.002 pp
Hang Seng Index	1%	ms_noreset	7	-6.365%	-4.731%	-1.634 pp
Hang Seng Index	1%	garch	6	-6.915%	-5.006%	-1.909 pp
Hang Seng Index	1%	ewma	6	-6.915%	-4.246%	-2.669 pp
Hang Seng Index	5%	hybrid	21	-4.016%	-3.187%	-0.828 pp
Hang Seng Index	5%	ms_noreset	33	-3.252%	-2.905%	-0.348 pp
Hang Seng Index	5%	garch	31	-3.396%	-3.203%	-0.194 pp
Hang Seng Index	5%	ewma	28	-3.429%	-2.853%	-0.576 pp
KOSPI Composite Index	1%	hybrid	26	-5.020%	-4.214%	-0.807 pp
KOSPI Composite Index	1%	ms_noreset	21	-5.498%	-5.173%	-0.326 pp

Market	α	Model	Breaches	Realized mean on breaches	Mean predicted ES	Calibration gap
KOSPI Composite Index	1%	garch	27	-5.295%	-5.102%	-0.193 pp
KOSPI Composite Index	1%	ewma	28	-4.482%	-3.822%	-0.660 pp
KOSPI Composite Index	5%	hybrid	60	-3.823%	-2.997%	-0.826 pp
KOSPI Composite Index	5%	ms_noreset	57	-3.790%	-3.294%	-0.496 pp
KOSPI Composite Index	5%	garch	68	-3.514%	-2.998%	-0.517 pp
KOSPI Composite Index	5%	ewma	60	-3.660%	-3.072%	-0.588 pp
S&P/ASX 200	1%	hybrid	17	-2.085%	-1.994%	-0.091 pp
S&P/ASX 200	1%	ms_noreset	12	-2.399%	-2.658%	0.259 pp
S&P/ASX 200	1%	garch	17	-2.274%	-2.475%	0.201 pp
S&P/ASX 200	1%	ewma	23	-2.123%	-1.830%	-0.293 pp
S&P/ASX 200	5%	hybrid	50	-1.712%	-1.583%	-0.129 pp
S&P/ASX 200	5%	ms_noreset	56	-1.650%	-1.644%	-0.006 pp
S&P/ASX 200	5%	garch	56	-1.644%	-1.575%	-0.069 pp
S&P/ASX 200	5%	ewma	68	-1.560%	-1.406%	-0.154 pp

Note. Expected Shortfall results are descriptive calibration measures and do not constitute formal ES hypothesis tests.

Appendix G. Benchmark Stability

Repeated walk-forward estimation can occasionally generate unstable ARMA solutions. The final implementation records convergence validity and fallback behaviour at every forecast origin.

Table G1. ARMA(1,1) Stability Audit

Market	OOS forecasts	Fallbacks	Valid states	Maximum absolute forecast	All finite
SENSEX	1000	0	100%	0.3799%	Yes
Nikkei 225	1000	0	100%	0.5065%	Yes
Hang Seng	1000	0	100%	0.2559%	Yes
KOSPI	1000	0	100%	0.6425%	Yes
ASX 200	1000	0	100%	0.6919%	Yes

No fallback forecasts are required in any market, and every forecast remains finite.

Appendix H. Additional Robustness

H.1 SENSEX Estimation-Window Sensitivity

Table H1. Maximum Estimation-Window Sensitivity

Maximum window	RMSE	MAE	QLIKE	CRPS	95% coverage	1% breaches	1% CC p	5% breaches	5% CC p
1000	0.7957%	0.5806%	-8.6542	0.004287	94.8%	15	0.266	49	0.601
1500	0.7957%	0.5806%	-8.6542	0.004287	94.8%	15	0.266	49	0.601
2000	0.7957%	0.5806%	-8.6542	0.004287	94.8%	15	0.266	49	0.601

The equality of these results is mechanical rather than substantive. The actual SENSEX adaptive estimation window ranges from 252 to 731 observations, with a mean of approximately 416, so none of the tested maximum-window restrictions binds.

H.2 Methodological Notes

The robustness tests are intentionally asymmetric: detailed parameter-sensitivity analysis is performed for the primary SENSEX application, while external robustness is supplied by applying the baseline design unchanged to the other four markets. Alternative specifications are not selected according to which produces the most favourable hybrid performance.

Appendix I. Controlled Simulation Detail

I.1 Regime Classification

Table I1. Raw Regime-Recovery Confusion Matrix

Unnamed: 0	inferred_0	inferred_1	inferred_2
true_0	2606	62	0
true_1	117	953	19
true_2	5	64	373

Overall aligned classification accuracy is 93.64%.

Table I2. State-Specific Recovery

State	True observations	Predicted observations	Precision	Recall
Low volatility	2668	2728	95.53%	97.68%
Medium volatility	1089	1079	88.32%	87.51%
High volatility	442	392	95.15%	84.39%

I.2 Parameter Recovery

Table I3. True and Estimated Parameters in the Simulated Regime DGP

State	True intercept	Estimated intercept	True AR(1)	Estimated AR(1)	True daily volatility	Estimated daily volatility	True stay probability	Estimated stay probability
Low	0.0350%	0.0429%	0.0300	0.0331	0.5000%	0.4921%	97.60%	97.50%
Medium	0.0050%	-0.0644%	-0.0400	-0.0332	1.5000%	1.4837%	92.00%	89.86%
High	-0.0700%	0.2749%	0.0800	0.0823	5.5000%	6.0369%	90.00%	84.73%

I.3 Structural-Break Recovery

Table I4. Structural-Break Localization

True break	Nearest detected	Signed error	Absolute error	Within ± 30 days
1400	1400	0	0	Yes
2800	2790	-10	10	Yes

LWZ selects two breaks, matching the true number, while BIC selects three. BOCPD generates 29 alarms in the structural-break DGP, consistent with its role as a sequential adaptation mechanism rather than a permanent-break counter.

Appendix J. Reproducibility and No-Look-Ahead Audit

J.1 Data Integrity

Table J1. Market-Specific Data Audit

Market	Return observations	Calendar coverage	Missing expected sessions	Flagged extremes	SHA-256
SENSEX	4898	99.53%	23	3	07eef6e4ba20062b6e64ff55cbc0307837f0fda61401d19789042e8048f7c7f2
Nikkei 225	4874	99.94%	3	4	01add6651c7e7d726629b6d45be8a60f1e37de70b672e98ec65b3c932794aa57
Hang Seng	4908	99.94%	3	4	8688cad4841c4d0fb93d5f60ddb0ad9d6904f8736ead73588b6b575eb95bfe1e
KOSPI	4909	99.84%	8	10	bc3d5ecfc8da2744605db3e43b44ad207e32fb2e49fe448bcc69cd7ebf2c0e9a
ASX 200	5035	99.76%	12	2	ce5c15fe0e9d4b66dcb45bd8a7a9631c55c3d40efee53a3acc0adfbaf0b40083

J.2 Computational Environment

Table J2. Software Environment

Component	Version
Python	3.13.15
NumPy	2.1.3
pandas	2.2.3
SciPy	1.16.3
statsmodels	0.15.0
Matplotlib	3.10.0
arch	8.0.0
yfinance	0.2.66
exchange_calendars	4.13.2
Random seed	42

J.3 Principal Configuration**Table J3. Reproducibility Parameters**

Parameter	Final specification	
Annualization	252 trading days	
Diagnostic lags		20
ARCH lags		20
Break minimum segment		126
Maximum retrospective breaks		12
Break candidate spacing	5 observations	
Primary break criterion	LWZ	
BOCPD expected run length		250
BOCPD alarm threshold		0.2
Minimum alarm gap		20
Maximum BOCPD run length		1200
Candidate Markov regimes		2, 3, 4
Regime-selection criterion	Pre-OOS BIC	
Markov maximum iterations		300
OOS forecasts per market		1000
Minimum initial training sample		750
Regular refit frequency	21 trading observations	
Maximum estimation window		1500
Minimum adaptive window		252
EWMA decay		0.94
VaR / ES levels		1%, 5%
Simulation observations		4200
True simulated breaks	1,400; 2,800	
Isolated simulated crash		3,200
Random seed		42

J.4 No-Look-Ahead Audit**Table J4. OOS Causality and Refit Audit**

Market	OOS forecasts	Information date < target	First OOS target	Hybrid refits	No-reset refits	Benchmark refits	ARMA fallbacks
SENSEX	1000	1000/1000	03 Aug 2022	51	48	48	0
Nikkei 225	1000	1000/1000	26 Jul 2022	52	48	48	0
Hang Seng	1000	1000/1000	01 Aug 2022	52	48	48	0
KOSPI	1000	1000/1000	22 Jul 2022	49	48	48	0
ASX 200	1000	1000/1000	13 Sep 2022	50	48	48	0

Every one of the 5,000 market-level forecasts satisfies information end date < forecast target date. The regime count is selected independently for each market before the corresponding test period begins. Forecasting uses filtered regime probabilities only; full-sample smoothed probabilities are restricted to retrospective figures. BOCPD information is similarly restricted to alarms already available before each target observation.