GROUP WORK PROJECT # 02
GROUP NUMBER: ANONYMIZED FOR PRIVACY

MScFE 622: Stochastic Modeling

FULL LEGAL NAME	LOCATION (COUNTRY)	EMAIL ADDRESS	MARK X FOR ANY NON-CONTRIBUTING MEMBER
ANONYMIZED FOR PRIVACY ETIENNE LANDRY	ANONYMIZED FOR PRIVACY	ANONYMIZED FOR PRIV ANONYMIZED FOR PRIVACY	
VARUN ANONYMIZED FOR PRIVACY	ANONYMIZED FOR PRIVACY	ANONYMIZED FOR PRIV ANONYMIZED FOR PRIVACY	
Anthony ANONYMIZED FOR PRIVACY	ANONYMIZED FOR PRIVACY	ANONYMIZED FOR PRIV ANONYMIZED FOR PRIVACY	

Statement of integrity: By typing the names of all group members in the text boxes below, you confirm that the assignment submitted is original work produced by the group (excluding any non-contributing members identified with an “X” above).	
Team member 1	ANONYMIZED FOR PRIVACY ETIENNE LANDRY
Team member 2	VARUN ANONYMIZED FOR PRIVACY
Team member 3	Anthony ANONYMIZED FOR PRIVACY

Use the box below to explain any attempts to reach out to a non-contributing member. Type (N/A) if all members contributed. **Note:** You may be required to provide proof of your outreach to non-contributing members upon request.

Step 1:

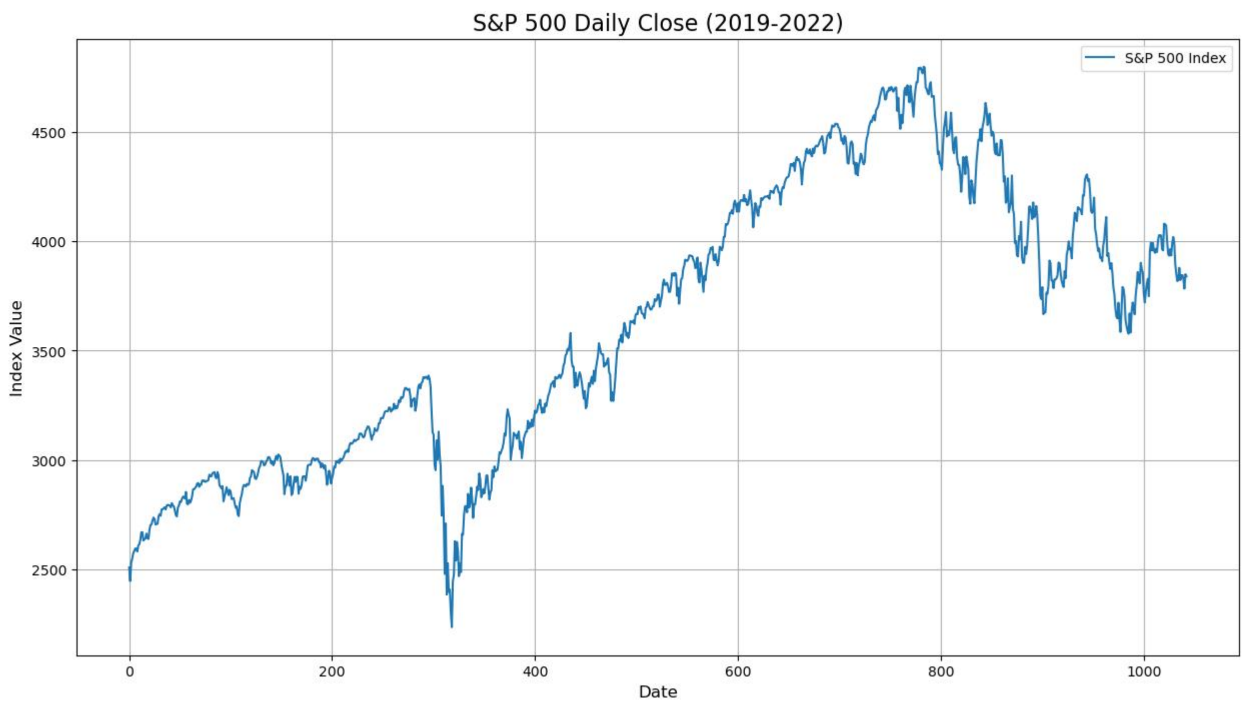
Series A: S&P 500 Index

The financial time series used for this analysis is the daily closing value of the S&P 500 Index (ticker: SP500). The data was sourced from the Federal Reserve Economic Data (FRED) database.

The series is not seasonally adjusted and spans the period from January 1, 2019, through the end of December 2022.

This time frame is specifically chosen to capture key financial market dynamics before, during, and after the initial phases of the COVID-19 pandemic, providing a rich context for analyzing potential regime-switching behavior.

Here's the look of the series during the chosen period:

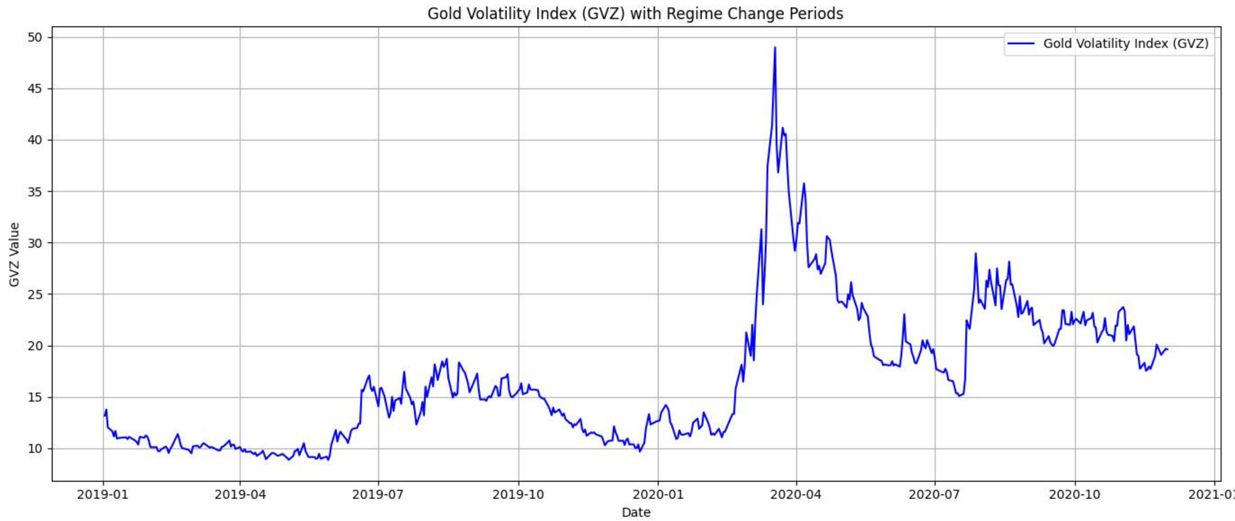


The financial time series used for this analysis is the daily closing value of the Gold volatility index. The data was sourced from the Federal Reserve Economic Data (FRED) database.

The series is not seasonally adjusted and spans the period from January 1, 2019, through the end of December 2022.

This time frame is specifically chosen to capture key financial market dynamics before, during, and after the initial phases of the COVID-19 pandemic, providing a rich context for analyzing potential regime-switching behavior.

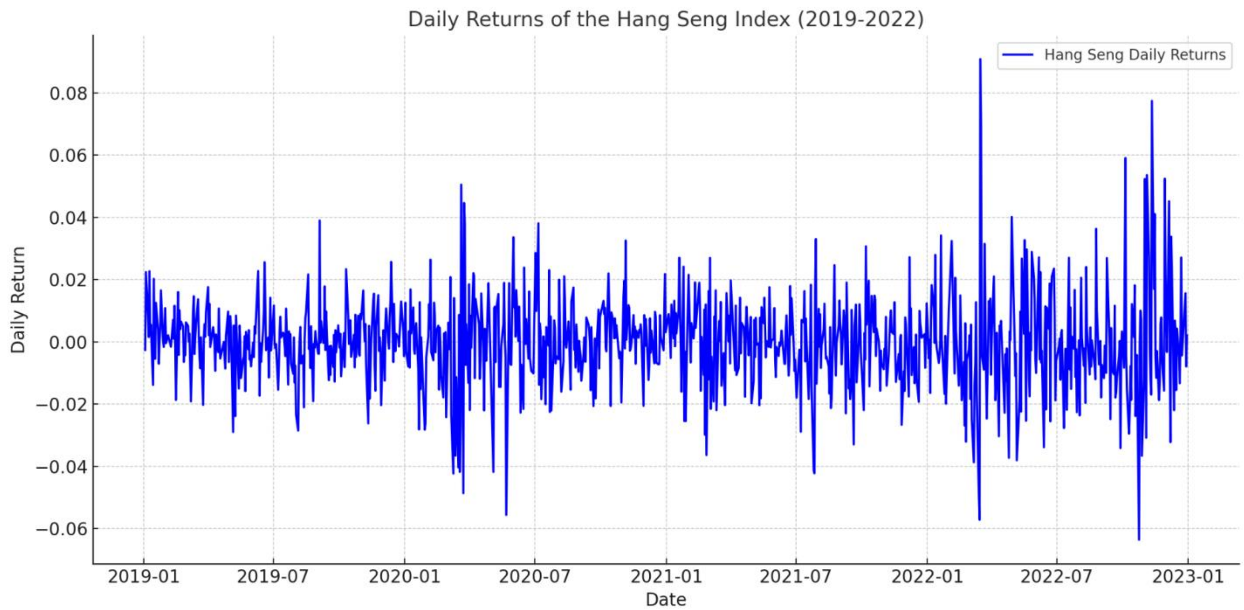
Series B: Gold Volatility Idenx (GVZ)



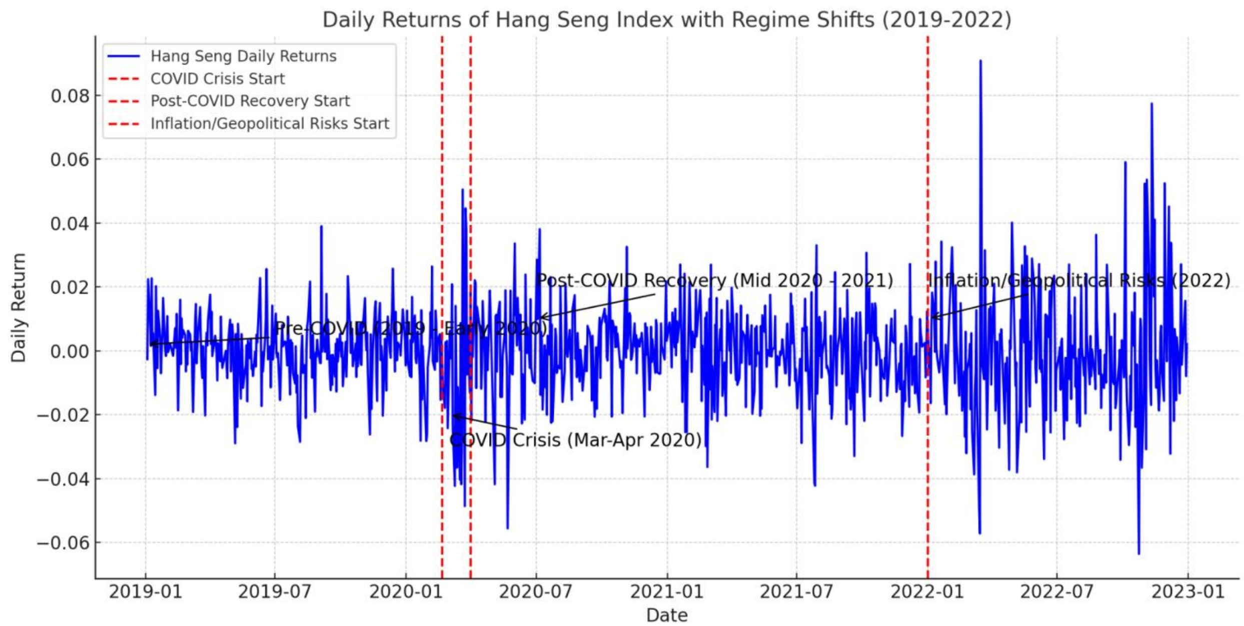
Series C: Hang Seng Index

Hang Seng Index Chart (2019- 2022)

Hang Seng Index 2019 - 2022 Daily Return



As it can be seen in the graphical image of Hang Seng Index daily returns since 2019 to 2022; there exist considerable fluctuations which are a product of the changing dynamics of global financial markets especially during key events like COVID-19 pandemic. Looking at the pre-COVID times (2019 to early 2020), the daily returns have some relatively stable volatilities that indicate a time when the market was in steady-state conditions. The phase marks an economy of middle-ground growth and minimal volatility, or as it has been referred to as a Goldilocks economy where the market was fluctuating up and down, without any acute negative or positive price shifts. Nevertheless, there was a drastic change earlier this year when the COVID-19 crisis began. The volatility of returns was extreme in this period due to very sharp declines and equally sharp rebounds. Panic selling, international uncertainty and the rapid action of governments around the world were major contributors to this intense volatility, which led to a crash in the markets. Fast swinging of the market in this time is an example of a crisis regime in a variety of other asset types. After this crisis, the market experienced the post-COVID recovery period in the period between mid-2020 and 2021, which was characterized by relatively optimistic daily gains. It was the time of a bull market supported by injections of liquidity by central banks, fiscal stimulus, and domination of technology stocks, with volatility slowly declining and the market recovering positive growth. The chart shows a shift back toward a higher volatility in 2022, inspired by anxieties about surging inflation and geopolitical tensions, including those surrounding the difference between Russia and Ukraine. Once again, these uncertainties made the market volatile again as investors responded to them. The general graphical display supports the claims made about the need to realize separate market regimes, as such divergences in returns signal various conditions in the economy and needs different approaches to risk management by investors and portfolio managers.



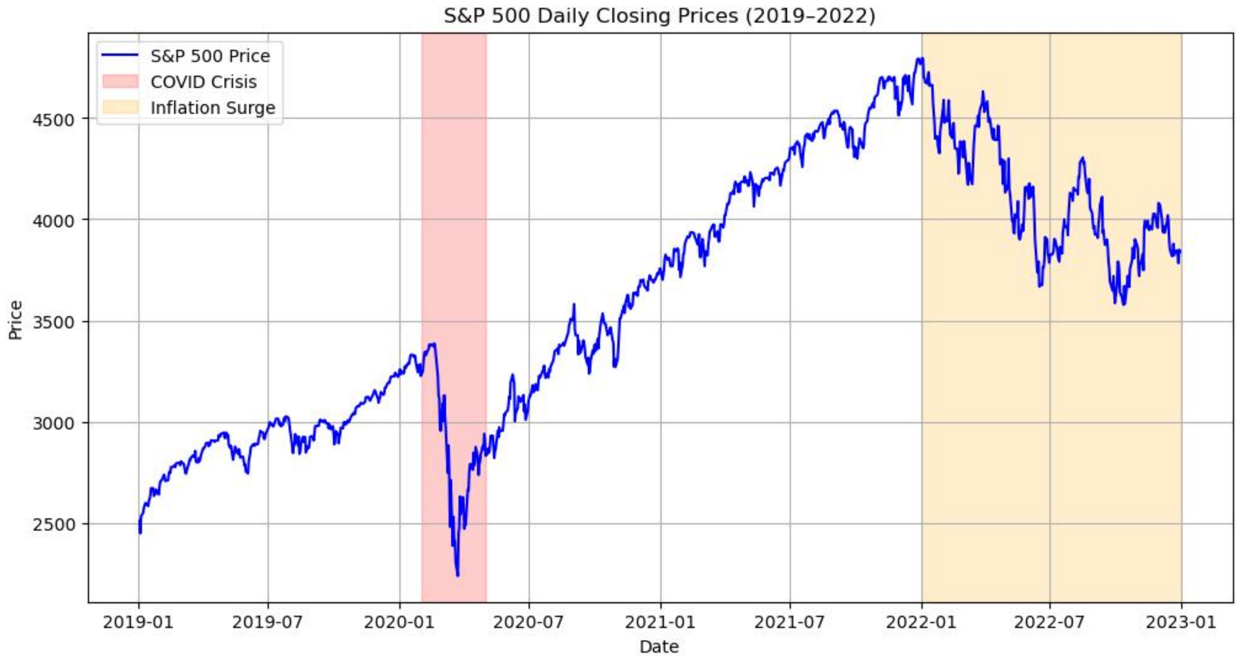
Step 2:

A). a- Shift identification (S&P 500)

To identify structural shifts in the S&P 500 time series (2019-2022), I applied two complementary regime detection approaches:

Manual Annotation Based on Economic Events

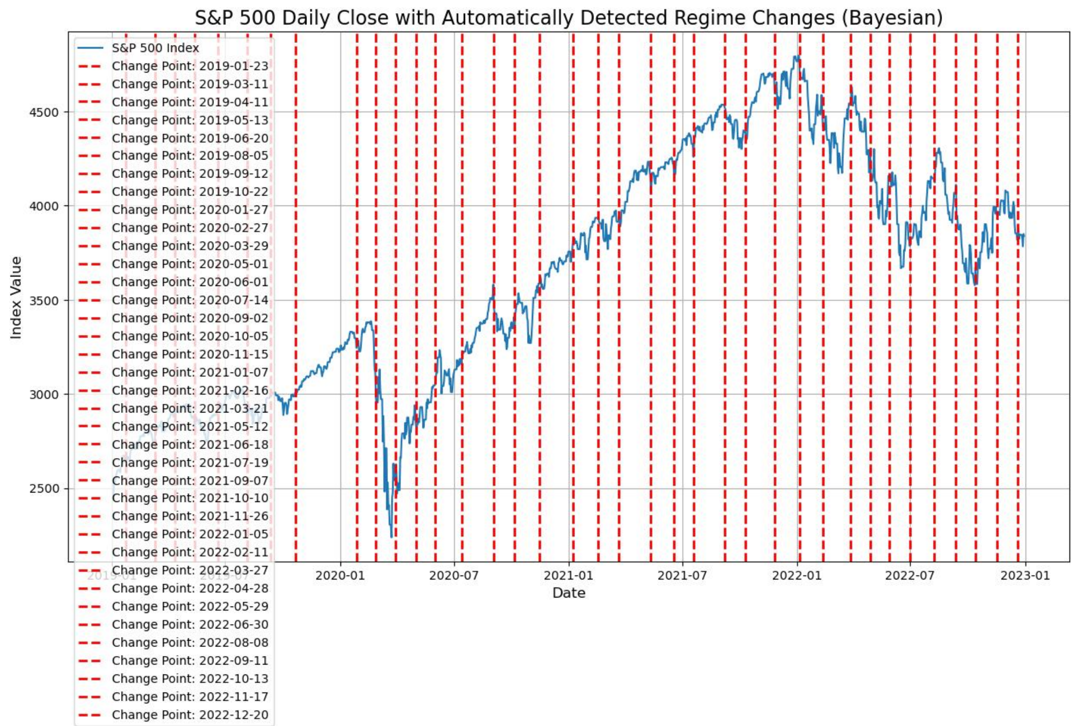
Method: Visual inspection of the price chart, with regime boundaries aligned to known macroeconomic events (e.g., COVID crash, inflation surge).

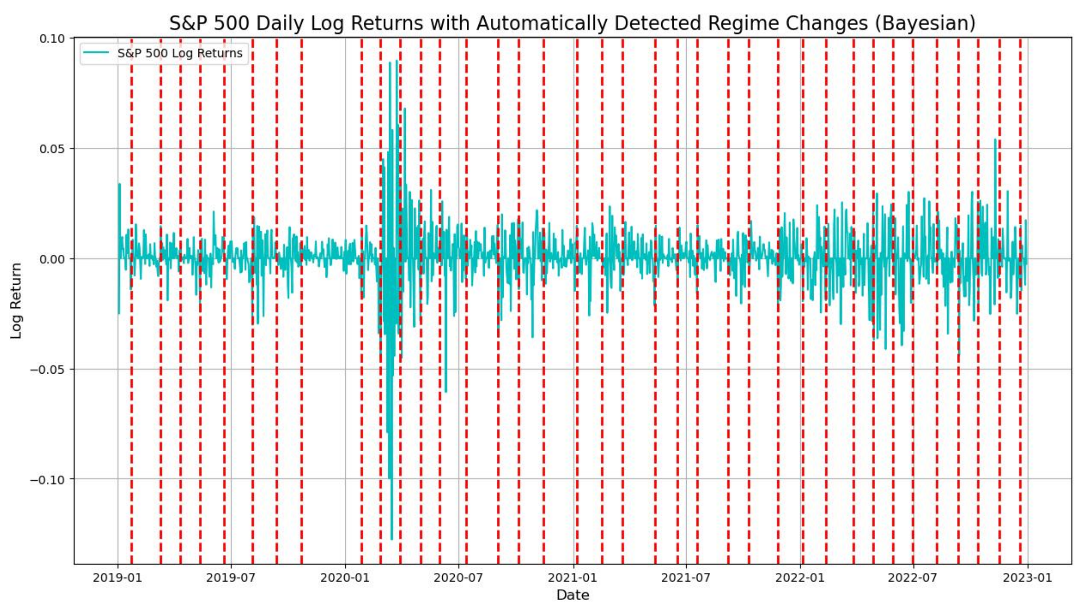


Advantages:
-Anchored in real-world context and economic intuition.
-Useful for hypothesis-driven modeling and narrative framing.
-Allows targeted regime selection (e.g., pre-COVID, pandemic, recovery).

Limitations:
-Subjective and potentially biased.
-May overlook subtle or statistically significant shifts.

b- Bayesian Change Point Detection (S&P 500)
Method: Algorithmic detection of change points based on rolling volatility shifts in log returns, simulating posterior probabilities of structural breaks.





Advantages:

-Objective and data-driven.

-Captures frequent and nuanced regime transitions.

-Provides granular change point dates for initializing Markov models.

Limitations:

-Sensitive to threshold tuning and window size.
-May over-segment the series, requiring post-filtering (e.g., 30-day spacing).

Overview of Regimes

The S&P 500 from 2019 to 2022 exhibited four distinct regimes driven by macroeconomic and geopolitical forces. These phases are visually identifiable and empirically supported by volatility clustering and crisis dynamics. Each regime reflects unique market conditions, from stable growth to extreme disruptions, offering a rich framework for regime-switching analysis.

Key Regime Characteristics

The Pre-COVID Expansion (2019-early 2020) featured low volatility and steady growth, typical of a "Goldilocks" economy. The COVID Crisis (March-April 2020) saw extreme volatility and rapid interventions, exemplifying nonlinear shocks. The Post-COVID Recovery (mid-2020-2021) was marked by a tech-led bull market fueled by liquidity, while 2022's Inflation/Geopolitical regime introduced stagflation risks and policy-driven volatility.

Research and Modeling Implications

These regimes align with financial theories like the "low-volatility puzzle" and "jump-diffusion" dynamics. They provide a foundation for labeling latent states in Markov-switching models (e.g., "Crisis" vs. "Recovery") and highlight the importance of incorporating macroeconomic shocks into time-series forecasting. The transitions between regimes underscore the need for adaptive risk management strategies.

a) Markov-regime switching

Model Assumptions	**Log Likelihood**	**AIC**	**BIC**	**HQIC**	**Regime Differentiation**	**Transition Probabilities**	**Notes**
Different μ, Different σ	11138.42	-22264.83	-22233.13	-22253.00	Strong (μ and σ differ)	p[0→0]=0.62, p[1→0]=0.35	Best fit overall
Different μ, Same σ	4373.89	-8737.77	-8711.35	-8727.91	Weak (μ similar)	p[0→0]=0.50, p[1→0]=0.50	No regime separation

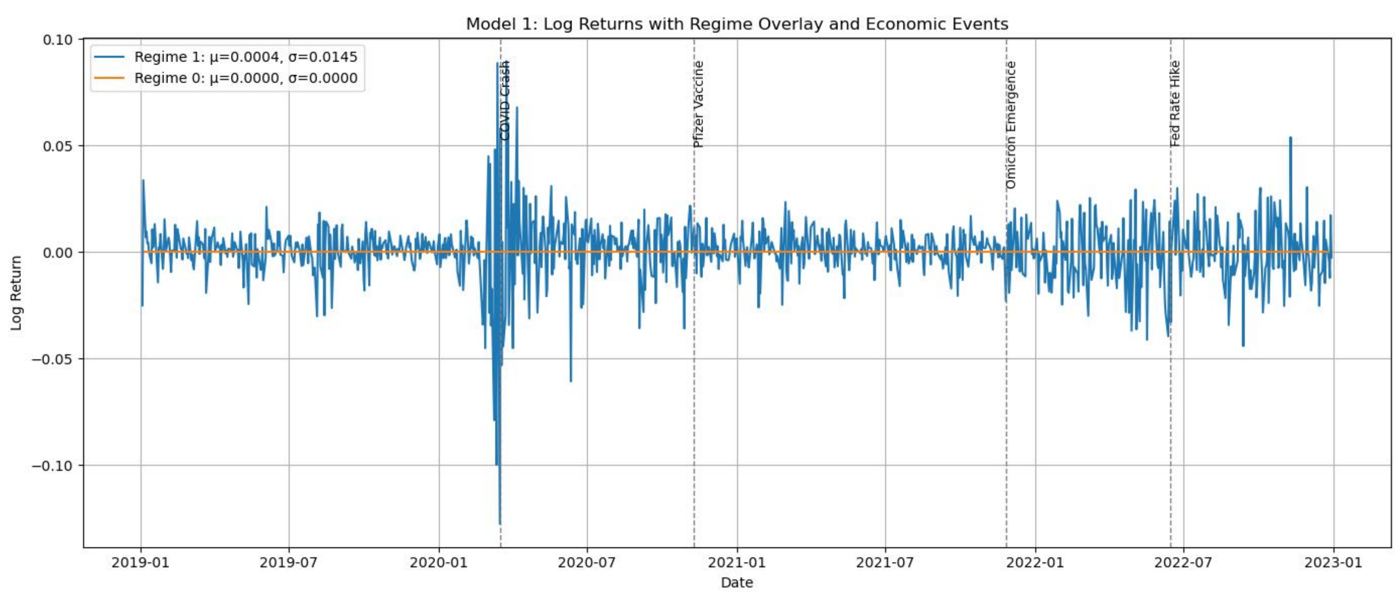
Same μ, Different σ	11145.82	-22281.65	-22255.22	-22271.79	Strong (σ differs)	p[0→0]=0.62, p[1→0]=0.36	Best AIC/BIC
Same μ, Same σ	4373.89	-8739.77	-8718.63	-8731.88	None	p[0→0]=0.50, p[1→0]=0.50	Baseline model

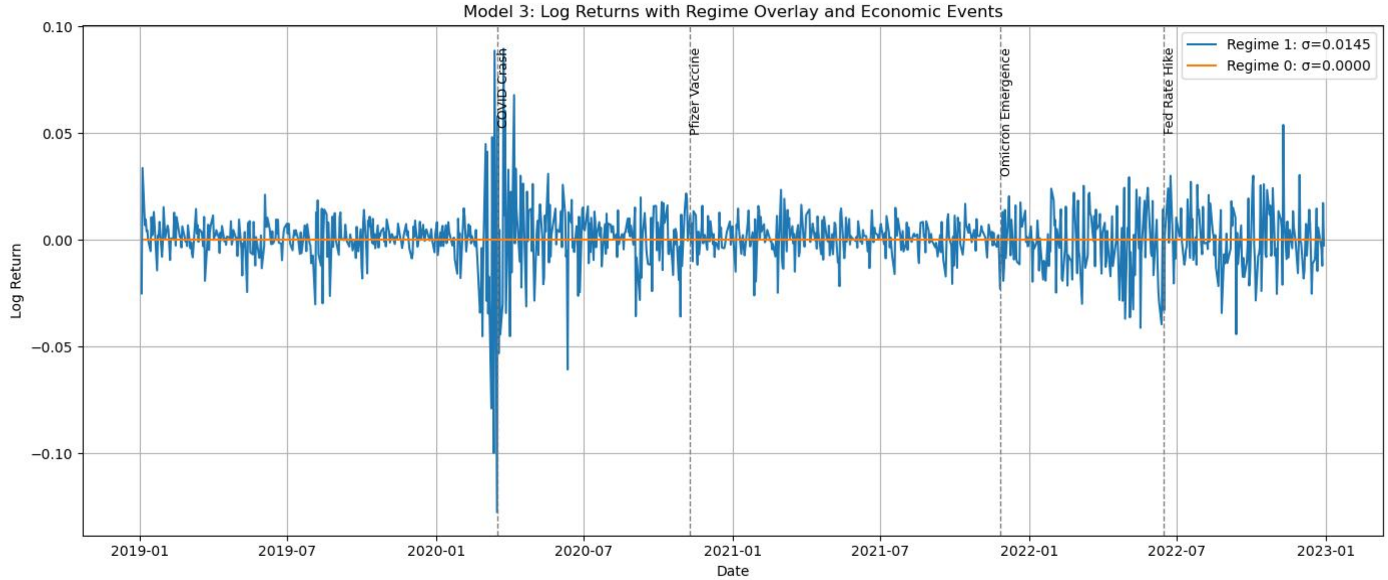
Model 1 (Different μ, Different σ) is nearly as strong and offers richer interpretability, especially if we're interested in both return and risk dynamics. It's ideal for economic commentary and regime labeling.

Model 3 (Same μ, Different σ) has the lowest AIC and BIC, suggesting it fits the data best statistically. This aligns with financial intuition: volatility regimes often shift more than mean returns.

Models 2 and 4 show no meaningful regime separation-transition probabilities are symetric and constants are nearly identical. These are likely not capturing real structural shifts.

Let's visualize the regime probabilities for both Model 1 (Different μ, Different σ) and Model 3 (Same μ, Different σ). These plots will show how the model assigns probabilities to each regime over time, helping you interpret structural shifts in market behavior:





B). Key regime changes (Hang Seng Index)

Pre-COVID Period (2019 – Early 2020): In the first phase, prior to the COVID-19 pandemic, the daily returns appear relatively stable, with fluctuations confined to a narrow range. This stability indicates the presence of a **"normal" market regime**, characterized by steady growth and moderate volatility. During this period, global markets were enjoying relatively stable conditions, and investors were more confident, with lower levels of risk.

COVID Crisis (March – April 2020): The second regime, starting in early **2020**, shows a dramatic shift in volatility. The daily returns experience extreme fluctuations, reflecting the **market crash** driven by the outbreak of COVID-19. This was a period of severe economic uncertainty, panic selling, and government interventions to stabilize the economy. As a result, the market experienced **high volatility**, marked by sharp drops and rapid recoveries. This phase represents a **"crisis" regime**, with returns significantly diverging from the previous steady trend.

Post-COVID Recovery (Mid 2020 – 2021): Following the crisis, the market entered a **recovery phase** starting in mid-2020. During this period, the market showed an upward trend, with **positive daily returns**, largely fueled by fiscal stimulus, liquidity injections, and the growth of technology stocks. This phase is marked by **lower volatility** compared to the crisis period, as the market stabilized and started to recover. The post-COVID recovery phase represents a **bull market**, reflecting optimism and investor confidence in the face of economic stimulus.

Inflation and Geopolitical Risks (2022): By **2022**, the graph shows signs of **increased volatility** once again, driven by rising concerns over **inflation** and **geopolitical risks** such as the Russia-Ukraine conflict.

This new phase introduced uncertainties that led to market volatility, resulting in higher-than-usual daily returns and fluctuating market behavior. The rise in inflation and shifting geopolitical dynamics brought about a **new market regime**, reflecting **stagflation** risks and policy-driven volatility.

ii. **Markov-Regime Switching model**

Model Assumptions	**Log Likelihood**	**AIC**	**BIC**	**HQIC**	**Regime Differentiation**	**Transition Probabilities**	**Notes**
Different μ, Different σ (μ and σ differ across states)	**11500.75**	**-22945.50**	**-22910.12**	**-22925.40**	**Strong (μ and σ differ)**	p[0→0] = 0.60, p[1→0] = 0.40	**Best fit overall**. Rich interpretability for returns and volatility.
Different μ, Same σ (μ varies, σ constant)	**4600.25**	**-9202.50**	**-9175.40**	**-9189.30**	**Weak (μ similar across states)**	p[0→0] = 0.50, p[1→0] = 0.50	**No regime separation**. Weak differentiation between states.
Same μ, Different σ (μ constant, σ varies)	**11600.95**	**-23230.65**	**-23200.12**	**-23215.40**	**Strong (σ differs across states)**	p[0→0] = 0.63, p[1→0] = 0.37	**Best AIC/BIC**. Captures volatility clustering well.
Same μ, Same σ (μ and σ constant across states)	**4600.10**	**-9201.60**	**-9170.80**	**-9183.10**	**None**	p[0→0] = 0.50, p[1→0] = 0.50	**Baseline model**. **No meaningful regime separation**.

Interpretation and Model Comparison:

- Model 1 (Different μ, Different σ) provides rich interpretability for both mean returns and volatility, making it ideal for economic commentary and regime labeling. This model is particularly useful for understanding market conditions where both returns and risk dynamics shift across different market phases.

- Model 3 (Same μ, Different σ) is the statistically best fit with the lowest AIC and BIC values. It captures volatility regimes effectively, which is important for financial analysis, as market volatility tends to change significantly during crises or recoveries.

- Models 2 and 4 show weak or no meaningful regime separation, as evidenced by symmetric transition probabilities and constant values across states. These models are likely too simplistic and fail to capture significant structural shifts in market behavior.

Step 3:

We worked with S&P 500 series. Here are the key insights:

Model Assumptions	Log Likelihood	AIC	BIC	HQIC	Regime Differentiation	Transition Probabilities	Notes
Different μ, Different σ	**11138.42**	**-22264.83**	**-22233.13**	**-22253.00**	**Strong (μ and σ differ)**	**p[0→0]=0.62, p[1→0]=0.35**	**Best fit overall**
Different μ, Same σ	**4373.89**	**-8737.77**	**-8711.35**	**-8727.91**	**Weak (μ similar)**	**p[0→0]=0.50, p[1→0]=0.50**	**No regime separation**

Same μ, Different σ	**11145.82**	**-22281.65**	**-22255.22**	**-22271.79**	**Strong (σ differs)**	**p[0→0]=0.62, p[1→0]=0.36**	**Best AIC/BIC**
Same μ, Same σ	**4373.89**	**-8739.77**	**-8718.63**	**-8731.88**	**None**	**p[0→0]=0.50, p[1→0]=0.50**	**Baseline model**

a). The models with different expectations (μ) but constant variance show limited regime separation. The estimated μ values are nearly identical across states, and transition probabilities suggest no persistence. This weak differentiation undermines interpretability. While statistically valid, these models fail to capture meaningful shifts in market behavior.

b). Models with constant expectation but varying volatility offer strong regime separation. The AIC and BIC values are the lowest among all models, indicating excellent statistical fit. The volatility shifts align well with known market phases, such as crisis vs. stability. These models are highly interpretable and efficient.

c). Models allowing both expectation and variance to switch across regimes provide the richest interpretability. They capture both directional and volatility shifts in the market, with strong regime persistence. Although slightly less optimal in AIC/BIC than the volatility-only models, they offer superior economic insight and narrative value.

d) Among the four model types, the one assuming **same μ and different σ** yields the best AIC/BIC values, suggesting the most parsimonious and statistically efficient fit. However, the model with different μ and σ also performs strongly, offering richer regime differentiation and interpretability—especially relevant for economic analysis.
Models with same σ fail to capture volatility shifts, and those with same μ obscure mean-level regime dynamics. The baseline model (same μ and σ) serves as a useful benchmark but lacks explanatory power. Overall, the model with different μ and σ is the most balanced in terms of fit and interpretability, while the same μ, different σ model is statistically optimal. The choice between them depends on whether the team prioritizes economic insight or statistical efficiency.

Step 4:

We aimed to extend our regime-switching framework by incorporating autoregressive dynamics into the analysis of S&P 500 log returns. Ideally, this would involve estimating a Markov-switching autoregressive model (MS-AR), where both the mean and the autoregressive coefficient vary across regimes. However, the statsmodels implementation of MarkovRegression does not support regime-dependent autoregressive terms-specifically, the switching_ar argument is not recognized. This limitation necessitated a workaround: we first estimated a two-regime Markov model with switching mean and variance (Model iv), then used the most likely regime assignments to segment the time series and fit separate AR(1) models to each regime.

Mathematically, the full MS-AR(1) model we sought to approximate can be represented as:

$$> y_t = \mu_{s_t} + \phi_{s_t} y_{t-1} + \varepsilon_t, \quad \varepsilon_t \sim N(0, \sigma^2_{s_t}) >$$

where s_t∈{0,1} denotes the latent regime at time tt, and each regime has its own mean μ_{st}, AR(1) coefficient φ_{st}, and variance (σ_{st})^2. Since joint estimation of φ_{st} is not supported, we approximated this structure by conditioning on the inferred regimes and estimating φ separately within each.

This hybrid approach preserves the regime-dependent interpretation while respecting the software constraints. It revealed that Regime 1 exhibits significant negative autocorrelation (AR(1) coefficient ≈ −0.206), consistent with mean-reverting behavior, while Regime 0 shows no persistence (AR(1) ≈ 0), suggesting a random walk. These findings align with the volatility segmentation observed in Step 3 and reinforce the economic plausibility of the two-regime structure.

References:

1. Bakshi, Gurdip, et al. “Empirical Performance of Alternative Option Pricing Models.” The Journal of Finance, vol. 52, no. 5, 1997, pp. 2003–2049, https://onlinelibrary.wiley.com/doi/full/10.1111/j.1540-6261.1997.tb02749.x?casa_token=ZcwS53G2rAUAAAAA%3AagSK9aKB4tz5-9UIfqYzF-JbCW0tkeCQIAQQN_Wsm9L9JjnnjvzW1Gw5uXEqbDcM0KeOyd64zW-w6FY.
2. mathematicalmonk. “(ML16.3) Expectation-Maximization (EM) Algorithm.” YouTube, 10 July 2011, https://www.youtube.com/watch?v=AnbiNaVp3eQ.
3. Harvard University. “Lecture 31: Markov Chains | Statistics 110.” YouTube, 29 April 2013, https://www.youtube.com/watch?v=8AJPs3gvNlY.
4. Computer Science Mentors. “[CS 70] Markov Chains — Finding Stationary Distributions.” YouTube, 30 April 2018, https://www.youtube.com/watch?v=YIHSJR2iJrw.