Bayesian Hierarchical GARCH Modeling of Tech Stock Volatility

Muhammad Shahbaz Murtaza

University of British Columbia

Abstract. This study presents a Bayesian Hierarchical approach to model volatility dynamics of major and influential technology stocks using an AR(1)-GARCH(1,1) model with Student-t likelihood. By analyzing daily returns of Meta, Apple, Amazon, Netflix, Google, and the tech sector ETF(XLK) from January 2020 to the present, we examine individual effects and differences in volatility between these companies and their shared sector-specific effects. The model is implemented and documented using Stan and R, with further model diagnostics performed on the posterior approximations. Our results convey different volatility profiles, with differences in baseline volatility and persistence parameters signifying company-specific risk characteristics. This analysis provides insights for risk management, portfolio optimization, and understanding of tech-sector dynamics in a period characterized by market changes including the COVID-19 pandemic and recent US market tariff impositions.

Keywords: GARCH · Bayesian Modeling · Volatility · Hierarchical· Student t-Distributionn · Stan · Hamiltonian MC.

1 Introduction

1.1 Motivation

With the current financial instability in US markets with tariff impositions and the COVID-19 pandemic earlier in the decade, the technology sector has tolerated several disturbances which has led to a more volatile financial market. In the wake of this instability, market makers are continuously forecasting volatility, which in financial terms is coined to explain the variation of a trading price series over time and is often modeled as the standard deviation of logarithmic returns, to better cushion their investments and simultaneously optimize their portfolios. However, to curtail these giant risk appetites, historically informed and data-supported decisions are pertinent for successful risk management.

1.2 Our Research

The technology sector is a major player in the US Stock Markets, especially the FAANG indices: Meta (META), Apple (AAPL), Amazon (AMZN), Netflix (NFLX), and Alphabet (GOOGL) which are modeled alongside TECH ETF (XLK).

This study applies a Bayesian Modeling Approach modeling individual stocks as well as a hierarchical model for the tech sector using the GARCH parameters. An AR(1) component is introduced to capture the autocorrelation between volatilities across days, acknowledging the correlative nature of the prevalent time series. We employ a Student t-distribution to model the fat-tailed log-returns with learnable degrees of freedom.

With the above in mind, our primary research questions: **how well the Student t-distribution captures the fat-tailed nature of the tech stocks** as well as **how volatility persistence differs across individual stocks.**

2 Methods

2.1 Literature Review

Financial time series exhibit both time-varying volatility and volatility clustering [10]. The Generalized Autoregressive Conditional Heteroskedasticity (GARCH) models, first introduced by Bollerslev [1] as an extension of Engle's ARCH model [3], have become fundamental in financial econometrics, capturing volatility clusters and persistence effectively. Hansen and Lunde [5] compared 330 volatility models and found GARCH(1,1) to outperform others, as did Poon and Granger [8], who surveyed 93 studies on volatility forecasting and identified GARCH(1,1) as both simple and effective.

To offer a novel perspective, Shephard and Pitt [9] utilized improved Markov chain Monte Carlo (MCMC) methods to estimate volatility over blocks of time, rather than single periods as done by Jacquier, Polson, and Rossi [6] in their stochastic volatility (SV) model. While modeling volatility at weekly or monthly intervals could yield different patterns, we retained daily log-returns to preserve granularity and maximize volatility information.

In related literature, SV models based on continuous-time stochastic processes have gained popularity [4]. These models often incorporate heavy-tailed errors to better capture financial return behavior, as shown in studies like Jacquier et al. [7] and Chib, Nardari, and Shephard [2], who modeled SV processes using Student-t distributed errors. Inspired by this, our model also uses a Student-t likelihood to increase robustness to outliers and extreme price movements.

Finally, the specification of informative, well-behaved priors remains a key challenge in Bayesian modeling, which approximates the posterior distribution as a product of the prior and the likelihood. Section 2.2 elaborates on our model design and the rationale behind our choice of priors.

2.2 Model Specification

We model the daily log-returns $r_{t,i}$ for stock $i \in \{1, \ldots, N\}$ on day $t \in \{1, \ldots, T\}$, where each stock belongs to one of S sectors. Let $s_i \in \{1, \ldots, S\}$ denote the sector of stock i. The model structure is as follows:

Volatility Process

$$
\begin{aligned}
h_{1,i} &= \frac{a_{0,i}}{1 - a_{1,i} - b_i}, \quad \text{(stationary init.)} \\
h_{t,i} &= a_{0,i} + a_{1,i} r_{t-1,i}^2 + b_i h_{t-1,i}, \quad \text{for } t = 2, \ldots, T.
\end{aligned}
$$

Observation Model

$$
\begin{aligned}
r_{1,i} &\sim \text{Student-}t(\nu_i, 0, \sqrt{h_{1,i}}), \\
r_{t,i} &\sim \text{Student-}t(\nu_i, \phi_i r_{t-1,i}, \sqrt{h_{t,i}}), \quad \text{for } t = 2, \ldots, T.
\end{aligned}
$$

Hierarchical Structure Let stock-level parameters $\theta_i = \{a_{0,i}, a_{1,i}, b_i, \phi_i, \nu_i\}$ be drawn from sector-level distributions, indexed by sector $s = s_i$:

$$
\begin{aligned}
a_{0,i} &\sim \text{Truncated Normal}^+(\mu_{a0}^{(s)}, \tau_{a0}^{(s)}), \\
a_{1,i} &\sim \mathcal{N}(\mu_{a1}^{(s)}, \tau_{a1}^{(s)}), \\
b_i &\sim \mathcal{N}(\mu_b^{(s)}, \tau_b^{(s)}), \\
\phi_i &\sim \mathcal{N}(\mu_\phi^{(s)}, \tau_\phi^{(s)}), \\
\nu_i &= 2 + \exp(\mu_\nu^{(s)} + \tau_\nu^{(s)} \cdot z_{\nu,i}), \quad z_{\nu,i} \sim \mathcal{N}(0, 1).
\end{aligned}
$$

Where:

Parameters on the unit interval (e.g., $a_{1,i}, b_i, \phi_i$) are modeled via inverse-logit transforms:

$$
\mu_{a1}^{(s)} = \text{logit}^{-1}(\mu_{a1,\text{logit}}^{(s)}), \quad \text{etc.}
$$

Priors Sector-level hyperpriors are specified as:

$$
\begin{aligned}
\mu_{a0}^{(s)} &\sim \mathcal{N}(0.001, 0.001), & \tau_{a0}^{(s)} &\sim \mathcal{N}^{+}(0, 0.005),\\
\mu_{a1,\text{logit}}^{(s)} &\sim \mathcal{N}(\text{logit}(0.05), 0.3), & \tau_{a1}^{(s)} &\sim \mathcal{N}^{+}(0, 0.05),\\
\mu_{b,\text{logit}}^{(s)} &\sim \mathcal{N}(\text{logit}(0.85), 0.3), & \tau_{b}^{(s)} &\sim \mathcal{N}^{+}(0, 0.025),\\
\mu_{\phi,\text{logit}}^{(s)} &\sim \mathcal{N}(\text{logit}(0.45), 0.5), & \tau_{\phi}^{(s)} &\sim \mathcal{N}^{+}(0, 0.05),\\
\mu_{\nu}^{(s)} &\sim \mathcal{N}(\log(18), 1), & \tau_{\nu}^{(s)} &\sim \mathcal{N}^{+}(0, 1).
\end{aligned}
$$

Soft Stationarity Prior To favor stationarity in the GARCH process, a soft prior penalizes the sum $a_{1,i} + b_i$:

$$a_{1,i} + b_i \sim \mathcal{N}(0.95, 0.02).$$

The exploratory plots allowed us to make our decisions. Plot 1 shows the daily log returns over the 5 years where volatility patterns across stocks are visible - which help us understand the scale of the baseline volatility. The autocorrelative nature up to lag-5 is significant in our ACF plots in Plots 3.1-3.6, informing AR[1]'s inclusion. Plot 2 shows log returns' bell-curved distributions, helping us set the priors for the degrees of freedom in the Student-t distribution. Plot 4 reinforces the presence of volatility clusters and our decision to use GARCH processes; logit-transformed normal priors are used to constraint the parameters to interval [0,1] for GARCH processes. Inclusion of a Soft Stationary Prior is to help keep the process stationary (mean and variance do not change) and penalizes if it deviates too far from the target value.

The hyperpriors are chosen to thoughtfully balance conservative assumptions about sector-level variation and still allow for flexibility to capture real-world financial data. Normal priors with small standard deviations, keeping sector-specific parameters close to the baseline values. Furthermore, truncation is used to ensure that the parameters remain physically meaningful.

3 Results

After fitting the Stan model using Hamiltonian Monte Carlo (HMC) with 3 chains and 2000 iterations each, we obtained a well-mixed posterior sample that demonstrated strong convergence diagnostics. We primarily focused on the volatility parameters: $\alpha_{0,i}$, $\alpha_{1,i}$, and β_i for each stock. Trace plots indicated good mixing behavior, with characteristic "wiggly" patterns suggesting no sticking or divergent transitions. Among these, $\mu_{\alpha_0,i}$ and $\tau_{\alpha_0,i}$ showed particularly stable behavior. Since these parameters influence the volatility dynamics, we conducted additional diagnostics. Pairs plots showed circular, uncorrelated blobs—evidence that the chains explored the posterior space efficiently. $\hat{R}$ values for all parameters were under 1.05, reinforcing convergence confidence. We also checked effective sample size (ESS), finding that most parameters exceeded 300. However, some parameters such as $\mu_{\alpha_0,i}$, $\tau_{\alpha_0,i}$, and a few $h_{t,i}$ had ESS values below 100, indicating that future runs might benefit from improved tuning.

We then summarized the sector-level hyperparameter posterior means:

$$\alpha_0 = 8.17 \times 10^{-6},$$

suggesting a low baseline variance.

$$\alpha_1 = 0.0463,$$

implying that past shocks have a small positive impact on current volatility.

$$\beta = 0.899,$$

indicating strong volatility persistence.

$$\phi = 0.117,$$

reflecting weak but positive autocorrelation in the daily log-returns.

We also analyzed individual stock behavior. Posterior means and 95% credible intervals for each stock's AR(1) parameter ϕ_i were computed (see Appendix for R code). To visualize volatility dynamics, we plotted the posterior mean of latent volatility over time (Plot 5). Volatility spikes were observed in 2020, 2022, and 2025. The 2020 spike aligns with the onset of the COVID-19 pandemic, with AAPL and XLK showing particularly high volatility. In 2022, META and NFLX led during the post-COVID market crash. The 2025 spike may reflect macroeconomic uncertainty or a new US tariff policy, again with AAPL and XLK showing pronounced effects. Autocorrelation in volatility was also observed, supported by ACF plots of the posterior volatilities.

Limitations The complexity of the model led to slow convergence, with chains taking 3-4 hours, which impacted debugging and resulted in a simpler model than intended. Low ESS for some variables suggests inefficient exploration of the posterior. The GARCH(1,1) model may underfit sharp volatility spikes, as seen in 2022. Due to the large dataset, we couldn't explore sector-specific cross-effects, limiting the use of more hierarchical modeling. As GARCH models and quantitative finance were new, gaps in knowledge may exist. Time constraints and heavy processing also restricted the depth of model exploration and diagnostics.

Conclusion Each stock's posterior showed different behaviour with AAPL and XLK being more volatile in 2020 and 2025; META and NFLX being more volatile in 2022 market crash and GOOGL and AMZN being stable throughout. However, overall sector-wise, they exhibited similar behavior which was what we expected. Depending on the investor's trading strategy, these insights can be utilized to aid them in managing their risks.

References

1. Bollerslev, T.: Generalized autoregressive conditional heteroskedasticity. Journal of Econometrics **31**(3), 307–327 (1986)
2. Chib, S., Nardari, F., Shephard, N.: Markov chain monte carlo methods for stochastic volatility models. Journal of Econometrics **108**(2), 281–316 (2002)
3. Engle, R.F.: Autoregressive conditional heteroscedasticity with estimates of the variance of united kingdom inflation. Econometrica **50**(4), 987–1007 (1982)
4. Ghysels, E., Harvey, A.C., Renault, E.: Stochastic volatility. In: Handbook of Financial Econometrics. Elsevier (2002)
5. Hansen, P.R., Lunde, A.: A forecast comparison of volatility models: Does anything beat a garch(1,1)? Journal of Applied Econometrics **20**(7), 873–889 (2005)
6. Jacquier, E., Polson, N.G., Rossi, P.E.: Bayesian analysis of stochastic volatility models. Journal of Business & Economic Statistics **12**(4), 371–389 (1994)
7. Jacquier, E., Polson, N.G., Rossi, P.E.: Bayesian analysis of stochastic volatility models with fat tails and correlated errors. Journal of Econometrics **122**(1), 185–212 (2004)
8. Poon, S.H., Granger, C.W.J.: Forecasting volatility in financial markets: A review. Journal of Economic Literature **41**(2), 478–539 (2003)
9. Shephard, N., Pitt, M.K.: Likelihood analysis of non-gaussian measurement time series. Biometrika **84**(3), 653–667 (1997)
10. Tsay, R.S.: Analysis of Financial Time Series. John Wiley & Sons (2002)

4 Appendix

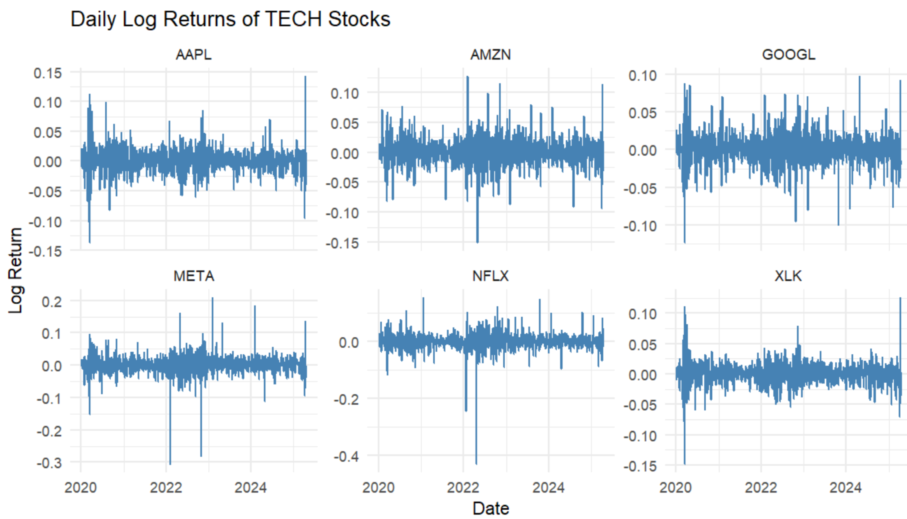

Fig. 1. Daily Log Returns for all Tech Stocks. AAPL and XLK suggest moderate returns with high volatility at ends. NFLX reports high average log returns with steepest volatility jump in 2022. META exhibits similar behavior with stability overall except the 2022 period. AMZN and GOOGL seem to experience the most average volatility, with AMZN reaching sharp peaks and lows in 2022, and GOOGL in 2020

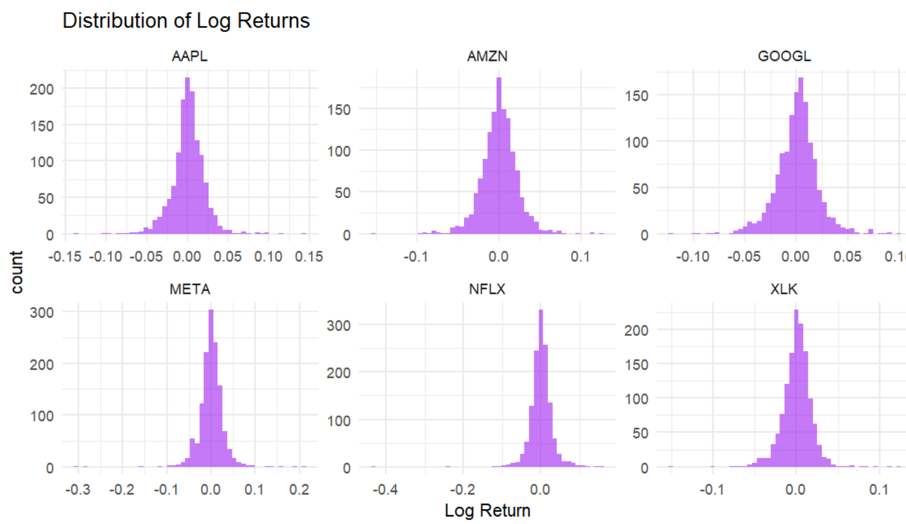

Fig. 2. Distributions of Log Returns for each our of tech stocks are presented. We observe high density bell-shaped curves with tails. NFLX and META seem to be shifted slightly to the right suggesting a left skew.

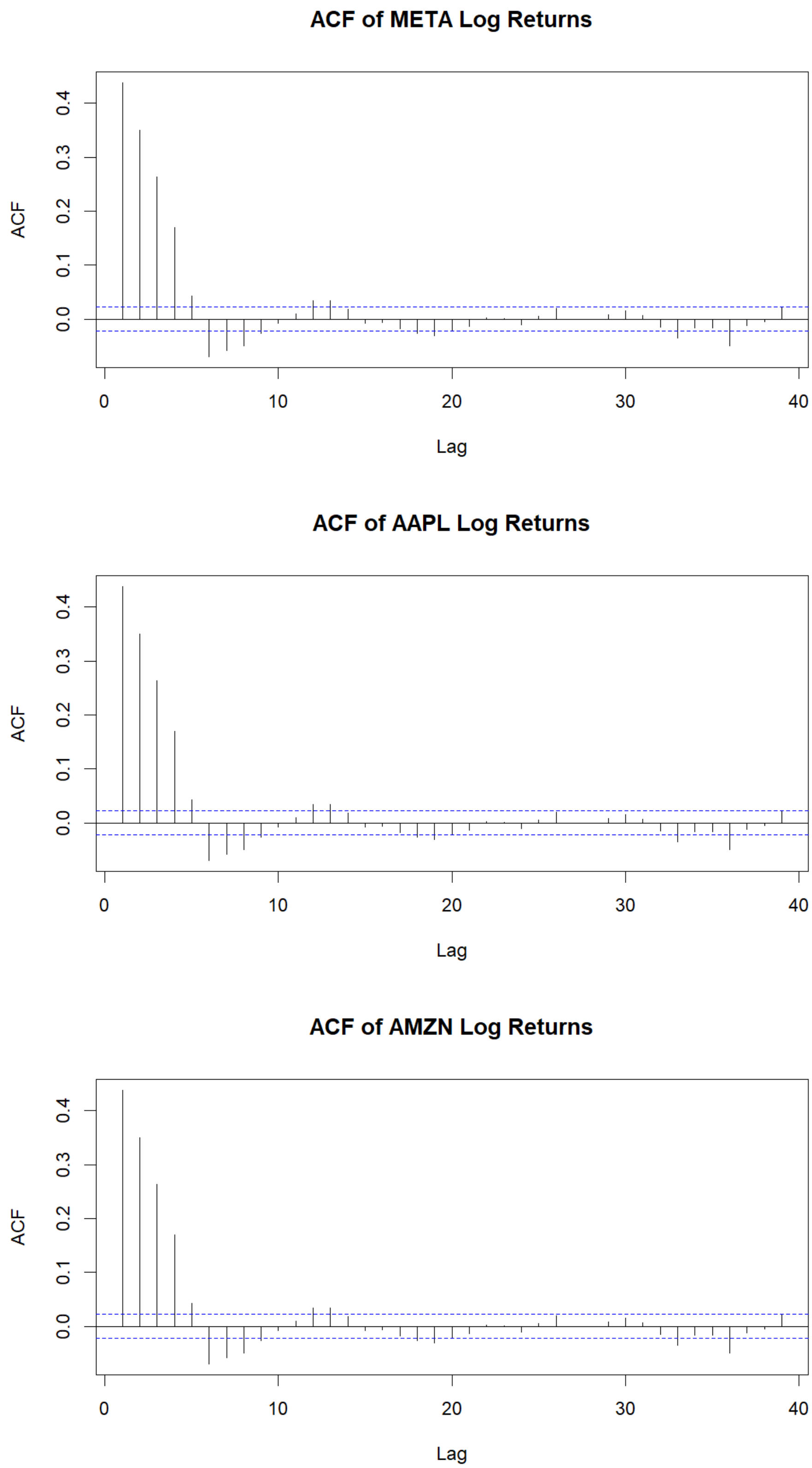

Fig. 3. Autocorrelation Function Plot for AAPL, AMZN, META showing lag up to 4 are significant with lingering effects till lag of 8. Other stocks mirrored the same plot, hence were excluded.

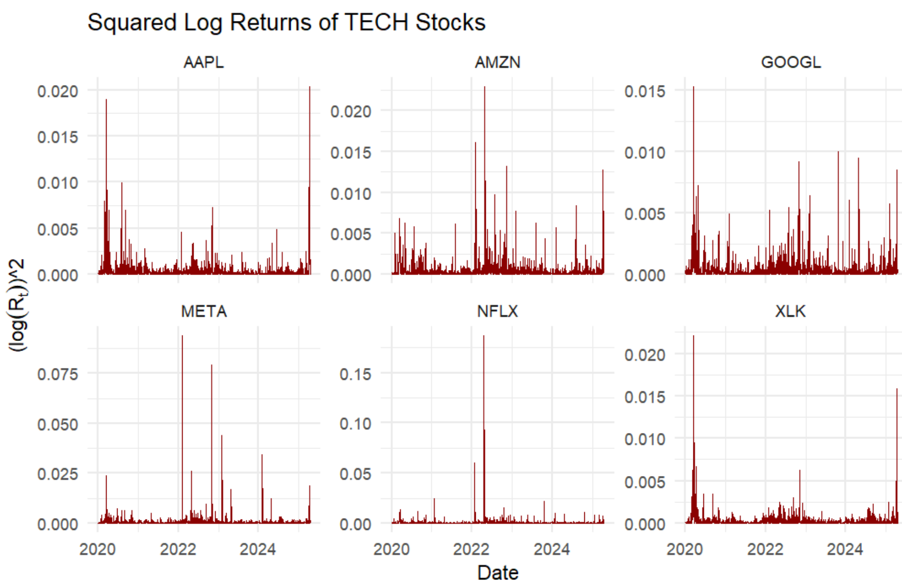

Fig. 4. Squared Log Returns of each of our tech stocks. Here we can see continued periods of volatility (volatility clusters) at 2022 especially in AMZN and GOOGL, and on a lesser extent in AAPL. META, NFLX seem to have sharper, more volatile peaks which could be due to macroeconomic factors. XLK seems to be the most stable with high volatility in 2020 and 2025; it seems that XLK was still stable through the 2022 Stock Market Crash.

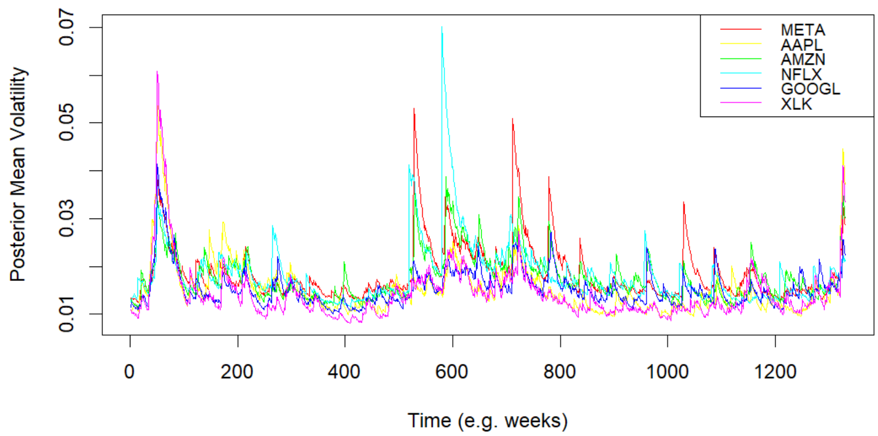

Fig. 5. Posterior Mean Volatility is plotted against time for each Stock after the model was run. These plots are good for visualizing each stock in the technology sector and how it performs against the whole curve. In the beginning in 2020, there is a sharp peak with the whole tech sector following but XLK and AAPL seem to be the most volatile. Around 2020 between weeks 500 to 800, there are sharp sudden peaks throughout suggesting volatility clustering with NFLX having the sharpest peak at 600 weeks. At the end of the plot in 2025, there seem to be increasing volatility suggesting unstable periods ahead for the technology sector.

```r
library(quantmod)
library(dplyr)
library(tidyr)
library(lubridate)
library(tidyverse)
library(ggplot2)
library(forecast)
library(rstan)
library(posterior)
library(tidybayes)

## the tech stacks we will be working with
tickers <- c("META", "AAPL", "AMZN", "NFLX", "GOOGL", "XLK")

## retrieving data from yahoo finance for this current decade
getSymbols(tickers, src = "yahoo", from = "2020-01-01", to = Sys.Date())

# extract close prices and work with them accordingly
prices <- do.call(merge, lapply(tickers, function(tk) Ad(get(tk))))
colnames(prices) <- tickers

# data wrangling, removing na's and converting it into the dataframe; obtaining
# log_return by using using lag function to obtain yesterday's price and
# calculating the appropriate log ratio; dropped NAs instead of imputing.
returns_long <- prices %>%
  na.omit() %>%
  as.data.frame() %>%
  rownames_to_column(var = "date") %>%
  mutate(date = as.Date(date)) %>%
  pivot_longer(-date, names_to = "ticker", values_to = "price") %>%
  group_by(ticker) %>%
```

Fig. 6. R Script. More available at: https://github.com/shahbazayaz/stat447/tree/main

```stan
data {
  int<lower=1>    T;                // number of days
  int<lower=1>    N;                // number of stocks
  int<lower=1>    S;                // number of sectors
  int<lower=1,upper=S> sector_id[N];
  matrix[T, N]    r;                // log-returns
}

parameters {
  //-- Sector-level hyperparameters ------------------------------
  vector<lower=0>[S]      mu_a0;              // baseline variance
  vector<lower=0>[S]      tau_a0;             // spread around mu_a0

  vector[S]               mu_a1_logit_raw;    // unconstrained for logit(a1)
  vector<lower=0>[S]      tau_a1;             // spread around mu_a1

  vector[S]               mu_b_logit_raw;     // unconstrained for logit(b)
  vector<lower=0>[S]      tau_b;              // spread around mu_b

  vector[S]               mu_phi_logit_raw;   // unconstrained for logit(phi)
  vector<lower=0>[S]      tau_phi;            // spread around mu_phi

  vector[S]               mu_nu_log_raw;      // unconstrained for log(nu-2)
  vector<lower=0>[S]      tau_nu;             // spread around mu_nu

  //-- Stock-level raw deviations --------------------------------
  vector[N]               a0_raw;
  vector[N]               a1_raw;
  vector[N]               b_raw;
  vector[N]               phi_raw;
```

Fig. 7. Stan Model. More available at: https://github.com/shahbazayaz/stat447/tree/main