AI Programming project - Group 16

MÁTÉ CSIKÓS (12535753) and NICCOLÒ SIGNORELLI (12502991)

1 Introduction

Financial returns exhibit volatility clustering and heavy tails, features that standard linear models struggle to capture [1]. This study applies Probabilistic Programming to the S&P 500 index to test the hypothesis that "fat tails" are primarily artifacts of unmodeled volatility states. By progressing from long-term trend analysis to high-frequency volatility forecasting, we aim to develop a robust framework capable of identifying and adapting to systemic market crises.

2 Phase 1: Trend analysis

2.1 Data and Preprocessing

The analysis focuses on the S&P 500 index (ĜSPC), leveraging monthly closing prices to capture long-term trends and volatility dynamics while minimizing high-frequency noise.

- **Timeline:** The dataset spans from January 2007 to December 2025. The period from 2007-2022 serves as the training set, while the remaining data (2023-2025) is reserved for out-of-sample forecasting.
- **Returns:** Log-returns were calculated as $r_t = \ln(P_t/P_{t-1})$, where P_t is the closing price at the t-th timepoint.

2.2 Methodology

A Bayesian workflow using `PyMC` was employed to model the time series, progressing from simple trend detection to segmented volatility dynamics. Four main model configurations were evaluated, all assuming a piecewise linear trend structure:

(1) **Regularized Piecewise Linear Trend (Model 1.2b):** The assumption of a constant linear trend was relaxed by introducing changepoints, allowing the slope and intercept of the market trend $\mu(t)$ to adapt over time. This structure captured structural shifts, such as the post-2008 recovery, while maintaining the assumption of constant Gaussian volatility ($\epsilon_t \sim N(0, \sigma)$).
(2) **Student-t Likelihood (Model 1.2d):** To address potential "fat tails," the Gaussian likelihood was replaced with a Student-t distribution. This introduced a degrees-of-freedom parameter ($\nu \sim \text{Exp}(0.1)$), allowing the model to handle extreme observations without varying the volatility.
(3) **Segmented Volatility (Model 1.3):** The constant volatility assumption was relaxed by allowing σ to vary across time segments defined by the trend changepoints. A unique volatility parameter σ_i was assigned to each segment (annually), drawn from a hierarchical prior ($\sigma_i \sim$ HalfNormal). This allows the model to adapt to periods of varying stability.
(4) **Segmented Volatility with Student-t Likelihood (Model 1.3t):** Finally, the segmented volatility approach was combined with a Student-t likelihood to assess whether heavy tails were more accurately captured by this formulation.

2.3 Model Selection

The models were compared using the Widely Applicable Information Criterion (WAIC) to assess the trade-off between goodness-of-fit and complexity. The results, summarized in Table 1, reveal a clear hierarchy in model performance.

Authors' Contact Information: Máté Csikós (12535753); Niccolò Signorelli (12502991).

Model	Rank	elpd_waic	Weight	SE
1.3: Regime σ + Normal	0	343.65	1.00	12.02
1.3t: Regime σ + Student-t	1	341.53	0.00	12.15
1.2d: Constant σ + Student-t	2	303.17	0.00	13.64
1.2b: Constant σ + Normal	3	294.86	≈ 0	14.37

Table 1. WAIC comparison of Phase 1 models. Model 1.3 is the clear winner.

The most significant finding is that Model 1.3 (Segmented Volatility with Normal errors) outperformed all other models, including the Student-t variants. This result confirms a fundamental hypothesis of financial time series: apparent fat tails are often an artifact of volatility clustering [2]. Thus, once time-varying volatility was explicitly modeled (Model 1.3), the Gaussian likelihood became sufficient. The mathematical formulation of such model is:

$$y_t \sim \mathcal{N}(\mu_t, \sigma^2_{r(t)}) \qquad \text{Likelihood} \tag{1}$$

$$\mu_t = (k + \mathrm{A}(t)^\top \boldsymbol{\delta})t + (m + \mathrm{A}(t)^\top \boldsymbol{\gamma}) \qquad \text{Trend model} \tag{2}$$

where

- $\delta_j \sim \text{Laplace}(0, \tau)$ models the change in slope, with $\tau \sim \text{HalfNormal}(0.5)$.
- $\gamma_j = -s_j\delta_j$, ensures the continuity of the trend line at each changepoint.
- $\sigma_{r(t)} = \sigma_j \sim \text{HalfNormal}(1.0)$ is the Regime-Switching Volatility for each changepoint t in the j-th regime.
- $k \sim \mathcal{N}(0, 1)$ is the trend intercept.
- $m \sim \mathcal{N}(0, 0.1)$ is the trend slope.
- $\mathrm{A}(t)$ is the activation vector where the j-th element is 1 if $t \geq s_j$ (time t is past changepoint s_j) and 0 otherwise.

2.4 Phase 1 Out-of-Sample Forecast (2023–2025)

To assess Model 1.3's predictive capability, a recursive forecast was performed on the test data. Instead of refitting the full model monthly, an online update scheme was used: the volatility estimate (σ) was updated at each step t based on the prediction error from $t-1$, using an exponentially weighted moving average to adapt the Bayesian posterior samples. The results are presented in Figure 1.

The mean forecast (dashed blue line) tracks the price path with a visible one-step lag. This is intrinsic to the recursive process: the model receives information about a price shock only after the month closes, adjusting its starting level for the next period. This confirms the forecast is strictly causal.

While directional prediction is limited by the random walk nature of monthly prices, the model succeeds in uncertainty quantification. The shaded 90% CI successfully encapsulates the true price path for the majority of the test period. This validates that while the trend is hard to predict, the risk envelope (volatility) is well-calibrated by the regime-switching priors learned with the training set.

3 Phase 2: Regime-Switching Volatility Modeling

Given the critical role of volatility clustering observed in the first part of the project, the second phase of this analysis uses volatility as the primary object of study. The goal is to explicitly model its behavior during crisis periods compared to normal market conditions.

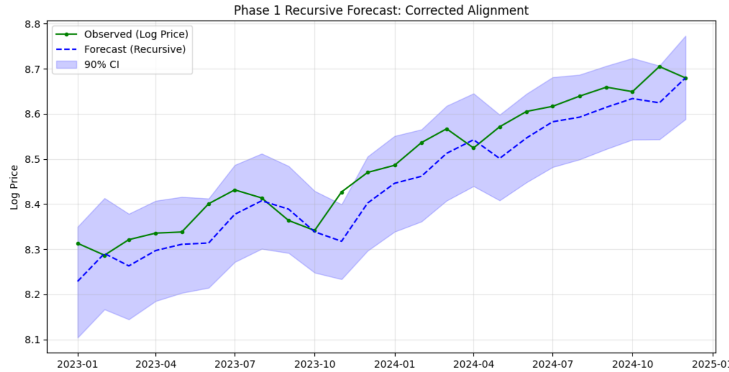

Fig. 1. **Phase 1 Recursive Forecast (2023–2025).** Observed cumulative log-prices (green) vs. recursive forecast (blue) with 90% Credibility Intervals (CI). The CI adapts via online volatility updates, successfully capturing the price path despite the expected random-walk lag.

3.1 Data Preprocessing

Daily closing prices of the S&P 500 index (ĜSPC) were retrieved via the `yfinance` API [3]. The dataset spans from January 1, 2007, to January 1, 2025, allowing the analysis of both the 2008 Global Financial Crisis and the 2020 COVID-19 crash.

Daily log-returns were calculated as:

$$r_t = (\ln P_t - \ln P_{t-1}) \times 100$$

where the percent scale was utilized to ensure numerical stability. Zero-return days were removed to prevent numerical issues in logarithmic transformations.

The data was split to allow fitting and testing:

- *Training Set*: Jan 2007 – Dec 2022 ($\approx$ 4000 days). Used for Bayesian estimation and defining regime thresholds.
- *Test Set*: Jan 2023 – Jan 2025. Reserved for evaluating forecast performance.

Since volatility is latent, an observable feature, Realized Volatility (RV_t), was constructed using the standard deviation of returns over a 5-day rolling window (representing one trading week):

$$RV_t = \text{std}(r_{t-4}, \dots, r_t)$$

The analysis then modeled the *log-volatility* h_t. To ensure no data leakage, the test set was centered using the mean of the training set log-volatility:

$$h_t^{\text{centered}} = h_t - \mu_{\text{train}}$$

3.2 Methodology

Attempts were initially made to fit a canonical Stochastic Volatility (SV) model [4] within a Bayesian framework, with the intention of progressively incorporating leverage effects as formulated by Yu (2005) [5]. However, despite testing various prior settings (both uninformative and weakly informative) and simplifying the model structure, the NUTS[1] chains failed to converge reliably.

Consequently, a simpler, more robust, and observable approach was adopted based on the Self-Exciting Threshold Autoregressive model introduced by Tong (1980) [7]. This framework assumes that the market switches between two distinct latent states: *Normal* ($S_t = 0$) and *Crisis* ($S_t = 1$).

[1]See [6].

Let h_t denote the centered log-volatility at time t. Its dynamics follow a regime-dependent AR(1) process:

$$h_t = \mu_{S_t} + \phi_{S_t} h_{t-1} + \epsilon_t, \quad \epsilon_t \sim t_\nu(0, \sigma_{S_t})$$

where the degrees of freedom ν are shared across regimes and estimated via $\nu \sim \text{Exp}(0.1)$, capturing the global heavy-tailed nature of volatility shocks. The latent state $S_t \in \{0, 1\}$ is determined deterministically by whether the current Realized Volatility (RV_t) exceeds a critical threshold γ (set to the 90th percentile of the training data):

$$S_t = \begin{cases} 0 & \text{if } RV_t \leq \gamma \quad \text{(Normal Regime)} \\ 1 & \text{if } RV_t > \gamma \quad \text{(Crisis Regime)} \end{cases}$$

The model parameters are estimated separately for each regime. We adopt weakly informative priors following the recommendations of Gelman (2006) [8] and Kim et al. (1998) [9]:

- **Normal Regime ($S_t = 0$):** Represents stable market conditions.
  - $\mu_N \sim \mathcal{N}(-0.5, 1)$: Centers the baseline log-volatility at a low level.
  - $\phi_N \sim \text{Beta}(20, 2)$: Enforces strong mean reversion (large persistence) typical of calm periods.
  - $\sigma_N \sim \text{HalfNormal}(0.3)$: Assumes small volatility shocks.
- **Crisis Regime ($S_t = 1$):** Represents high-stress market conditions.
  - $\mu_C \sim \mathcal{N}(0.5, 1)$: Shifts the baseline log-volatility significantly higher.
  - $\phi_C \sim \text{Beta}(15, 3)$: Allows for slightly lower persistence, as crises can be erratic.
  - $\sigma_C \sim \text{HalfNormal}(0.5)$: Permits larger, more volatile shocks.

3.3 Out-of-Sample Forecasting

To evaluate the model's predictive power, a fixed-parameter recursive forecast was performed on the unseen test period (Jan 2023 – Jan 2025). The posterior distributions of the parameters, estimated solely on the 2007–2022 training data, were used to generate one-step-ahead predictions. This simulates a real-time forecasting scenario where the structural dynamics learned from past crises (2008, 2020) are applied to future data.

For each time step t, the active regime was identified using the *lagged* realized volatility RV_{t-1} to prevent look-ahead bias. The predictive distribution for the log-volatility h_t was then generated by propagating the full Bayesian parameter uncertainty:

$$\hat{h}_t^{(i)} \sim t_\nu \left(\mu_{S_{t-1}}^{(i)} + \phi_{S_{t-1}}^{(i)} h_{t-1}^{\text{obs}}, \sigma_{S_{t-1}}^{(i)} \right)$$

where i indexes the MCMC posterior samples and S_{t-1} denotes the regime determined at time $t-1$.

The forecasting performance is visualized in Figure 2. The model achieved a Mean Absolute Error (MAE) of **0.187** on the log-volatility scale. As shown in the middle panel of Figure 2, the prediction errors remain centered around zero, indicating that the model successfully avoids systematic bias even during volatile transitions.

The top panel demonstrates the value of the Bayesian approach: the 90% credibility interval correctly hugs the observed data. Moreover, the regime classification mechanism (bottom panel) proved to be effective to identify **August 1, 5, and 6, 2024**. This corresponds precisely to the "Black Monday" crash triggered by the unwinding of the Yen carry trade [10]. By correctly switching to the high-variance state during this event, the model demonstrated that the crisis thresholds learned from 2008 and 2020 generalize robustly to new, unseen market shocks.

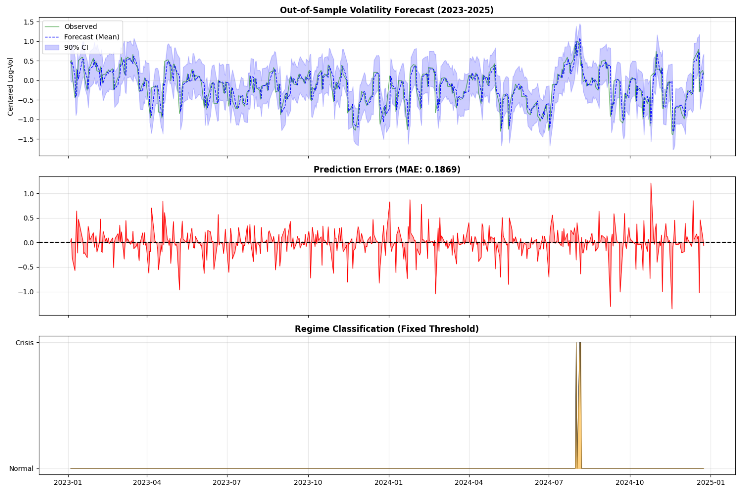

Fig. 2. **Out-of-Sample Forecast Results (2023–2025).** *Top Panel:* Observed centered log-volatility (green) vs. Bayesian forecast mean (dashed blue) with 90% CI. *Middle Panel:* Prediction errors. *Bottom Panel:* Regime classification, correctly identifying the "Black Monday" crash (August 2024) as the sole crisis event.

4 Conclusion

This study demonstrates that S&P 500 dynamics are more effectively captured by modeling discrete volatility regimes than by relying on static heavy-tailed distributions. Phase 1 analysis confirmed that segmented volatility models significantly outperform constant-volatility Student-t formulations, suggesting that the "fat tails" often observed in financial returns are primarily artifacts of volatility clustering.

Building on this, the Phase 2 regime-switching framework proved robust in out-of-sample forecasting. By explicitly modeling the transition between "Normal" and "Crisis" states, the model not only achieved high predictive accuracy (MAE 0.187) but also successfully identified systemic risk during the August 2024 "Black Monday" event. These results validate the Bayesian approach for risk management, offering an alternative mechanism to quantify uncertainty in markets.

References

[1] Rama Cont. Empirical properties of asset returns: stylized facts and statistical issues. *Quantitative Finance*, 1(2):223–236, 2001.

[2] Rama Cont. Empirical properties of asset returns: stylized facts and statistical issues. *Quantitative Finance*, 1(2):223–236, 2001.

[3] Ran Aroussi. yfinance: Yahoo! finance market data downloader. https://github.com/ranaroussi/yfinance, 2023. Python package version 0.2.36.

[4] Stephen J Taylor. *Modelling Financial Time Series.* John Wiley & Sons, Chichester, 1986.

[5] Jun Yu. On leverage in a stochastic volatility model. *Journal of Econometrics*, 127(2):165–178, 2005.

[6] Matthew D Hoffman and Andrew Gelman. The no-u-turn sampler: adaptively setting path lengths in hamiltonian monte carlo. *Journal of Machine Learning Research*, 15(1):1593–1623, 2014.

[7] Howell Tong and K. S. Lim. Threshold autoregression, limit cycles and cyclical data. *Journal of the Royal Statistical Society: Series B (Methodological)*, 42(3):245–292, 1980.

[8] Andrew Gelman. Prior distributions for variance parameters in hierarchical models. *Bayesian Analysis*, 1(3):515–533, 2006.

[9] Sangjoon Kim, Neil Shephard, and Siddhartha Chib. Stochastic volatility: Likelihood inference and comparison with arch models. *Review of Economic Studies*, 65(3):361–393, 1998.

[10] Matteo Aquilina, Andreas Schrimpf, and Hyung Song. The market turbulence and carry trade unwind of august 2024. *BIS Bulletin*, (90), 2024.